



**HAL**  
open science

# Apprentissage génératif pour la synthèse d'images médicales dynamiques 4D

Yi-Heng Cao

► **To cite this version:**

Yi-Heng Cao. Apprentissage génératif pour la synthèse d'images médicales dynamiques 4D. Imagerie médicale. Université de Bretagne occidentale - Brest, 2024. Français. NNT : 2024BRES0004 . tel-04632074

**HAL Id: tel-04632074**

**<https://theses.hal.science/tel-04632074>**

Submitted on 2 Jul 2024

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

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

# THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

ÉCOLE DOCTORALE N° 637

*Sciences de la Vie et de la Santé*

Spécialité : *Analyse et Traitement de l'Information et des Images Médicales*

Par

**Yi-Heng CAO**

## Apprentissage génératif pour la synthèse d'images médicales dynamiques 4D

Thèse présentée et soutenue à Brest, le 08/02/2024

Unité de recherche : Laboratoire de Traitement de l'Information Médicale – UMR 1101

### Rapporteurs avant soutenance :

David SARRUT Directeur de recherche CNRS, CREATIS  
Vincent NOBLET Ingénieur de recherche CNRS HDR, ICube, Université de Strasbourg

### Composition du Jury :

Président :	Ulrike SCHICK	Professeur des Universités Praticien Hospitalier, LaTIM, CHRU Brest
Examineurs :	Ulrike SCHICK	Professeur des Universités Praticien Hospitalier, LaTIM, CHRU Brest
	David SARRUT	Directeur de recherche CNRS, CREATIS
	Vincent NOBLET	Ingénieur de recherche CNRS HDR, ICube, Université de Strasbourg
	Hadi FAYAD	Assistant Professor of Research, Weill Cornell Medicine
	Dimitris VISVIKIS	Directeur de recherche INSERM, LaTIM, UBO
Dir. de thèse :	Vincent JAOUEN	Maitre de conférences, LaTIM, IMT Atlantique
	Dimitris VISVIKIS	Directeur de recherche INSERM, LaTIM, UBO
Co-dir. de thèse :	Vincent JAOUEN	Maitre de conférences, LaTIM, IMT Atlantique



# REMERCIEMENTS

---

Je tiens tout d'abord à remercier mon directeur de thèse, Dimitris Visvikis, et mon co-encadrant, Vincent Jaouen, pour m'avoir donné l'opportunité de travailler sur cette thèse au sein du LaTIM.

Je tiens à exprimer ma reconnaissance envers mes rapporteurs, David Sarrut et Vincent Noblet, pour le temps qu'ils ont consacré à lire et évaluer mon manuscrit en pleine période de fêtes de fin d'année et en un temps très court. Je remercie également les membres du jury qui ont aussi pris le temps de venir évaluer mon travail, Ulrike Schick et Hadi Fayad.

Je remercie également mes collègues du laboratoire, Guillaume, Clément, Aziliz, Sundar, Mateo, Gustavo, Baptiste, Wistan, Noël, Thore, Fang, Zhihan, Hui, Yihao, Jing et Philippe pour leur amitié, leur soutien moral et leur bonne humeur, ainsi que toutes les sorties restaurants et buffets à volonté asiatiques, qui m'ont permis de me détendre et de prendre un peu de recul sur ma thèse (et du poids).

Je suis également reconnaissant envers Jean-François Clément, Isabelle Sanchez et Christelle Besnier pour leurs aides précieuses et leur disponibilité.

Je remercie avec amour ma femme, Xiaoyun Huang, pour m'avoir épaulé durant ces années de thèse, et qui m'a donné le courage de faire de mon mieux aussi bien à distance qu'en personne. #COVID-19

Je souhaiterais aussi ajouter des remerciements pour les cas suivants :

- À nouveau Vincent Jaouen, pour avoir bien voulu me répêcher auprès de Pierre-Henri et d'avoir eu foi en moi pour travailler sur ce sujet. Pour son soutien indéfectible tout au long de cette aventure, son exigence, sa rigueur et sa bienveillance.
- À Pierre-Henri Conze pour avoir donc transmis mon dossier de candidature qui était initialement destiné à son sujet de thèse sur le foie. Sans quoi j'aurais très probablement suivi une autre voie que celle actuelle.
- À la ville de Brest, pour sa météo plus clémente que ce que mon entourage a pu dire avant mon arrivée, pour avoir pu me marier au sein du magnifique hôtel de ville de Brest, ainsi que pour ses différentes ressources naturelles (chataignes, champignons, pommes, ail des ours).
- À la team E126, Guillaume, Gustavo et Amine, pour les bons moments passés à



---

discuter sur le sujet de l'IA bien sûr mais aussi sur des bêtises en tout genre. Sans oublier Clément pour ses bonnes blagues et Aziliz pour son dynamisme et sa joie. Enfin, je remercie enfin mes parents, ma sœur, mes cousins et mes amis pour tous les moments de joie passés ensemble, ainsi que toutes les personnes qui ont contribué à ma réussite.

# TABLE DES MATIÈRES

---

<b>Remerciements</b>	<b>3</b>
<b>Liste des figures</b>	<b>9</b>
<b>Liste des tableaux</b>	<b>12</b>
<b>Liste des sigles et acronymes</b>	<b>13</b>
<b>Introduction</b>	<b>15</b>
Motivation . . . . .	15
Aperçu de la thèse . . . . .	17
Organisation de la thèse . . . . .	17
<b>1 Tomodensitométrie quadridimensionnelle pour la planification de la radiothérapie</b>	<b>19</b>
1.1 Radiothérapie pour le cancer du poumon . . . . .	19
1.1.1 Généralités sur le cancer et la radiothérapie . . . . .	19
1.1.2 Définition des volumes en radiothérapie . . . . .	20
1.2 Principe de la tomodensitométrie quadridimensionnelle . . . . .	22
1.2.1 Introduction à la tomodensitométrie . . . . .	23
1.2.2 Composante temporelle . . . . .	25
1.3 Limitations de la modalité . . . . .	31
1.3.1 Protocole lourd . . . . .	31
1.3.2 Durée d'exposition . . . . .	32
1.3.3 Variabilité des acquisitions . . . . .	32
1.4 Discussion . . . . .	33
1.5 Conclusion . . . . .	34
<b>2 Apprentissage profond pour l'imagerie TDM thoracique en radiothérapie</b>	<b>35</b>
2.1 Introduction . . . . .	35

2.1.1	Synthèse d'image à image . . . . .	35
2.1.2	Domaine d'application de l'apprentissage profond en radiothérapie .	39
2.2	Correction de mouvement . . . . .	43
2.2.1	Introduction au recalage d'images médicales . . . . .	44
2.2.2	Recalage par deep learning non supervisé . . . . .	47
2.2.3	Recalage par deep learning supervisé . . . . .	49
2.3	Estimation de mouvement . . . . .	51
2.3.1	Introduction à la synthèse de vidéo . . . . .	52
2.3.2	Exemple en imagerie médicale . . . . .	53
2.4	Conclusion . . . . .	54
<b>3</b>	<b>Génération d'images dynamiques : synthèse directe par conditionnement respiratoire scalaire</b>	<b>57</b>
3.1	Introduction . . . . .	58
3.2	Méthode de synthèse directe . . . . .	59
3.2.1	Conditionnement additionnel par injection et fonction de coût focalisée aux zones d'intérêt . . . . .	59
3.2.2	Contrôle explicite du recalage . . . . .	63
3.3	Protocole expérimental . . . . .	65
3.3.1	Jeux de données . . . . .	65
3.3.2	Implémentations . . . . .	67
3.4	Résultats . . . . .	69
3.4.1	Images synthétiques . . . . .	69
3.4.2	Images cliniques . . . . .	72
3.5	Discussion . . . . .	74
3.6	Conclusion . . . . .	75
<b>4</b>	<b>Synthèse indirecte par apprentissage conjointement supervisé et adversaire</b>	<b>77</b>
4.1	Introduction . . . . .	78
4.2	Méthode . . . . .	79
4.2.1	Formulation du problème et notation . . . . .	81
4.2.2	Fonction de coût pour les DVFs . . . . .	82
4.3	Protocole expérimental . . . . .	83
4.3.1	Métriques quantitatives . . . . .	83

4.3.2	Jeux de données . . . . .	84
4.3.3	Expériences additionnelles . . . . .	86
4.3.4	Implémentation et entraînement . . . . .	87
4.4	Résultats . . . . .	90
4.4.1	Modèle synthétique . . . . .	90
4.4.2	Analyses globales des images et des DVFs . . . . .	91
4.4.3	Évaluation locale . . . . .	94
4.5	Discussion . . . . .	102
4.6	Conclusion . . . . .	103
<b>5</b>	<b>Conditionnement respiratoire surfacique par nuages de points</b>	<b>105</b>
5.1	Introduction . . . . .	105
5.2	Simulation de mouvement surfacique . . . . .	107
5.3	Protocole expérimental . . . . .	108
5.3.1	Jeux de données . . . . .	108
5.3.2	Extraction de nuage de point . . . . .	109
5.3.3	Méthode d'évaluation . . . . .	109
5.4	Résultats . . . . .	111
5.4.1	Effet de chaque disques . . . . .	111
5.4.2	Évaluation par zones d'intérêts . . . . .	113
5.4.3	Nombre de couches cachées pour le MLP . . . . .	118
5.5	Discussion . . . . .	119
5.6	Conclusion . . . . .	120
<b>6</b>	<b>Application exploratoire pour un autre objectif : la vessie</b>	<b>123</b>
6.1	Introduction . . . . .	123
6.2	Implémentation . . . . .	124
6.2.1	Jeu de données . . . . .	125
6.2.2	Interpolation . . . . .	126
6.2.3	Autres détails d'implémentations . . . . .	127
6.3	Résultats . . . . .	128
6.4	Discussion . . . . .	130
6.5	Conclusion . . . . .	130
	<b>Conclusion générale et perspectives</b>	<b>133</b>

TABLE DES MATIÈRES

---

<b>Annexes</b>	<b>137</b>
<b>Bibliographie</b>	<b>143</b>

# LISTE DES FIGURES

---

1.1	Définition des volumes cibles . . . . .	21
1.2	Fonctionnement d'un scanner . . . . .	24
1.3	Schéma explicatif de la reconstruction d'image 3D . . . . .	25
1.4	Ceinture respiratoire pour la synchronisation . . . . .	26
1.5	Application et composants d'un système RPM . . . . .	27
1.6	Acquisition du mouvement respiratoire par caméra temps de vol . . . . .	28
1.7	Exemple d'acquisition par groupement par amplitude . . . . .	29
1.8	Exemple d'acquisition par groupement par phase . . . . .	30
1.9	Exemples de variation entre plusieurs acquisitions . . . . .	33
2.1	Architecture U-Net. . . . .	37
2.2	Illustration de l'architecture d'un GAN. . . . .	37
2.3	Illustration de l'architecture d'un cGAN. . . . .	39
2.4	Architecture pix2pix. . . . .	40
2.5	Application de l'IA dans le flux de travail de la radiothérapie . . . . .	41
2.6	Vue d'ensemble des 104 structures anatomiques segmentées par TotalSegmentator . . . . .	42
2.7	Illustration de différents types de transformations. . . . .	44
2.8	Organigramme du fonctionnement du recalage d'images médicales. . . . .	46
2.9	Architecture d'un réseau transformateur spatial. . . . .	48
2.10	Exemple de recalage non supervisé de deux images synthétiques de poumons. . . . .	49
2.11	Exemple de recalage supervisé de deux images synthétiques de poumons. . . . .	50
3.1	Fonctionnement global de l'approche de synthèse directe. . . . .	60
3.2	Aperçu du fonctionnement de la couche AdaIN initiale. . . . .	61
3.3	Fonctionnement de la fonction de coût pondérée par zone d'intérêt. . . . .	62
3.4	Comparaison d'une image générée et de l'image cible. . . . .	63
3.5	Exemples d'images synthétiques et simulation de dynamique. . . . .	65
3.6	Illustration des phases respiratoires d'une TDM 4D. . . . .	66

3.7	Calcul de la variation de volume entre deux phases respiratoires. . . . .	67
3.8	Exemple de résultats de synthèse d’images synthétiques. . . . .	69
3.9	Résultats quantitatifs des jeux de données synthétiques. . . . .	70
3.10	Exemple de résultats de synthèse d’images synthétiques avec et sans $TA$ . . . . .	71
3.11	Exemple de résultats de synthèse d’images cliniques. . . . .	72
3.12	Exemple de résultats sur les données cliniques avec et sans fonction focalisée sur la tumeur. . . . .	73
4.1	Architecture proposée pour la génération par déformation . . . . .	80
4.2	Exemple d’une image synthétique en coupe coronale et sagittale. . . . .	84
4.3	Caractéristiques des deux jeux de données cliniques utilisés pour nos expérimentations. . . . .	86
4.4	Étape de prétraitement des images. . . . .	88
4.5	Extraction du mouvement surfacique. . . . .	89
4.6	Exemple d’un cas de synthèse d’image des données Phantom. . . . .	90
4.7	Aperçu qualitatif de la capacité de synthèse de mouvement respiratoire du modèle. . . . .	92
4.8	Influence des termes de coût sur le DVF généré. . . . .	93
4.9	Mouvement pulmonaire global sur les plans coronal et sagittal. . . . .	95
4.10	L’évaluation de la synthèse du mouvement de la tumeur . . . . .	95
4.11	Comparaison du chevauchement des organes entre les images . . . . .	97
4.12	Comparaison de l’erreur de CoM de la tumeur (en mm) entre les phases dans les jeux de données 4D-Lung (A) et CHRU (B). . . . .	98
4.13	Précision de la synthèse du mouvement tumoral par phases respiratoires dans le jeu de données 4D-Lung . . . . .	100
5.1	Illustration d’un mouvement surfacique. . . . .	107
5.2	Illustration de la conversion d’un objet 3D en nuage de points. . . . .	109
5.3	Illustration des disques en surface utilisés pour les expérimentations. . . . .	110
5.4	Résultats de DSC et de COM par conditionnement de zones surfaciques . . . . .	116
5.5	Résultats de superposition des ITVs avec un conditionnement par zone composée. . . . .	117
5.6	Résultats de superposition des ITVs en fonction du nombre de couches cachées du MLP . . . . .	119

6.1	Illustration du niveau de remplissage de la vessie. . . . .	125
6.2	Illustration des étapes de prétraitement des données cliniques sur la vessie.	126
6.3	Augmentation de données à l'aide d'interpolation. . . . .	127
6.4	Exemple de génération de vessies pleines à partir de vessie vide. . . . .	128



# LISTE DES TABLEAUX

---

3.1	Comparaison de modèle avec ou sans la fonction pondérée sur les images synthétiques. . . . .	71
3.2	Résultats quantitatifs sur les données cliniques. . . . .	73
3.3	Comparaison des résultats avec ou sans la fonction pondérée sur les images cliniques. . . . .	74
4.1	Correction de la position de la tumeur entre les multiples acquisitions pour la phase expiratoire. . . . .	89
4.2	Résultats du modèle avec les données Phantom . . . . .	91
4.3	Analyses de l'impact des fonctions de coûts dans la fonction totale . . . . .	93
4.4	Analyses globales des images après synthèse pour les données 4D-Lung. . . . .	94
4.5	Analyses globales des images après synthèse pour les données CHRU. . . . .	94
4.6	Résultats centrés sur la tumeur avant (sans synthèse de mouvement) et après la synthèse des données 4D-Lung. . . . .	96
4.7	Résultats centrés sur la tumeur avant (sans synthèse de mouvement) et après la synthèse des données CHRU. . . . .	96
4.8	Étude de variabilité des TDM 4D. . . . .	99
4.9	Comparaison des performances du modèle RMSim et de notre approche proposée sur les différents OARs. . . . .	101
5.1	Répartitions des zones. . . . .	110
5.2	Résultats des 25 zones utilisés indépendamment comme conditionnement sur l'ensemble des phases respiratoires . . . . .	112
5.3	Résultats des 25 zones utilisés indépendamment comme conditionnement pour la phase EOI . . . . .	114
5.4	Résultats des zones composées sur l'ensemble des phases respiratoires . . . . .	115
5.5	Résultats de l'influence du nombre de couches cachées dans le MLP pour le conditionnement . . . . .	118
6.1	Résultats quantitatifs sur les données de vessies. . . . .	129

# LISTE DES SIGLES ET ACRONYMES

---

- AdaIN** normalisation d'instance adaptative (*adaptive instance normalization*)
- CBNPC** cancers non à petites cellules
- cGAN** réseau antagoniste génératif conditionnel (*conditional generative adversarial network*)
- COM** centre de masse (*center of mass*)
- CTV** volume cible clinique (*clinical target volume*)
- DL** apprentissage profond (*deep learning*)
- DSC** coefficient de Dice (*Dice similarity coefficient*)
- DVF** champ vectoriel de déformation (*deformation vector field*)
- EOE** fin de l'expiration (*end of exhale*)
- EOI** fin de l'inspiration (*end of inhale*)
- GAN** réseau antagoniste génératif (*generative adversarial network*)
- GPU** processeur graphique (*graphics processing unit*)
- GTV** volume tumoral macroscopique (*gross target volume*)
- HU** hounsfield
- I2I** synthèse d'image à image (*image-to-image translation*)
- IA** intelligence artificielle
- IRM** image par résonance magnétique
- ITV** volume cible interne (*internal target volume*)
- MAE** erreur absolue moyenne (*mean absolute error*)
- MLP** perceptron multicouche (*multilayer perceptron*)
- MSE** erreur quadratique moyenne (*mean squared error*)

**OAR** organe à risque

**PTV** volume cible de planification (*planning target volume*)

**PVD** pourcentage de différence de volume (*percentage volume difference*)

**RPM** contrôle des mouvements respiratoires en temps réel (*real-time position management*)

**SSIM** similarité structurale (*structural similarity*)

**STN** réseau transformateur spatial (*spatial transformer network*)

**TDM** tomodensitométrie

**TDM 3D** tomodensitométrie tridimensionnelle

**TDM 4D** tomodensitométrie quadridimensionnelle

**TOF** temps de vol (*time of flight*)

**TRE** erreur de recalage de la cible (*target registration error*)

**TV** variation totale (*total variation*)

**VV** variation de volume

# INTRODUCTION

---

## Motivation

L'imagerie médicale est devenue indispensable pour le diagnostic et la planification des traitements du cancer. Lorsqu'un patient est diagnostiqué avec un cancer et est indiqué pour un traitement par radiothérapie, il est nécessaire d'établir un plan de traitement ayant pour objectif de maximiser l'efficacité du traitement tout en minimisant les effets indésirables sur les tissus sains avoisinants. Dans le cadre de cette planification, les oncologues-radiothérapeutes utilisent la tomодensitométrie (TDM) pour estimer la densité électronique des tissus. Dans le cas spécifique du traitement du cancer du poumon, il est également impératif de prendre en compte le mouvement respiratoire du patient dans l'élaboration du plan de traitement. En effet, ce mouvement peut entraîner des variations significatives de la position et, dans une moindre mesure, de la forme de la tumeur, avec un risque important de sous-dosage. Dans ce contexte, la tomодensitométrie quadridimensionnelle (TDM 4D) fournit des informations spatiales et temporelles sur le mouvement respiratoire (3D+t), permettant ainsi aux médecins de mieux définir les zones à irradier. Cependant, l'acquisition des images TDM 4D nécessite une exposition prolongée aux radiations, jusqu'à six fois plus importante qu'une acquisition tomодensitométrie tridimensionnelle (TDM 3D) conventionnelle [1], ce qui est potentiellement néfaste pour la santé du patient. De plus, cela peut entraîner un certain inconfort et des contraintes logistiques lors du processus d'imagerie. Il est donc pertinent de trouver des solutions pour limiter l'exposition aux rayonnements induite par l'imagerie TDM 4D.

Les avancées et les résultats de l'intelligence artificielle (IA) dans le domaine de la vision par ordinateur suscitent un intérêt particulier au sein de la communauté médicale. Les approches d'apprentissage profond (DL, *deep learning*) ont révolutionné la capacité des ordinateurs à comprendre, analyser et reproduire des images avec une précision remarquable, s'intégrant ainsi dans diverses applications médicales telles que la segmentation, la reconstruction, le débruitage et la synthèse d'images [2]. En ce qui concerne la synthèse d'images, les modèles génératifs, en particulier les réseaux antagonistes génératifs (GANs, *generative adversarial network*), permettent désormais de générer des images de très haute

qualité. Ces avancées ont ouvert de nouvelles possibilités, et notamment dans le domaine de l'imagerie médicale. Par exemple, il est maintenant possible de générer des pseudo images TDMs à partir d'images par résonance magnétique (IRMs) pour un grand nombre d'applications [3]. Ces techniques offrent des perspectives prometteuses pour améliorer l'efficacité et la précision des examens médicaux. Cependant, la plupart des applications actuelles se concentrent principalement sur les modalités d'images volumétriques statiques (3D), tandis que peu d'attention est portée aux images dynamiques (3D+t).

Cette thèse se concentre sur le développement de modèles génératifs adaptés à l'imagerie dynamique, en particulier pour l'imagerie TDM 4D. La question centrale qui a guidé notre travail est : « Peut-on générer des mouvements réalistes à partir d'une image diagnostique et d'une mesure externe de la respiration du patient (dont la nature peut varier), tout en atteignant un degré de reproductibilité comparable aux acquisitions TDM 4D réelles ? ». En réponse à cette question, et étant donné les performances remarquables de ces approches en synthèse d'images, notre travail s'est principalement concentré sur des méthodes d'apprentissage supervisé basées sur des réseaux convolutionnels et notamment avec une composante non supervisée de type GAN afin d'améliorer la qualité et le réalisme des images générées [3].

Nous avons tout d'abord étendu des approches existantes, telles que pix2pix [4], à l'imagerie dynamique avec l'ajout d'un mécanisme d'injection d'un conditionnement supplémentaire basé sur l'amplitude respiratoire du patient. Bien que nous ayons obtenu des résultats prometteurs, cette approche de génération d'image, que nous qualifions de « directe », ne s'est pas révélée entièrement satisfaisante, comme nous l'expliquerons par la suite. Nous avons ensuite orienté notre travail vers une solution « indirecte » qui intègre un apprentissage supervisé non seulement sur les images déformées, mais également sur les champs vectoriel de déformation (DVF, *deformation vector field*). La particularité de l'approche retenue pour notre travail principal réside dans l'utilisation de deux stratégies complémentaires : l'apprentissage supervisé sur les DVFs et un coût adversaire sur leur amplitude. Comme nous l'expliquerons plus en détail, c'est la contrainte liée à l'amplitude de ces champs qui permet à notre approche de générer des déformations larges et précises grâce à la corrélation avec l'information respiratoire que nous utilisons en tant que conditionnement supplémentaire.

## Aperçu de la thèse

Le but premier de cette thèse est de proposer une nouvelle approche de génération d'images dynamiques pour la planification de traitement en radiothérapie. Cette méthode alternative aux acquisitions TDM 4D basée sur un modèle génératif permet de générer des images dynamiques à partir de l'image TDM 3D de diagnostic ainsi que des données respiratoires du patient. Les contributions principales de cette thèse sont les suivantes :

- L'utilisation d'un modèle de DL simple, un GAN, pour la génération d'images dynamiques.
- Le conditionnement par une information d'amplitude respiratoire relativement aisée à acquérir en routine clinique.
- Une méthode de génération basée sur des déformations afin d'estimer un mouvement réaliste et local.
- L'exploration de conditionnements plus sophistiqués par l'intégration d'informations surfaciques.
- Une étude exploratoire avec l'application d'une architecture similaire pour une application différente : la modélisation du remplissage vésical dans le contexte du cancer de l'utérus.

## Organisation de la thèse

La thèse est organisée en six chapitres principaux et une conclusion, accompagnés d'annexes et d'une liste de références :

Le chapitre 1 présente la modalité de la TDM 4D. Une introduction à la TDM statique et dynamique y est faite. Nous décrivons aussi les avantages ainsi que les limitations de l'imagerie TDM 4D pour la planification de traitement en radiothérapie.

Le chapitre 2 fournit un état de l'art sur les méthodes d'apprentissage profond pour l'estimation et la correction de mouvement pour des applications en radiothérapie. Nous présentons des exemples de méthodes basées DL ainsi que des approches existantes en regard de l'imagerie dynamique.

Le chapitre 3 établit une première méthode de synthèse d'image à image (I2I, *image-to-image translation*) directe doublement conditionnée par l'image et l'amplitude respiratoire, tirant également profit d'une fonction de coût permettant d'aider le réseau à se focaliser sur les tumeurs. Des résultats sur des images synthétiques et des images cliniques

sont présentés.

Le chapitre 4, cœur de notre travail, décrit le modèle vers lequel nos travaux ont convergé. Il s'agit d'une méthode de synthèse dynamique basée sur la combinaison d'une information supervisée par apprentissage sur des DVFs et par un nouveau coût adversaire sur l'amplitude de ces champs. Nous montrons en particulier la capacité de notre modèle à produire des mouvements plus locaux et réalistes que l'état de l'art du domaine, dans le cadre d'une évaluation plus exhaustive que la plupart des méthodes existantes : au niveau global de l'image, mais aussi d'un point de vue plus local de l'organe et de la tumeur. En particulier, nous démontrons que l'incertitude de notre modèle n'est pas plus importante que celle existant entre deux acquisitions TDM 4D réelles, ce qui est un élément clé vers l'acceptabilité clinique de ces modèles.

Le chapitre 5 explore l'utilisation de données surfaciques plus complexes en tant que conditionnement. Des informations de mouvements surfaciques sont extraites depuis les acquisitions afin de simuler un mouvement respiratoire d'un patient. Différentes versions sont évaluées et les résultats préliminaires sont présentés et comparés à un conditionnement scalaire simple.

Le chapitre 6 généralise la méthode proposée à une autre localisation et une autre application de la synthèse de mouvement : le remplissage vésical dynamique dans le contexte du cancer de l'utérus. Il s'agit davantage d'une preuve de concept où nous présentons des résultats expérimentaux et une analyse préliminaire de performance.

Le dernier chapitre résume les résultats de la recherche et leur niveau de maturité dans le domaine de la radiothérapie, en rappelant notamment les limites de nos travaux et évoquant des pistes potentielles pour de futures recherches. Il conclut la thèse en donnant une vue d'ensemble des travaux et des contributions.

# TOMODENSITOMÉTRIE QUADRIDIMENSIONNELLE POUR LA PLANIFICATION DE LA RADIOTHÉRAPIE

---

## Résumé

Ce chapitre décrit l'imagerie de tomodensitométrie quadridimensionnelle (TDM 4D), où les projections sont reconstruites en plusieurs composantes temporelles (TDM 3D+t) permettant de suivre le mouvement des organes. Après avoir brièvement introduit le cancer, les enjeux de la radiothérapie et de sa planification, nous présentons la TDM 4D en décrivant le principe de la tomodensitométrie, ainsi que les méthodes de synchronisation aux données respiratoires. Nous finissons avec les limitations qui nous poussent à proposer une alternative aux protocoles actuels.

## 1.1 Radiothérapie pour le cancer du poumon

### 1.1.1 Généralités sur le cancer et la radiothérapie

#### Cancer

Les cellules du corps se divisent et se multiplient de manière ordonnée pour remplacer les cellules endommagées ou vieillissantes. Cependant, dans le cas d'un cancer, ce processus de régulation est perturbé. Le cancer se caractérise par une croissance anormale et non régulée de cellules mettant en danger la vie des patients. Il peut se développer dans n'importe quelle partie du corps. Chaque type de cancer est caractérisé par la cellule d'origine et son comportement spécifique. Par exemple, le cancer du sein se développe à partir de cellules mammaires, tandis que le cancer du poumon se forme à partir de cellules pulmonaires. En 2020, le cancer du sein était la forme de cancer la plus courante dans le monde avec plus de 2,26 millions de cas, suivi par le cancer du poumon avec 2,21 millions



de cas. Cependant, il est important de noter que le cancer du poumon est la forme de cancer la plus mortelle, ayant entraîné 1,80 million de décès, tandis que le cancer du sein a causé 0,68 million de décès. [5].

## **Radiothérapie**

Pour traiter le cancer, différents traitements sont possibles en fonction de la nature du cancer, de son stade et des caractéristiques individuelles du patient. Parmi les principales options de traitement, nous retrouvons notamment la chimiothérapie avec des traitements médicamenteux, la chirurgie avec l'ablation de la tumeur et la radiothérapie avec l'irradiation des cellules cancéreuses. Dans 50% des cas de cancer, les médecins recommandent la radiothérapie pour traiter les tumeurs et parmi ces cas, 75% concernent des patients atteints d'un cancer du poumon [6].

La radiothérapie consiste à irradier par des rayonnements ionisants les cellules cancéreuses dans le but de les éliminer et d'arrêter leur prolifération. L'objectif principal est ainsi de diriger la plus grande dose possible de rayonnement vers la tumeur tout en préservant au maximum les tissus sains qui l'entourent. La quantité de rayonnement délivrée est mesurée en Gray (Gy), ce qui représente une quantité d'énergie par unité de masse du tissu irradié (1 Gy équivaut à 1 J/kg). Pour avoir une idée des ordres de grandeur caractérisant les effets biologiques, 1 Gy entraîne la mort d'environ 15 à 25% des cellules irradiées, 2 Gy 40 à 55%, et 3 Gy 50 à 70%. Au-delà de cette valeur, chaque Gy supplémentaire détruit environ 50% des cellules restantes. Cependant, les rayonnements émis lors des traitements n'ont pas la capacité de distinguer les cellules cancéreuses des cellules saines, ce qui entraîne la destruction de toutes les cellules exposées. Il est donc essentiel de planifier avec précision la zone à irradier pour maximiser l'efficacité du traitement tout en réduisant les effets secondaires liés à une exposition excessive aux rayonnements. Ces considérations sont étudiées lors de la planification de traitement qui s'effectue en amont des séances de radiation. Cette étape de planification est cruciale et est composée de nombreuses étapes telles que l'acquisition des données du patient et la définition du plan de dose.

### **1.1.2 Définition des volumes en radiothérapie**

Lors de la planification du traitement, il est essentiel de connaître avec précision la position de la tumeur afin de délimiter les volumes d'intérêt et de calculer la dose à

administrer au patient [7], [8]. Cette étape permet de focaliser les rayonnements sur la tumeur tout en préservant les organes à risque (OARs) environnants. C'est au radiologue ou à l'oncologue-radiothérapeute d'effectuer les différentes délimitations permettant de définir les volumes d'intérêt (figure 1.1).

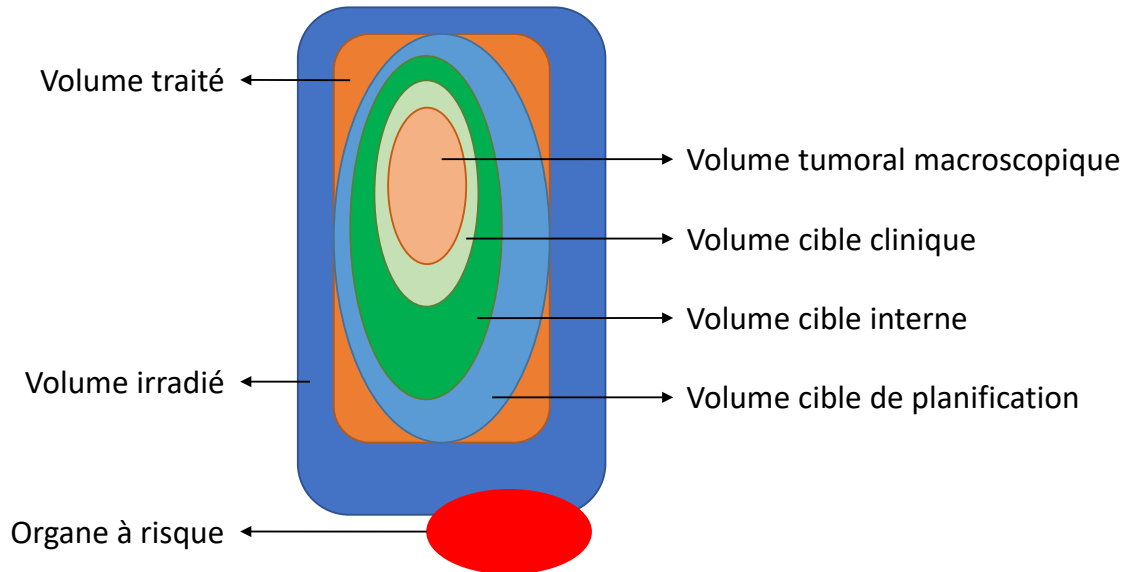


FIGURE 1.1 – Définition des volumes cibles.

Le volume tumoral macroscopique (GTV, *gross target volume*) correspond à la délimitation précise du volume de la tumeur observable sur les images médicales. Il constitue la cible principale du traitement en radiothérapie et reçoit la dose de radiation la plus élevée. Une délimitation précise du GTV est cruciale pour garantir l'efficacité du traitement et la préservation des tissus sains avoisinants.

Le volume cible clinique (CTV, *clinical target volume*) englobe le GTV ainsi que des tissus adjacents à forte probabilité tumorale, même si ces tissus ne sont pas visibles à l'imagerie. La définition du CTV repose sur l'expertise et l'expérience du radiologue ou de l'équipe médicale, en prenant en compte les informations disponibles sur le patient et la maladie. La définition du CTV peut varier d'un praticien à l'autre, mais des protocoles de consensus sont utilisés pour assurer une approche cohérente dans certains cas spécifiques.

Le volume cible interne (ITV, *internal target volume*) représente une expansion du CTV qui tient compte des variations de position et de forme de la tumeur due aux mouvements internes du patient. Il est important de considérer l'ITV pour garantir une

couverture adéquate de la tumeur.

Le volume cible de planification (PTV, *planning target volume*) est dérivé de l'ITV et représente une expansion supplémentaire pour tenir compte des incertitudes géométriques, des variations de position du patient, des erreurs de traitement et d'autres facteurs qui peuvent influencer la précision de l'administration du traitement. Il s'agit d'une marge de sécurité autour de l'ITV qui garantit que la tumeur reçoit la dose de radiation prescrite, même en tenant compte des variations et des incertitudes. Le PTV est déterminé en fonction des caractéristiques spécifiques du traitement, des équipements utilisés et des protocoles de traitement établis. Il est important de noter que la taille du PTV est généralement plus grande que celle de l'ITV, afin de compenser les incertitudes et d'assurer une couverture adéquate de la tumeur.

Un organe à risque (OAR) est une structure anatomique adjacente ou proche du volume à traiter qui est sensible aux rayonnements ionisants et peut être endommagée par la radiothérapie. Il peut s'agir d'organes vitaux tels que le cœur, les poumons, le foie, les reins ou d'autres tissus sains critiques. La protection des OARs est une considération essentielle dans la planification du traitement en radiothérapie. L'objectif est de minimiser la dose de radiation reçue par les OARs tout en assurant une dose adéquate au volume à traiter.

La définition des volumes permet ainsi d'optimiser et sécuriser au maximum le déroulement du traitement. Cependant, dans certains cas, des acquisitions TDM 3D ne suffisent pas pour visualiser avec précision les tumeurs à irradier. En effet, par exemple, dans le cas du cancer du poumon, en raison des mouvements respiratoires du patient, la position de la tumeur peut varier, entraînant ainsi des incertitudes. Ces incertitudes sont particulièrement importantes lorsque la tumeur est située dans les lobes inférieurs des poumons, où le mouvement est le plus prononcé [9]. Il est donc crucial de capturer avec précision les différentes positions de la tumeur en fonction des mouvements respiratoires. C'est pourquoi l'acquisition TDM 4D est largement utilisée en radiothérapie pour la planification des traitements de ce type de cancer.

## 1.2 Principe de la tomodensitométrie quadridimensionnelle

L'acquisition tomodensitométrie quadridimensionnelle (TDM 4D) est une technique d'imagerie médicale utilisée pour visualiser les structures anatomiques en mouvement

telles que les battements du cœur ou encore les mouvements internes des organes liés à la respiration du patient. Elle est composée d'une série d'images TDM 3D qui permettent ainsi de rendre compte de la composante temporelle [10]. Elle permet ainsi de délimiter avec plus de précision la forme et la position des organes cibles pendant tout le cycle respiratoire, ce qui permet de concevoir des plans de traitement plus précis et d'optimiser la délivrance de la dose de rayonnement aux tissus cibles tout en minimisant les effets sur les tissus sains environnants.

### 1.2.1 Introduction à la tomодensitométrie

#### Fonctionnement général

La tomодensitométrie (TDM) est une technique d'imagerie médicale non invasive qui permet de reconstituer des images 2D et 3D des structures anatomiques du corps humain. Elle repose sur le principe de l'absorption des rayons X par le corps humain : les différents tissus (os, muscles, organes, etc.) absorbent les rayons X à des degrés variables. L'appareil équipé de détecteurs et d'un tube émetteur de rayons X mesure l'absorption de ces rayons après qu'ils ont traversé la région du corps à étudier. La TDM offre des avantages majeurs, notamment la visualisation des organes internes, ainsi que la détection sensible de lésions et d'anomalies. Elle constitue donc une méthode d'imagerie largement utilisée en médecine et en fait un outil essentiel dans le diagnostic, la planification des traitements et le suivi des patients [11].

#### Composition d'un scanner

L'appareil TDM, ou scanner, comprend un tube à rayons X qui tourne autour du patient pendant que la table se déplace. Cette méthode dite hélicoïdale ou spirale permet d'acquérir une série d'images en continu sans interruption de l'acquisition. Le tube émet des faisceaux de rayons X traversant le corps du patient, et les récepteurs situés en opposition de ce dernier mesurent l'intensité des rayons X après qu'ils ont été partiellement absorbés durant leur passage à travers le corps (figure 1.2). Ces mesures sont utilisées pour produire une image en coupe par traitement informatique analytique ou par des méthodes algorithmiques itératives (étape dite de reconstruction).

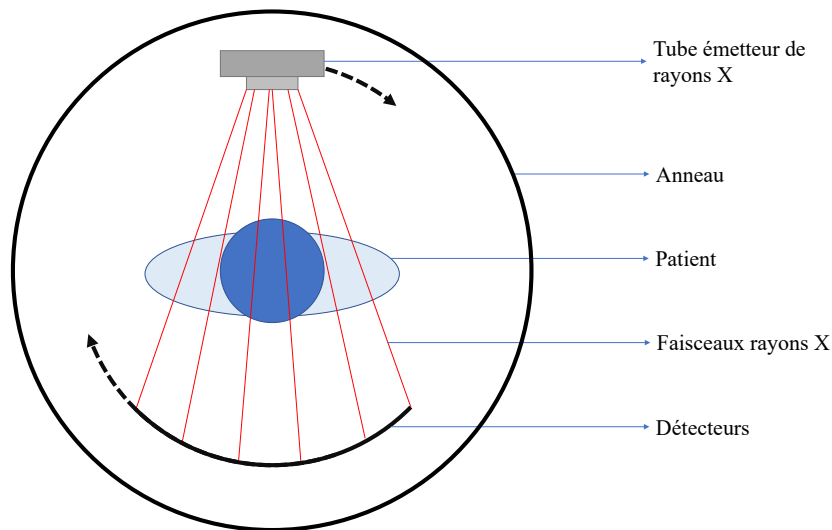


FIGURE 1.2 – Fonctionnement d'un scanner. Pendant une acquisition scanner, un tube émetteur de rayons X et un détecteur rotatif sont utilisés pour capturer les données nécessaires à la reconstruction d'une image 3D. Le tube émetteur tourne autour du patient pendant que le détecteur, positionné en face du tube, mesure l'intensité des rayons X après avoir traversé le corps du patient.

## Reconstruction de l'image

Les détecteurs mesurent l'intensité des rayons X après leur atténuation. Ces données, appelées projections, sont constituées de milliers de mesures correspondant à différents angles à travers lesquels les rayons ont traversé le corps. Ces mesures sont généralement stockées sous forme de sinogrammes, qui sont des représentations graphiques des mesures effectuées à différents angles autour du patient. À partir des projections de sinogramme, l'ordinateur procède à la reconstruction de l'image 3D. Différents algorithmes de reconstruction, tels que la rétroprojection filtrée ou des méthodes itératives (*maximum likelihood expectation maximization* [12], *weighted least squares* [13], etc.), sont utilisés pour reconstruire une image 3D à partir des données brutes du sinogramme (figure 1.3). Ces algorithmes prennent en compte l'atténuation des rayons X et la géométrie de l'appareil scanner pour reconstruire l'image avec une résolution et une qualité appropriées. En général, plusieurs coupes transversales sont obtenues pour créer une représentation en 3D de la zone examinée. Les images résultantes sont exprimées en unités de hounsfield (HU) sur une échelle allant par exemple de -1000 à +1000. La densité de l'eau est choisie comme référence et a pour valeur 0 HU. Ainsi, les structures qui absorbent davantage les rayons X se traduisent par des valeurs positives, tandis que les structures moins absorbantes

affichent des valeurs négatives.

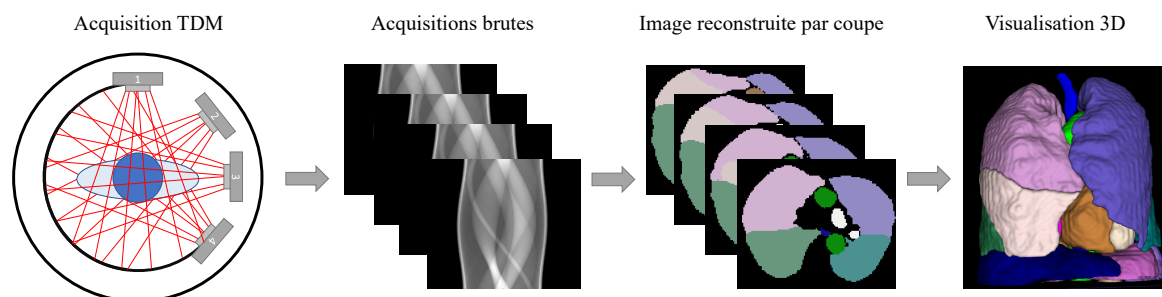


FIGURE 1.3 – Étape de la reconstruction d’images. Les données brutes collectées par le scanner sont enregistrées sous forme de sinogramme. À l’aide d’algorithme de reconstruction spécifique, l’ordinateur analyse les données du sinogramme pour reconstruire les différentes coupes transversales de l’image. Une fois que toutes les coupes transversales ont été reconstruites, elles sont finalement regroupées pour former une image 3D complète de la région examinée.

## 1.2.2 Composante temporelle

L’un des principaux avantages de l’acquisition TDM 4D est la capacité à suivre le mouvement des structures internes et à connaître leur position selon la phase respiratoire du patient. Cette technique permet notamment d’obtenir une estimation de la position de la tumeur, qui peut varier en fonction du mouvement respiratoire. Pour cela, il est nécessaire de synchroniser les images avec la respiration du patient.

### Synchronisation aux données respiratoires

Des techniques de suivi respiratoire sont utilisées en utilisant par exemple des capteurs placés sur le patient, tels que des ceintures respiratoires ou des capteurs optiques, qui enregistrent les mouvements respiratoires du patient. Les informations de suivi respiratoire sont ensuite utilisées pour associer chaque phase de l’acquisition à la phase respiratoire correspondante. Une fois que toutes les images TDMs sont reconstruites et associées à leurs phases respiratoires respectives, elles peuvent être utilisées pour analyser les variations anatomiques liées à la respiration et pour la planification des traitements en radiothérapie, par exemple.

**Ceinture respiratoire :** Une ceinture respiratoire est un dispositif élastique que le patient porte autour de sa poitrine ou de son abdomen pendant l’examen. La ceinture

est équipée de capteurs qui mesurent les variations de la circonférence thoracique ou abdominale pendant la respiration. Ces capteurs peuvent être basés sur différents principes de détection, tels que les capteurs de pression (figure 1.4).

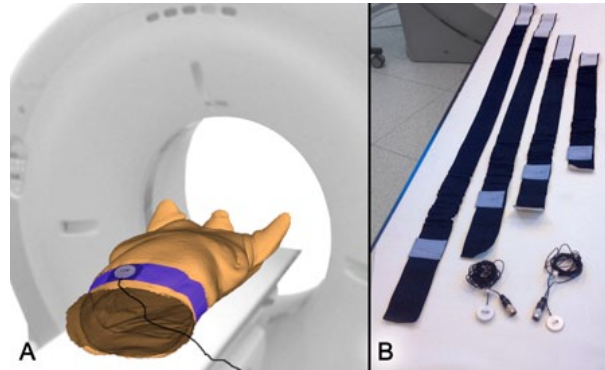


FIGURE 1.4 – Application et composants du système Anzai : (a) Transducteur de pression placé au niveau de l'abdomen du patient, attaché sur la ceinture; (b) Ceintures de différentes tailles et deux transducteurs de pression. Source : [14]

Lorsque le patient respire, la ceinture respiratoire détecte les mouvements de la paroi de la zone concernée et enregistre les variations correspondantes de la circonférence. Ces signaux respiratoires sont ensuite transmis à l'appareil pour marquer chaque image TDM avec la phase respiratoire correspondante.

Le processus de synchronisation commence généralement par une procédure de calibrage, au cours de laquelle le patient est invité à effectuer quelques respirations spécifiques pour que le système puisse enregistrer les variations de la ceinture respiratoire qui correspondent à différentes phases respiratoires. Cela permet d'établir une relation entre les signaux de respiration et les phases respiratoires spécifiques.

Pendant l'acquisition des images, la ceinture respiratoire continue de mesurer les mouvements respiratoires du patient. Les variations de la circonférence sont utilisées pour déterminer la phase respiratoire actuelle, en comparant les signaux en temps réel aux enregistrements de calibrage.

La synchronisation des images avec la respiration à l'aide d'une ceinture respiratoire offre une méthode pratique et non invasive pour capturer les variations respiratoires du patient pendant l'acquisition d'images médicales.

**Système de contrôle des mouvements respiratoires :** Un système de contrôle des mouvements respiratoires en temps réel (RPM, *real-time position management*) est

un dispositif utilisé pour la synchronisation respiratoire lors de l'acquisition d'images médicales. Il utilise des capteurs optiques pour suivre les mouvements du patient pendant la respiration [15].

Le système RPM se compose de deux composants principaux : un boîtier de suivi et des réflecteurs ou marqueurs corporels. Les réflecteurs sont généralement fixés sur la poitrine ou l'abdomen du patient, et ils sont fabriqués à partir d'un matériau réfléchissant la lumière infrarouge. Lorsque le système est activé, le boîtier de suivi émet un faisceau de lumière infrarouge qui est réfléchi par les marqueurs corporels. Le boîtier de suivi détecte ensuite la position des marqueurs corporels en fonction des variations du signal infrarouge réfléchi. Cela permet de mesurer les mouvements respiratoires du patient (figure 1.5). Les variations de position des marqueurs sont enregistrées et utilisées pour marquer chaque image médicale avec la phase respiratoire correspondante. Le système RPM avec capteur optique offre une méthode fiable et précise de synchronisation respiratoire, sans contact direct avec le patient.

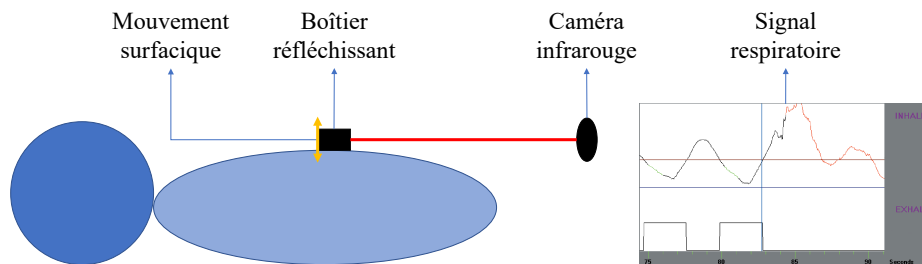


FIGURE 1.5 – Application et composants d'un système RPM. Pour le suivi du mouvement respiratoire, un faisceau de lumière infrarouge est émis par le boîtier de suivi qui est réfléchi par le boîtier réfléchissant. Cela permet de connaître la position du boîtier en fonction des variations du signal réfléchi.

**Caméra :** Les caméras temps de vol (TOF, *time of flight*) sont des dispositifs optiques avancés qui utilisent la mesure du temps de vol des photons pour capturer les mouvements du patient pendant la respiration. Ces caméras sont capables de fournir des informations 3D en temps réel, ce qui les rend particulièrement adaptées à la synchronisation respiratoire [16].

Dans ces méthodes, une ou plusieurs caméras TOF sont placées à des positions stratégiques autour du patient. Les caméras émettent des impulsions de lumière infrarouge et mesurent le temps que mettent ces impulsions pour se propager jusqu'au patient et revenir à la caméra. En utilisant cette mesure, les caméras peuvent reconstruire une carte de



profondeur en temps réel, représentant la géométrie 3D de la surface corporelle du patient (figure 1.6). Pour la synchronisation respiratoire, les variations de la surface corporelle pendant la respiration sont utilisées. Les mouvements respiratoires du patient entraînent des changements dans la forme et la position de la surface corporelle, ce qui est capturé par les caméras TOF. Les informations de profondeur sont extraites à partir des images capturées par les caméras et sont utilisées pour estimer les déformations et les mouvements respiratoires du patient.

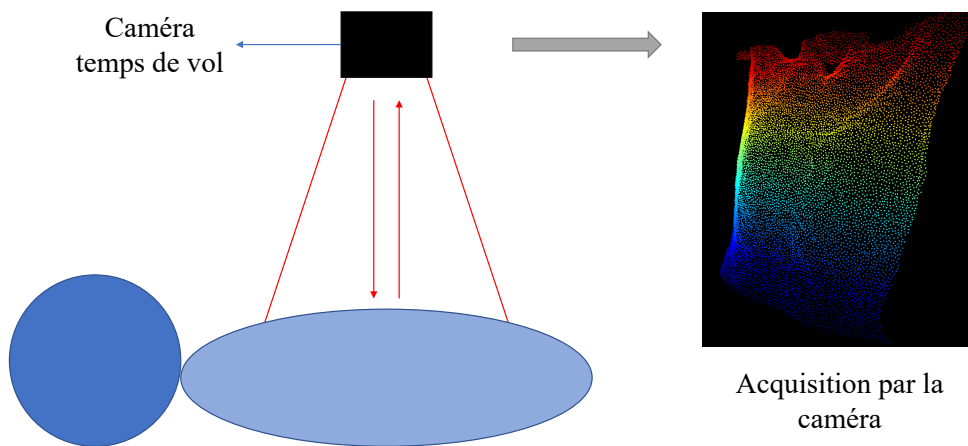


FIGURE 1.6 – Acquisition du mouvement respiratoire par caméra temps de vol. La caméra envoie une impulsion lumineuse qui est réfléchiée par la surface du patient et est renvoyée vers la caméra. La caméra mesure ensuite le temps que met l'impulsion lumineuse pour effectuer l'aller-retour, ce qui permet de déterminer la distance entre la caméra et la surface. Les données obtenues peuvent ainsi reconstruire une carte de profondeur représentant la surface du patient.

Une fois que les mouvements respiratoires sont estimés, ces informations peuvent être utilisées pour synchroniser les images médicales avec les phases respiratoires correspondantes. L'utilisation de caméras TOF pour la synchronisation respiratoire présente plusieurs avantages. Ces caméras permettent une acquisition en temps réel des mouvements respiratoires, ce qui est essentiel pour une synchronisation précise. De plus, les informations 3D fournies par les caméras TOF permettent une analyse plus détaillée des mouvements respiratoires et une meilleure compréhension des déformations du corps.

## Méthode d'acquisitions

Deux approches couramment utilisées pour l'imagerie TDM 4D sont le groupement (*binning* en anglais) par amplitude et le groupement par phase. Le groupement par amplitude regroupe les images en fonction de l'amplitude du signal respiratoire correspondant, qui est étroitement liée à l'amplitude du mouvement du diaphragme. En revanche, dans le groupement par phase, chaque image est attribuée à un groupe spécifique en fonction de la phase du signal respiratoire au moment de la génération de l'image. Il est important de noter que ces techniques de groupement peuvent aboutir à des résultats différents selon l'approche utilisée [17].

**Amplitude :** L'acquisition basée sur l'amplitude repose sur la variation de l'amplitude des mouvements respiratoires du patient. Les images TDM 4D sont acquises de manière continue tout au long du cycle respiratoire complet du patient. Les images acquises sont ensuite triées et regroupées en fonction de l'amplitude du signal respiratoire correspondant. (figure 1.7)

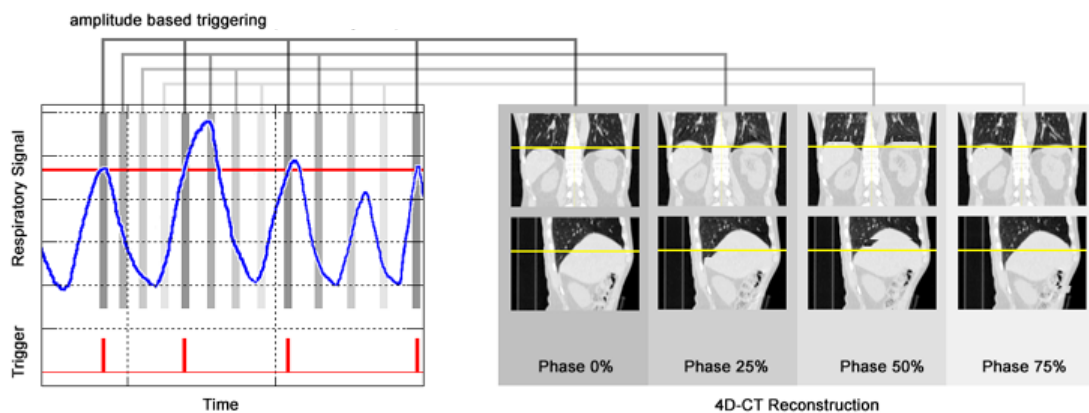


FIGURE 1.7 – Exemple d'acquisition par groupement par amplitude. L'acquisition se déclenche lorsque le signal respiratoire atteint un certain seuil (ligne rouge supérieure). Dans cet exemple, la reconstruction est incorrecte en raison d'un déclenchement de l'acquisition faussé visible sur le deuxième cycle respiratoire. Source : [14]

L'acquisition basée sur l'amplitude présente plusieurs avantages. Elle permet une représentation plus complète et détaillée des changements morphologiques liés à la respiration. De plus, elle offre une meilleure résolution temporelle en capturant les variations d'amplitude du mouvement respiratoire tout au long du cycle respiratoire. Cela permet

d'obtenir des informations précieuses sur les mouvements internes du corps et leur impact sur les structures anatomiques.

Cependant, il convient de noter que l'acquisition basée sur l'amplitude peut être plus sensible aux mouvements respiratoires irréguliers ou incohérents du patient. Des artefacts de mouvement peuvent apparaître dans les images, ce qui peut rendre l'interprétation difficile.

**Gating :** Le groupement par phase, également connu sous le nom de *gating* en anglais, est une technique utilisée lors de l'acquisition d'images médicales pour synchroniser l'acquisition avec le cycle respiratoire du patient.

Le principe fondamental du *gating* consiste à sélectionner des moments spécifiques du cycle respiratoire pendant lesquels les images sont acquises. Un capteur de respiration détecte les phases spécifiques du cycle respiratoire, telles que l'inspiration ou l'expiration, en mesurant les changements de position ou de volume du thorax ou de l'abdomen. Une fois que les phases respiratoires d'intérêt ont été identifiées, l'acquisition des images est déclenchée uniquement lorsque le patient se trouve dans l'une de ces phases spécifiques. Les images acquises à chaque phase sont ensuite associées à cette phase particulière du cycle respiratoire (figure 1.8).

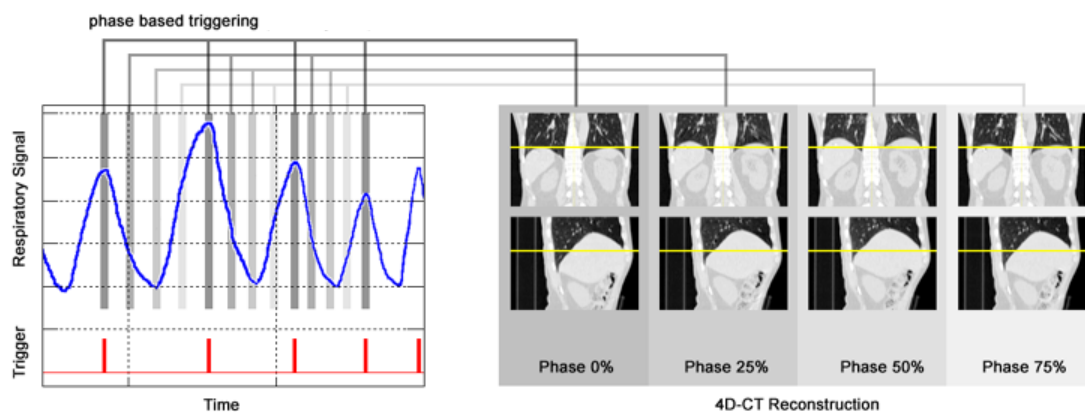


FIGURE 1.8 – Exemple d'acquisition par groupement par phase. L'acquisition se déclenche dans cet exemple lorsque le signal respiratoire atteint un pic maximal représentant la phase fin de l'inspiration (EOI, *end of inhale*). Source : [14]

L'acquisition basée sur le *gating* présente plusieurs avantages. Elle permet de figer le mouvement respiratoire, ce qui réduit les artefacts de mouvement et améliore la netteté des images. Cela est particulièrement bénéfique lors de l'imagerie des régions thoraciques

ou abdominales, o  le mouvement respiratoire peut entra ner un flou ou une d formation des structures d'int r t. De plus, le *gating* permet d'obtenir une correspondance pr cise entre les images acquises et les phases respiratoires sp cifiques, ce qui est important pour l'analyse quantitative et la comparaison entre les diff rentes phases.

Cependant, il convient de noter que l'acquisition bas e sur le *gating* n cessite une synchronisation pr cise entre le capteur de respiration et l'appareil d'imagerie. Des erreurs de synchronisation peuvent entra ner des erreurs dans l'attribution des images aux phases respiratoires correspondantes. De plus, l'utilisation de l'acquisition bas e sur le *gating* peut prolonger le temps d'acquisition global, ce qui peut  tre contraignant pour le patient et n cessite sa coop ration.

## 1.3 Limitations de la modalit 

### 1.3.1 Protocole lourd

L'acquisition TDM 4D est un protocole complexe qui demande  norm ment de ressources humaines et mat rielles et qui prend du temps   ex cuter.

Durant la phase d'acquisition, cela n cessite une collaboration active de la part du patient, impliquant la r alisation d'exercices respiratoires sp cifiques pendant la s ance d'imagerie. Ces exercices peuvent inclure des instructions pour retenir sa respiration pendant des p riodes sp cifiques. Ces mesures visent   minimiser les artefacts respiratoires et   obtenir des images de meilleure qualit . Cependant, cela peut  tre difficile  tant donn  que les patients sont souvent d j  malades et peuvent pr senter des difficult s respiratoires [18].

L'obtention d'une image de qualit  optimale peut  tre entrav e par divers facteurs tels qu'une mauvaise reconstruction, la pr sence d'artefacts, et autres. Dans de tels cas, il devient n cessaire d'effectuer de nouvelles acquisitions afin de rem dier   ces limitations. Ainsi, l'utilisation de marqueurs corporels devient incontournable pour faciliter l'alignement pr cis du patient lors de chaque s ance de traitement ou d'imagerie. Ces marqueurs jouent un r le crucial en permettant un calibrage pr cis de la machine, garantissant ainsi un positionnement coh rent du patient d'une s ance   l'autre. Mais un mouvement ind sirable lors de l'acquisition peut survenir pour diverses raisons, telles que la fatigue du patient, l'anxi t  ou des mouvements involontaires. Ces mouvements peuvent entra ner une baisse de qualit  de l'image, voire rendre l'image inutilisable.

### 1.3.2 Durée d'exposition

L'une des principales limitations de l'acquisition TDM 4D est la durée d'exposition du patient aux rayons X. En effet, l'acquisition de plusieurs phases respiratoires nécessite une acquisition prolongée, ce qui entraîne une exposition prolongée aux rayons X [1]. La durée d'exposition dépend de plusieurs facteurs tels que le nombre de phases respiratoires souhaitées et la vitesse d'acquisition de l'appareil de tomographie. Plus le nombre de phases respiratoires est élevé, plus la durée d'exposition sera longue. De plus, la résolution temporelle souhaitée peut également influencer la durée totale de l'acquisition [19].

La durée d'exposition est importante à prendre en compte, car une exposition prolongée aux rayonnements peut entraîner des effets néfastes pour la santé du patient. Il est donc essentiel de trouver un équilibre entre la nécessité d'acquérir des informations précises sur le mouvement respiratoire et la réduction de l'exposition aux rayonnements qui n'est pas négligeable dans les examens prescrits aux patients, qui est d'autant plus importante si plusieurs acquisitions sont nécessaires [20].

### 1.3.3 Variabilité des acquisitions

Pour garantir une précision optimale dans la délimitation des volumes cibles, il est courant de réaliser plusieurs acquisitions TDM 4D. Ces multiples acquisitions visent à obtenir une connaissance plus précise de la position et du mouvement des tumeurs, en particulier dans les lobes inférieurs des poumons où le mouvement respiratoire est le plus important [9]. L'objectif principal de cette approche réside dans la réduction des incertitudes liées aux mouvements respiratoires. En effet, des différences dans le rythme respiratoire du patient entre les différentes séances peuvent influencer le positionnement des tumeurs [21], [22]. Ces variations peuvent être observées dans des conditions cliniques où la respiration du patient peut être altérée en raison de facteurs tels que la maladie pulmonaire ou d'autres problèmes de santé. La réalisation de multiples acquisitions TDM 4D permet d'obtenir une représentation plus complète et précise des variations de la tumeur tout au long du cycle respiratoire, contribuant ainsi à une meilleure évaluation des déplacements et des changements morphologiques potentiels de la tumeur (1.9).

Malgré la mise en place de protocoles rigoureux et de techniques avancées, la variabilité entre les acquisitions reste une considération importante. Il est crucial de comprendre ces variations afin de pouvoir ajuster les approches de traitement en conséquence. Des méthodes d'analyse et de correction peuvent être appliquées pour tenir compte de la va-

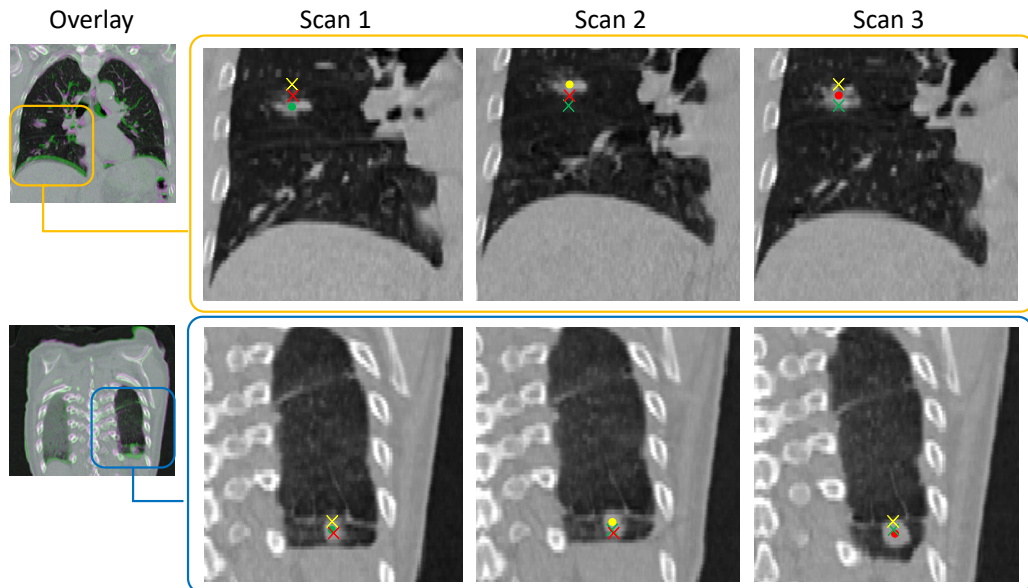


FIGURE 1.9 – Exemples de variation entre plusieurs acquisitions. Deux patients ont effectué chacun trois acquisitions TDM 4D. Les points verts, jaunes et rouges représentent les centres de masse des tumeurs pour chaque scan respectivement. Les croix représentent la position des centres de masse de la tumeur sur les autres scans. Données fournies par le service de radiothérapie du CHRU de Brest.

riabilité entre les acquisitions, contribuant ainsi à une délimitation plus précise des volumes cibles avec, par exemple, l’union des zones à irradier de chaque acquisition TDM 4D [23].

Il est important de souligner que chaque nouvelle acquisition TDM 4D ajoute une dimension supplémentaire à la prise en charge du patient et accroît son exposition aux rayonnements. La répétition de ces acquisitions peut entraîner une augmentation cumulative de l’exposition aux rayonnements au fil du temps. Cette considération souligne l’importance d’explorer des approches alternatives qui pourraient réduire le nombre d’acquisitions nécessaires tout en maintenant la précision diagnostique souhaitée.

## 1.4 Discussion

Dans le cas du cancer du poumon, le développement de la TDM 4D a permis d’améliorer la délimitation des volumes cibles en fournissant une visualisation des mouvements. Les images obtenues permettent également de mieux prévenir les OARs de l’exposition aux rayons X lors du traitement, réduisant ainsi les effets secondaires liés aux radiations [24].

Cependant, ce protocole d’acquisition présente des inconvénients qui incitent à rechercher des alternatives plus intéressantes et plus sûres pour la santé des patients. Différentes approches sont étudiées pour éviter la surexposition aux radiations, telles que le recalage entre les phases inspiratoires et expiratoires pour simuler une acquisition TDM 4D [25], ou la combinaison de différentes modalités, comme l’utilisation d’une acquisition PET 4D standard avec une seule image TDM, qui offre la possibilité de générer une TDM 4D par déformation [26].

Plus récemment, de nouvelles alternatives basées sur du DL se développent pour générer, par exemple, des pseudos TDM 4D à partir d’images IRM 4D [27]. Il est donc intéressant d’explorer ces méthodes alternatives afin de générer des TDM 4D, améliorant ainsi la santé et la prise en charge des patients (réduction de l’exposition aux rayonnements, des examens et accès au soin dans les hôpitaux non équipés).

## 1.5 Conclusion

Nous avons exposé l’imagerie TDM 4D dans le contexte de la planification du traitement du cancer du poumon par radiothérapie. Les acquisitions TDM 4D sont distinguées par leur composante temporelle, permettant le suivi du mouvement interne des organes et des tumeurs. Cette représentation temporelle améliore la délinéation de la tumeur tout en prenant en considération les structures anatomiques critiques à éviter, conduisant ainsi à un calcul de dose plus précis pour le traitement de la tumeur.

Ces acquisitions demandent des coûts humains importants et nécessitent une collaboration active de la part du patient pour contrôler sa respiration. De plus, elles exposent les patients à des doses de rayonnement supplémentaires, engendrant de possibles effets secondaires. Cette exposition accrue résulte du temps d’acquisition plus long qu’une acquisition TDM 3D, ainsi que de la nécessité de répéter ces acquisitions afin de mieux comprendre les variations éventuelles liées à la respiration du patient.

Face à ces limitations, il devient impératif de rechercher des alternatives moins coûteuses et moins contraignantes pour les patients. Le chapitre suivant aborde des approches de synthèse d’images basées sur le DL qui présentent une capacité à générer des images réalistes ainsi que des notions de recalage d’images pour l’estimation ou la synthèse de mouvement.

# APPRENTISSAGE PROFOND POUR L'IMAGERIE TDM THORACIQUE EN RADIOTHÉRAPIE

---

## Résumé

Le fond méthodologique de cette thèse repose sur des approches de synthèse d'images basées sur l'apprentissage profond (DL, *deep learning*). Dans ce contexte, ce chapitre introduit ces concepts en s'attachant à citer un certain nombre d'applications dans le domaine de la radiothérapie. Nous concentrons ainsi nos travaux sur les modèles synthèse d'image à image (I2I, *image-to-image translation*) par réseau antagoniste adversaire (GAN, *generative adversarial network*), constituant la grande partie de l'état de l'art en synthèse d'images. Nous introduisons ensuite des notions fondamentales de recalage d'images pour l'estimation ou la synthèse de mouvement. Nous mentionnons également quelques travaux connexes en vision par ordinateur. Nous relevons ainsi les limitations des modèles actuels en imagerie TDM qui ont motivé nos travaux développés dans les chapitres suivants.

## 2.1 Introduction

### 2.1.1 Synthèse d'image à image

L'apprentissage profond (DL, *deep learning*), est une branche de l'IA qui utilise des modèles de réseaux neuronaux profonds pour apprendre et extraire des relations hautement non linéaires à partir de données. Ces méthodes ont suscité un vif intérêt ces dernières années dans le domaine de l'imagerie médicale, y compris en radiothérapie.

L'objectif de la synthèse d'image à image (I2I, *image-to-image translation*) est d'apprendre à transférer des images d'un domaine source vers un domaine cible tout en pré-



servant le contenu du domaine source. De ce fait, il existe de nombreuses applications, telles que la synthèse de photos à partir de cartes labellisées, la reconstruction d'objets à partir d'images ne comportant que des contours, la transformation d'images de jour en images de nuit, etc. [28]-[30]. Un point intéressant de cette tâche de vision par ordinateur est la possibilité de générer des étapes intermédiaires entre deux images, ce qui permet de visualiser, par exemple, une transition graduelle d'un paysage de jour à un paysage de nuit, ou encore d'un visage neutre à un visage souriant [31]-[33]. Nous décrivons ensuite deux architectures majeures qui ont fortement contribué au développement de cette tâche : les U-Nets et les GANs.

### Architecture U-Net

Le « U-Net » est un réseau qui a été conçu à l'origine pour la segmentation d'images médicales [34]. Il est structuré en trois parties distinctes : la contraction, le goulot d'étranglement, et l'expansion, formant visuellement un « U » (figure 2.1).

La première partie du U-Net, l'encodeur, diminue progressivement la résolution spatiale de l'image, tandis que la dernière partie, le décodeur, rétablit la résolution initiale de l'image. L'encodage permet au réseau de se concentrer sur des détails plus fins qui seront reconstitués par le décodeur. Pour conserver des informations à un niveau plus global, des connexions résiduelles sont établies entre les différents niveaux de l'encodeur et du décodeur. Cette approche permet de fusionner des informations spatiales à la fois de haut niveau et de bas niveau lors de la reconstruction effectuée par le décodeur. Ainsi, le réseau est capable d'apprendre à la fois les caractéristiques globales et locales de l'image, et il s'est avéré robuste même avec des volumes de données d'entraînement limités.

### Architecture des réseaux antagonistes génératifs

Un réseau antagoniste adversaire (GAN, *generative adversarial network*) est un type de modèle d'apprentissage automatique qui est utilisé pour générer de nouvelles données réalistes à partir d'un ensemble de données existant [35]. Il se compose de deux composants principaux : un générateur  $G$  et un discriminateur  $D$ , tous deux des réseaux de neurones (figure 2.2).

Le générateur  $G$  est responsable de la création de nouvelles données de manière à ce qu'elles ressemblent le plus possible aux données réelles  $x$ . Il prend généralement en entrée un vecteur aléatoire  $z$  (appelé vecteur latent) et génère des données synthétiques  $G(z)$  qui

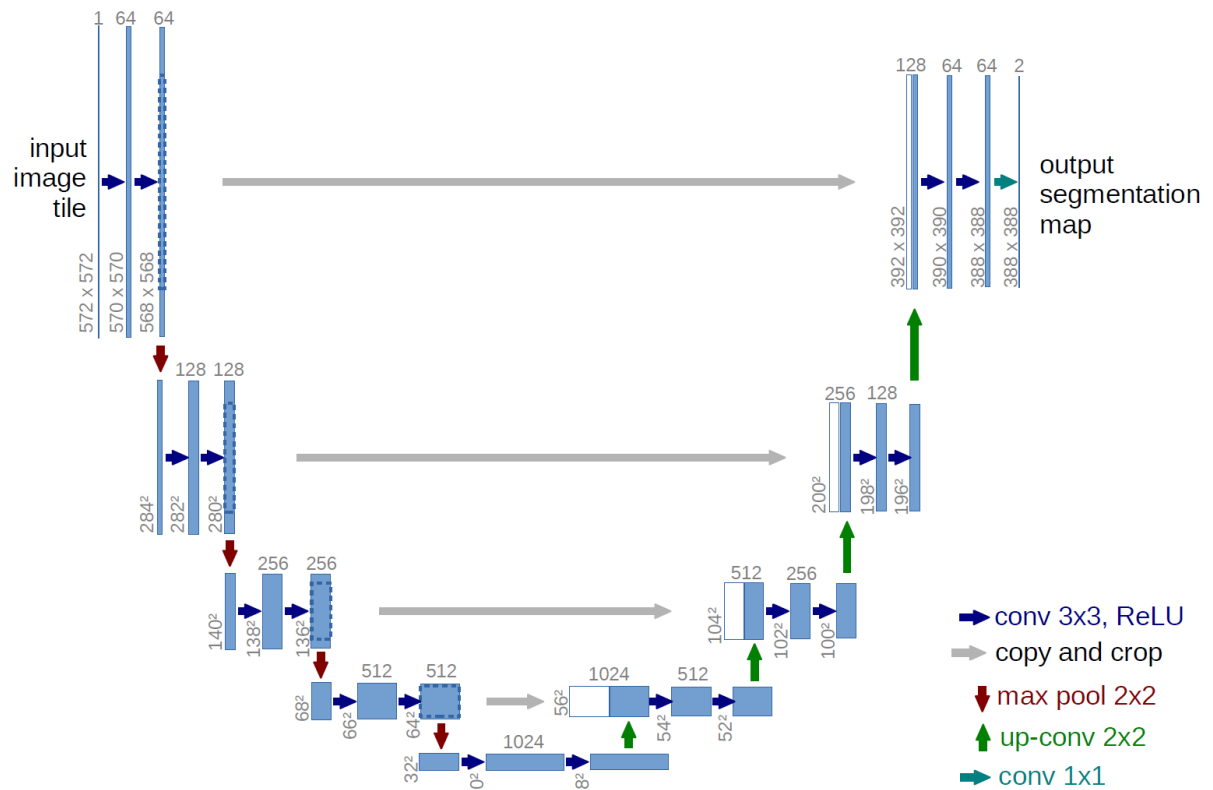


FIGURE 2.1 – Architecture U-Net. Chaque boîte bleue représente une carte de caractéristiques à canaux multiples. Le nombre de canaux est noté en haut de chaque boîte, tandis que les dimensions en termes de coordonnées x-y sont précisément spécifiées dans l'angle inférieur gauche. Les flèches indiquent les connexions résiduelles entre l'encodeur et le décodeur. Les boîtes blanches représentent les cartes de caractéristiques copiées de la partie de contraction fusionnées à celles de l'expansion. Source : [34]

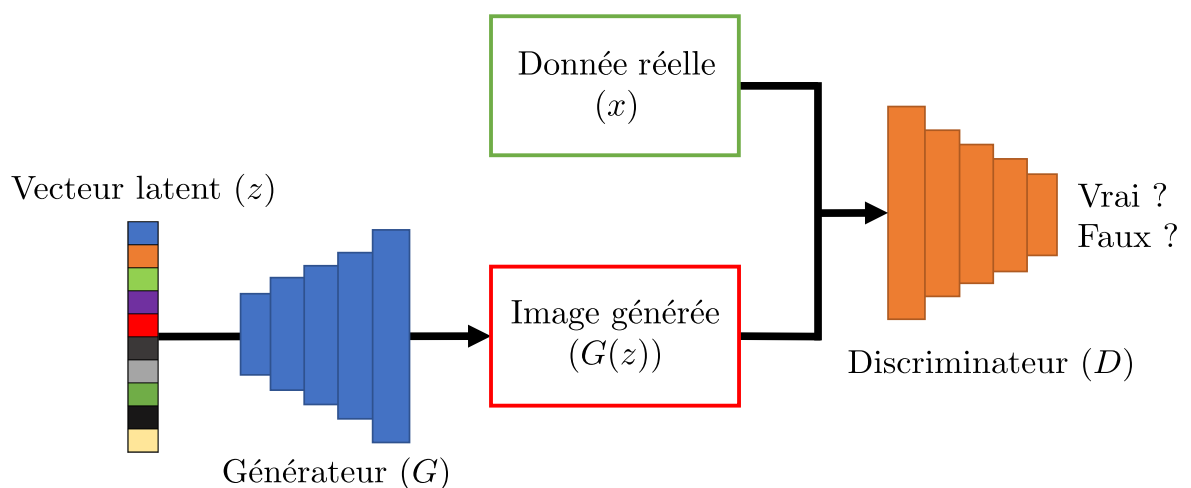


FIGURE 2.2 – Illustration de l'architecture d'un GAN.

sont ensuite soumises au discriminateur  $D$ . Ce code latent contrôle ainsi la variabilité de la génération.

Le discriminateur  $D$ , quant à lui, agit comme un « juge » et essaie de distinguer les données réelles des données générées par le générateur. Il prend en entrée des échantillons de données réelles et des échantillons de données générées, et essaie de classifier correctement ces échantillons en tant que réels ou faux.

Les deux composants du GAN sont entraînés de manière antagoniste. Le générateur cherche à tromper le discriminateur en générant des données qui sont indiscernables des données réelles, tandis que le discriminateur cherche à améliorer sa capacité à distinguer les vraies données des données générées. En d'autres termes,  $G$  et  $D$  s'affrontent dans un jeu à deux joueurs *minimax* avec la fonction de valeur  $V(G, D)$  :

$$\min_G \max_D V(G, D) = \mathbb{E}[\log D(x)] + \mathbb{E}[\log(1 - D(G(z)))] \quad (2.1)$$

Ce processus d'entraînement itératif permet au générateur d'apprendre progressivement à générer des données de plus en plus réalistes. L'objectif final du GAN est donc d'obtenir un générateur capable de produire des données qui sont si réalistes qu'elles sont indiscernables des données réelles.

Un modèle étendu d'un GAN, le réseau antagoniste adversaire conditionnel (cGAN, *conditional generative adversarial network*), proposé par Mirza et coll. peut être considéré comme une version supervisée d'un GAN [36]. En complément du vecteur latent  $z$ , une information supplémentaire  $y$  est ajoutée à la fois au générateur et au discriminateur en tant que condition (figure 2.3).

De manière similaire au GAN, la fonction objectif est basée sur l'algorithme *minimax* avec une probabilité conditionnelle en plus :

$$\min_G \max_D V(G, D) = \mathbb{E}[\log D(x|y)] + \mathbb{E}[\log(1 - D(G(z|y)))] \quad (2.2)$$

Un exemple important d'un cGAN dans le domaine du I2I est le modèle pix2pix d'Isola et coll. [4]. Le modèle utilise un U-Net en tant que générateur qui permet au modèle d'apprendre à transférer une image d'un domaine source vers un domaine cible et parvient à préserver un réalisme de haute qualité grâce au coût adversaire apporté par l'architecture GAN. Le générateur du modèle encode une image  $y$  d'un domaine source en un vecteur latent non aléatoire, puis la reconstruit en une image synthétique  $G(x)$  d'un domaine cible. Quant au discriminateur, il apprend à classifier le couple d'images composé

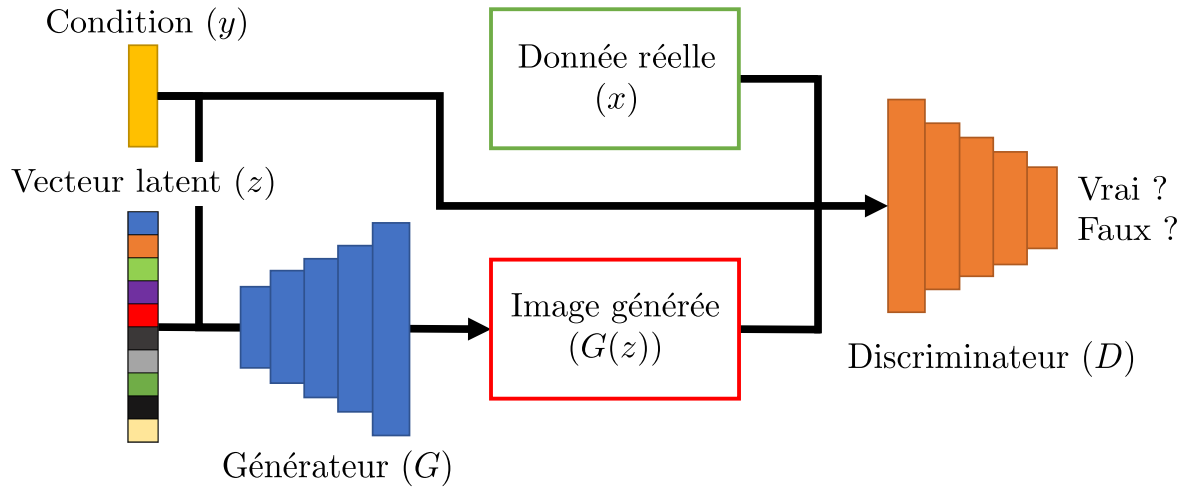


FIGURE 2.3 – Illustration de l’architecture d’un cGAN.

de vraies images  $\{x, y\}$  du couple composé de la fausse image  $\{x, G(x)\}$  (figure 2.4).

L’objectif cGAN de pix2pix se définit comme :

$$cGAN(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))] \quad (2.3)$$

L’utilisation d’une fonction de perte constituée uniquement de la norme  $\ell_1$  conduit à des résultats raisonnables, mais flous ; d’autre part, la cGAN seule donne des résultats nets, mais introduit des artefacts dans certaines applications de I2I [4]. Le modèle pix2pix démontre qu’avec une utilisation conjointe d’un coût adversaire et d’une norme  $\ell_1$ , les images générées sont plus nettes avec une faible quantité d’artefacts. L’objectif final du modèle s’écrit sous la forme :

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{\ell_1}(G) \quad (2.4)$$

### 2.1.2 Domaine d’application de l’apprentissage profond en radiothérapie

L’IA a le potentiel de s’intégrer dans la routine clinique afin d’aider et d’améliorer le travail du corps médical. Par exemple, l’utilisation de DL peut être intégrée à différents niveaux de la planification de traitement de la radiothérapie (acquisition d’images, post-traitements, etc.) pour améliorer l’efficacité et la qualité des traitements. Ces méthodes

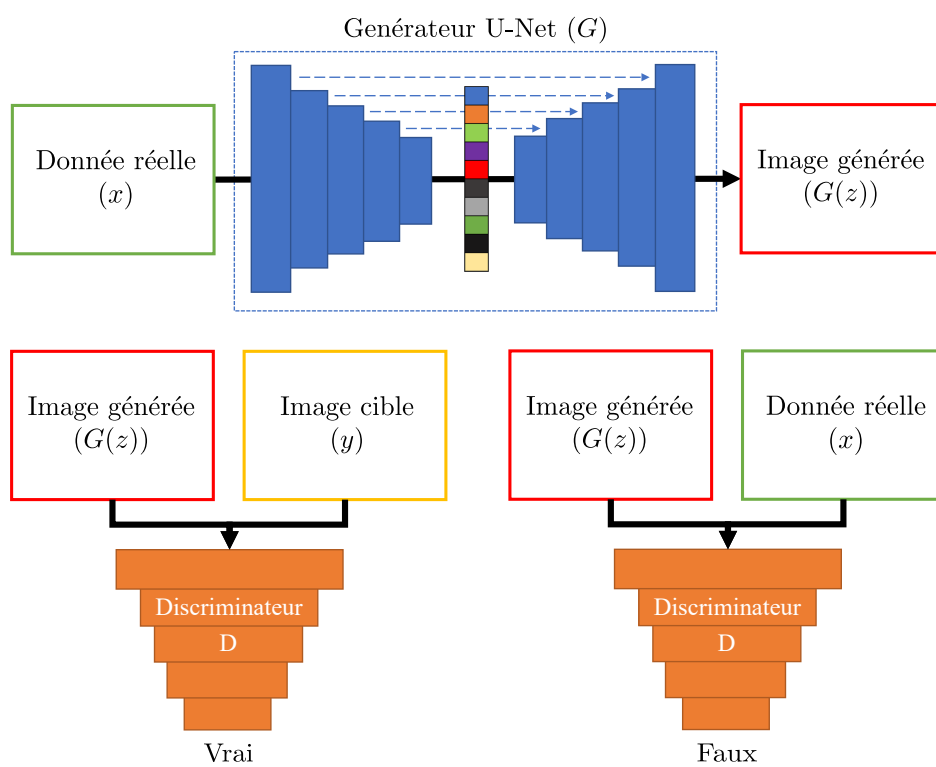


FIGURE 2.4 – Architecture pix2pix.

peuvent aussi fournir des informations complémentaires aux informations cliniques et peuvent aussi aider à prédire des réponses aux traitements (figure 2.5) [37].

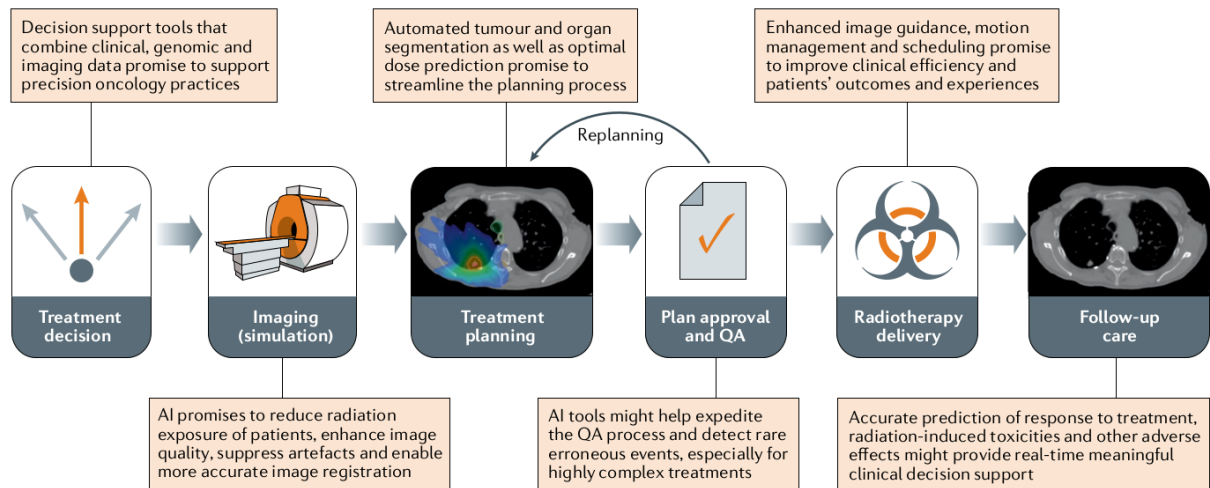


FIGURE 2.5 – Application de l'IA dans le flux de travail de la radiothérapie. Le flux de travail commence par la prise de décision concernant l'utilisation ou non à la radiothérapie, puis une séance de simulation est programmée pour acquérir des images médicales nécessaires à la planification du traitement. Ensuite, un plan de traitement spécifique au patient est élaboré, soumis à une approbation, examiné et soumis à des mesures de contrôle qualité. Après le traitement, le patient bénéficie d'un suivi médical. Source : [37].

Par exemple, des méthodes I2I permettent d'aider à la segmentation d'OARs. Waserthal et coll. développent un outil permettant de segmenter 104 structures anatomiques incluant les os, le système cardiovasculaire, le cerveau, les poumons, etc. (figure 2.6) [38]. Leur méthode utilise la structure nnU-Net qui se base sur un U-Net qui configure automatiquement tous les hyperparamètres en fonction des caractéristiques de l'ensemble de données [39]. Cirillo et coll. présentent dans leurs travaux une méthode basée sur une version 3D de pix2pix, baptisée vox2vox, pour la segmentation des tumeurs cérébrales, tandis que Jain et coll. proposent un GAN pour la segmentation des nodules pulmonaires [40], [41]. L'utilisation du coût adversaire dans ces cas d'études permet de pénaliser les résultats de segmentation qui ne semblent pas réalistes. Ces outils d'aide à la segmentation contribuent à faciliter la délimitation précise des volumes cibles par le radiothérapeute, afin de permettre l'administration optimale de la dose de radiation spécifique à chaque patient.

Un autre exemple d'application de modèle I2I est celle de la synthèse d'image lors de l'acquisition d'images médicales pour la planification de traitement. Il est courant d'acquérir différentes modalités d'images, par exemple, une IRM pour une meilleure visualisation

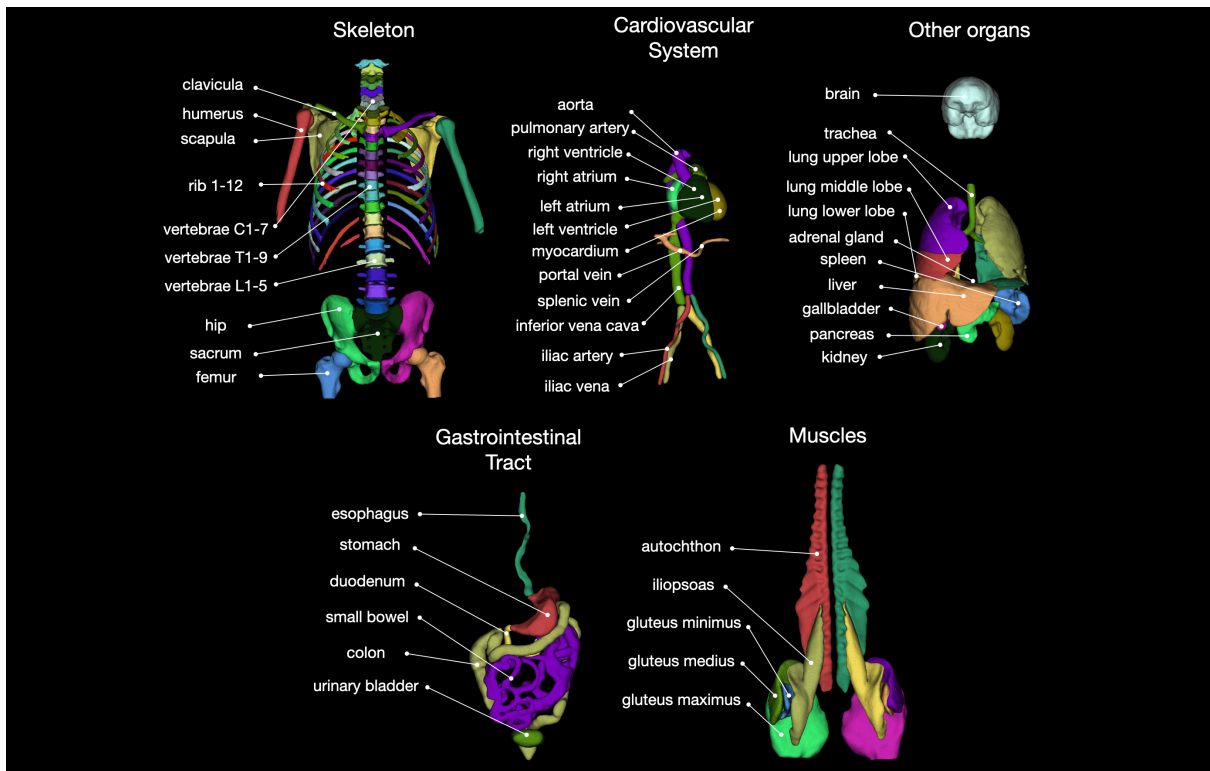


FIGURE 2.6 – Vue d'ensemble des 104 structures anatomiques segmentées par TotalSegmentator. Source : [38].

de la tumeur dans certains cas et donc une meilleure segmentation, suivie d'une TDM. Mais il existe donc un décalage entre les acquisitions dû à un positionnement différent du patient ou une différence de temps entre les acquisitions. Une étape de recalage est nécessaire entre les modalités, notamment avec la TDM car elle est utilisée pour effectuer une conversion des unités HU en densité électronique afin de réaliser le calcul de dose pour le traitement. Des travaux tels que ceux de Maspero et coll. démontrent l'intérêt de l'utilisation d'un cGAN de type pix2pix pour générer des images TDMs synthétiques du bassin à partir des images IRMs, réduisant ainsi les incertitudes liées à l'alignement des images et diminuant l'exposition aux rayonnements due aux acquisitions TDMs [42].

Cependant, toutes ces méthodes sont principalement dédiées à des contextes d'imagerie statique. Dans un contexte dynamique, la prise en compte des mouvements internes du corps, tels que la respiration et les mouvements des organes, revêt une grande importance, c'est notamment le cas pour la planification du traitement en radiothérapie. Ces mouvements peuvent entraîner des erreurs de positionnement de la tumeur cible et des OARs, ce qui compromet l'efficacité du traitement et accroît les risques pour les tissus sains environnants. Avant de passer au traitement, il est nécessaire de réduire les incertitudes liées à des problèmes de positions ou à des mouvements internes. C'est pourquoi des méthodes de corrections et d'estimations du mouvement qui visent à garantir la précision et l'efficacité du traitement sont des aspects cruciaux de la radiothérapie.

Dans la suite de ce chapitre, nous nous concentrerons d'abord sur la correction du mouvement en présentant les méthodes de recalage d'images médicales. Ensuite, nous aborderons les perspectives de l'utilisation de méthodes de DL pour l'estimation du mouvement, en illustrant plusieurs exemples d'imagerie médicale. Enfin, nous analyserons les limites des approches présentées.

## 2.2 Correction de mouvement

La correction de mouvement en imagerie médicale est une technique utilisée pour compenser les mouvements indésirables ou involontaires qui peuvent survenir lors de l'acquisition d'images médicales. Ces mouvements peuvent être causés par divers facteurs tels que la respiration, les battements cardiaques, les tremblements, les mouvements volontaires du patient ou encore une erreur de positionnement entre plusieurs acquisitions. Ainsi, en imagerie médicale, des techniques de recalage peuvent être appliquées pour aligner les différentes acquisitions et corriger les mouvements. Cela permet de superposer



les images de différentes phases du mouvement ou de les aligner avec d'autres modalités d'imagerie [43].

### 2.2.1 Introduction au recalage d'images médicales

Les méthodes de recalage peuvent être classées en deux catégories principales : le recalage rigide et le recalage non rigide.

Le recalage rigide consiste à appliquer des transformations rigides telles que la translation et la rotation pour aligner les images. Ces transformations sont généralement utilisées lorsque les images présentent des différences géométriques simples (figure 2.7b). Le recalage rigide est rapide et efficace, mais il ne permet pas de prendre en compte les déformations locales nécessaires pour corriger des différences plus complexes. Le recalage non rigide, également appelé recalage élastique ou déformable, permet de capturer des déformations locales plus complexes entre les images. Il utilise des transformations non linéaires, qui peuvent être représentées par des champs vectoriel de déformation (DVF, *deformation vector field*), pour modéliser les variations spatiales et les déformations locales par voxel entre les images (figure 2.7c) [44]. Les méthodes de recalage conventionnelles utilisent des algorithmes d'optimisation pour estimer les paramètres de transformation qui minimisent une mesure de similarité entre les images.

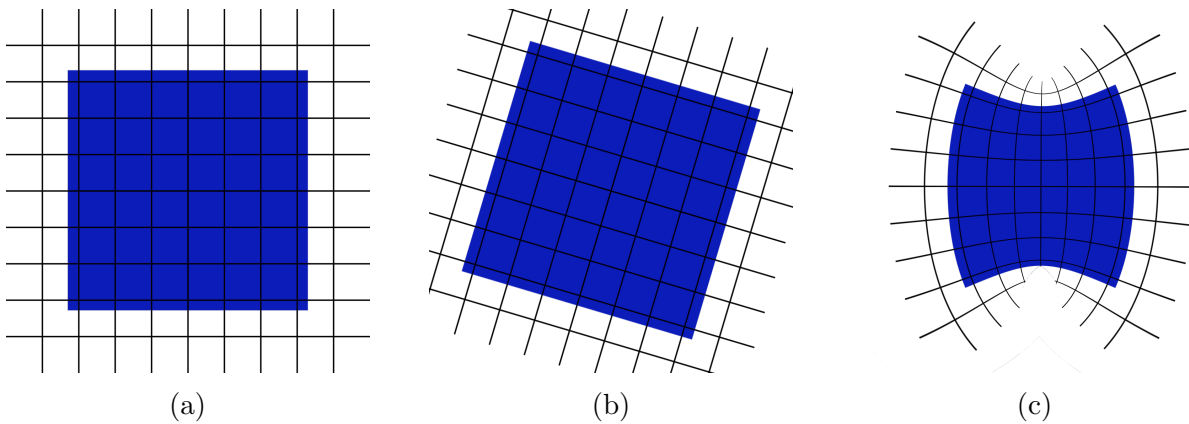


FIGURE 2.7 – Illustration de différents types de transformations : (a) Image source ; (b) Transformation rigide ; (c) Transformation non rigide.

Le processus de recalage consiste à trouver une transformation géométrique entre deux images qui permet d'aligner ces images avec des méthodes rigides, non rigides ou la combinaison des deux [43]. Soient  $I_F$  une image 3D fixe,  $I_M$  une image 3D mobile et  $T$  la transformation spatiale qui fait correspondre les voxels de  $I_M$  à ceux de  $I_F$ . Le

recalage d'image peut être formulé comme un problème d'optimisation, où l'objectif est de trouver la meilleure transformation  $T$  qui maximise  $\mathcal{S}()$ , une métrique de similarité (ou de dissimilarité) entre  $I_F$  et l'image mobile déformée  $T(I_M)$  :

$$\hat{T} = \arg \min_T \mathcal{S}(I_F; T(I_M)) \quad (2.5)$$

L'objectif du processus de recalage consiste donc à itérativement améliorer les estimations de la transformation  $T$  de manière à optimiser la fonction de coût définie par  $\mathcal{S}()$ . Pour ce faire, nous cherchons à ajuster les paramètres de transformation de manière à minimiser la différence entre  $I_F$  et  $T(I_M)$  et cible. Cette différence est quantifiée par une mesure de similarité qui évalue à quel point les deux images sont alignées. Parmi les mesures de similarité couramment utilisées, plusieurs métriques sont populaires, dont la corrélation croisée [45] et les mesures basées sur l'information mutuelle [46]. Ces métriques sont conçues pour quantifier la similarité entre les intensités des voxels, les contours ou les caractéristiques structurelles des images [47], [48].

De manière traditionnelle, les algorithmes de recalage d'images médicales se composent de trois éléments distincts : un modèle de transformation, une mesure de similarité et un algorithme d'optimisation (figure 2.8) [49]. Le processus global de recalage peut être résumé en 5 étapes :

1. Choisir un modèle de transformation approprié (rigide, affine ou non rigide) et initialiser ses paramètres associés ;
2. Utiliser le modèle de transformation pour déformer  $I_M$  ;
3. Évaluer la différence entre  $T(I_M)$  et  $I_F$  ;
4. Mettre à jour les paramètres du modèle de transformation en optimisant la fonction de coût formulée à partir de la métrique de similarité, en utilisant un algorithme d'optimisation approprié ;
5. Répéter les étapes (2) à (4) jusqu'à ce qu'un critère de convergence satisfaisant soit atteint (généralement basé sur le changement de la métrique de similarité ou des paramètres de transformation entre les itérations) ;

Parmi les méthodes traditionnelles les plus usuelles, nous pouvons citer les approches *multiresolution* [50]-[52], le logiciel *statistical parametric mapping* [53], les méthodes *free-form deformations* par *B-spline* [54], la méthode des *Démons* de Thirion [55] ou encore les méthodes impliquant des transformations difféomorphiques, qui préservent la topologie, avec par exemple, le *Large Diffeomorphic Distance Metric Mapping* [56], [57] et

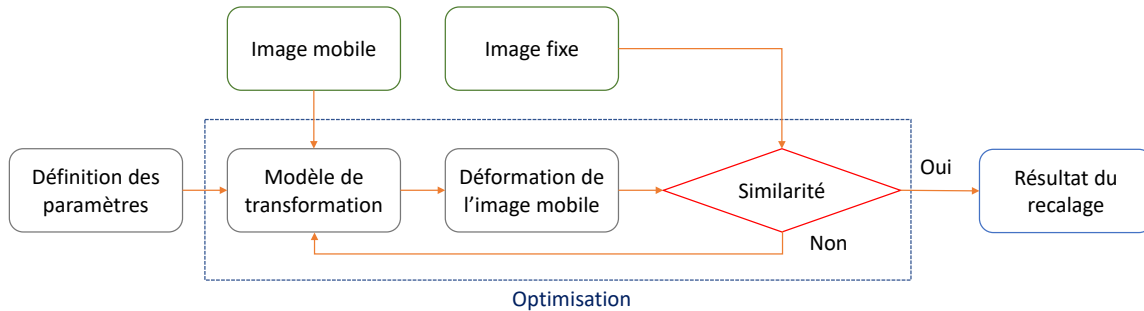


FIGURE 2.8 – Organigramme du fonctionnement du recalage d’images médicales. Modifié de : [49].

les *diffeomorphic demons* [58]. Toutes ces méthodes, qui ne reposent pas sur l’apprentissage, optimisent une fonction d’énergie pour chaque paire d’images. Plusieurs algorithmes d’optimisation sont couramment employés dans ce contexte, parmi lesquels nous pouvons citer la descente de gradient [59], [60], la méthode de Powell sans gradient [61], [62] et la recherche stochastique [63]. Lors de ce processus d’optimisation, il est également essentiel de garder un contrôle sur le recalage à l’aide de termes de régularisation. Ces termes de régularisation sont utilisés pour prévenir des transformations non physiques ou irréalistes. Parmi les termes de régularisation les plus connus, nous retrouvons la régularisation  $L_2$  [64], la régularisation variation totale (TV, *total variation*) [65], la régularisation par le Laplacien [66] et la régularisation anisotropique [67]. Cependant, en général, ces algorithmes d’optimisation engendrent des temps de calcul longs pour chaque paire d’images, ce qui se traduit par des recalages lents, de l’ordre de plusieurs minutes, même pour des ordinateurs actuels, pour les dimensions typiquement rencontrées en imagerie médicale TDM.

Pour améliorer les performances, des méthodes basées apprentissage par DL sont explorées. Ces méthodes révolutionnent l’approche traditionnelle en utilisant des réseaux de neurones préalablement entraînés pour réaliser le recalage. Contrairement aux méthodes conventionnelles, les méthodes basées sur le DL n’optimisent pas itérativement les transformations image par image. Au lieu de cela, elles exploitent la capacité des réseaux de neurones à apprendre des représentations abstraites et à généraliser à partir d’un ensemble d’apprentissage. Une fois le réseau entraîné, il peut être utilisé pour recalage de nouvelles paires d’images en utilisant des transformations similaires à celles apprises lors de l’entraînement. Cela permet d’obtenir des résultats de recalage avec des temps d’inférence beaucoup plus courts qu’en recalage conventionnel, ce qui est très avantageux pour les

applications en temps réel ou pour le traitement rapide de grands ensembles de données. Le DL a su démontrer de bonnes performances dans la segmentation d’images médicales. Cependant, en ce qui concerne le recalage, bien qu’en plein essor ces dernières années [68], les résultats n’ont pas encore atteint le niveau escompté, et cette problématique fait l’objet d’une exploration active dans la communauté.

### 2.2.2 Recalage par deep learning non supervisé

Dans le recalage non supervisé, le modèle DL apprend à aligner les images sans nécessiter la connaissance de déformations en tant que labels vérité terrain, de manière similaire au recalage conventionnel. Le modèle est ainsi capable d’apprendre les caractéristiques intrinsèques des images et de les utiliser pour estimer la transformation, en minimisant une fonction de coût de similarité. Les méthodes de recalage non supervisées permettent d’éviter de restreindre le réseau à apprendre certains types de transformation. Cependant, il est difficile de définir une fonction de perte appropriée du réseau sans reposer sur des transformations vérité terrain de référence, se heurtant aux mêmes contraintes que le recalage conventionnel, avec notamment l’imposition de contrainte sur la lisseté des champs produit par des mesures de régularité de type  $\ell_2$ . Ces méthodes reposent essentiellement sur la notion de réseau transformateur spatial (STN, *spatial transformer network*), des opérateurs différentiables (et donc permettant la rétropropagation du gradient) de transformation spatiale composée d’un réseau de localisation (*localisation net*), d’un générateur de grille (*grid generator*) et d’un échantillonneur (*sampler*) (figure 2.9). Les STNs ont inspiré de nombreuses méthodes non supervisées [69], [70].

Le réseau de localisation analyse l’image d’entrée et estime les paramètres nécessaires pour réaliser une transformation spatiale. Ces paramètres englobent les informations concernant la transformation spatiale à appliquer requis pour aligner l’image d’entrée sur une référence spécifique. Ensuite, le générateur de grille crée une grille de coordonnées basée sur les paramètres prédits correspondant à un ensemble de point de l’image d’entrée qui doit être échantillonné pour produire une image transformée. Cette grille est ensuite utilisée pour extraire des échantillons de l’image d’entrée. Enfin, l’échantillonneur prend en entrée l’image d’entrée et la grille générée pour appliquer la transformation spatiale à l’image d’entrée. Ce processus d’application d’une transformation à une image est connu sous le nom de déformation, *warping* en anglais. Ce module étant différentiable, l’image déformée en sortie peut être utilisée pour le calcul des fonctions de coût et l’optimisation du réseau par rétropropagation du gradient. Ainsi, de manière similaire aux méthodes

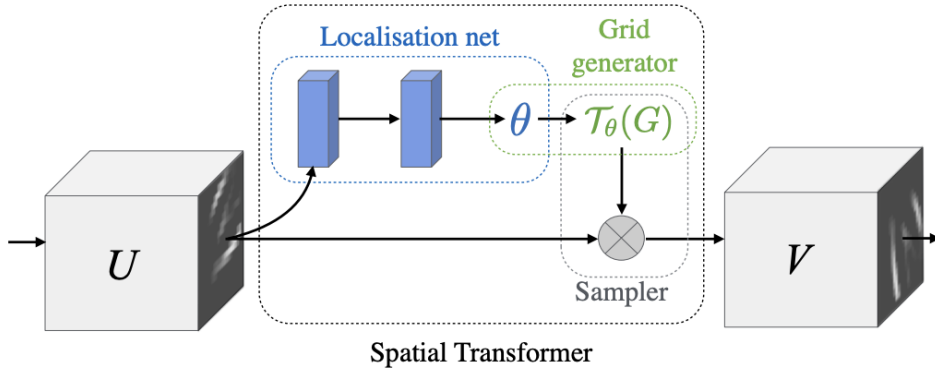


FIGURE 2.9 – Architecture d’un réseau transformateur spatial. L’entrée  $U$  est transmise à un réseau de localisation qui calcule les paramètres de transformation  $\theta$ . Le générateur de grille génère une grille d’échantillonnage  $T_\theta(G)$ , qui est appliquée à  $U$ , produisant ainsi la sortie déformée  $V$ . La combinaison du réseau de localisation et du mécanisme d’échantillonnage définit un transformateur spatial. Source : [69]

conventionnelles de recalage basées sur la similarité entre les intensités des voxels, le réseau DL apprend à recalibrer des images en optimisant une métrique de similarité entre une image fixe et l’image déformée par le réseau. L’avantage de l’approche DL est la capacité au réseau de réaliser un recalage d’images de manière non itérative lors de l’inférence sur des nouvelles données, réduisant le temps nécessaire à la tâche.

Un réseau typique de prédiction de transformation non supervisée pour le recalage prend une paire d’images en entrée et produit directement un DVF, utilisé par le STN pour l’appliquer sur l’image mobile et générer ainsi une image déformée. Les images déformées sont ensuite comparées aux images fixes pour calculer une fonction de coût basée sur la similarité des images (figure 2.10).

Balakrishnan et coll. proposent VoxelMorph, un réseau de recalage basé sur un U-Net qui s’est imposé comme la référence (*baseline*) en recalage par des méthodes d’apprentissage non supervisées[71]. Lors de l’entraînement, le réseau pénalise les différences d’apparence des images fixes et des images transformées par le DVF généré par le réseau. Une contrainte de régularité est utilisée pour pénaliser les variations spatiales locales dans la transformation prédite. Par la suite, le modèle est amélioré en donnant la possibilité d’exploiter les segmentations auxiliaires disponibles dans les données d’entraînement [70], permettant une certaine forme de supervision. Une fonction de coût de segmentation, le coefficient de Dice (DSC, *Dice similarity coefficient*), y est ajoutée aux fonctions de

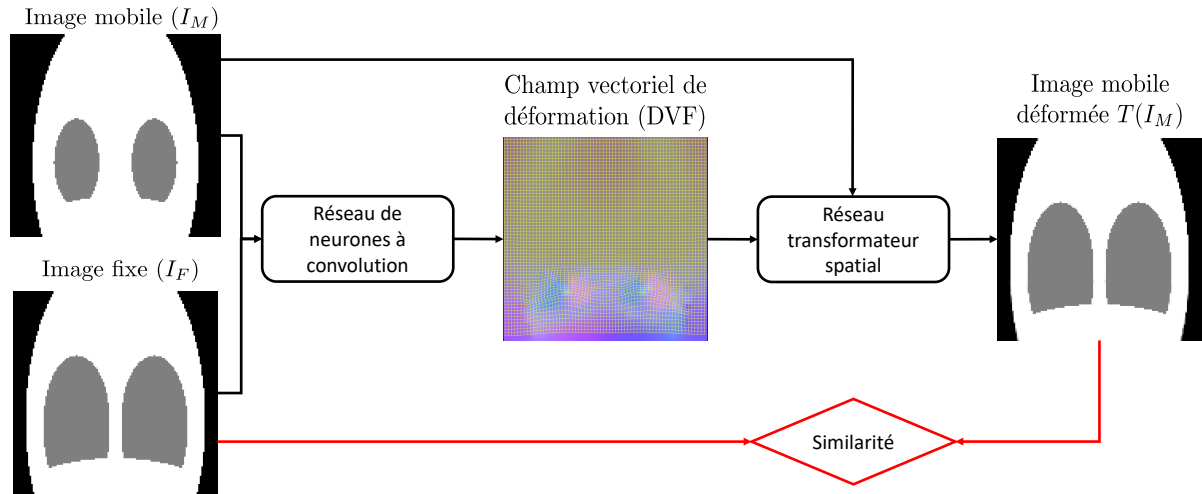


FIGURE 2.10 – Exemple de recalage non supervisé de deux images synthétiques de poumons.

coût d'origine lors de l'étape d'entraînement. En utilisant la fonction de coût DSC, le modèle est encouragé à produire des prédictions de segmentation qui sont en accord avec les véritables régions d'intérêt. Cela favorise la segmentation précise des objets malgré le déséquilibre de classe souvent présent dans le domaine médical [72]. Le modèle Voxel-Morph est comparé avec le modèle non supervisé, avec et sans la fonction de coût DSC. Leurs résultats montrent ainsi que la fonction de coût de segmentation peut contribuer à améliorer les résultats, et démontre ainsi l'intérêt d'une supervision partielle dans les méthodes non supervisées. Pour citer un exemple appliqué au recalage d'image TDM 4D de poumon, Fu et coll. présentent une méthode appelée LungRegNet [73] basée sur une architecture GAN. Le modèle est basé sur une approche en deux étapes : un réseau CoarseNet pour effectuer un recalage global sur une version sous échantillonnée de l'image dans un premier temps, puis un réseau FineNet basé sur des patchs pour affiner le recalage à un niveau plus local. Le discriminateur est présent pour assurer la cohérence des images déformées et améliorer l'exactitude du recalage, soulignant l'intérêt des coûts GAN non supervisés pour ce type d'objectifs.

### 2.2.3 Recalage par deep learning supervisé

Dans le recalage supervisé, un modèle DL est entraîné à apprendre une fonction de transformation qui permet d'aligner les images en utilisant des transformations de références typiquement représentées sous la forme d'un DVF. Ces transformations sont

généralement connues et associées à des paires fixes et mobiles. Un réseau typique de prédiction de transformation supervisée pour le recalage prend une paire d'images en entrée et produit un DVF qui est comparé avec une transformation obtenue par une méthode de recalage conventionnelle (figure 2.11).

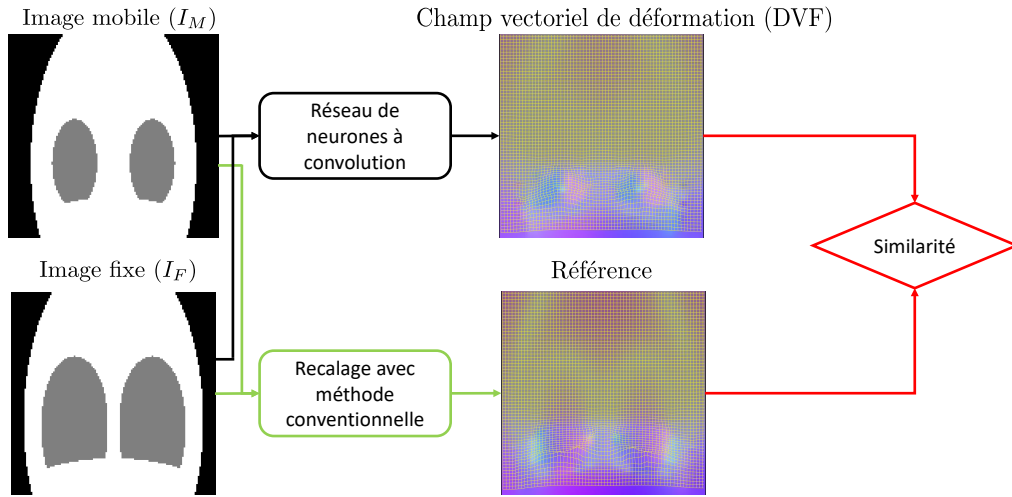


FIGURE 2.11 – Exemple de recalage supervisé de deux images synthétiques de poumons.

Dans leur étude, Sentker et coll. utilisent des DVFs issus de méthode de recalages d'images traditionnels tels que *PlastiMatch* [74], *NiftyReg* [75] et *VarReg* [76] comme référence pour un entraînement supervisé [77]. La fonction de coût utilisée pour l'entraînement du réseau de recalage d'images pulmonaires en 3D est basée sur l'erreur quadratique moyenne (MSE, *mean squared error*) entre les DVFs prédits par le réseau et les DVFs obtenus par ces méthodes conventionnelles. Le modèle est ensuite évalué avec des données du challenge DIR-Lab [78]. Les challenges sont des compétitions ouvertes auxquelles participent des chercheurs, des scientifiques et des professionnels de la santé pour résoudre des problèmes spécifiques. Ces défis sont organisés pour diverses raisons telles que la stimulation de la recherche ou la comparaison des performances avec une base commune d'évaluation. L'objectif du challenge DIR-Lab est de parvenir au meilleur recalage entre deux phases maximales du cycle respiratoire d'une acquisition TDM 4D. Pour cela, il y a à disposition des données TDM 4D publiques et une liste de coordonnées de repères anatomiques qui ont été identifiés manuellement par un expert en imagerie thoracique. Les ensembles de points servent de référence pour évaluer l'exactitude spatiale de la déformation d'images à l'intérieur des poumons pour chaque cas. Pour quantifier

la qualité du recalage, l'erreur de recalage de la cible (TRE, *target registration error*) est calculée. Il s'agit de mesurer la distance entre les points de repères anatomiques des différentes phases respiratoires après recalage. L'erreur moyenne pour chaque repère est ensuite utilisée comme valeur d'évaluation finale. Ainsi, le modèle de recalage supervisé avec la méthode VarReg comme référence obtient la TRE la plus faible par rapport à PlastiMatch et NiftyReg. Les résultats démontrent que la méthode de recalage basée sur les réseaux de neurones convolutifs est comparable au recalage traditionnel en termes de TRE mais avec une réduction importante en temps de calcul.

Fan et coll. introduisent BIRNet, un modèle de recalage d'images cérébrales avec une double supervision [79]. Pour obtenir les transformations réelles, BIRNet utilise les méthodes de recalage *Advanced normalisation tools* (ANTs) [80] et *Local correlation coefficient* (LCC-Demons) [81]. Pour résoudre le problème du manque de données de référence, BIRNet utilise à la fois des DVFs, préalablement obtenus par des méthodes de recalage conventionnelles, qui sont comparés à ceux générés par le réseau. Autrement dit, l'image fixe et l'image mobile déformée par un STN (non supervisé) guident la phase d'entraînement, permettant ainsi au modèle d'améliorer les résultats.

Dans le cas du recalage des différentes phases des images TDM 4D, Teng et coll. proposent une méthode capable de générer des DVFs entre les phases à partir des paires de patchs d'images en entrée du réseau [82]. Le réseau est entraîné en utilisant la MSE comme fonction de coût, afin de reproduire les champs produit par le logiciel Velocity qui se base sur des déformations de type *B-spline* et la maximisation de l'information mutuelle. Les résultats montrent que le réseau a atteint une précision comparable à celle de Velocity. La méthode proposée a l'avantage d'être plus rapide et entièrement automatique, sans dépendance de l'utilisateur, ce qui la rend préférable pour des applications cliniques, selon les auteurs.

## 2.3 Estimation de mouvement

L'estimation de mouvement en radiothérapie est une étape cruciale dans la prise en charge des patients atteints de cancer pour certaines localisations. Lors du traitement par radiothérapie, les tumeurs et les OARs peuvent subir des mouvements internes en raison de la respiration ou d'autres facteurs physiologiques. Ces mouvements peuvent entraîner une incertitude dans la délivrance précise de la dose de radiation à la tumeur cible et peuvent également exposer les organes sains environnants à des doses indésirables [7].



Pour assurer l'efficacité et la sécurité du traitement, il est essentiel de connaître ces mouvements. L'estimation de mouvement en radiothérapie vise à déterminer la position et la forme des tumeurs et des OARs permettant ainsi d'adapter le plan de traitement en conséquence. Avec l'aide des méthodes en vision par ordinateur, il est possible de proposer des solutions qui aideront à estimer le mouvement interne réel.

### 2.3.1 Introduction à la synthèse de vidéo

Compte tenu du contexte applicatif de notre thèse en imagerie TDM 4D, nous évoquons également ici des approches de synthèse vidéo qui partagent des objectifs voisins.

La synthèse de vidéos se concentre sur la création de contenus vidéo plutôt que d'images statiques. Une catégorie courante de la synthèse vidéo est celle de la synthèse de vidéo conditionnelle, qui repose sur une autre vidéo en tant qu'entrée du réseau. Une telle configuration est souvent appelée synthèse vidéo à vidéo [83]. Les travaux liés à cette catégorie ont montré des résultats prometteurs dans diverses tâches, telles que l'animation de personnages avec de nouvelles expressions ou de nouveaux mouvements [84], [85]. Par exemple, des travaux sont réalisés pour transférer les pas de danse d'une personne d'une vidéo à une autre [86]. Ces travaux utilisent un réseau pré-entraîné pour générer des poses à partir de la première vidéo, puis un autre réseau apprend à les transférer à la vidéo cible, créant ainsi une vidéo dans laquelle la personne cible exécute les pas de danse.

Une autre catégorie est celle de la prédiction des images futures, c'est-à-dire que l'objectif est de prédire la prochaine image d'une séquence en se basant sur les images passées [87], [88]. Toujours dans le domaine de transfert de pose, Walker et coll. combinent plusieurs architectures de réseaux de neurones afin de générer une vidéo d'une personne dansante à partir de seulement d'une photo de la personne [89]. Leur modèle prédit d'abord les différentes poses, puis les combine avec la photo de la personne pour obtenir la vidéo finale.

En outre, il est également possible de conditionner la génération de vidéos par d'autres types de données, tels que des signaux audio, qui sont donc des informations unidimensionnelles. Ce type de génération permet par exemple de faire parler ou chanter des visages ou encore faire danser des personnes au rythme d'une musique [90]-[92]. Vougioukas et coll. utilisent un GAN pour animer un visage à partir d'un discours audio [93]. Le générateur prend en entrée un signal audio et une image, puis génère plusieurs images consécutives pour synthétiser la vidéo. Deux discriminateurs sont utilisés : le premier fonctionne image par image pour déterminer si chaque image est réelle ou non, tandis que le deuxième

travaille sur l'ensemble de la séquence pour évaluer la cohérence globale de la vidéo.

Ces différentes approches témoignent des avancées récentes en matière de synthèse de vidéos conditionnelles, qui ouvrent la voie à de nombreuses applications, notamment dans le domaine de l'imagerie dynamique.

### 2.3.2 Exemple en imagerie médicale

Grâce aux méthodes de DL dans le domaine de vision par ordinateur, de multiples applications se sont développées dans les dernières années en imagerie dynamique pour répondre à des problèmes de temps d'exposition, de radiation ou encore de prédiction de mouvement dans le cadre de la radiothérapie.

Une première approche 2D par Jaouen et coll. est proposée afin de générer différentes phases respiratoires avec l'aide de plusieurs réseaux pix2pix [94]. Chaque réseau est entraîné à générer une phase spécifique à partir d'une phase respiratoire spécifique et le coût adversaire permet d'améliorer la qualité visuelle des images des phases générées. Les travaux démontrent ainsi qu'avec une architecture simple, il est possible de générer une vidéo directement à partir d'une image. Mori et coll. proposent une méthode similaire en 3D basée sur un auto-encodeur avec la particularité d'avoir autant de décodeurs que de phases à reconstruire [95]. Une spécificité de la méthode est la partie indirecte de la synthèse d'image. En effet, chaque décodeur génère un DVF pour chaque phase, qui est ensuite utilisé pour déformer l'image initiale. Cette méthode aboutit finalement à la synthèse d'images spécifiques à chaque phase, permettant ainsi de reconstruire une TDM 4D après les déformations. D'autres travaux se concentrent aussi sur la génération d'images TDM 4D de manière indirecte. Similaire à de la synthèse de vidéo, Duetschler et coll. proposent une méthode permettant de générer des DVFs à partir des phases respiratoires d'une acquisition IRM 4D pour ensuite les appliquer à une image TDM 3D pour obtenir un pseudo TDM 4D [27]. D'une autre manière, Chang et coll. proposent une méthode basée sur un GAN et un régresseur latent en utilisant pour générer des fantômes 4D-XCAT (*4D Extended Cardiac – Torso*) réalistes [96], [97]. Avec l'aide d'un modèle de mouvement basé sur de l'apprentissage automatique, le modèle parvient à générer des fantômes 4D-XCAT avec des mouvements respiratoires réalistes et contrôlables par l'utilisateur. Une limitation majeure des méthodes présentées est le manque de personnalisation des phases générées par rapport au mouvement respiratoire réel du patient. Ces méthodes reposent essentiellement sur la reproduction d'un modèle de mouvement préalablement appris au sein d'une population donnée, à savoir celle du jeu de données utilisé.

Pour prédire un mouvement se rapprochant à du temps réel, Romaguera et coll. présentent une méthode prédictive non supervisée pour générer des volumes d'IRM 4D de foie en apprenant une distribution de DVF sur un ensemble de données de population [98]. Le modèle utilise des images 2D et un volume 3D statique préopératoire afin de conditionner leur modèle. Simultanément, un mécanisme temporel inspiré par l'architecture seq2seq agit sur les images 2D pour produire des représentations extrapolées dans le temps. Outre la synthèse basée sur une population donnée, la méthode permet aussi de personnaliser le modèle avec des IRM 4D spécifique au patient si disponible et ainsi améliorer la prédiction. Une autre approche décrite par Huang et coll. permet de générer des TDM 4D de poumons à partir des composantes principales extraites de DVFs et des informations respiratoires extraites de la surface du patient [99]. Ces champs sont calculés à partir des TDM 4D acquises lors de la planification du traitement et avec une méthode de recalage conventionnelle [65]. Toutefois, ces méthodes spécifiques à un patient ne sont possibles que lorsque des acquisitions TDM 4D sont disponibles pour personnaliser le réseau à l'inférence.

Plus récemment, Sang et Ruan réalisent un modèle de mouvement respiratoire continu en 4D à partir d'une paire d'images de fin d'expiration et de fin d'inspiration. À l'aide d'une variable conditionnelle, le réseau peut générer des interpolations intermédiaires à l'aide de DVFs entre les deux phases, permettant ainsi un contrôle précis [100]. La méthode permet ainsi d'obtenir des phases intermédiaires spécifiques grâce aux phases extrêmes d'un patient. Toutefois, la variable utilisée est abstraite, sans réalité physique, et ne peut donc pas représenter des irrégularités respiratoires si présentes dans le cycle respiratoire. Lee et coll. proposent RMSim, un simulateur de mouvement respiratoire basé sur une architecture seq2seq 3D qui apprend à partir d'images TDM 4D et prédit les phases respiratoires futures à partir d'une image TDM 3D. Le modèle RMSim est conditionné par des données respiratoires 1D afin de mieux contrôler la respiration du patient [101]. Ces données respiratoires reposent sur le suivi du mouvement d'un point placé manuellement à l'apex du diaphragme. Par conséquent, en pratique clinique, il est difficile d'obtenir ces informations sans avoir effectué des acquisitions TDM 4D.

## 2.4 Conclusion

Les méthodes de synthèse par DL issues du domaine de la vision par ordinateur ont démontré leur efficacité dans de nombreux cas en imagerie médicale. Néanmoins, à l'heure

actuelle, peu d'études explorent le potentiel de ces méthodes dans le domaine de l'imagerie dynamique où la composante temporelle ne peut être ignorée, la plupart des approches se concentrant sur des modalités d'imagerie statique. Pourtant, ces méthodes de synthèse offrent des perspectives prometteuses, notamment pour l'amélioration du protocole en radiothérapie en facilitant par exemple l'estimation de mouvement. Dans la suite de ce travail de thèse, nous formulons l'hypothèse qu'il est possible de générer des acquisitions TDM 4D synthétiques à partir de données TDM 3D sur la base d'une architecture GAN conditionnée par un signal de respiration externe. Ces TDM 4D synthétiques pourraient alors être exploitées pour améliorer la précision du diagnostic et la planification des traitements, en se passant idéalement d'acquisitions TDM 4D réelles.

Dans le chapitre 3 suivant, nous présentons une première méthode pour générer des acquisitions TDM 4D synthétiques à l'aide de GAN, exploitant relativement directement les architectures I2I existantes, que nous nommons méthode directe. Le chapitre 4 traite d'une approche plus élaborée, indirecte, cœur de notre travail de thèse.



# GÉNÉRATION D'IMAGES DYNAMIQUES : SYNTHÈSE DIRECTE PAR CONDITIONNEMENT RESPIRATOIRE SCALAIRE

## Résumé

Dans ce chapitre, nous proposons une première méthode de synthèse « directe » afin de générer des pseudo acquisitions TDM 4D. Le terme « directe » vient du fait que nous exploitons une architecture I2I pour générer directement des images des différentes phases respiratoires à partir d'une phase unique. Nous faisons le choix d'une architecture I2I conventionnelle en incorporant un coût adversaire GAN à un backbone U-Net. Pour conditionner la synthèse par l'amplitude respiratoire, nous proposons d'exploiter un mécanisme de normalisation d'instance adaptative (*AdaIN*, *adaptive instance normalization*) au niveau du goulot d'étranglement de l'architecture convolutive. Le niveau de résolution et de réalisme insuffisant nous conduisent à proposer un recalage de la phase initiale vers la phase générée pour préserver des caractéristiques importantes de l'image. Après avoir décrit l'approche de synthèse, nous validons la méthode sur des objets synthétiques et sur des données cliniques TDM 4D. Nous montrons et comparons les résultats de mouvements respiratoires synthétiques à des acquisitions réelles.

Nous mentionnons qu'une présentation de ce travail préliminaire a été sélectionné parmi les 8 meilleurs *student papers* de la conférence IEEE NSS/MIC 2021 (Yokohama, Japon).

## 3.1 Introduction

En imagerie médicale, il existe peu de travaux basés sur des approches I2I afin de générer une dynamique respiratoire. Pourtant, la synthèse d’images dynamiques à partir d’une image statique à l’aide d’un réseau de neurones peut s’entendre comme une application (mapping) d’une phase respiratoire à l’instant  $t$  vers d’autres temps. Ainsi, en considérant un instant  $t$  comme un domaine source et les autres temps comme des domaines cibles, il est possible de proposer une méthode basée sur des approches I2I conventionnelles.

Les réseaux I2I tels que pix2pix [4] démontrent des capacités remarquables de synthèse d’images, notamment avec des images synthétisées réalistes grâce au coût adversaire. Parmi ces réseaux, de nombreux modèles I2I récents ont été développés pour générer des images dans les multiples domaines cibles avec, par exemple, la manipulation des attributs d’images synthétisées. Ces modèles permettent notamment un contrôle explicite du style en apprenant des attributs spécifiques [102]-[104], ainsi qu’une gestion précise de l’interpolation entre différents domaines dans l’espace latent [32], [105]. Ces modèles proposent un conditionnement basé sur des images de référence afin de transférer le style d’une image à une autre. D’autres approches conditionnent leurs réseaux avec des valeurs simples et des modèles physiques. Une architecture proposée par Pizzati et coll. introduit des transitions continues contrôlées [33]. Elle permet d’interpoler les images dans un domaine à l’autre en conservant le réalisme des images à l’aide de modèles physiques simples.

Dans un contexte de réduction de la dose en radiothérapie, il est essentiel qu’une approche de synthèse respiratoire soit spécifique à chaque patient. Dans ce chapitre, nous introduisons une nouvelle architecture de réseau I2I à doubles conditionnements pour synthétiser des acquisitions TDM 4D. Cette application se place dans le cadre de la radiothérapie interventionnelle pour le cancer du poumon. Les contributions majeures de ce premier travail sont :

- Nous proposons de contrôler la génération avec un conditionnement double, l’image TDM 3D et une mesure scalaire basée sur l’amplitude respiratoire du patient concerné. Cela permet d’obtenir des phases spécifiques en fonction de la respiration du patient. Nous montrons que cette contrainte scalaire permet de générer des phases respiratoires arbitraires.
- Nous définissons une approche de synthèse directe pour générer des images de poumons à différentes phases respiratoires et utilisons les avantages d’une architecture GAN pour améliorer les images générées. Cette approche est couplée à un recalage

conventionnel afin de garder un contrôle explicite des déformations sur l'image générée. Nous montrons que les images générées par cette approche présentent un réalisme satisfaisant au regard d'un certain nombre de métriques TDM 4D.

Nous montrons ainsi que l'approche permet de générer des dynamiques respiratoires en adéquation avec des informations cliniques extérieurs faciles à obtenir (ceinture externe).

Dans la suite de ce chapitre, nous décrivons cette première méthode de synthèse. Nous détaillons le contenu des deux jeux de données utilisés pour les expérimentations : données synthétiques et données cliniques réelles. Nous validons notre méthode qualitativement et quantitativement sur ces jeux de données.

## 3.2 Méthode de synthèse directe

Dans cette section, nous expliquons notre méthode de génération de mouvement respiratoire spécifique à un patient. Nous décrivons tout d'abord la méthode de conditionnement et une stratégie de pondération de la fonction de coût pour mieux contraindre le réseau sur les zones d'intérêt. Un schéma général de la méthode proposée est montré en figure 3.1.

### 3.2.1 Conditionnement additionnel par injection et fonction de coût focalisée aux zones d'intérêt

#### Formalisme

Les images TDM 4D sont obtenues par groupement par phase (*gating*), produisant ainsi une image dynamique  $\mathcal{J}^* : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$ , où  $\mathcal{J}^*(x, g)$  correspond à la valeur de l'image en position  $x$  pour la phase  $g \in \{1, 2, \dots, N_g\}$ .

Soient  $\mathcal{I}(x)$  une image TDM 3D de référence, où  $x \in \mathbb{R}^3$  est le domaine de l'image et  $\mathcal{S}(\mathcal{I})$  la segmentation de la tumeur. En utilisant une architecture I2I, notre objectif est d'apprendre une mise en correspondance implicite  $\varphi_g(x, \alpha, \mathcal{S})$ , qui est conditionnée par trois différents termes : 1) l'image  $\mathcal{I}$ , 2) sa segmentation de tumeur  $\mathcal{S}(\mathcal{I})$  et 3) une valeur  $\alpha$ . Cette mise en correspondance doit permettre à l'image générée  $\mathcal{I} \circ \varphi_g(x, \alpha, \mathcal{S}) = \mathcal{J}^*(x, \alpha)$  de partager la même dynamique respiratoire que l'image réelle  $\mathcal{J}(x, g)$ . La valeur  $\alpha$  est une mesure d'amplitude respiratoire scalaire, qui dans le cas présent sera identifié à la différence de volume pulmonaire entre deux phases respiratoires.



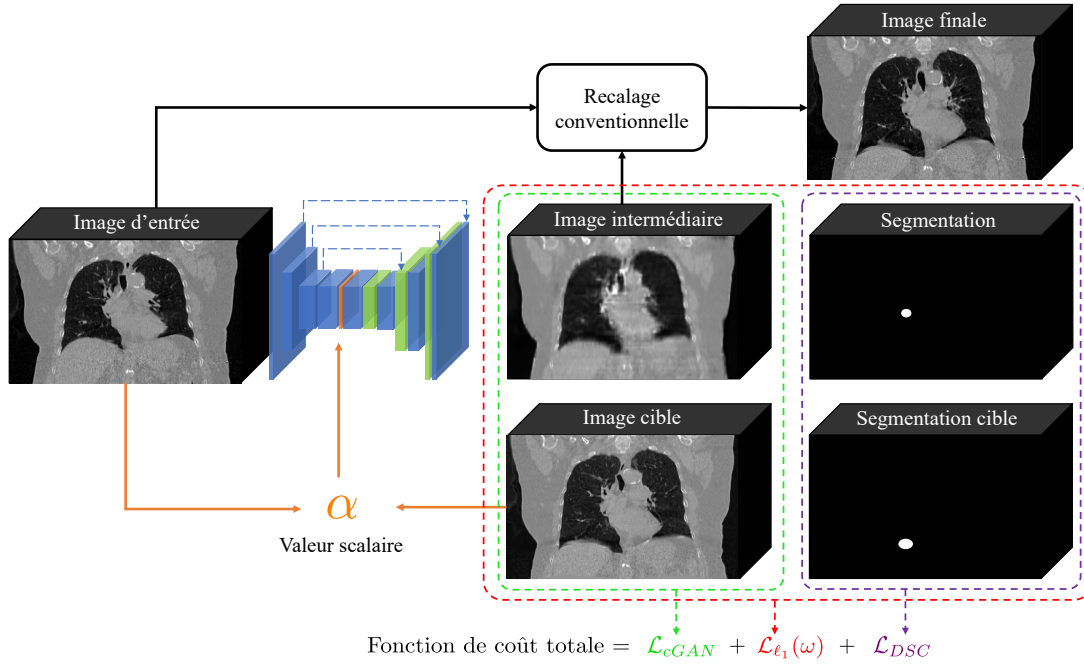


FIGURE 3.1 – Fonctionnement global de l’approche de synthèse directe. Un modèle est doublement conditionné par 1) une image fixe de phase déterminée et 2) l’amplitude  $\alpha$  correspondant calculée entre deux phases respiratoires. À l’inférence, un algorithme de recalage conventionnelle *B-spline* est utilisé pour recalibrer l’image d’entrée, considérée comme image mobile, vers l’image générée par le modèle, considérée comme image fixe. Cela permet de contraindre le réalisme de la déformation et d’obtenir une image finale de meilleure résolution.

### Conditionnement scalaire

Notre modification principale dans l’architecture GAN réside dans le générateur  $G$ , auquel nous intégrons une condition supplémentaire en plus de l’image  $\mathcal{I}$ . Pour pouvoir conditionner notre modèle avec l’amplitude respiratoire scalaire  $\alpha(t)$ , nous ajoutons une couche de neurones au niveau du goulot d’étranglement du générateur U-Net afin de pouvoir y injecter  $\alpha$  dans l’espace latent (figure 3.1). Cette couche de neurones est une version modifiée de la normalisation d’instance adaptative (*AdaIN*, *adaptive instance normalization*) [106] habituellement utilisée dans les tâches de transfert de style pour la synthèse d’images. Cette couche permet de conditionner la génération d’une image en lui transférant le style d’une image de référence à une image cible (figure 3.2). Dans notre cas, l’image de référence est remplacée par notre valeur scalaire d’amplitude  $\alpha$ . Elle est ainsi définie par :

$$\text{AdaIN}(x, \alpha) = \gamma(\alpha) \left( \frac{x - \mu(x)}{\sigma(x)} \right) + \beta(\alpha), \quad (3.1)$$

où  $\gamma$  et  $\beta$  sont des paramètres estimés à partir de la valeur  $\alpha$  à travers plusieurs couches de neurones. Ce mécanisme AdaIN permet donc une normalisation spécifique à l'amplitude respiratoire et qui permet ainsi au réseau de discriminer chaque phase de manière implicite.

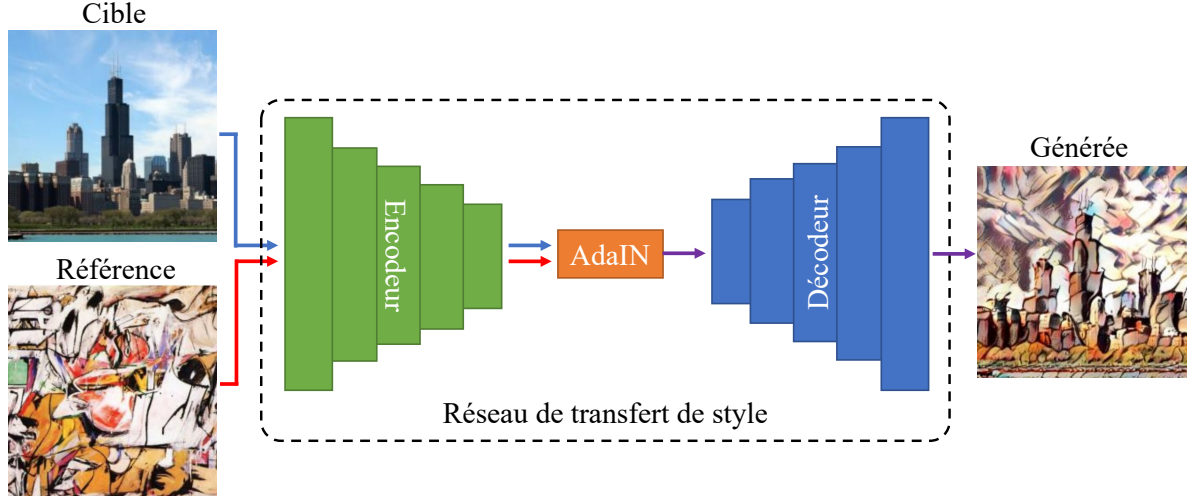


FIGURE 3.2 – Aperçu du fonctionnement de la couche AdaIN initiale. Le contenu de l'image de référence et le style de l'image cible sont tout d'abord encodés en des vecteurs latents. La couche AdaIN effectue le transfert de style dans l'espace latent en adaptant le premier vecteur par le deuxième. Le décodeur apprend à reconstruire le vecteur latent modifié en une image. Modifié de : [106]

### Fonction de coût

La fonction du générateur devient  $G : \{\mathcal{I}(x), \alpha\} \rightarrow \mathcal{J}^*(x, g)$ . Dans notre cas, la fonction de coût cGAN se définit comme :

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{\mathcal{I}, \mathcal{J}^*} [\log D(\mathcal{I}, \mathcal{J}^*)] + \mathbb{E}_{\mathcal{I}, \alpha} [\log(1 - D(\mathcal{I}, G(\mathcal{I}, \mathcal{J}^*)))], \quad (3.2)$$

où  $G$  essaie de minimiser cet objectif tandis que  $D$  essaie de le maximiser, c'est-à-dire :

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) \quad (3.3)$$

Dans un objectif de synthèse d'image, nous ajoutons une fonction de coût de reconstruction  $\ell_1$  à notre coût adversaire :

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{\ell_1}(G) \quad (3.4)$$

Pour améliorer le réalisme des mouvements locaux des structures d’intérêts de petite échelle comme les tumeurs, nous proposons également d’ajouter une contrainte avec une fonction de coût pondérée spatialement afin d’inciter le réseau à se concentrer sur ces zones d’intérêts. Pour cela, nous proposons tout d’abord une fonction de coût de reconstruction  $\ell_1$  pondérée, ou  $\ell_1(\omega)$  avec une pondération variable en fonction de la région considérée (tumeur ou tissus environnants), où  $\omega$  est traitée comme un hyperparamètre et dont l’estimation est empirique (figure 3.3).

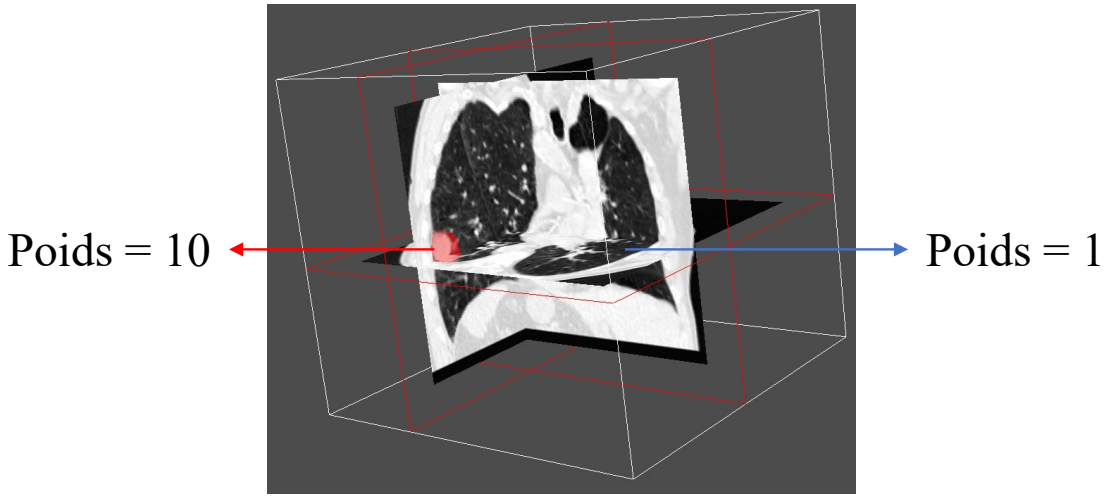


FIGURE 3.3 – Fonctionnement de la fonction de coût pondérée par zone d’intérêt. Le paramètre  $\omega$  aura différentes valeurs en fonction de la zone concernée. Ici, le poids sera plus important au niveau de la tumeur et sera normal dans le reste des poumons.

L’équation de la fonction pondérée est :

$$\mathcal{L}_{\ell_1}(\omega) = \frac{\sum_{i=1}^n |y_{true} - y_{predicted}| \times \omega}{\sum_{i=1}^n \omega}, \quad (3.5)$$

où  $y_{true}$  et  $y_{predicted}$  correspondent aux images de références et générées respectivement.

À cela, nous y ajoutons une fonction de coût basée sur la superposition des segmentations, la fonction DSC, qui va permettre de comparer la segmentation générée par le réseau ainsi qu’à la segmentation de référence  $\mathcal{S}(\mathcal{I})$ . Ainsi notre fonction de coût focalisée sur la tumeur  $TA$  (*tumor-awareness*) composée de  $\mathcal{L}_{\ell_1}(\omega)$  et de la fonction DSC  $\mathcal{L}_{DSC}$  est :

$$TA = \mathcal{L}_{\ell_1}(\omega) + \mathcal{L}_{DSC} \quad (3.6)$$

Notre objectif final peut se définir tel que :

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + TA(G) \quad (3.7)$$

En pratique, le rôle du générateur  $G$  est de prendre en entrée de manière aléatoire une paire composée d'une image  $\mathcal{I}$  et sa segmentation  $\mathcal{S}(\mathcal{I})$  associée et de générer une image cible  $\mathcal{J}^*(x, \alpha)$ , tout en s'appuyant sur la valeur scalaire  $\alpha$  injectée dans le goulot d'étranglement du générateur via le mécanisme d'AdaIN. Parallèlement, le discriminateur  $D$  est entraîné pour distinguer l'image réelle de la phase, c'est-à-dire  $\mathcal{J}(x, g)$ , de l'image synthétisée  $\mathcal{J}^*(x, \alpha)$ . Ainsi, l'entraînement du modèle lui permet de générer un mouvement respiratoire spécifique en se basant sur des images volumétriques TDM 3D tout en utilisant un scalaire correspondant à l'amplitude respiratoire et de garantir un mouvement réaliste au niveau local grâce à la segmentation.

### 3.2.2 Contrôle explicite du recalage

Les images générées issues d'un modèle type GAN peuvent présenter des déformations peu réalistes en raison du manque de contrainte explicite sur la nature de la déformation. Les images générées présentent un niveau de détail relativement faible, avec quelques aberrations locales (figure 3.4).

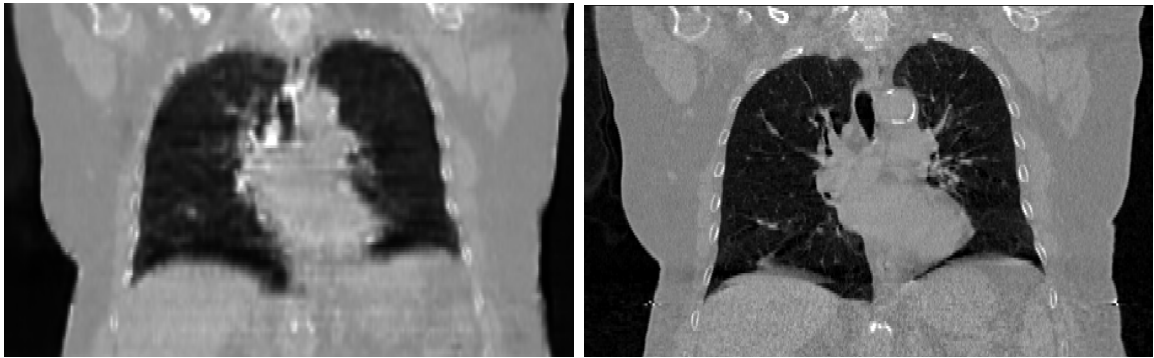


FIGURE 3.4 – Comparaison d'une image générée par le modèle et l'image de référence associée.

Il est donc souhaitable d'imposer une contrainte plus explicite sur cette déformation afin d'améliorer les résultats. Dans cette optique, nous proposons d'utiliser, lors de l'inférence, une méthode de recalage conventionnelle entre l'image d'entrée TDM 3D et la sortie générée par notre modèle afin de régulariser implicitement cette sortie et de prés-

ver les structures locales de la phase initiale. Pour ce faire, nous appliquons une méthode largement utilisée en imagerie TDM 4D basée sur un algorithme de recalage *B-spline* à l’aide de la librairie elastix [107]. Nous cherchons ainsi une transformation paramétrique  $\varphi_\mu^\dagger$  :

$$\hat{\mu} = \arg \min_{\mu} \mathcal{C}(\varphi_\mu^\dagger; \mathcal{J}^*, \mathcal{I}), \quad (3.8)$$

paramétrée par le vecteur de paramètres  $\mu$ , où  $\mathcal{C}$  est la fonction de coût (appelée métrique de recalage) qui définit la qualité de l’alignement.

La métrique utilisée est basée sur le coefficient de corrélation normalisée (*Normalised Correlation Coefficient* – NCC), qui s’avère adapté pour le recalage d’images monomodales et qui peut compenser les variations globales d’intensité dues aux différences de niveau d’inspiration. Elle est définie par :

$$NCC(\mu; \mathcal{J}^*, \mathcal{I}) = \frac{\sum_{x_i} (\mathcal{J}^*(x_i) - \bar{\mathcal{J}}^*) (\mathcal{I}(\varphi_\mu^\dagger(x_i)) - \bar{\mathcal{I}})}{\sqrt{\sum_{x_i} (\mathcal{J}^*(x_i) - \bar{\mathcal{J}}^*)^2 \sum_{x_i} (\mathcal{I}(\varphi_\mu^\dagger(x_i)) - \bar{\mathcal{I}})^2}}, \quad (3.9)$$

où  $x_i$  sont des patchs échantillons des images avec  $\bar{\mathcal{I}}$  et  $\bar{\mathcal{J}}^*$  les valeurs d’intensités moyennes pour  $\mathcal{I}$  et  $\mathcal{J}^*$ .

Pour optimiser l’équation 3.8, nous utilisons une procédure itérative, l’algorithme du gradient stochastique (*Adaptive Stochastic Gradient Descent* – ASGD) [108] :

$$\mu_{k+1} = \mu_k - \gamma(t_k) \tilde{g}_k, \quad (3.10)$$

où  $k$  est le nombre d’itérations et  $\tilde{g}_k$  une approximation de la dérivée de la fonction de coût  $\partial \mathcal{C} / \partial \mu_k$ . La valeur  $\gamma(t_k)$  détermine la taille du pas où  $\gamma(\cdot)$  est une fonction décroissante et  $t_k$  est définie par  $t_{k+1} = [t_k + f(-\tilde{g}_k^T \tilde{g}_{k-1})]^+$  avec  $[\cdot]^+ = \max(\cdot, 0)$  et  $f(\cdot)$  une fonction sigmoïde. L’utilisation du produit  $\tilde{g}_k^T \tilde{g}_{k-1}$  pour déterminer la taille du pas rend la procédure d’optimisation adaptative et le recalage plus robuste.

Cette combinaison entre une approche conventionnelle de recalage et le DL nous permet de tirer parti de la capacité de génération de notre architecture, qui est conditionnée de manière double, tout en préservant les propriétés souhaitables des méthodes de recalage classique avec une paramétrisation explicite de la déformation par des champs *B-spline* dont nous contrôlons la régularité. Nous obtenons ainsi une image finale de meilleure résolution tout en conservant un contrôle explicite sur la déformation de l’image.

## 3.3 Protocole expérimental

### 3.3.1 Jeux de données

#### Données synthétiques

Nous avons généré divers jeux de données synthétiques afin de valider la robustesse de notre méthode dans un contexte dans lequel l'amplitude est parfaitement connue. Ces jeux de données représentant différents niveaux de complexité sont composés de boules, d'ellipsoïdes uniques, de paires d'ellipsoïdes et de pseudo-poumons présentant différentes intensités de bruit. Les pseudo-poumons sont une modification des paires d'ellipsoïdes où le diaphragme est représenté par une coupe en arc de cercle au niveau inférieur. Concernant les jeux de données bruités, nous appliquons un bruit additif gaussien d'intensité élevée et variable. Nous générons aussi un jeu de données qui simule la présence d'une tumeur, pour cela, nous insérons un objet sphérique de taille aléatoire à l'intérieur des pseudo-poumons (figure 3.5a).

Afin d'introduire des variations de volume et de générer des dynamiques similaires à un mouvement, nous générons 200 objets par jeu de données, dont chacun subit plusieurs dilatations morphologiques à l'aide d'éléments structurants de taille variable et aléatoire (figure 3.5b). Ensuite, chaque cas est dilaté de manière aléatoire quatre fois successivement pour un total de 1000 images. Le centre de masse de l'objet sphérique représentant une tumeur se déplace en accord avec les dilatations. Pour l'entraînement des modèles, nous utilisons 4/5 des jeux de données et le reste est utilisé pour l'évaluation.

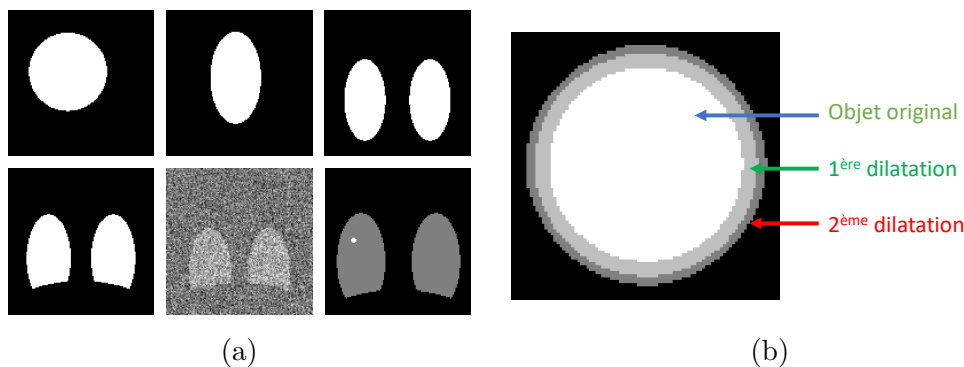


FIGURE 3.5 – Exemples d’images synthétiques et simulation de dynamique. (a) Illustrations des différents types de données synthétiques avec une boule, un ellipsoïde, une paire d’ellipsoïdes, des pseudo-poumons non bruités et bruités ainsi que des pseudo-poumons avec un objet sphérique simulé. (b) Un objet original est généré puis dilaté morphologiquement par un élément structurant pour augmenter sa taille et simulé un mouvement.

## Données cliniques

Notre base de données clinique est composée de 47 patients, dont les acquisitions ont été effectuées au service de radiothérapie du CHRU de Brest. Chaque patient possède une à quatre acquisitions TDM 4D. Une partie des images ont été exclues manuellement du jeu de données d’entraînement, car elles n’étaient pas utilisables pour diverses raisons telles que des artefacts trop importants ou des acquisitions partielles des poumons. Après une sélection manuelle basée sur des critères visuels, nous avons retenu 66 images TDM 4D pour nos expérimentations. Les images initiales en format DICOM ont été converties en format Nifti. Les tailles initiales étaient de  $512 \times 512 \times Z$  où  $Z$  représente la composante axiale avec des valeurs différentes en fonction des acquisitions. Pour homogénéiser les images et limiter le coût en mémoire, les images ont toutes été ré-échantillonnées en  $256 \times 256 \times 64$  voxels par interpolation linéaire.

À partir d’un total de 66 images TDM 4D, nous utilisons 48 images TDM 4D de 26 patients pour constituer notre jeu de données d’entraînement. Chaque acquisition est composée de huit phases respiratoires : de 0% à 100% pour passer de la phase fin de l’expiration (EOE, *end of exhale*) à la phase EOI et de 75% à 25% pour retourner à la phase EOE, avec un pas de 25% (figure 3.6). La validation du modèle clinique s’effectue avec un jeu de données composé de 18 images TDM 4D provenant de neuf patients supplémentaires.

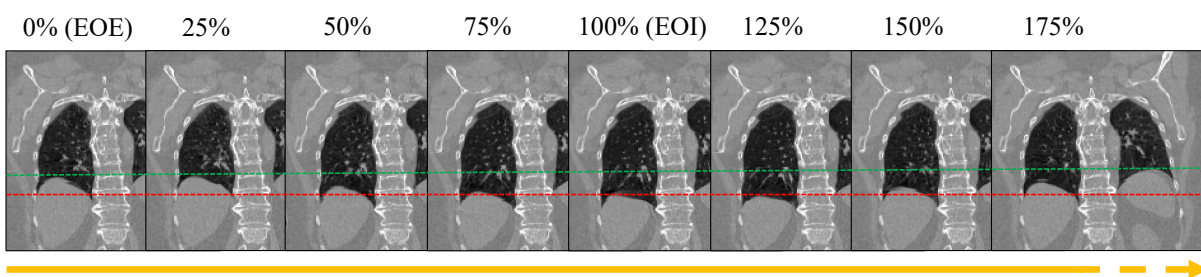


FIGURE 3.6 – Illustration des phases respiratoires d’une TDM 4D. Les acquisitions TDM 4D fournies par le CHRU de Brest sont composées de huit phases. Elles commencent par la phase EOE pour aller progressivement vers la phase EOI puis finissent par revenir vers la EOE. Les lignes vert et rouge représentent le niveau du diaphragme en phase EOE et en phase EOI respectivement.

### 3.3.2 Implémentations

Nous avons procédé à l’entraînement de notre réseau à partir de zéro en utilisant une carte graphique NVIDIA GeForce GTX 1080Ti avec 11 Go de mémoire, avec une taille de batch de 1. Pour les données synthétisées et cliniques, nous avons réalisé respectivement 100 et 200 itérations d’entraînement en utilisant l’algorithme d’optimisation ADAM [109].

Nous avons employé comme conditionnement scalaire le ratio de volume pulmonaire entre l’image en entrée  $V_I$  et la phase cible  $V_J$  (pourcentage de différence de volume (PVD, *percentage volume difference*)), défini par l’équation suivante :

$$PVD = \frac{V_J - V_I}{V_I} \times 100 \quad (3.11)$$

Pour les pseudo-poumons bruités, nous avons ajouté un bruit additif gaussien avec des variances  $\sigma^2$  égales à  $0,50^2$  et  $0,80^2$ . Quant aux données cliniques, nous avons utilisé une méthode de segmentation automatique robuste des poumons basée sur l’apprentissage profond [110] pour obtenir les masques pulmonaires, qui ont ensuite été utilisés pour calculer les volumes pulmonaires correspondants pour le conditionnement scalaire (figure 3.7).

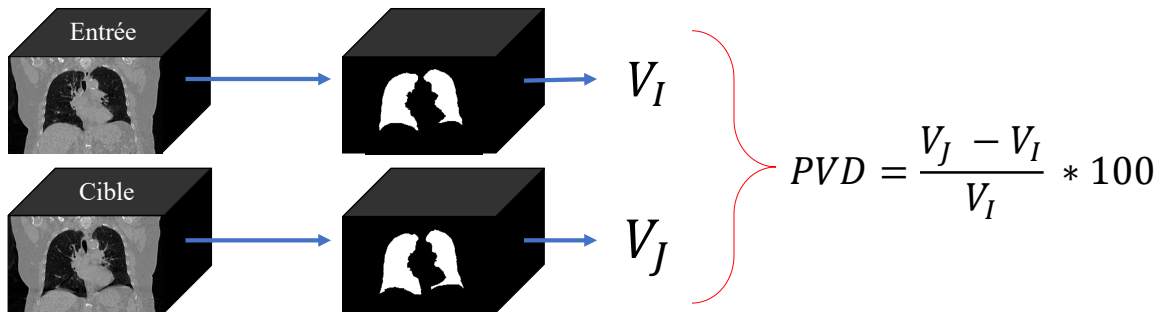


FIGURE 3.7 – Calcul de la variation de volume entre deux phases respiratoires.

Les segmentations des tumeurs n’étant pas fournies, nous avons utilisé le modèle nnU-Net [39] pré-entraîné avec la configuration gagnante du challenge *Medical Segmentation Decathlon* pour la segmentation des nodules pulmonaires [111]. Une étape de correction manuelle a été effectuée à la suite si nécessaire. Le recalage d’image effectué en sortie du réseau se base sur les paramètres elastix décrits par Staring et coll. [107].

Pour évaluer quantitativement les résultats, nous avons calculé la similarité entre le



volume généré et le volume de référence en utilisant les mesures suivantes :

1. la variation de volume (VV) pulmonaire :

$$VV = \frac{|V_J - V_I|}{V_I} \times 100, \quad (3.12)$$

où  $V_I$  et  $V_J$  représentent respectivement les volumes pulmonaires de la phase source et de la phase cible.

2. le DSC qui mesure le chevauchement d’organes :

$$\text{Dice} = \frac{2|\mathcal{S}_G \cap \mathcal{S}_{\mathcal{J}^*}|}{|\mathcal{S}_G| + |\mathcal{S}_{\mathcal{J}^*}|}, \quad (3.13)$$

où  $\mathcal{S}_G$  et  $\mathcal{S}_{\mathcal{J}^*}$  sont les masques de segmentations de l’image générée et de l’image de référence respectivement.

3. l’erreur absolue moyenne (MAE, *mean absolute error*), une métrique couramment employée en imagerie scanner pour comparer voxel par voxel une image :

$$\text{MAE} = \frac{1}{D \times H \times W} \sum_{i=1}^D \sum_{j=1}^H \sum_{k=1}^W |G(\mathcal{I})(i, j, k) - \mathcal{J}^*(i, j, k)|, \quad (3.14)$$

où  $D$ ,  $H$ ,  $W$  représentent respectivement la profondeur, la hauteur et la largeur des images, et  $G(\mathcal{I})(i, j, k)$  est le voxel de l’image générée et  $\mathcal{J}^*(i, j, k)$  est la valeur du voxel de la référence.

Idéalement, la valeur du DSC devrait être égale à 1 et inversement les valeurs de MAE et VV devraient être égales à 0, signifiant ainsi une parfaite similitude entre les deux échantillons. En plus du DSC, nous avons mesuré l’erreur de distance euclidienne entre les centres de masses (COMs, *center of mass*) des tumeurs afin de mieux appréhender son déplacement après génération :

$$COM = |\mathcal{C}_G(x, y, z) - \mathcal{C}_{\mathcal{J}^*}(x, y, z)|, \quad (3.15)$$

où  $\mathcal{C}_G$  et  $\mathcal{C}_{\mathcal{J}^*}$  sont les centres de masses des masques de segmentations, ayant pour coordonnées  $x, y, z$ , des images générée et de référence respectivement.

## 3.4 Résultats

### 3.4.1 Images synthétiques

Nous comparons à l'inférence les images dilatées morphologiquement du jeu de données de test aux images générées par notre méthode à partir des images initiales (figure 3.8). Nous pouvons observer que les images générées sont similaires aux images vérités terrains. Les images conservent la même forme et la même taille après génération.

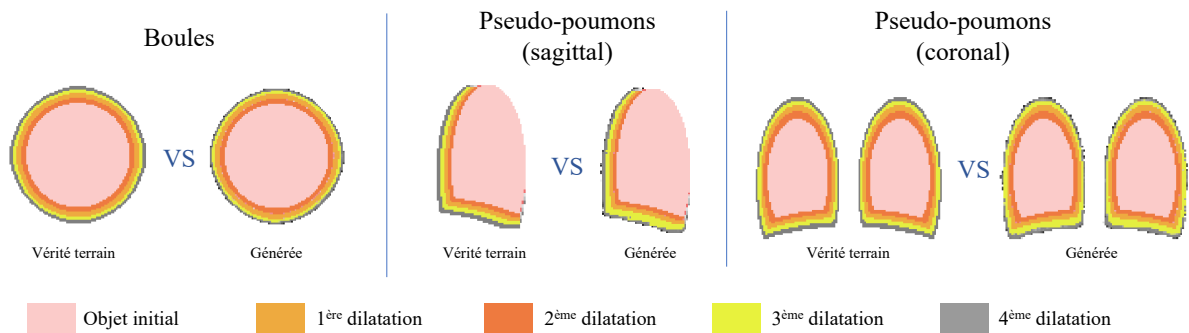


FIGURE 3.8 – Exemple de résultats de synthèse d'images synthétiques de boules et de pseudo-poumons (plan sagittal et coronal).

Pour les résultats quantitatifs, les boîtes à moustaches de la figure 3.9 montrent les erreurs initiales entre les images d'entrées et les images vérités terrains et les erreurs après synthèses d'images. Globalement, les erreurs sont relativement faibles pour tous les types d'objets présents dans les jeux de données. Par exemple, les erreurs moyennes de la VV sont de  $4,31 \pm 2,04\%$ ,  $5,92 \pm 5,20\%$ , et  $3,72 \pm 1,71\%$  pour les boules, les ellipsoïdes simples et les paires d'ellipsoïdes respectivement. Nous pouvons aussi observer une erreur de  $3,16 \pm 1,33\%$  pour les pseudo-poumons sans bruits et des erreurs plus élevées pour les versions avec du bruit ( $15,35\%$  pour 50% de bruit et  $21,47\%$  pour 80% de bruit), mais les erreurs restent acceptables. Pour le DSC et la MAE, nous observons un comportement similaire à celle de la VV. Malgré quelques valeurs aberrantes et une sensibilité à des bruits importants, les résultats montrent que la méthode est capable de générer une dynamique à partir d'une image 3D et d'une valeur scalaire.

La figure 3.10 présente des résultats visuels sur le jeu de données composé des pseudo-poumons avec tumeur et avec ou sans l'utilisation de la fonction de coût focalisée sur la tumeur  $TA$ . Avec ou sans  $TA$ , un abaissement global du diaphragme correspondant

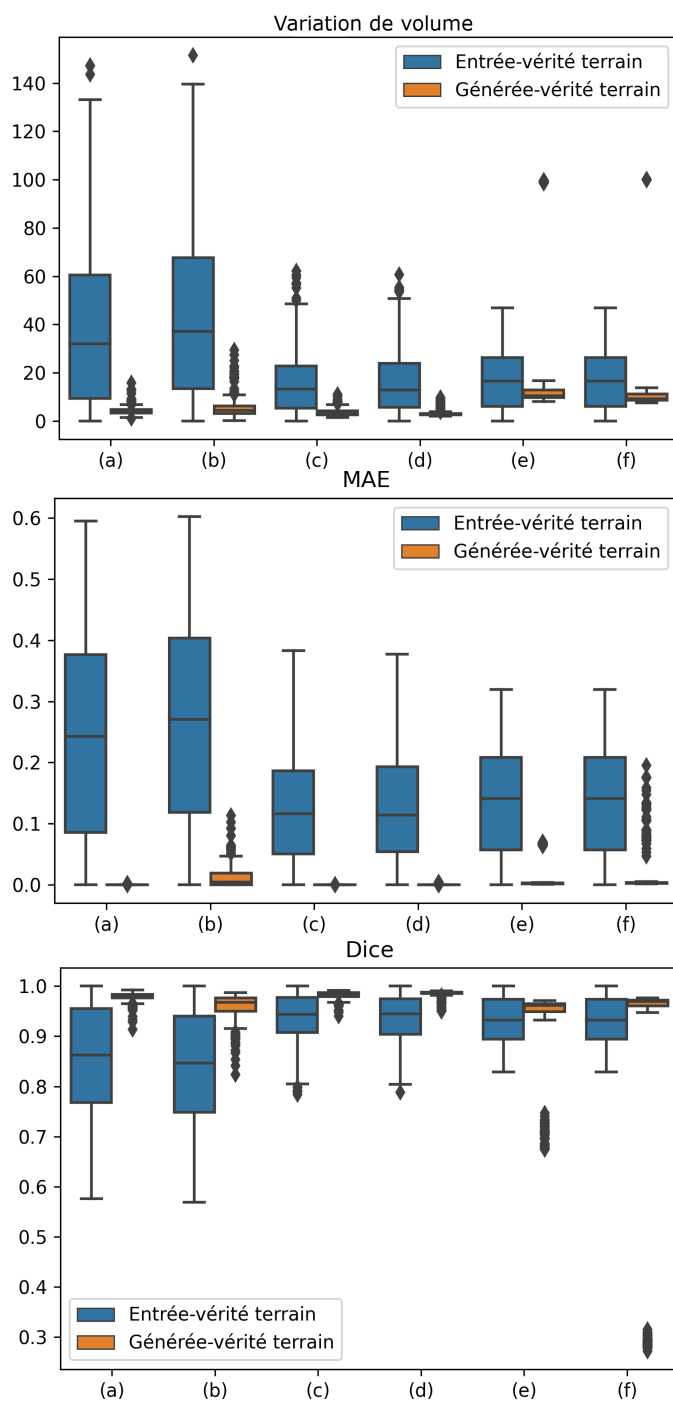


FIGURE 3.9 – Résultats quantitatifs des jeux de données synthétiques. (a) Boules (b) Ellipsoïdes (c) Paires d'ellipsoïdes (d) Pseudo-poumons sans bruit (e) Pseudo-poumons avec bruit ( $\sigma^2 = 0,50^2$ ) (f) Pseudo-poumons avec bruit ( $\sigma^2 = 0,80^2$ ).

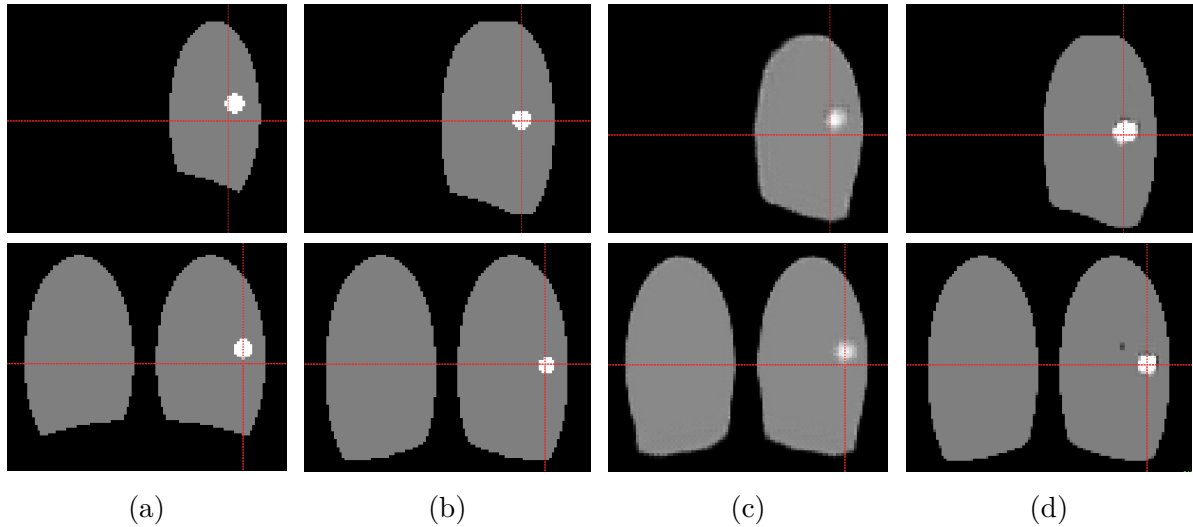


FIGURE 3.10 – Exemple de résultats de synthèse d’images synthétiques avec et sans  $TA$  (plan sagittal et coronal). Le croisement des lignes rouges représente le centre de masse de la tumeur de l’image cible. (a) Image d’entrée (b) Image cible (c) Image générée sans  $TA$  (d) Image générée avec  $TA$ .

à la phase cible (figure 3.10b) par la méthode est réalisé avec succès. Cependant, nous pouvons constater que la tumeur ne se déplace pas de manière appropriée dans le cas où la fonction  $TA$  n’est pas utilisée (figure 3.10c). L’utilisation de cette fonction  $TA$  permet de corriger ce problème grâce à la prise en compte de la tumeur (figure 3.10d).

Ces observations qualitatives sont confirmées quantitativement sur le jeu de données de test avec une nette amélioration de la similarité avec la vérité terrain en utilisant la fonction  $TA$  proposée (tableau 3.1). Les erreurs initiales montrent qu’il y a un faible chevauchement de la tumeur entre les phases, avec un DSC de 0,28 et une erreur COM de 3,73 mm. Lorsque seul l’objectif  $\mathcal{L}_{\ell_1}$  est utilisé, le mouvement de la tumeur n’est pas bien reproduit. Cependant, en intégrant un mécanisme de prise en compte de la tumeur, une nette amélioration est obtenue, donnant ainsi les meilleurs résultats, avec un DSC de 0,67 et une erreur COM de 1,08 mm en moyenne.

Modèle	DSC	COM (en mm)
Erreur initiale entre les phases	0,28	3,73
$\mathcal{L}_{\ell_1}$ seule	0,29	3,36
$\mathcal{L}_{\ell_1}(\omega)$ seule	0,53	1,94
$TA : \mathcal{L}_{\ell_1}(\omega)$ et $\mathcal{L}_{DSC}$	0,67	1,08

TABLE 3.1 – Comparaison de modèle avec ou sans la fonction pondérée sur les images synthétiques.

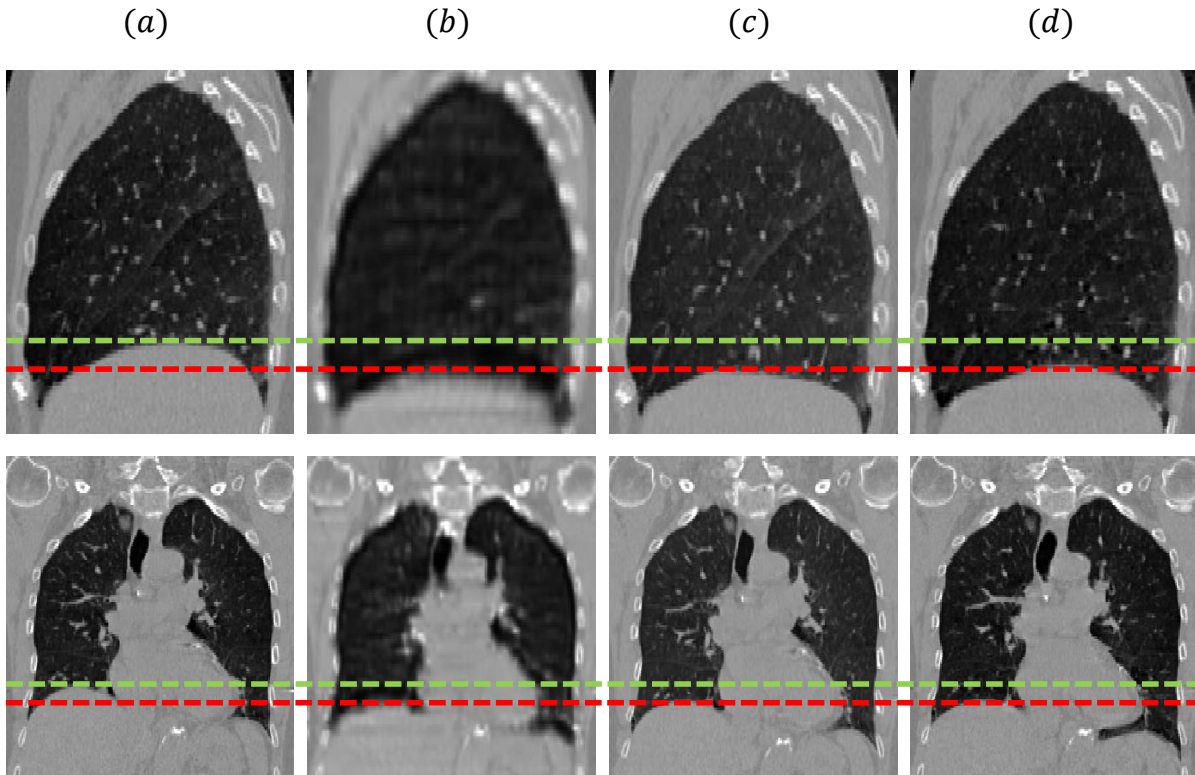


FIGURE 3.11 – Exemple de résultats de synthèse d’images cliniques (plan sagittal et coronal). Les lignes vert et rouge représentent la position initiale et finale du diaphragme respectivement. (a) Image d’entrée (b) Image générée intermédiaire faible résolution (c) Image finale après recalage (d) Image vérité terrain

### 3.4.2 Images cliniques

Concernant les données cliniques TDM 4D, nous observons une bonne correspondance visuelle entre les phases respiratoires réelles et synthétiques, restituant globalement la dynamique du mouvement respiratoire (figure 3.11).

À partir d’une image d’entrée (figure 3.11a), nous générons avec le modèle une image intermédiaire (figure 3.11b), puis grâce au recalage, nous déformons la phase d’entrée vers l’image générée pour obtenir une image plus fidèle qu’en sortie de réseau (figure 3.11c). La figure 3.11c montre un exemple représentatif de phase synthétique, où l’on peut observer un déplacement diaphragmatique entre les phases en accord avec la vérité terrain (figure 3.11d), aussi bien en coupe coronale qu’en coupe sagittale.

Les résultats quantitatifs sont présentés dans le tableau 3.2. Une amélioration significative a été obtenue après la synthèse conditionnelle. Nous observons que le volume

	Phase 25%	Phase 50%	Phase 75%	EOI
Volume moyen de référence	3560, 21 ± 1305, 15	3647, 51 ± 1312, 37	3716, 41 ± 1321, 91	3715, 50 ± 1334, 81
Volume moyen généré	3569, 63 ± 1283, 72	3657, 97 ± 1291, 30	3716, 11 ± 1299, 46	3712, 47 ± 1313, 12
VV initiale	7, 91%	10, 73%	12, 63%	12, 33%
VV après génération	1, 06%	1, 21%	1, 01%	1, 01%
DSC poumons initial	0, 95 ± 0, 02	0, 94 ± 0, 02	0, 93 ± 0, 03	0, 93 ± 0, 02
DSC poumons après génération	0, 96 ± 0, 01	0, 96 ± 0, 01	0, 96 ± 0, 02	0, 96 ± 0, 01
MAE initiale	152, 71 ± 28, 79	175, 21 ± 32, 94	189, 43 ± 36, 11	183, 57 ± 32, 39
MAE après génération	123, 77 ± 19, 87	129, 93 ± 18, 34	137, 69 ± 18, 50	129, 62 ± 17, 58

TABLE 3.2 – Résultats quantitatifs sur les données cliniques.

moyen généré est très proche du volume moyen de référence dans toutes les phases. Cela est aussi visible avec la VV qui passe de 12,33% à 1,01% après génération pour la phase EOI. Le DSC entre les poumons confirme aussi les résultats et la capacité au modèle de faire respirer les poumons. De la même manière, les valeurs de MAE après génération sont plus faibles que les erreurs initiales. Ces résultats démontrent ainsi la capacité de notre modèle proposé à reproduire une phase respiratoire correspondant à une phase réelle.

La figure 3.12 présente des résultats visuels de deux patients de notre jeu de données cliniques. Nous pouvons observer dans le premier cas que le mouvement initial entre l'image d'entrée et l'image cible est faible. Lorsque l'image est générée par la méthode sans la fonction  $TA$ , la tumeur est effacée contrairement au modèle avec  $TA$ . Malgré un mouvement initial important de la tumeur pour le second patient, les méthodes proposées ne permettent pas de corriger cette erreur, et ce, même si un mouvement de diaphragme satisfaisant est observé.

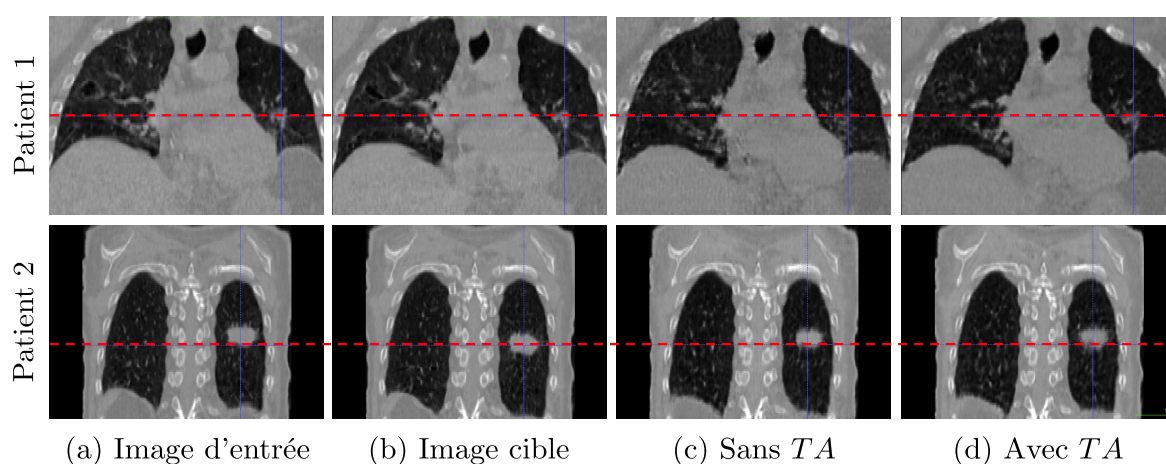


FIGURE 3.12 – Exemple de résultats sur les données cliniques avec et sans fonction focalisée sur la tumeur. Les pointillés rouges représentent le centre de masse des tumeurs de l'image cible.

L’évaluation quantitative des tumeurs est présentée sur le tableau 3.3. Nous pouvons observer que l’erreur initiale DSC moyenne entre les images d’entrées et les images vérités terrains est de 0,54 et que le COM est de 7,13 mm. Après synthèse des images, sans *TA*, nous observons une très légère amélioration avec, par exemple, en phase EOI un DSC de 0,55 au lieu de 0,49 et un COM de 6,36 mm au lieu de 8,54 mm. Dans le cas avec *TA*, les résultats ne sont pas très différents du cas où nous n’utilisons pas la fonction. Ces résultats montrent que la fonction pondérée sur la tumeur ne permet pas d’améliorer son déplacement au niveau local dans l’ensemble du jeu de données de test.

		Phase 25%	Phase 50%	Phase 75%	EOI
DSC	Initial	0,66 ± 0,25	0,53 ± 0,27	0,48 ± 0,27	0,49 ± 0,29
	sans <i>TA</i>	0,69 ± 0,22	0,57 ± 0,21	0,50 ± 0,28	0,55 ± 0,25
	avec <i>TA</i>	0,70 ± 0,24	0,60 ± 0,25	0,55 ± 0,26	0,58 ± 0,24
COM	Initial	4,71 ± 4,29	7,12 ± 5,69	8,19 ± 5,81	8,54 ± 6,21
	sans <i>TA</i>	4,44 ± 4,95	6,48 ± 6,19	7,05 ± 5,90	6,36 ± 5,20
	avec <i>TA</i>	4,13 ± 4,06	5,97 ± 5,77	6,66 ± 4,69	6,33 ± 4,87

TABLE 3.3 – Comparaison des résultats avec ou sans la fonction pondérée sur les images cliniques.

### 3.5 Discussion

Dans ce chapitre, nous avons montré qu’il est possible de générer une dynamique respiratoire à partir d’une architecture GAN conditionnée par deux facteurs. Contrairement aux méthodes I2I traditionnelles où le transfert de style ou la manipulation d’attributs se fait à l’aide d’une autre image [4], [102], [106], nous avons démontré qu’il est aussi possible d’influencer la génération par une valeur scalaire respiratoire unidimensionnelle. La méthode est testée dans un premier temps sur des jeux de données synthétiques afin de vérifier sa pertinence. Les résultats montrent qu’il est possible de faire « respirer » les différents objets, et même en présence de bruit. L’apport d’une fonction pondérée sur une zone d’intérêt prouve aussi son efficacité quant à la précision du mouvement de l’objet d’intérêt dans ces jeux de données synthétiques.

Cette première validation a permis ensuite d’appliquer la méthode avec des données cliniques réelles. Les résultats obtenus sont satisfaisants concernant le mouvement global des poumons et notamment du mouvement du diaphragme.

Néanmoins, le degré de réalisme obtenu en sortie de réseau nous est paru largement insuffisant, compte tenu de la tendance des GANs à halluciner certaines caractéristiques de l'image. C'est ainsi que nous avons été amenés à recalibrer les phases initiales vers les phases respiratoires générées pour contrôler explicitement la transformation par une paramétrisation *B-spline* : l'utilisation d'une méthode de recalage conventionnelle sur l'image en sortie de réseau permet de garder un contrôle sur les déformations et de garantir des images de haute qualité tout en permettant une dynamique respiratoire.

Bien qu'illustrée sur des images synthétiques, l'utilisation d'une pondération spatiale de la fonction de coût de reconstruction  $\ell_1$  n'a pas démontré un apport bénéfique en situation réelle sur des images TDM 4D. Cette baisse de performance peut s'expliquer par différentes raisons telles qu'un mauvais choix des hyperparamètres et donc d'un manque d'optimisation. Mais aussi des conditions d'entraînement qui complexifie l'apprentissage rendant la tâche plus difficile : données plus complexes que de simples images synthétiques (des tumeurs trop petites qui sont donc noyées dans trop d'information), choix des paires d'images aléatoires lors de l'entraînement, composition des jeux de données non équilibrée avec, par exemple, plus de tumeurs statiques que de tumeurs avec une grande amplitude de mouvement. Nous pensons que ces différentes raisons conduisent le réseau à sous-estimer globalement l'amplitude du mouvement de la tumeur. Une autre limitation de la méthode est la synthèse d'image directe qui génère des images de faible résolution que nous tentons de corriger à l'aide d'un recalage.

Malgré ces contraintes, cette première approche a démontré la possibilité de synthétiser un mouvement respiratoire et de la lier à une contrainte scalaire externe, et donc fournit une première preuve de concept sur la capacité d'une architecture de ce type à simuler un mouvement respiratoire tel que mesuré par une acquisition TDM 4D.

## 3.6 Conclusion

La nouvelle architecture I2I proposée génère différentes phases respiratoires et est contrôlable par un conditionnement scalaire. Les phases respiratoires sont générées par une approche basée GAN afin de conserver un haut réalisme grâce au coût adversaire et le choix de la phase cible générée est possible grâce au conditionnement scalaire qui adapte le vecteur latent dans l'espace latent du générateur. Cette approche permet ainsi de générer des phases respiratoires spécifiques à un patient grâce aux informations internes et externes fournies par l'image TDM 3D et l'information respiratoire respectivement.



Les approches I2I sont utilisées en imagerie médicale pour leurs capacités à transférer des informations d'une modalité d'image à une autre et plus généralement dans le cadre d'imagerie statique. Les expérimentations effectuées ont permis de montrer la possibilité de passer d'un domaine statique à un domaine dynamique à l'aide de ce type d'architecture. Le modèle parvient à générer un mouvement interne global et spécifique au patient. Ce type de synthèse d'image peut également être complexifié afin d'améliorer les performances et la précision des mouvements internes des patients.

Nous verrons dans le chapitre suivant une approche plus aboutie basée sur une synthèse d'image « indirecte », c'est-à-dire, par génération de champs de déformation.

# SYNTHÈSE INDIRECTE PAR APPRENTISSAGE CONJOINTEMENT SUPERVISÉ ET ADVERSAIRE

## Résumé

Ce chapitre propose la contribution principale de la thèse, qui porte sur une méthode de synthèse TDM 4D indirecte par génération de champs de déformations DVF. Il est le fruit de réflexions évoquées au chapitre précédent sur les limitations inhérentes de la synthèse d'image directe par GAN, qui nous avaient déjà incité à imposer un recalage conventionnel vers l'image produite pour contrôler la nature des déformations sur l'image. Si d'autres méthodes de synthèse TDM 4D basées sur la génération de DVF existent, elles reposent toutes sur des méthodes de recalage par DL non supervisé, où les déformations produites sont contrôlées par des fonctions de coût explicite sur la régularité des champs produits. Or, le recalage pulmonaire n'est pas un problème résolu en DL. Dans notre approche, nous nous affranchissons de la responsabilité de ce recalage et entraînons notre modèle de manière supervisée avec des champs obtenus par une méthode de recalage externe robuste. Nous contrôlons le réalisme des champs par un coût adversaire sur leur magnitude, ce qui permet de prendre en compte l'amplitude respiratoire dans le coût adversaire. Cette approche permet de générer des champs avec des mouvements plus réalistes et permet d'exploiter des méthodes de recalage conventionnelles plus abouties, contrairement à des approches non supervisées où les déformations sont souvent trop lisses. Après avoir décrit la méthode, nous l'évaluons à une échelle plus fine que les approches existantes de l'état de l'art, notamment avec des métriques locales pour évaluer les mouvements de tumeurs et des OARs. Une autre analyse originale et capitale est de nous comparer à la variabilité intrinsèque des TDM 4D, en exploitant des acquisitions répétées d'un même patient. Ce travail est en relecture dans *Physics in Medicine and Biology*.

## 4.1 Introduction

Comme nous l’avons vu dans le chapitre précédent, il est possible de générer un mouvement global de poumons à partir d’une image TDM 3D et d’une valeur scalaire calculée à partir de la variation de volume pulmonaire entre deux phases respiratoires avec un modèle de DL basé sur une architecture GAN. Cette première approche de synthèse directe ne permet pas de réaliser des mouvements locaux, alors qu’un mouvement réaliste dans les zones d’intérêts est crucial pour la planification du traitement en radiothérapie. Il est donc nécessaire de réussir à synthétiser des images avec une meilleure fidélité de mouvements aux niveaux des zones d’intérêt.

En imagerie médicale, les DVFs sont principalement utilisés pour des applications de recalage d’images, mais ils peuvent aussi être utilisés afin d’estimer des mouvements. Fayad et coll. ont généré des TDM 4D synthétique en interpolant des DVFs entre les phases maximales EOE et EOI obtenus à partir des méthodes de recalages [26]. Récemment, Sang et Ruan proposent une méthode basée DL pour interpoler les phases respiratoires grâce à un conditionnement avec la paire d’images et un conditionnement scalaire supplémentaire [100]. Ces approches nécessitent toutefois un minimum de deux acquisitions statiques ou d’une acquisition TDM 4D pour interpoler. De plus, les interpolations présentées sont linéaires et ne sont donc pas spécifiques au patient. D’autres approches similaires de génération de dynamisme à partir d’une image statique sont étudiées. Chang et coll. parviennent à générer des fantômes 4D-XCAT de poumons avec des mouvements respiratoires plus réalistes [96]. La méthode est composée de deux blocs principaux, un modèle DL de génération de DVF et un modèle respiratoire basé sur l’apprentissage automatique. Ce dernier est entraîné sur une population donnée et permet aux images de synthèse 4D-XCAT de reproduire des mouvements respiratoires réalistes et est contrôlable par l’utilisateur si nécessaire. Romaguera et coll. proposent une architecture basée sur un seq2seq afin de générer des images 4D IRM de foie [98]. La méthode apprend aussi sur une population donnée et à la particularité d’être spécifique à un patient si lors de l’inférence des acquisitions 2D du patient sont disponibles. Ces méthodes sont entraînées de manière non supervisée et permettent ainsi de produire des résultats précis sans avoir besoin de données de recalage prédéfinies. Toutefois, des approches non supervisées ont tendance à générer des champs lisses et non représentatifs du mécanisme respiratoire plus complexe. Dans une approche supervisée, Lee et coll. proposent un modèle basé sur une architecture seq2seq 3D, RMSim [101]. La méthode permet de prédire les phases respiratoires futures à

partir d'une image TDM 3D et de données respiratoires 1D. Le modèle RMSim est conditionné par des données respiratoires 1D afin de mieux simuler la respiration du patient. Ces données respiratoires sont extraites à partir des informations disponibles sur des acquisitions TDM 4D en choisissant le déplacement d'un point situé à l'apex du diaphragme. Outre la difficulté d'obtenir ces données respiratoires en routine clinique sans acquisitions TDM 4D, les résultats présentés ne montrent des résultats qu'avec des métriques basées sur la similarité globale des images et ne prennent pas en compte les zones d'intérêts.

Dans ce chapitre, nous explorons une approche de synthèse indirecte dans le but d'obtenir un déplacement interne des organes plus précis. Nous proposons de générer des DVFs au lieu de générer directement les phases respiratoires et, pour nous rapprocher d'une utilisation clinique, nous remplaçons le scalaire basé sur le volume respiratoire par une amplitude respiratoire extraite à partir de la surface du patient. Le modèle est entraîné de manière supervisée avec des DVFs calculés à partir d'un algorithme robuste de recalage d'images. De plus, nous proposons un coût adversaire sur la magnitude des champs générés pour améliorer le réalisme des amplitudes respiratoires. L'approche proposée permet ainsi de générer des mouvements plus précis localement en corrélation avec la respiration du patient tout en conservant une haute résolution grâce à une déformation de la phase respiratoire initiale par le champ généré.

Dans la suite de ce chapitre, nous présentons les modifications apportées au conditionnement et à la génération. Nous évaluons dans un premier temps avec un modèle synthétique la robustesse du modèle puis avec des données cliniques réelles. Afin de mieux évaluer la qualité des images générées, nous proposons également d'évaluer les résultats de manière locale. Dans un but final d'utilisation en routine clinique pour la radiothérapie, le point critique reste les zones d'intérêts incluant les tumeurs et les OARs. Il est donc important d'évaluer ces zones d'intérêts avec des mesures appropriées. Nous nous comparons à la méthode RMSim et nous proposons en plus d'évaluer le déplacement des OARs et des tumeurs avec des métriques classiques telles que le DSC. Du fait de la variation intrinsèque de la respiration pour un même patient, nous proposons aussi d'évaluer la variabilité entre les multiples acquisitions d'un même patient et les images générées.

## 4.2 Méthode

La figure 4.1 illustre le fonctionnement global de la méthode proposée. Pendant l'entraînement, le modèle apprend à établir une correspondance entre la phase EOE et les

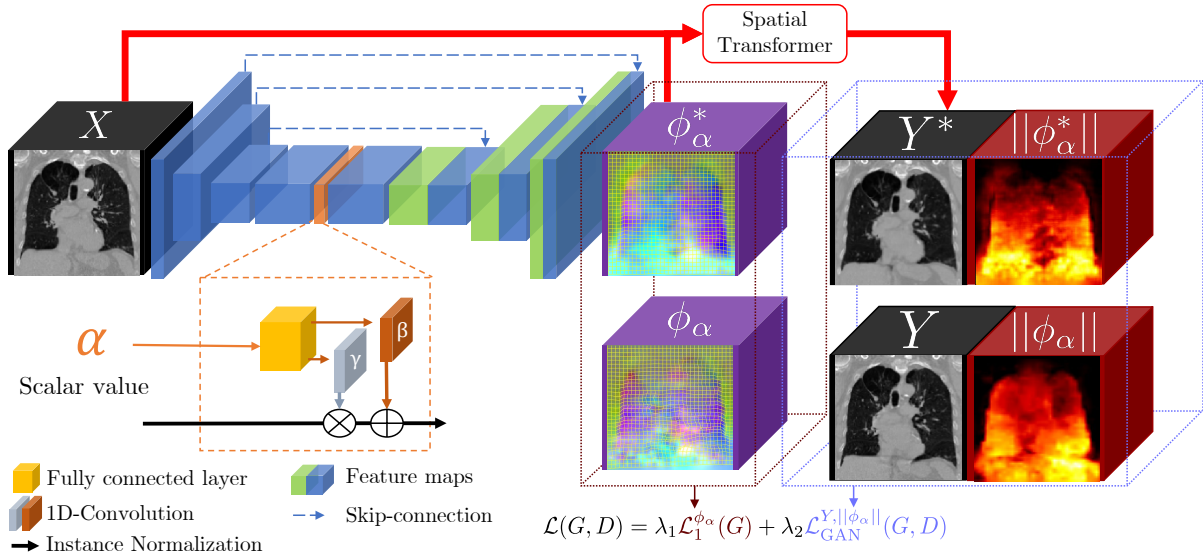


FIGURE 4.1 – Architecture proposée pour la génération par déformation. L’amplitude respiratoire est injectée à travers un mécanisme AdaIN au niveau du goulot d’étranglement. Étant donné une image d’entrée  $X$  et une amplitude respiratoire  $\alpha$ , le modèle apprend un DVF  $\phi_\alpha^*$  à partir des étiquettes supervisées  $\phi_\alpha$ . Ensuite, l’image d’entrée est déformée à l’aide d’un transformateur spatial pour obtenir la phase synthétique  $Y^* = X \circ \phi_\alpha^*$ . Le coût adversaire basé sur la concaténation de l’image déformée et de la magnitude du DVF  $\|\phi_\alpha\|$  est utilisée pour améliorer le réalisme du DVF sans contraindre explicitement sa régularité.

différentes phases des images TDM 4D réelles. La génération est conditionnée par la phase d’entrée et une estimation basée sur la surface externe  $\alpha(t)$ , qui peut être extraite de différents dispositifs de surveillance respiratoire utilisés cliniquement. Le générateur associe à la paire d’entrées image- $\alpha$  un DVF à trois composantes de manière supervisée en utilisant une méthode de recalage conventionnelle robuste [65]. Pendant l’entraînement, les DVFs de sortie sont comparés à leur vérité terrain en suivant un objectif de minimisation  $\ell_1$ . Une contribution essentielle est d’améliorer davantage le réalisme de la transformation grâce à un terme adversaire basé non seulement sur l’image d’entrée déformée, comme c’est généralement le cas dans les méthodes de recalage basées sur l’apprentissage profond, mais également sur la magnitude du DVF généré lui-même.

Au moment du test, étant donné une image d’entrée TDM 3D et les valeurs correspondantes de  $\alpha$ , le modèle reproduit des images synthétiques de phase respiratoire qui imitent la respiration avec les amplitudes spécifiées.

### 4.2.1 Formulation du problème et notation

Soit  $X(x)$  une image TDM 3D, où  $x \in \mathbb{R}^3$  est le domaine de l'image. Soit également  $\mathcal{Y}(x)$  une image TDM 4D, c'est-à-dire une série dynamique de  $N_g$  phases :

$$\mathcal{Y}(x) = \left( Y_0(x), \dots, Y_k(x), \dots, Y_{N_g}(x) \right), \quad (4.1)$$

où  $k \in \mathbb{N}$  est le numéro de la phase, avec  $k = 0$  correspondant à la phase EOE.

Notre objectif est de synthétiser une image 4D dynamique, de valeur vectorielle, imitant la dynamique respiratoire d'une acquisition TDM 4D à partir de  $X(x)$ , c'est-à-dire un ensemble de  $N$  images 3D :

$$\mathcal{Y}^*(x) = \left( Y_0^*(X, \alpha_0), \dots, Y_k^*(X, \alpha_k), \dots, Y_N^*(X, \alpha_N) \right), \quad (4.2)$$

où  $\alpha_k$  est la variation relative d'amplitude respiratoire (par exemple, en mm) par rapport à  $X(x)$ . Ce faisant, nous faisons la simplification qu'il existe une correspondance un à un (bijective) entre l'image réelle  $Y_k$  et l'amplitude respiratoire  $\alpha_k$ , c'est-à-dire que pour chaque  $\alpha_k$ , il ne peut y avoir qu'une seule image TDM 3D associée.

À cette fin, lors de l'entraînement, nous apprenons à partir d'un ensemble d'acquisitions réelles de TDM 4D  $\{\mathcal{Y}^i\}$ . Pour chaque image TDM 4D d'entraînement, nous apprenons à reproduire un ensemble de DVF :

$$\left( \phi_1, \dots, \phi_k, \dots, \phi_{N_g} \right) : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (4.3)$$

entre la phase EOE  $Y_0(x)$  et les autres phases  $N_g$ . Pour apprendre cette transformation, nous utilisons un générateur  $G$  de type U-Net de manière supervisée, dont la dernière couche est une couche à trois canaux, un pour chaque dimension du DVF. Pour réaliser une synthèse d'image dépendante de l'amplitude,  $G$  est conditionné par une valeur scalaire  $\alpha_k$  qui mesure l'amplitude respiratoire relative entre les phases  $Y_0$  et  $Y_k$ , c'est-à-dire que  $\phi_{\alpha_k}^* = G(Y_0, \alpha_k)$  sont les DVFs générés pendant l'entraînement.

Lors du déploiement, les DVFs sont générés pour n'importe quelle valeur  $\alpha$  désirée à partir d'une image TDM 3D pour obtenir l'image souhaitée après déformation :

$$Y^*(X, \alpha) = X \circ \phi_\alpha^*, \quad (4.4)$$

où  $X \circ \phi_\alpha^*$  représente  $X$  déformé par  $\phi_\alpha^*$ .

Pour conditionner la génération d’images selon le schéma respiratoire du patient, nous utilisons la méthode décrite dans le chapitre 3, l’ajout d’une couche AdaIN [106] pour adapter les caractéristiques du vecteur latent de l’image à l’amplitude respiratoire  $\alpha$  au niveau du goulot d’étranglement de  $G$ .

## 4.2.2 Fonction de coût pour les DVFs

La fonction de coût de notre modèle de synthèse de DVF est une fonction composée :

$$\mathcal{L}(G, D) = \lambda_1 \mathcal{L}_{\ell_1}^{\phi_\alpha}(G) + \lambda_2 \mathcal{L}_{\text{GAN}}^{Y, \|\phi_\alpha\|}(G, D), \quad (4.5)$$

où  $\lambda_1$  et  $\lambda_2$  sont des paramètres contrôlant l’équilibre entre ces deux termes.

En se distinguant des méthodes non supervisées qui contraignent explicitement la régularité des DVFs, nous entraînons  $G$  de manière supervisée en ajoutant un coût de reconstruction  $\ell_1$  sur les composantes marginales du DVF :

$$\mathcal{L}_{\ell_1}^{\phi_\alpha} = \sum_{xyz} \|\phi_{\alpha_k}^* - \phi_k\|_1, \quad (4.6)$$

où chaque composante du champ est traitée indépendamment.  $\phi_k$  est la sortie de pTVreg [65]. Dans le cas d’application aux poumons, cet algorithme basé sur la minimisation isotropique de la variation totale a prouvé son efficacité et sa précision en atteignant des résultats de TRE submillimétriques lors du challenge DIR-Lab [78] et, jusqu’à présent, aucune autre méthode conventionnelle ou basée sur du DL a réussi à atteindre ces performances [99], [112].

Pour encourager la génération de DVFs réalistes  $\phi_\alpha^*$  sans modéliser explicitement la régularité du champ, nous proposons d’entraîner le modèle selon un objectif adversaire commun qui concatène à la fois l’image déformée  $Y_0 \circ \phi_\alpha^*$  et l’amplitude du DVF lui-même :

$$\mathcal{L}_{\text{GAN}}^{Y, \|\phi_\alpha\|} = \mathbb{E}[\log D(X, Y_k, \|\phi_k\|)] + \mathbb{E}[\log(1 - D(X, Y_k^*, \|\phi_{\alpha_k}^*\|))]. \quad (4.7)$$

Le coût GAN évalue l’alignement de la distribution entre l’image cible  $Y_k$  et l’image générée  $Y_k^*$ , c’est-à-dire la phase EOE  $X$  déformée par  $\phi_k^*$ . Pour améliorer davantage la synthèse, la magnitude du DVF est incluse dans le coût adversaire. Le raisonnement derrière cela est que l’information contenue dans la magnitude est liée à l’amplitude du mouvement du schéma respiratoire du patient, qui est associée à la valeur de conditionnement scalaire  $\alpha$ . Les valeurs plus élevées de  $\alpha$  devraient donc être corrélées avec des

amplitudes plus élevées. Nous démontrons dans la section expérimentale que l'ajout de ce terme est, en effet, essentiel pour obtenir une synthèse de mouvement plus fidèle.

## 4.3 Protocole expérimental

### 4.3.1 Métriques quantitatives

Différentes métriques ont été utilisées pour la validation du modèle de synthèse du mouvement respiratoire proposé.

**Qualité de la reconstruction :** Nous avons d'abord utilisé les métriques globales de reconstruction, à savoir la MAE et la similarité structurale (SSIM, *structural similarity*), pour évaluer la qualité globale des phases de TDM 4D synthétisées  $Y_k^*$  par rapport à la phase de référence  $Y_k$ . La SSIM est définie par l'équation suivante :

$$\text{SSIM} = \frac{(2\mu_{Y_k^*}\mu_{Y_k} + C_1)(2\sigma_{Y_k^*}\sigma_{Y_k} + C_2)}{(\mu_{Y_k^*}^2 + \mu_{Y_k}^2 + C_1)(\sigma_{Y_k^*}^2 + \sigma_{Y_k}^2 + C_2)}, \quad (4.8)$$

où  $\mu_{Y_k^*}$  et  $\mu_{Y_k}$  sont les valeurs moyennes des pixels des images  $Y_k^*$  et  $Y_k$ ,  $\sigma_{Y_k^*}$  et  $\sigma_{Y_k}$  sont les variances, et les termes  $C$  sont des constantes de régularisation, où  $C_1$  et  $C_2$  sont égales à  $(0.01 \times 2000)^2$  et  $(0.03 \times 2000)^2$  respectivement, et 2000 est la plage dynamique des images. Une valeur de SSIM proche de 1 indique une très forte similarité.

**Évaluation par région :** Ensuite, nous avons mesuré la variation des volumes pulmonaires globaux en comparant à la fois les phases synthétiques et les phases réelles à la phase initiale EOE. Comme déjà mentionné dans l'introduction, en plus de l'évaluation du mouvement pulmonaire global, il est également essentiel pour la planification du traitement de radiothérapie d'évaluer l'exactitude du modèle au niveau de la tumeur, un problème négligé par les études précédents sur la synthèse de TDM 4D [100], [101]. À cette fin, nous avons utilisé le DSC entre les GTVs délimités sur  $Y_k^*$  et la phase réelle  $Y_k$ . Étant donné la grande variabilité des tailles de tumeurs, nous avons également utilisé l'erreur entre les COMs des tumeurs comme une mesure supplémentaire indépendante du volume.

La respiration affecte non seulement les poumons, mais aussi les OARs, dont la localisation est primordiale pour une planification précise du traitement. Par conséquent, nous avons également évalué le DSC volumétrique et les erreurs de COM pour chaque organe, notamment le cœur, le foie, l'estomac, la rate, les reins et le pancréas.



**Qualité des DVFs :** Pour évaluer la qualité des DVFs générés, nous avons quantifié la TV des composantes du DVF, comme dans [65]. Le raisonnement derrière est que les DVFs réalistes, capables de reproduire des transitions nettes (par exemple, aux interfaces des organes), présentent une TV plus faible que les DVFs excessivement lisses. Conformément à la pratique courante en recalage d’image médicale, nous avons également calculé le déterminant négatif du jacobien de la transformation  $|J(\phi^*)|$  pour caractériser davantage la régularité de  $\phi^*$ .

### 4.3.2 Jeux de données

Nous avons évalué et validé la qualité de notre modèle de synthèse d’images dynamiques proposé en utilisant un jeu de données synthétiques ainsi que deux jeux de données cliniques de TDM 4D qui présentent des différences de caractéristiques.

**Phantom :** De la même manière que dans le chapitre précédent, nous avons construit des images synthétiques, dénommées par la suite « Phantom », afin de vérifier si l’utilisation d’une valeur d’amplitude scalaire sur la surface est pertinente pour conditionner le modèle. Ces images simplifiées de poumons sont composées d’un pseudo-corps et de pseudo-poumons. Le pseudo-corps est un ellipsoïde de grande taille et les pseudo-poumons sont inclus dans le pseudo-corps avec les mêmes propriétés que présentés dans le chapitre précédent. Pour simuler un mouvement surfacique, nous appliquons plusieurs dilatations morphologiques à l’aide d’éléments structurants de taille variable et aléatoire dans la direction de l’axe antéro-postérieur (figure 4.2).

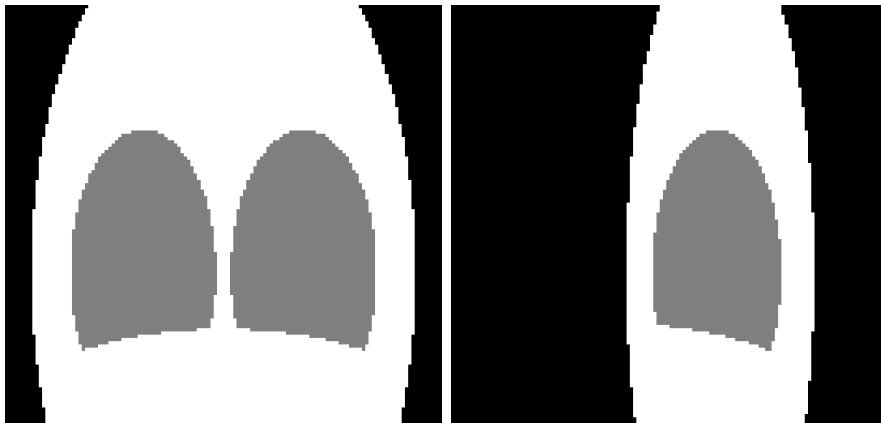


FIGURE 4.2 – Exemple d’une image synthétique en coupe coronale (gauche) et sagittale (droite)

**4D-Lung :** Nous avons utilisé un jeu de données public provenant de *The Cancer*

*Imaging Archive* (TCIA), appelé ci-après « 4D-Lung » [113] et mentionné dans d'autres études de synthèse de TDM 4D [99], [100]. Il est composé de 20 images de patients atteints de cancers non à petites cellules (CBNPC) et comprend des masques de segmentation de GTV. Chaque acquisition TDM 4D est composée de 10 phases couvrant le cycle respiratoire complet, débutant par la phase EOE (phase 0%) pour atteindre la phase EOI (phase 50%) et pour finalement revenir à une phase intermédiaire juste avant la phase EOE (phase 90%) avec un pas de 10%. Les images ont été acquises sur un scanner TDM hélicoïdal à 16 coupes (Brilliance Big Bore, Philips Medical Systems) avec une épaisseur de coupe de 3 mm et une résolution axiale de  $512 \times 512$  (taille de pixel d'environ 1 mm). Les tumeurs sont de taille plutôt importante, avec un volume moyen de  $77 \pm 96 \text{ cm}^3$ . Cela se traduit par une superposition relativement élevée des volumes de tumeurs entre les phases EOE et EOI, évaluée par un DSC de  $0,72 \pm 0,14$  et des erreurs de COM de tumeur de  $5,46 \pm 2,66 \text{ mm}$ .

**CHRU :** Les métriques rapportées sur les tumeurs telles que nous les trouvons dans le jeu de données 4D-Lung entravent le fait que la précision locale peut être insuffisante pour suivre des volumes plus petits, en raison, par exemple, de la dépendance connue du DSC à la taille de l'objet [114]. Pour cette raison, nous avons également pris en compte une base de données cliniques locales composée de 43 patients acquis au département de radio-oncologie du CHRU de Brest, en France. Ce jeu de données, appelé par la suite « CHRU », contient des patients présentant à la fois des tumeurs primaires et des métastases pulmonaires. Comparées au jeu de données 4D-Lung, les données CHRU présentent des GTVs beaucoup plus petits en moyenne et une plus grande variabilité des volumes, avec des volumes moyens de  $6,42 \pm 9,11 \text{ cm}^3$ , ce qui se traduit par un DSC initial beaucoup plus faible de  $0,39 \pm 0,31$  et des distances plus grandes entre les COM ( $8,44 \pm 6,04 \text{ mm}$ ). Une autre caractéristique notable est l'inclusion de patients ayant subi des acquisitions répétées de TDM 4D : 30 patients ont eu deux TDM 4D, cinq patients ont eu trois TDM 4D et un patient a eu quatre TDM 4D. Cela permet d'évaluer la variabilité inter TDM 4D, qui est une caractéristique clé pour évaluer l'intérêt clinique d'un modèle de synthèse de mouvement respiratoire. Chaque acquisition de TDM a été reconstruite en huit phases respiratoires : de la phase 0% à la phase 100% correspondant à la phase EOE et EOI respectivement. Puis de la phase EOI jusqu'à un retour en phase EOE, avec un pas de 25%. Les images ont été acquises avec un scanner Siemens Sensation Open, une épaisseur de coupe de 2 mm et une résolution axiale de  $512 \times 512$  (taille de pixel d'environ 1 mm).

La figure 4.3a montre la répartition des volumes de tumeurs dans les deux jeux de

données avec deux exemples typiques. Pour détailler davantage les différences de caractéristiques, la figure 4.3b montre la distribution des distances entre les phases EOE et EOI dans chaque acquisition de TDM 4D, ce qui est lié à l’amplitude de correction du mouvement requise par l’approche de synthèse de mouvement. Comparé à 4D-Lung, le jeu de données CHRU montre des différences de volumes plus importantes entre les phases EOE et EOI, des GTVs plus petits et une plus grande diversité d’aspects de tumeurs.

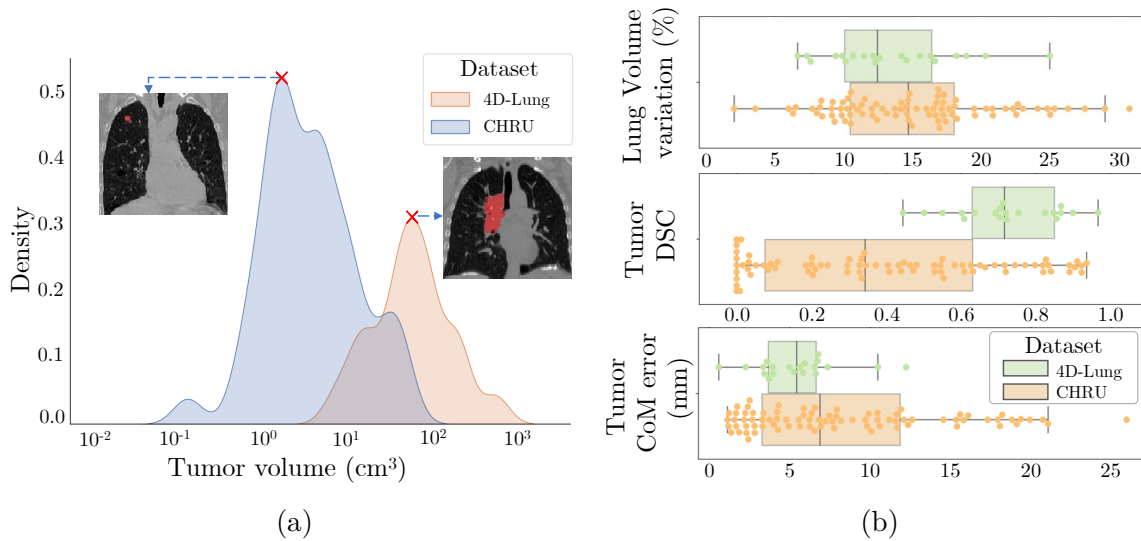


FIGURE 4.3 – Caractéristiques des deux jeux de données cliniques utilisés pour nos expérimentations. (a) Distributions des volumes des tumeurs dans les jeux de données 4D-Lung (rouge) et CHRU (bleu); (b) Différentes métriques d’évaluation entre les phases EOE et EOI.

### 4.3.3 Expériences additionnelles

**Étude de la variabilité des TDM 4D :** En raison de la variabilité intrinsèque des schémas respiratoires chez les patients acquis à différents moments, ce qui est lié aux limitations inhérentes de la planification de traitement en TDM 4D, une mesure utile pour valider la pertinence clinique de la méthode de synthèse du mouvement respiratoire proposée consiste à vérifier que la distance des zones d’intérêt entre la phase synthétique  $Y_k^*$  et la phase réelle  $Y_k$  n’est pas supérieure à la distance entre les acquisitions répétées de TDM 4D  $Y_k^1$  et  $Y_k^2$  d’un même patient. Les acquisitions répétées ont donc été comparées en utilisant le DSC et l’erreur COM pour évaluer cette variabilité.

**Validation de la fonction de coût composée :** Nous avons quantifié l’amélioration obtenue par la fonction objective proposée, à savoir la composition de la fonction par un

coût de reconstruction  $\ell_1$  sur les composantes du DVF et d'un coût adversaire sur la magnitude du DVF. À cette fin, nous avons également entraîné des modèles alternatifs dans lesquels le coût total incluait des combinaisons de termes de reconstruction et adversaires basés sur l'image déformée et/ou le DVF.

**Évaluation comparative :** Nous avons comparé notre approche à un travail récent, RMSim [101], qui génère également des séquences TDM 4D à partir d'images statiques et de signal respiratoire. RMSim utilise une architecture seq2seq modifiée basée sur l'apprentissage non supervisé des DVFs pour garantir la cohérence temporelle du mouvement. Contrairement à notre approche, qui repose sur des mesures respiratoires externes, l'amplitude a été obtenue à partir d'informations internes dérivées de l'image (déplacement maximal du diaphragme dans l'axe supérieur-inférieur), ce qui n'est pas réaliste en routine clinique. Cependant, comme seul le code d'inférence était disponible auprès des auteurs, pour garantir une comparaison équitable, nous avons ré-entraîné notre approche en utilisant un conditionnement similaire et avons utilisé le même scénario d'entraînement (c'est-à-dire que nous apprenons à mettre en correspondance la phase EOI vers la phase EOE, plutôt que de EOE vers EOI). Les auteurs ont entraîné leur modèle sur 100 patients avec des TDM 4D à 10 phases en utilisant quatre GPU A40 avec 48 Go de mémoire chacun. Leur modèle utilise 35,2 Go des 48 Go disponibles et le temps d'entraînement était de 72h approximativement.

#### 4.3.4 Implémentation et entraînement

Le modèle a été entraîné en 3D. Pour le modèle entraîné avec les données Phantom, nous avons utilisé 800 images pour l'entraînement et 200 pour la validation. Pour les données cliniques, le jeu de données d'entraînement était composé de 42 images TDM 4D provenant de 20 patients des données CHRU. Le jeu de données de test était composé de 44 autres images TDM 4D provenant de 23 patients supplémentaires des données CHRU et de 20 images TDM 4D provenant de 20 patients des données 4D-Lung. Le modèle a été entraîné pendant 200 époques en utilisant l'optimisation ADAM [109]. Les valeurs des hyperparamètres  $\lambda_1$  et  $\lambda_2$  sont fixées à 100 et 1 respectivement.

Les acquisitions initiales de TDM 4D étaient de dimensions  $512 \times 512 \times Z$ , où  $Z$  représente l'axe axial et varie en fonction des protocoles d'acquisition. Les images ont été rognées autour des segmentations poumons qui ont été extraites grâce à une méthode robuste de segmentation automatique des poumons [110]. Nous avons ensuite sélectionné la zone d'intérêt délimitée par la plus grande boîte englobante et nous avons étendu de 2

cm en dessous du diaphragme pour inclure les organes abdominaux d'intérêt. Ce choix a été fait en fonction de l'ensemble des données valides afin de capturer le plus d'organes abdominaux possibles dans le champ de vue de traitement. Enfin, chaque image a été ré-échantillonnée à une taille de  $128 \times 128 \times 128$  pour des raisons de contraintes de mémoire de processeur graphique (GPU, *graphics processing unit*) relativement modeste (figure 4.4). Dans toutes nos expériences, nous avons utilisé une carte graphique NVIDIA GeForce GTX 1080Ti avec 11 Go de mémoire et une taille de batch de 1. Le temps d'entraînement du modèle était d'environ 16h.

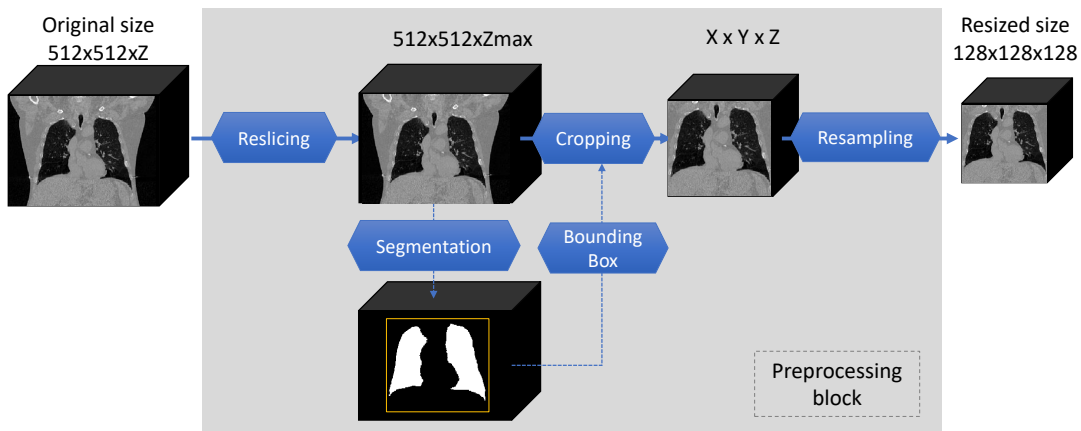


FIGURE 4.4 – Étape de prétraitement des images.

Les masques de tumeur disponibles pour chaque phase du jeu de données 4D-Lung ont été utilisés dans les expériences nécessitant des délimitations des GTVs. Pour le jeu de données CHRU, les GTVs ont été obtenus automatiquement dans toutes les phases en utilisant un modèle nnU-Net [39] entraîné avec les poids gagnants de la tâche de tumeur pulmonaire du *Medical Segmentation Decathlon* [111]. Les segmentations du cœur ont été réalisées à l'aide de PlatiPy, un logiciel de segmentation libre qui exploite nnU-Net avec des heuristiques de traitement d'image supplémentaires [115]. Les segmentations des organes restants (foie, estomac, reins et rate) ont été obtenues en entraînant un modèle nnU-Net sur le jeu de données d'entraînement du *Multi-Modality Abdominal Multi-Organ Segmentation Challenge (AMOS) 2022* [116]. Dans tous les cas, les masques de segmentation finaux ont été revus par un expert et corrigés manuellement si nécessaire.

Pour chaque acquisition répétée disponible dans l'étude de variabilité des TDM 4D, nous avons effectué un recalage rigide entre les phases EOE pour éviter les décalages dans les positions des tumeurs et des organes. L'erreur COM a ensuite été calculée entre les

deux segmentations de tumeurs. L'erreur de distance dans chaque axe a été utilisée pour effectuer un deuxième recalage avec la deuxième acquisition, corrigeant l'erreur COM initiale. Par la suite, la transformation a été appliquée aux autres phases et aux segmentations de la deuxième acquisition pour assurer la cohérence entre toutes les phases (tableau 4.1).

	Erreur initiale	Après recalage
Erreur du COM (mm)	$4,16 \pm 5,09$	$1,07 \pm 0,25$
Axe droite-gauche	$1,28 \pm 1,27$	$0,62 \pm 0,34$
Axe antéro-postérieur	$2,56 \pm 3,71$	$0,56 \pm 0,31$
Axe supérieur-inférieur	$2,55 \pm 3,65$	$0,44 \pm 0,32$
DSC	$0,64 \pm 0,18$	$0,72 \pm 0,19$

TABLE 4.1 – Correction de la position de la tumeur entre les multiples acquisitions pour la phase expiratoire.

La valeur de conditionnement scalaire  $\alpha$  a été calculée à partir des mesures d'amplitude respiratoire de manière entièrement automatisée. Pour chaque phase TDM 4D des deux jeux de données, nous avons calculé une amplitude respiratoire (en mm) à partir de l'image en suivant la méthode décrite dans [117], imitant ainsi les informations qui peuvent être obtenues à partir d'un système RPM. Chaque phase a été segmentée pour obtenir le corps du patient [118]. La différence en surface du patient entre les images binaires a été obtenue entre la phase EOE et les autres phases. La distance moyenne dans l'axe antéro-postérieur a été définie comme la valeur  $\alpha$  (figure 4.5).

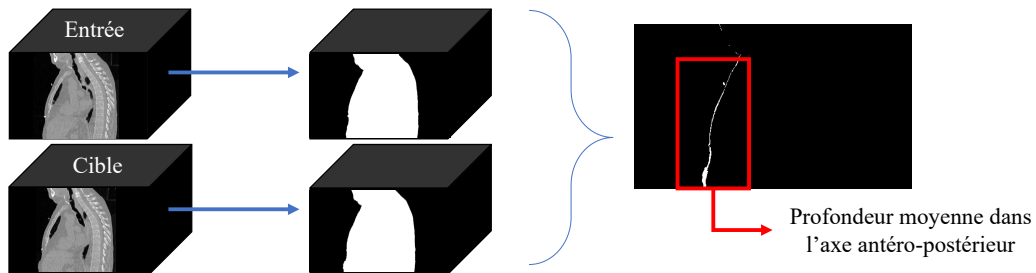


FIGURE 4.5 – Extraction du mouvement surfacique.

## 4.4 Résultats

### 4.4.1 Modèle synthétique

La figure 4.6 montre un cas de synthèse d'image avec la méthode proposée. Nous observons que l'image cible présente une augmentation de volume pulmonaire ainsi que celle du corps par rapport à l'image d'entrée (figure 4.6a-4.6b). La figure 4.6c montre l'image générée à partir de l'image d'entrée avec l'amplitude respiratoire correspondant à l'image cible. La figure 4.6d correspond à la superposition de l'image générée et l'image cible. Nous n'observons que de faibles erreurs au niveau du diaphragme, indiquant que le modèle parvient à reproduire fidèlement la dynamique des images.

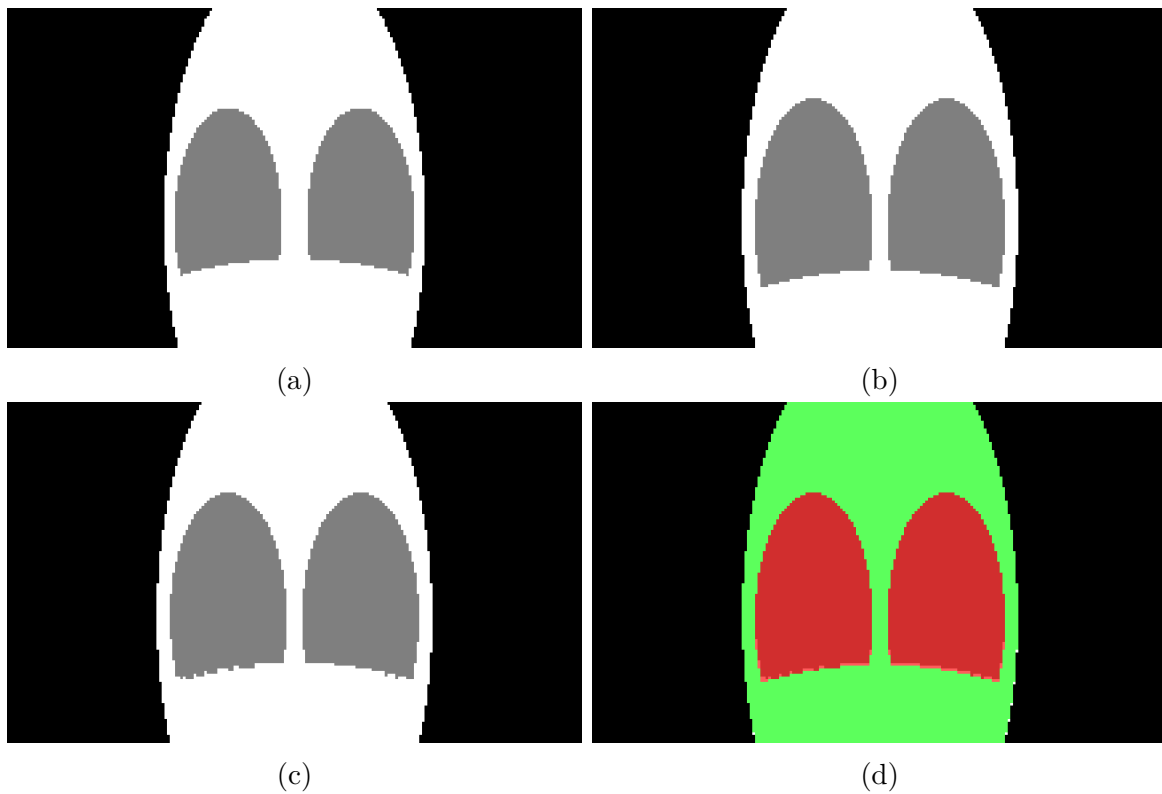


FIGURE 4.6 – Exemple d'un cas de synthèse d'image des données Phantom. (a) Image d'entrée; (b) Image cible; (c) Image générée; (d) Superposition des images cible et générée : le rouge indique le chevauchement entre les pseudo-poumons des images et le vert, le chevauchement des pseudo-corps.

Les résultats quantitatifs avec les données Phantom sont présentés dans le tableau 4.2. Nous pouvons observer une erreur VV moyenne de 25,53% et un DSC de 0,89 montrant ainsi une différence de taille initiale des poumons. Après génération, le modèle est capable

de faire respirer les images, démontré par une réduction de l'erreur VV et une augmentation du DSC pour l'ensemble des dilatations. Concernant la surface du corps, un meilleur chevauchement est observé en moyenne sur le DSC du corps après synthèse (0,99) malgré une erreur très faible initial (0,95). Les amplitudes initiales mesurées entre les dilatations sont plus importantes que celles mesurées après la synthèse avec, par exemple, pour la dernière dilatation, des valeurs passant de 3,21 à 0,26. Ces résultats démontrent que le modèle arrive à faire respirer correctement les données synthétiques tout en conservant une cohérence avec la mesure d'amplitude respiratoire.

	Moyenne globale	Dilatation 1	Dilatation 2	Dilatation 3	Dilatation 4
VV initial	25,53 ± 15,06	9,81 ± 3,16	20,03 ± 6,54	30,63 ± 10,14	41,63 ± 13,96
VV après synthèse	2,81 ± 2,02	0,82 ± 0,52	2,20 ± 1,52	4,21 ± 2,05	4,01 ± 1,45
DSC initial (poumons)	0,89 ± 0,06	0,95 ± 0,01	0,91 ± 0,03	0,87 ± 0,04	0,83 ± 0,05
DSC après synthèse (poumons)	0,99 ± 0,01	0,99 ± 0,01	0,99 ± 0,01	0,98 ± 0,01	0,98 ± 0,01
DSC initial (corps)	0,95 ± 0,02	0,98 ± 0,01	0,96 ± 0,01	0,94 ± 0,02	0,93 ± 0,02
DSC après synthèse (corps)	0,99 ± 0,01	0,99 ± 0,01	0,99 ± 0,01	0,99 ± 0,01	0,99 ± 0,01
Amplitude initiale (mm)	2,01 ± 1,13	0,80 ± 0,25	1,61 ± 0,51	2,41 ± 0,76	3,21 ± 1,01
Amplitude après synthèse (mm)	0,12 ± 0,23	0,04 ± 0,05	0,03 ± 0,02	0,14 ± 0,19	0,26 ± 0,38

TABLE 4.2 – Résultats du modèle avec les données Phantom

#### 4.4.2 Analyses globales des images et des DVFs

La figure 4.7 montre les phases synthétiques obtenues par notre générateur à partir d'une phase EOE d'un patient pour différentes valeurs de  $\alpha$ , qui ont été comparées à l'acquisition TDM 4D réelle correspondante. Visuellement, une descente du diaphragme a été obtenue avec l'augmentation des valeurs de  $\alpha$ , en accord avec les phases TDM 4D réelles. Nous pouvons noter que le modèle est également capable d'interpolation, c'est-à-dire de générer des images entre deux phases respiratoires réelles, et d'extrapolation, abaisser la position maximale du diaphragme au-delà de la phase réelle de EOI.

Pour illustrer qualitativement l'impact de la fonction de coût proposée, la figure 4.8 compare les DVFs générés obtenus en utilisant des modèles avec des fonctions de coût totales différentes. Lorsque seule le coût de reconstruction sur le DVF est utilisée (figure 4.8a), le DVF généré est plus lisse et principalement unidirectionnel le long de l'axe supérieur-inférieur, avec une magnitude absolue faible correspondant à un abaissement moins prononcé du diaphragme. Lorsque le coût adversaire est ajouté en ne considérant que l'image déformée  $Y^* = X \circ \phi_k^*$ , comme cela peut être fait par exemple dans les



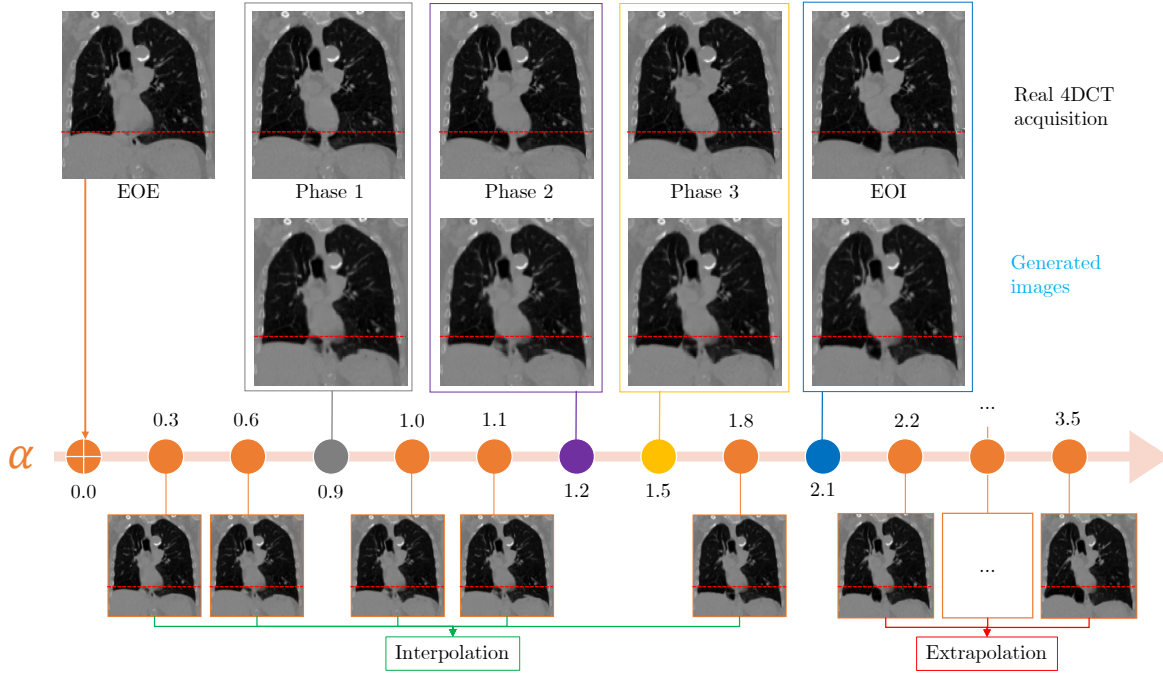


FIGURE 4.7 – Aperçu qualitatif de la capacité de synthèse de mouvement respiratoire du modèle. Une phase EOE d’une acquisition réelle TDM 4D (haut) est transmise au réseau avec des valeurs d’amplitude respiratoire  $\alpha$  variables. Le niveau initial du diaphragme de la phase EOE est représenté par la ligne rouge. En fonction de  $\alpha$ , le modèle peut reproduire les phases souhaitées (centre), mais aussi interpoler entre des phases réelles ou extrapoler au-delà de la phase EOI (bas).

méthodes de recalage non supervisées [73], l’amplitude du déplacement se rapproche du champ de référence, mais reste néanmoins excessivement lisse (figure 4.8b). Un résultat de DVF plus fidèle est obtenu grâce à l’utilisation conjointe du coût de reconstruction sur le champ ainsi que du coût adversaire prenant en compte la magnitude du DVF  $\|\phi^*\|$  (figure 4.8c). Des transitions plus nettes sont obtenues, en meilleure concordance avec le DVF obtenu avec la méthode de recalage de référence pTVreg (figure 4.8d).

Pour confirmer ces observations visuelles, le tableau 4.3 présente les résultats obtenus en utilisant d’autres combinaisons de termes de coût. Les entraînements en utilisant uniquement des coûts de reconstruction de manière supervisée ( $\mathcal{L}_{\ell_1}^\phi$ ) ou en imitant les méthodes de recalages non supervisées en utilisant l’image déformée ( $\mathcal{L}_{\ell_1}^Y$ ) conduisent à des erreurs de reconstruction relativement élevées (MAE moyenne  $\simeq 55$  et SSIM moyenne  $\simeq 0,57$ ). Lorsque le coût  $\ell_1$  est combiné à un coût GAN sur l’image déformée  $Y^*$  uniquement, des champs excessivement lisses sont mesurés par des valeurs élevées de TV et de déterminant de jacobien  $|J|$ . En revanche, les meilleurs scores quantitatifs par rapport à toutes les mesures considérées sont obtenus lorsque le coût adversaire prend conjointement en

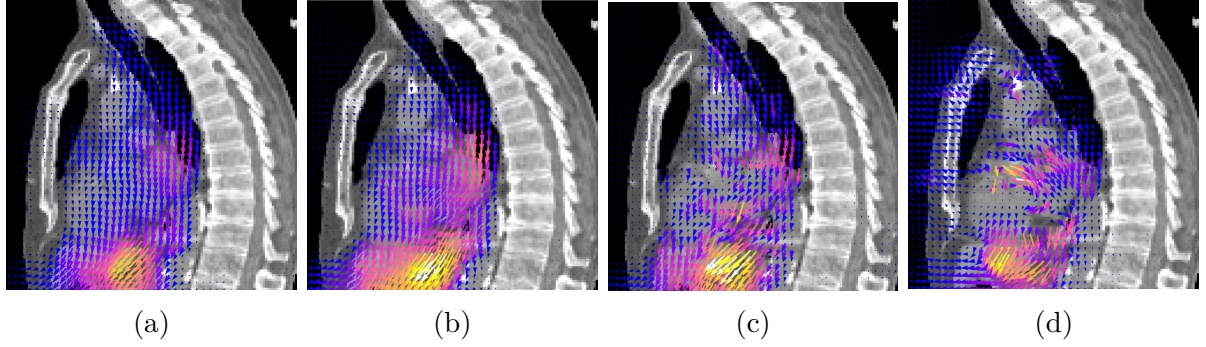


FIGURE 4.8 – Influence des termes de coût sur le DVF généré. (a) Coût de reconstruction uniquement :  $\ell_1$  sur le DVF ; (b) Ajout du coût adversaire sur l'image déformée ; (c) Fonction totale proposée :  $\ell_1$  sur le DVF + coût adversaire conjoint sur l'image déformée et la magnitude du DVF ; (d) DVF de référence obtenu par pTVreg [65]

compte la magnitude de  $\phi^*$ , avec une SSIM moyenne beaucoup plus élevée ( $0,74 \pm 0,05$ ) et une MAE plus faible ( $31,75 \pm 5,23$ ), ainsi que des mesures de champs plus rugueux, comparables au DVF de référence obtenu par pTVreg.

	Image déformée		DVF	
	MAE	SSIM	TV	$ J $
Pas de synthèse	$54,53 \pm 8,98$	$0,53 \pm 0,06$	N/A	N/A
$\mathcal{L}_{\ell_1}^Y$	$55,77 \pm 13,17$	$0,57 \pm 0,10$	$0,25 \pm 0,08$	$1,037 \pm 0,018$
$\mathcal{L}_{\ell_1}^\phi$	$56,29 \pm 13,79$	$0,57 \pm 0,10$	$0,26 \pm 0,08$	$1,037 \pm 0,016$
$\mathcal{L}_{\ell_1}^Y + \mathcal{L}_{\text{GAN}}^Y$	$43,00 \pm 10,64$	$0,57 \pm 0,07$	$0,41 \pm 0,08$	$1,028 \pm 0,010$
$\mathcal{L}_{\text{GAN}}^{Y, \ \phi_\alpha\ }$	$42,59 \pm 13,85$	$0,63 \pm 0,10$	$0,21 \pm 0,02$	$1,004 \pm 0,002$
$\mathcal{L}_{\ell_1}^\phi + \mathcal{L}_{\text{GAN}}^Y$	$37,39 \pm 12,12$	$0,68 \pm 0,10$	$0,12 \pm 0,04$	$1,008 \pm 0,005$
$\mathcal{L}_{\ell_1}^\phi + \mathcal{L}_{\text{GAN}}^{Y, \ \phi_\alpha\ }$	<b><math>31,75 \pm 5,23</math></b>	<b><math>0,74 \pm 0,05</math></b>	<b><math>0,12 \pm 0,04</math></b>	<b><math>1,006 \pm 0,004</math></b>
DVF de référence	N/A	N/A	$0,13 \pm 0,03$	$1,007 \pm 0,006$

TABLE 4.3 – Analyses de l'impact des fonctions de coûts dans la fonction totale.

Le tableau 4.4 résume les résultats de la reconstruction évalués par les métriques quantitatives MAE et VV des données de test de 4D-Lung. La VV moyenne est de 7,69% et la MAE moyenne est de 36,64 avant la synthèse du mouvement, ce qui suggère des changements de volume relativement modérés entre la phase EOE et les autres phases. Après la synthèse du mouvement, la VV est réduite en moyenne à 1,61%. Dans les résultats des données de CHRU (tableau 4.5), une VV pulmonaire initiale plus élevée de 11,90%

est réduite en moyenne à 4,96% après la synthèse du mouvement. La MAE est également considérablement améliorée. L’entraînement de notre modèle uniquement sur les données CHRU ne nuit pas aux performances obtenues sur les données 4D-Lung, suggérant une bonne capacité de généralisation même lorsque les ensembles de données ont des propriétés différentes.

	Phase 40%	Phase 30%	Phase 20%	Phase 10%	EOI
VV initial (%)	1,88±0,76	4,33±1,56	7,38±3,23	11,43±3,98	13,45±4,74
VV après synthèse (%)	1,02±0,61	1,67±0,85	2,23±1,17	1,65±1,41	1,51±1,32
MAE initiale	16,55±5,16	25,89±5,82	36,84±7,67	48,39±10,12	55,53±11,46
MAE après synthèse	16,15±4,92	23,53±4,69	31±5,72	37,37±6,71	41,69±7,17

TABLE 4.4 – Analyses globales des images après synthèse pour les données 4D-Lung.

	Phase 25%	Phase 50%	Phase 75%	EOI
VV initial (%)	7,23±3,01	10,95±4,09	14,14±5,19	15,22±6,04
VV après synthèse (%)	5,88±5,29	5,22±4,22	4,59±3,69	4,15±3,67
MAE initiale	46,35±6,71	55,58±7,92	62,99±9,58	65,01±11,66
MAE après synthèse	42,81±8,88	46,64±7,91	49,49±7,86	50,17±7,89

TABLE 4.5 – Analyses globales des images après synthèse pour les données CHRU.

### 4.4.3 Évaluation locale

La figure 4.9 montre une phase synthétique d’un patient des données CHRU avec une tumeur relativement petite (0,54 cm<sup>3</sup>) visible à la fois dans les plans sagittal et coronal. La position de la tumeur et du diaphragme après la synthèse du mouvement correspond bien à l’image vérité terrain dans les deux plans (figure 4.9b). Un autre exemple avec une tumeur plus grande (31,56 cm<sup>3</sup>) est présenté à la figure 4.10. Les images de différence entre les phases avant (figure 4.10b) et après la synthèse du mouvement (figure 4.10c) illustrent la capacité du modèle à reproduire fidèlement le mouvement de la tumeur.

La figure 4.11 montre des boîtes à moustaches des chevauchements volumétriques entre les phases EOI et EOE pour les organes considérés dans nos expériences dans les deux jeux de données. Pour tous les organes, le DSC s’est nettement amélioré après la synthèse du mouvement. Cela indique que notre modèle est non seulement capable de modéliser le mouvement à l’intérieur des poumons, mais peut également déformer les

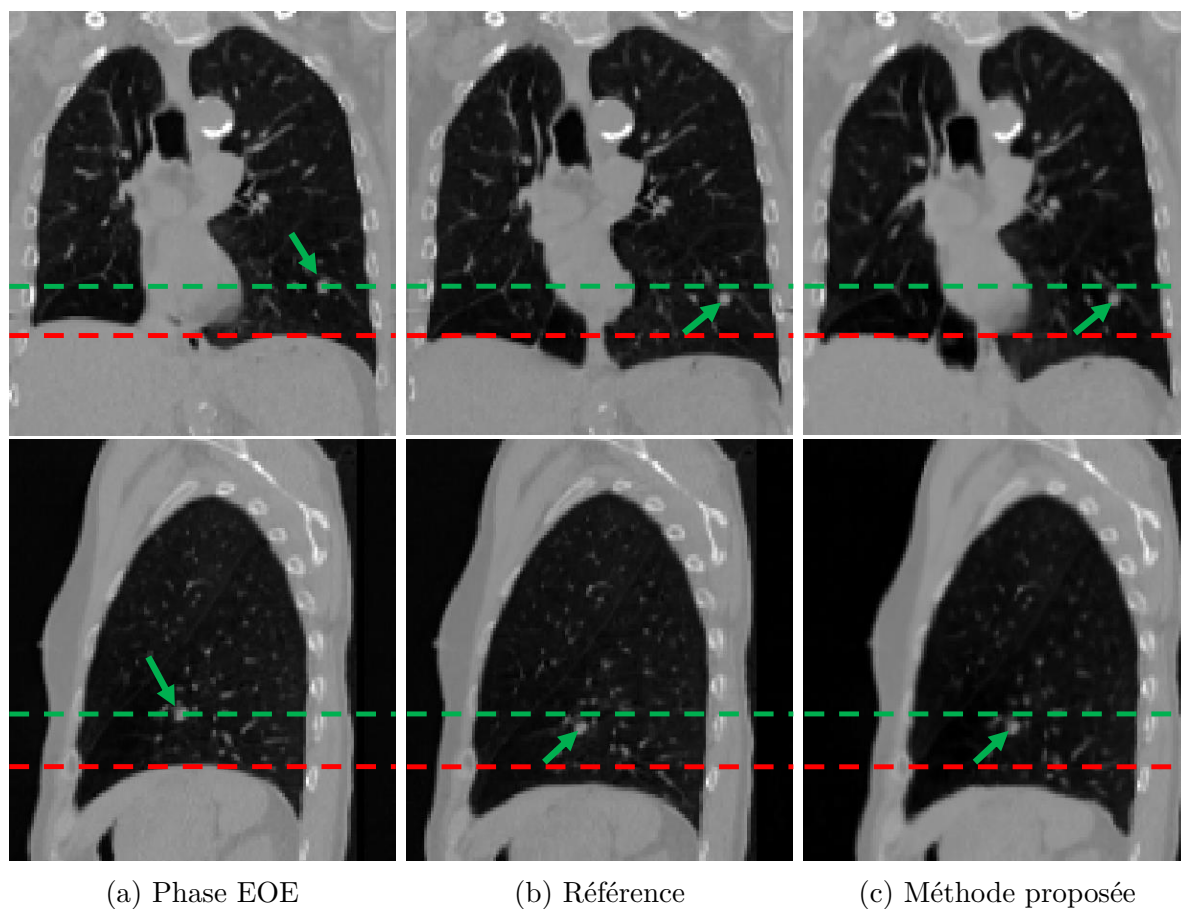


FIGURE 4.9 – Mouvement pulmonaire global sur les plans coronal et sagittal. Les lignes rouge et vert représentent la position initiale du diaphragme dans le poumon gauche et la tumeur respectivement. La flèche verte montre l'emplacement de la tumeur.

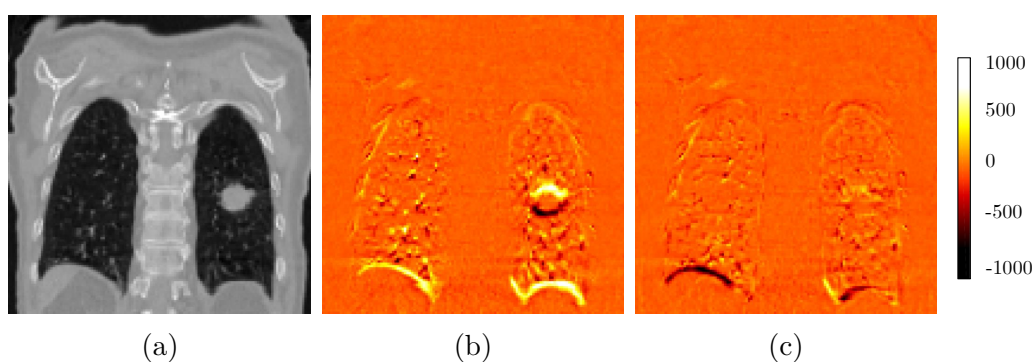


FIGURE 4.10 – L'évaluation de la synthèse du mouvement de la tumeur par les images de différence entre (a) Phases EOE et EOI; (b) Erreur initiale; (c) Erreur après la synthèse, où la suppression du signal de la tumeur montre une bonne correspondance entre les phases réelles et synthétiques. Un désaccord dû à une légère surestimation du mouvement du diaphragme peut être observé, marqué par des valeurs négatives dans la partie inférieure de l'image.

organes périphériques avec une bonne précision, avec un chevauchement plus élevé après la synthèse pour tous les organes ( $EOE \geq 0,65$ ).

Le tableau 4.6 présente les résultats de la synthèse du mouvement au niveau de la tumeur dans le jeu de données 4D-Lung. Comme discuté précédemment, l'ensemble de données 4D-Lung présente de plus faibles erreurs initiales de COM entre les tumeurs avec une moyenne de 2,83 mm et un DSC moyen de la tumeur de 0,85 en raison du fait que les tumeurs sont de taille plus importante et présente un mouvement relativement faible. Après la synthèse, l'erreur moyenne de COM est réduite à 1,45 mm pour un DSC moyen de 0,89 dans les données 4D-Lung. En revanche, les données CHRU comportent des erreurs initiales de COM moyennes plus grandes (4,86 mm) et un DSC initial moyen plus faible (0,61). L'erreur moyenne est de COM de 2,90 mm et un DSC moyen de 0,70 dans le jeu de données CHRU (tableau 4.7). Malgré les différences dans les caractéristiques des tumeurs, le modèle est capable de déplacer les tumeurs dans les deux jeux de données de manière réaliste et précise.

	Phase 40%	Phase 30%	Phase 20%	Phase 10%	EOI
DSC initial	0,99±0,03	0,93±0,07	0,85±0,10	0,77±0,12	0,73±0,14
DSC après synthèse	0,98±0,03	0,93±0,06	0,88±0,08	0,84±0,09	0,82±0,10
Erreur COM initiale (mm)	0,21±0,36	1,26±0,95	2,72±1,67	4,30±2,03	5,46±2,66
Erreur COM après synthèse (mm)	0,33±0,50	0,98±0,72	1,49±1,15	2,08±1,18	2,37±1,35

TABLE 4.6 – Résultats centrés sur la tumeur avant (sans synthèse de mouvement) et après la synthèse des données 4D-Lung.

	Phase 25%	Phase 50%	Phase 75%	EOI
DSC initial	0,69±0,29	0,61±0,31	0,55±0,33	0,57±0,37
DSC après synthèse	0,73±0,19	0,72±0,14	0,68±0,19	0,69±0,20
Erreur COM initiale (mm)	3,55±3,59	4,69±4,61	5,64±5,27	5,56±5,33
Erreur COM après synthèse (mm)	2,68±1,80	2,78±1,51	3,13±1,65	2,99±1,87

TABLE 4.7 – Résultats centrés sur la tumeur avant (sans synthèse de mouvement) et après la synthèse des données CHRU

La figure 4.12 fournit une analyse plus exhaustive des schémas de mouvement des tumeurs en montrant les résultats individuels pour chaque phase respiratoire de tous les sujets. Les boîtes à moustaches se concentrent sur la distribution des erreurs de COM initiales et après la synthèse, où le modèle a obtenu une réduction substantielle des erreurs de COM dans les deux jeux de données.

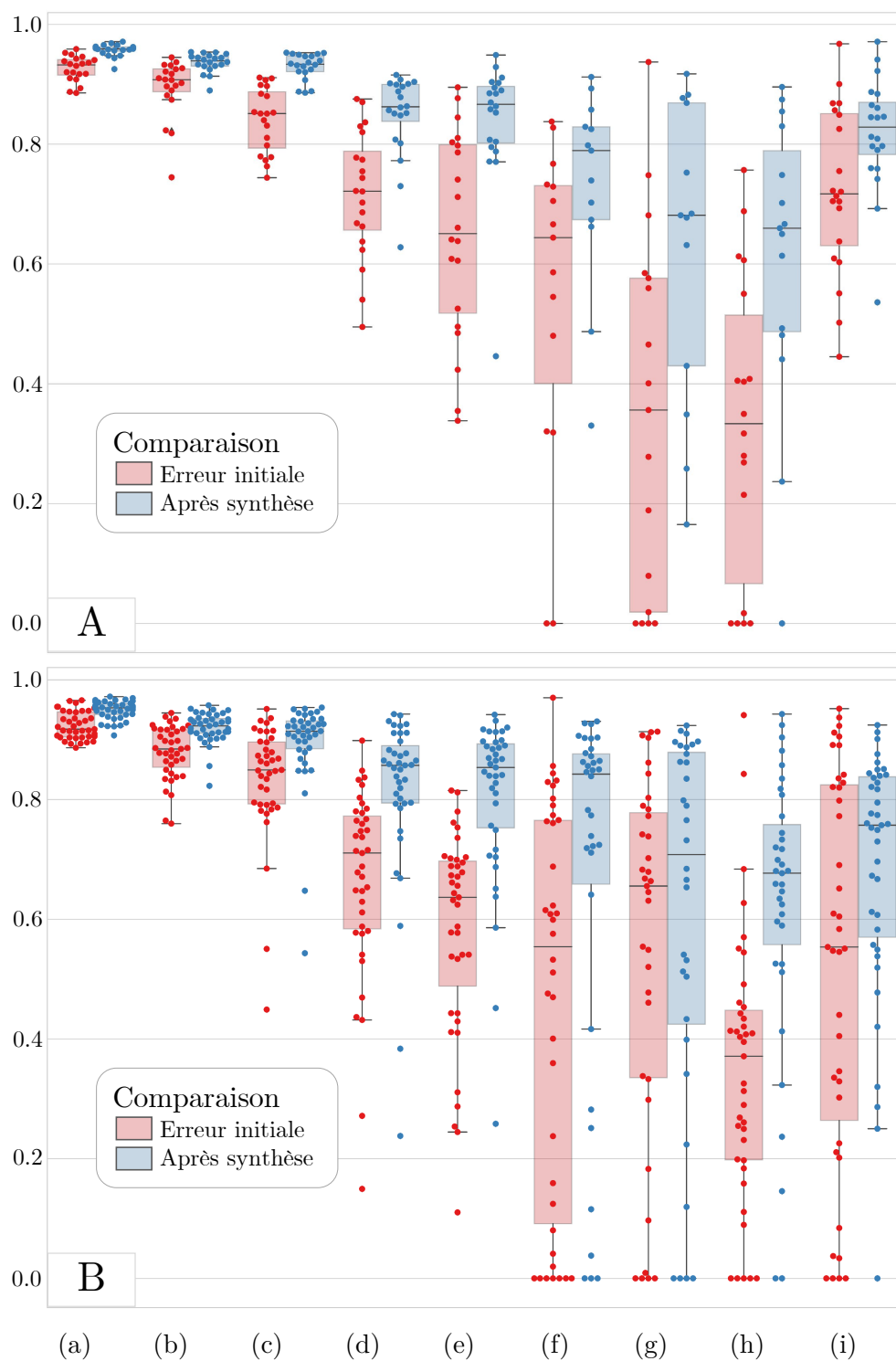


FIGURE 4.11 – Comparaison du chevauchement des organes entre les images pour la phase EOI dans les jeux de données 4D-Lung (A) et CHRU (B). (a) Poumon (b) Cœur (c) Foie (d) Estomac (e) Rate (f) Rein gauche (g) Rein droit (h) Pancréas (i) Tumeur.

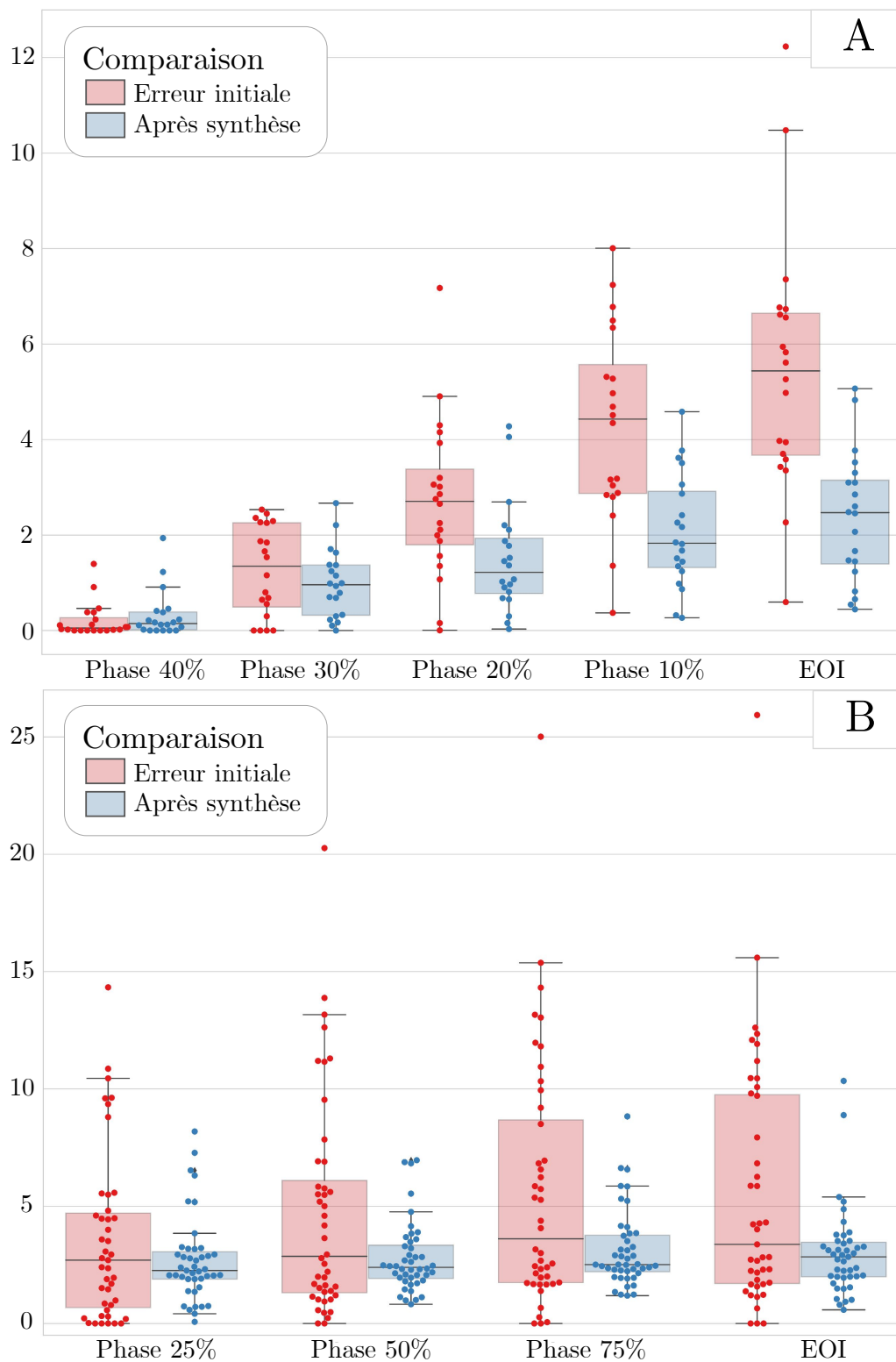


FIGURE 4.12 – Comparaison de l’erreur de COM de la tumeur (en mm) entre les phases dans les jeux de données 4D-Lung (A) et CHRU (B).

	Erreur COM (mm)		DSC						
	Tumeur	Tumeur	Cœur	Foie	Estomac	Rate	Rein droit	Rein gauche	Pancréas
4DCT1 vs 4DCT2	1,97±1,05	0,63±0,25	0,90±0,04	0,88±0,05	0,67±0,13	0,70±0,09	0,69±0,14	0,68±0,20	0,58±0,10
Ours1 vs 4DCT1	2,35±0,08	0,71±0,15	0,91±0,02	0,91±0,03	0,78±0,05	0,78±0,08	0,72±0,15	0,67±0,15	0,69±0,11
Ours2 vs 4DCT2	2,86±1,32	0,67±0,25	0,93±0,02	0,90±0,04	0,79±0,07	0,78±0,07	0,67±0,16	0,69±0,17	0,63±0,15

TABLE 4.8 – Étude de variabilité des TDM 4D.

Le tableau 4.8 présente les résultats quantitatifs correspondant à l'étude de variabilité TDM 4D, où nous évaluons l'accord entre deux acquisitions TDM 4D réalisées par rapport à l'accord entre les acquisitions réelles et synthétiques. Dans ce contexte, la phase EOI de la première acquisition TDM 4D (4DCT1) est comparée à la phase EOI de la deuxième acquisition (4DCT2). Les erreurs moyennes de COM entre les acquisitions répétées et le DSC sont respectivement de 1,97 mm et de 0,63, mettant en évidence la variabilité inhérente dans la position observée de la tumeur entre les acquisitions TDM 4D réelles répétées. La comparaison entre les phases réelles et synthétiques conduit à des erreurs moyennes de COM similaires de 2,35 mm et 2,86 mm et un DSC moyen de 0,71 et 0,67 pour la tumeur. Des résultats similaires sont obtenus pour les chevauchements volumétriques entre les OARs, avec une variabilité comparable entre les images des acquisitions répétées réelles et les images synthétiques.

La figure 4.13 compare la précision de la synthèse des tumeurs dans le jeu de données 4D-Lung entre notre approche et RMSim [101]. Les erreurs initiales du COM pour la phase EOI sont en moyenne de 5,46 mm et un DSC de 0,72 suggérant un mouvement tumoral maximal modéré. RMSim est capable de réduire cette erreur à 3,99 mm et un DSC de 0,75 tandis que notre méthode parvient à obtenir de meilleurs résultats avec une erreur COM de 2,51 mm et un DSC de 0,82. Notre modèle a constamment surpassé RMSim pour toutes les phases respiratoires avec des moyennes d'erreurs COM ou DSC plus favorables, malgré un ensemble de données d'entraînement plus petit et des exigences computationnelles plus faibles à la fois pendant l'entraînement et les tests.

Le tableau 4.9 présente les résultats sur le DSC entre le modèle RMSim et notre approche. Nous pouvons observer que dans l'ensemble notre méthode parvient à obtenir des DSC plus important que le modèle RMSim.



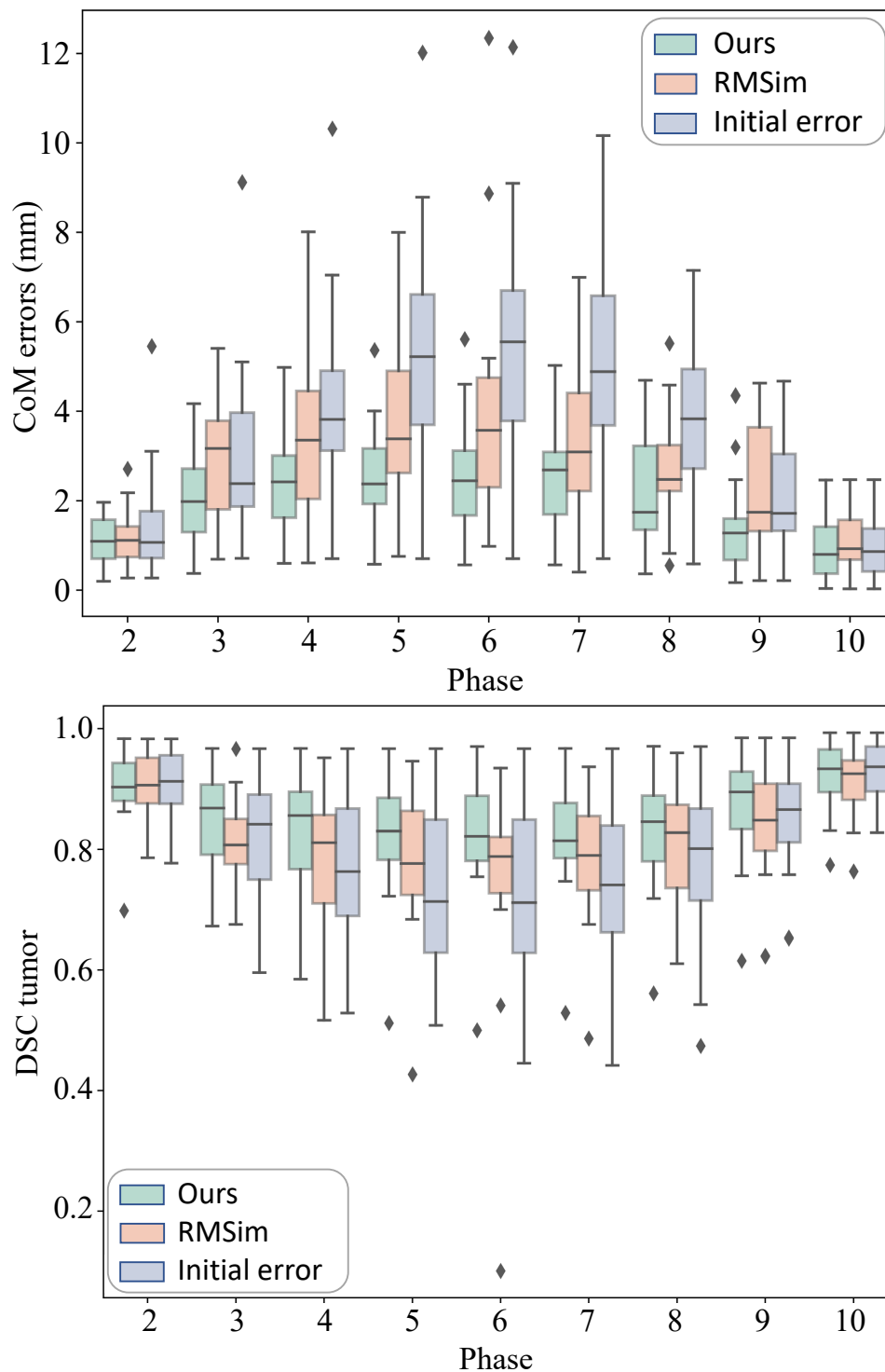


FIGURE 4.13 – Précision de la synthèse du mouvement tumoral par phases respiratoires dans le jeu de données 4D-Lung (évaluée par les erreurs COM en mm et de DSC). Notre approche est comparée à RMSim, un travail récent basé sur l’apprentissage non supervisé des champs DVFs avec une régularisation temporelle [101].

TABLE 4.9 – Comparaison des performances du modèle RMSim et de notre approche proposée sur les différents OARs.

	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6	Phase 7	Phase 8	Phase 9	Phase 10	
Heart	Initial error	0,96±0,02	0,94±0,03	0,92±0,03	0,90±0,04	0,89±0,05	0,89±0,05	0,91±0,03	0,94±0,02	0,97±0,02
	RMSim	0,96±0,02	0,94±0,02	0,92±0,02	0,92±0,02	0,91±0,04	0,91±0,03	0,92±0,03	0,94±0,02	0,96±0,02
	Ours	0,96±0,02	0,95±0,01	0,94±0,01	0,94±0,02	0,94±0,02	0,93±0,02	0,94±0,02	0,95±0,01	0,96±0,02
Spleen	Initial error	0,90±0,07	0,79±0,12	0,72±0,14	0,67±0,16	0,66±0,17	0,67±0,17	0,74±0,15	0,84±0,08	0,93±0,07
	RMSim	0,91±0,06	0,80±0,20	0,79±0,15	0,79±0,13	0,77±0,16	0,77±0,17	0,78±0,16	0,84±0,08	0,91±0,08
	Ours	0,91±0,06	0,85±0,10	0,82±0,11	0,82±0,12	0,81±0,12	0,81±0,13	0,83±0,12	0,87±0,07	0,93±0,07
Right kidney	Initial error	0,70±0,25	0,46±0,33	0,37±0,32	0,35±0,31	0,35±0,31	0,37±0,30	0,43±0,32	0,57±0,33	0,74±0,27
	RMSim	0,60±0,33	0,28±0,30	0,25±0,28	0,30±0,30	0,27±0,24	0,28±0,26	0,38±0,30	0,55±0,35	0,71±0,29
	Ours	0,70±0,25	0,45±0,33	0,36±0,31	0,35±0,30	0,32±0,26	0,35±0,26	0,42±0,32	0,56±0,34	0,72±0,29
Left kidney	Initial error	0,86±0,11	0,73±0,17	0,62±0,25	0,53±0,28	0,53±0,28	0,53±0,28	0,62±0,25	0,73±0,24	0,91±0,07
	RMSim	0,81±0,13	0,62±0,17	0,54±0,25	0,47±0,28	0,48±0,25	0,48±0,27	0,58±0,23	0,70±0,23	0,89±0,07
	Ours	0,81±0,16	0,64±0,19	0,52±0,22	0,46±0,26	0,44±0,28	0,45±0,27	0,53±0,25	0,64±0,27	0,82±0,14
Liver	Initial error	0,96±0,02	0,92±0,03	0,89±0,04	0,86±0,05	0,86±0,05	0,86±0,05	0,89±0,04	0,93±0,02	0,97±0,01
	RMSim	0,96±0,01	0,93±0,02	0,90±0,04	0,90±0,04	0,89±0,05	0,90±0,04	0,91±0,03	0,93±0,03	0,96±0,01
	Ours	0,96±0,02	0,93±0,02	0,91±0,02	0,91±0,02	0,90±0,02	0,90±0,03	0,91±0,03	0,94±0,02	0,97±0,01
Stomach	Initial error	0,92±0,04	0,83±0,08	0,77±0,09	0,73±0,10	0,72±0,10	0,73±0,11	0,77±0,10	0,86±0,07	0,94±0,03
	RMSim	0,92±0,04	0,82±0,11	0,79±0,10	0,77±0,09	0,77±0,08	0,75±0,11	0,78±0,10	0,85±0,07	0,93±0,03
	Ours	0,91±0,05	0,84±0,09	0,82±0,09	0,81±0,11	0,81±0,09	0,81±0,10	0,84±0,08	0,87±0,07	0,93±0,03
Pancreas	Initial error	0,75±0,19	0,47±0,31	0,38±0,27	0,33±0,25	0,33±0,25	0,33±0,26	0,40±0,29	0,57±0,29	0,86±0,11
	RMSim	0,68±0,22	0,34±0,30	0,31±0,26	0,28±0,25	0,27±0,25	0,29±0,25	0,38±0,28	0,56±0,28	0,84±0,14
	Ours	0,75±0,19	0,45±0,31	0,39±0,29	0,38±0,30	0,37±0,29	0,37±0,29	0,40±0,29	0,54±0,29	0,78±0,16

## 4.5 Discussion

Dans ce chapitre, nous avons entraîné notre modèle génératif explicitement sur des DVFs pulmonaires tout en encourageant le réalisme du mouvement sous un objectif adversaire qui contraint conjointement l'image déformée et la magnitude du DVF. La raison principale de cette stratégie est d'améliorer le mouvement entre phases et d'encourager indirectement les corrélations avec l'amplitude respiratoire qui est transmise simultanément au modèle, résultant ainsi en une amélioration de la qualité de la reconstruction par rapport aux coûts de reconstruction conventionnels. Ces résultats concordent avec les observations précédentes selon lesquelles les pertes adverses sont efficaces pour I2I uniquement lorsque la reconstruction n'est pas trop contraignante [119]. Cela explique probablement pourquoi nos efforts dans le chapitre précédent utilisant des pertes de reconstruction directes n'ont pas produit de résultats satisfaisants sur les tumeurs.

Nous avons validé notre méthode en utilisant deux jeux de données de tumeurs pulmonaires TDM 4D présentant des caractéristiques différentes à travers une variété d'images, d'OARs et de métriques de qualité au niveau des tumeurs basées sur le volume, mais aussi sur leur distance de déplacement. Ce type de validation est essentiel afin de garantir l'estimation précise du mouvement des tumeurs et des OARs voisins pour mieux calculer la dose délivrée au patient. Nous avons comparé les résultats obtenus avec une méthode récente, RMSim. Notre méthode présente des meilleurs résultats et des avantages par rapport à la méthode comparée tels qu'un temps de calcul inférieur et une intégration réaliste grâce à un conditionnement surfacique et non basé sur un point interne difficilement accessible en routine clinique.

Un résultat clé de notre travail est la démonstration que l'erreur associée à la synthèse du mouvement est comparable aux incohérences généralement obtenues entre les acquisitions TDM 4D répétées du même patient. Les erreurs moyennes COM tumorales atteintes étaient inférieures à 3 mm, cela est conforme à une étude récente recommandant des marges PTV dues au mouvement respiratoire entre 3 mm et 5 mm selon la localisation tumorale [23].

En ce qui concerne le conditionnement de la respiration, nous avons proposé de contraindre le modèle avec une valeur scalaire basée sur l'amplitude extraite des traces de signal respiratoire externe 1D spécifiques au patient au lieu d'une différence de volume. Cependant, cela reste une simplification du problème en ne considérant qu'une seule valeur simple pour représenter la complexité du mécanisme respiratoire.

## 4.6 Conclusion

Ce chapitre a présenté notre méthode principale de synthèse de phases respiratoires à partir d’images statiques et d’un signal respiratoire externe, en s’appuyant sur le même mécanisme d’injection scalaire que le modèle présenté au chapitre précédent. Plusieurs contributions importantes ont été faites dans ce travail, tant du point de vue méthodologique que du point de vue de la vérification expérimentale. D’une part, et contrairement aux approches existantes de la littérature qui produisent des champs de déformation par apprentissage non supervisé, nous proposons de ne pas nous intéresser explicitement au problème du recalage en exploitant des champs de déformation produits par une méthode de l’état de l’art en recalage thoracique, et d’apprendre de manière supervisée sur ces champs considérés comme des vérités terrains. La méthode est ainsi découplée du problème du recalage, et permet de s’appuyer par exemple sur des méthodes non issues du DL. Nous ajoutons durant l’apprentissage une contrainte non supervisée conjointement sur l’image déformée, mais aussi, et de manière critique, sur l’amplitude des DVFs. Ce faisant, nous évitons de forcer la régularité du champ de manière explicite comme dans les méthodes de recalage non supervisées, et encourageons les corrélations avec l’amplitude respiratoire du patient. Nous montrons que ce second aspect est déterminant dans l’amélioration des performances de la méthode.

La vérification expérimentale des performances de l’approche couvre un plus grand nombre d’aspects que les approches existantes, qui mesuraient des métriques globales ou des mesures de type DSC sur des tumeurs de grande taille. Cette analyse plus fine met en défaut les méthodes actuelles et montre que le modèle permet un mouvement local plus précis de la tumeur, notamment comparé à l’état de l’art ou à l’approche de synthèse directe proposée dans le chapitre précédent. De manière importante, nous comparons également ces images synthétisées à des acquisitions TDM 4D réelles répétées, ce qui n’avait encore jamais été proposé dans la littérature. Les résultats encourageants montrent que le modèle présente une variabilité proche de celle mesurée entre deux acquisitions réelles, ce qui, nous le croyons, est une étape primordiale vers l’acceptabilité clinique de telles approches de synthèse.

Notre approche permet de générer des mouvements locaux et réalistes, mais conditionnée par une simple valeur scalaire. Ce faisant, nous faisons l’hypothèse qu’il existe une bijection entre cette mesure d’amplitude externe et le mouvement interne, ce qui est une simplification du problème. La suite de ce travail de thèse s’intéresse à révéler des corrélations

lations plus complexes à travers des mesures respiratoires de dimension supérieure. Nous verrons ainsi dans le chapitre suivant une approche d'un conditionnement plus complexe basé sur des nuages de points surfaciques, constituant une information plus riche. L'inclusion de ce type d'information dans le conditionnement pourrait ainsi améliorer davantage les performances du modèle.

# CONDITIONNEMENT RESPIRATOIRE SURFACIQUE PAR NUAGES DE POINTS

---

## Résumé

Nous présentons dans ce chapitre nos travaux préliminaires concernant l'exploitation de conditionnements plus complexes qu'une simple amplitude respiratoire, à savoir un mouvement surfacique représenté par un nuage de points. L'objectif est d'incorporer des technologies de suivi respiratoire plus récentes, capables de fournir des informations plus riches qu'un simple signal respiratoire unidimensionnel, dans le cadre de la pratique clinique courante.

Les méthodes traditionnelles de synchronisation respiratoire pour l'acquisition de TDM 4D offrent des informations générales sur le cycle respiratoire. L'utilisation d'un nuage de points en surface permet une description plus précise du comportement respiratoire global d'un patient, offrant ainsi une meilleure caractérisation des mouvements.

Après avoir brièvement décrit le concept des nuages de points et des applications possibles en imagerie médicale, nous détaillons notre méthode de simulation de mouvement surfacique. Nous testons ce concept à travers divers modèles et de zones d'intérêt pour la surface. Nous finissons sur les résultats et sur les conclusions que nous pouvons tirer de nos expériences ainsi que des perspectives de cette méthode, qui est restée au stade de l'exploration préliminaire durant le temps de la thèse.

## 5.1 Introduction

Un nuage de points, dans le contexte de la vision par ordinateur et de la géométrie 3D, est un ensemble de coordonnées tridimensionnelles qui représentent les positions spatiales de points dans l'espace. Chaque point du nuage est défini par ses coordonnées en trois dimensions. Les nuages de points sont couramment utilisés pour capturer et représenter

la géométrie tridimensionnelle d’objets, de scènes ou d’environnements. Ils jouent un rôle essentiel dans divers domaines, tels que la modélisation 3D, la reconstruction de surfaces, la réalité augmentée, la cartographie, la robotique et bien d’autres applications basées sur du DL [120]-[122]. Il est possible d’obtenir ce type de données avec, par exemple, des caméras de profondeur ou encore caméras TOFs. Ces caméras ont la particularité de retourner à l’utilisateur des images sous forme de nuages de points indiquant notamment la distance entre la caméra et un objet.

En imagerie médicale, les caméras TOFs trouvent aussi leur place. Wentz et coll. proposent d’utiliser ces caméras comme un outil dans diverses applications médicales telles que l’acquisition de données respiratoire précise sur le mouvement respiratoire des patients lors d’acquisition TDM 4D [123], [124]. Nazir et coll. démontrent la possibilité de positionner précisément un patient à l’aide d’une caméra TOF et donc sans nécessiter de marquages sur le patient ni l’utilisation d’une dose de radiation supplémentaire pour le positionnement [16].

La manipulation de l’espace latent est une méthode bien connue du domaine de la vision par ordinateur afin de générer une variété d’images. Des nombreuses méthodes montrent des résultats très réalistes, notamment par des techniques de transfert de style à l’aide d’images [105], [106], [125]. Nous avons montré qu’il est aussi possible d’effectuer une manipulation simple à l’aide d’une valeur scalaire afin de générer la phase respiratoire souhaitée. Toutefois, l’apport d’information par une valeur simple peut paraître limité et peut se retrouver face à des difficultés lors de cas plus complexe et moins général de la respiration (changement de la manière de respirer au cours de l’acquisition, toux, irrégularité respiratoire, etc.). Ainsi, une réduction de dimension de la surface globale à une valeur simple peut sous-estimer ou surestimer la respiration en fonction de la manière de respirer du patient. Il est donc intéressant d’explorer si des informations plus complexes telles que des nuages de points permettent ou non d’améliorer le réalisme de la respiration.

Dans ce chapitre, nous décrivons donc une approche basée sur des nuages de points afin de conditionner notre modèle DL. Nous détaillons dans un premier temps l’étape de simulation et les modifications apportées afin d’intégrer ces informations dans le réseau de neurones. Nous continuons avec un protocole expérimental basé sur des comparaisons de modèles afin de mesurer et de comprendre le changement apporté par ces données.

## 5.2 Simulation de mouvement surfacique

Pour obtenir des informations sur le mouvement surfacique du patient en routine clinique, il est possible d'utiliser des caméras TOFs qui vont extraire les informations [16]. Dans notre cas, seules les images TDM 4D sont disponibles pour nos expérimentations. Nous proposons donc une méthode afin de simuler un mouvement respiratoire en surface à partir de nuage de points et des DVFs obtenus avec les acquisitions et de s'en servir pour conditionner notre modèle génératif.

Soient  $X$  une phase source et  $Y$  une phase cible d'une acquisition TDM 4D. Entre ces deux phases respiratoires, nous pouvons obtenir un DVF  $\phi$  avec une méthode de recalage. De plus, nous pouvons obtenir  $P(X)$  et  $P(Y)$ , les nuages de points en surfaces de l'image  $X$  et  $Y$  respectivement. À partir de ces données, il est possible d'extraire seulement les vecteurs d'intérêts de  $\phi$  pour représenter un mouvement respiratoire surfacique :

$$\exists \delta \in \phi, p + \delta = q, \quad (5.1)$$

où  $p$  et  $q$  sont des points de  $P(X)$  et  $P(Y)$  respectivement et  $\delta$  est un vecteur permettant de déplacer le point  $p$  vers le point  $q$ . L'ensemble des vecteurs  $\delta$  représentent ainsi un mouvement surfacique d'une phase  $X$  à une autre phase  $Y$  (figure 5.1).

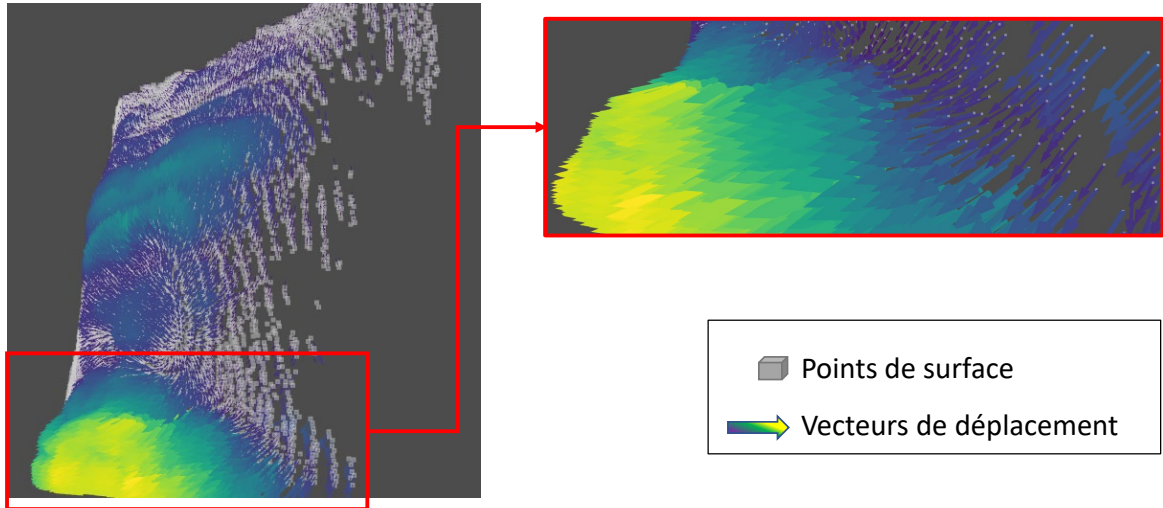


FIGURE 5.1 – Illustration d'un mouvement surfacique.

Pour rappel, notre objectif est de générer un DVF à partir de  $X$  ainsi qu'un condi-



tionnement sur l’amplitude respiratoire qui permet, après transformation, d’obtenir une phase  $Y^*$  qui partage la même dynamique que  $Y$ . Nous remplaçons notre ancienne valeur  $\alpha$ , une valeur scalaire basée sur l’amplitude respiratoire, par les vecteurs  $\delta$  pour simuler un déplacement en surface. En raison de la nouvelle dimensionnalité du conditionnement, nous avons apporté une modification sur la structure de notre méthode d’injection basée la couche AdaIN. Auparavant, seule une couche de neurones entièrement connectés était utilisée pour prendre en compte une valeur scalaire simple. Cette couche est remplacée par un perceptron multicouche (MLP, *multilayer perceptron*) avec plusieurs couches cachées, permettant ainsi au réseau d’apprendre des caractéristiques internes des données de conditionnement. L’adaptation des caractéristiques dans l’espace latent reste inchangée.

## 5.3 Protocole expérimental

### 5.3.1 Jeux de données

Dans nos expérimentations, nous utilisons une partie des données du jeu de données CHRU décrit dans le chapitre 4. Nous reprenons les acquisitions et appliquons un pré-traitement différent pour capturer au mieux le mouvement surfacique. Les acquisitions ne présentant que de très faible mouvement, voire aucun mouvement surfacique sont exclues. Lors des calculs de l’erreur COM, les acquisitions présentant des déplacements de tumeurs inférieurs à 2 mm faussent la précision concernant les tumeurs mobiles, nous décidons donc d’exclure ces acquisitions. Pour rappel, les acquisitions TDM 4D sont composées de 8 phases représentant un cycle respiratoire complet. Nous incluons ici l’ensemble des phases du cycle respiratoire des acquisitions TDM 4D.

Les fichiers sont acquis à partir d’un scanner Siemens Sensation Open avec une épaisseur de coupe de 2 mm et une résolution axiale de  $512 \times 512$  (taille de pixel d’environ 0,975 mm). Nous décidons d’abord de les rendre isotropes, c’est-à-dire que les acquisitions ont le même espacement dans toutes les directions. Après cette opération, les images ont donc un espacement entre les voxels de 2 mm et sont de tailles  $250 \times 250 \times Z$  où  $Z$  est la composante axiale avec des valeurs différentes en fonction des acquisitions. Nous segmentons les poumons des acquisitions avec une méthode de segmentation automatique [110].

À partir de la segmentation, nous nous plaçons au niveau supérieur des poumons et nous intégrons tous les organes se trouvant dans l’intervalle de 160 pixels dans l’axe

supérieur-inférieur. Pour l'axe antéro-postérieur, nous nous plaçons à l'arrière des poumons et nous incluons les poumons ainsi que la surface toujours dans un intervalle de 160 pixels. Enfin, dans l'axe gauche-droite, nous calculons le centre dans l'axe par rapport aux poumons et nous englobons une zone de 80 pixels des deux côtés. Ce choix fixe de taille de  $160 \times 160 \times 160$  ne permet pas de garantir l'intégralité des poumons dans certains cas et sont donc exclues. De plus, les images après traitement qui ne présentent pas la totalité de la surface sont aussi exclues. Nous obtenons donc un total de 69 acquisitions TDM 4D de 60 patients. Nous entraînons nos modèles à partir de 47 images de 40 patients et validons sur les 22 images des 20 autres patients.

### 5.3.2 Extraction de nuage de point

Nous proposons d'extraire des nuages de points pour chaque phase respiratoire des patients. Pour cela, nous commençons par obtenir une segmentation du corps du patient qui s'effectue avec à la méthode de segmentation automatique TotalSegmentator [38]. À partir de la segmentation, nous allons appliquer la méthode *marching cubes* qui permet de créer un objet polygonal, ou plus précisément un maillage (en anglais *mesh*), qui est composé de sommets, d'arêtes et de faces. Grâce à ce maillage, il est possible d'extraire des points de coordonnées qui vont constituer le nuage de points (figure 5.2).

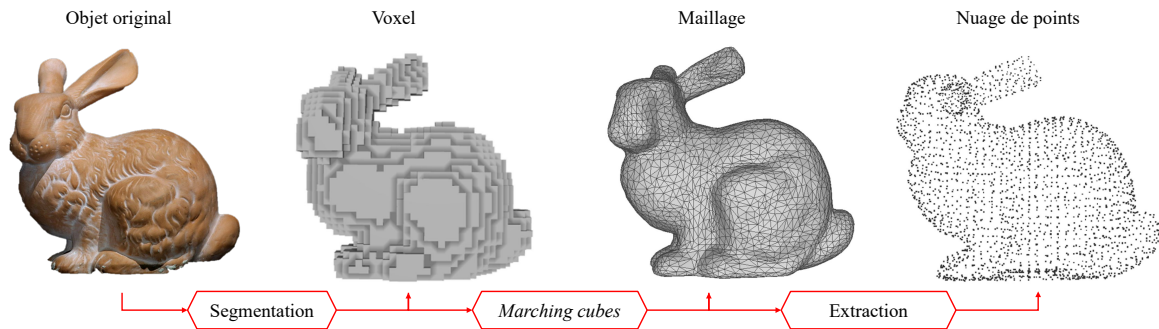


FIGURE 5.2 – Illustration de la conversion d'un objet 3D en nuage de points. Modifié de : [126]

### 5.3.3 Méthode d'évaluation

**Zones d'intérêts :** Pour simplifier l'entraînement, nous nous focalisons seulement dans certaines zones en surface plutôt que la surface totale. Cela a pour but d'éviter de

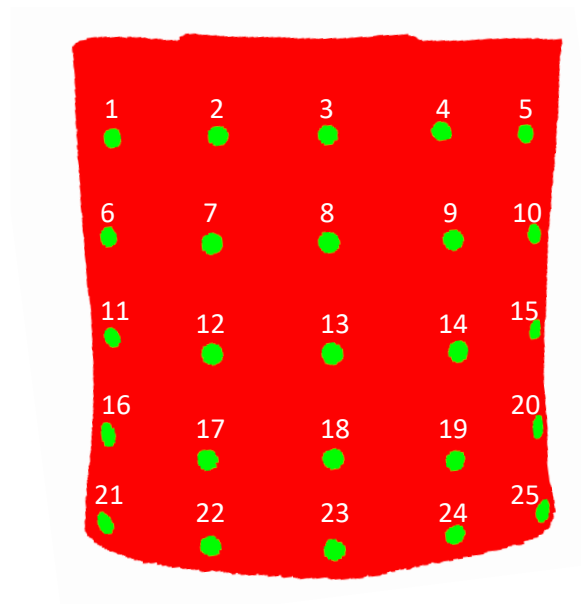


FIGURE 5.3 – Illustration des disques en surface utilisés pour les expérimentations.

Zones verticales		Zones horizontales		Zones composées	
Zone A	1,6,11,16,21	Zone F	1 à 5	Zone K	1 à 10
Zone B	2,7,12,17,22	Zone G	6 à 10	Zone L	1 à 15
Zone C	3,8,13,18,23	Zone H	11 à 15	Zone M	16 à 25
Zone D	4,9,14,19,24	Zone I	16 à 20	Zone N	11 à 25
Zone E	5,10,15,20,25	Zone J	21 à 25	All	1 à 25

TABLE 5.1 – Répartitions des zones.

complexifier l'apprentissage en incluant des zones non informatives. Ainsi, nous définissons 25 disques de surfaces de rayons 5 mm que nous avons choisis manuellement et en supposant qu'il y ait une corrélation avec le mouvement interne du patient (figure 5.3).

Nous entraînons le modèle avec chaque disque de manière indépendante afin d'évaluer leur impact sur les résultats et nous comparons les résultats avec le modèle de base, conditionné avec la moyenne de l'amplitude surfacique. De plus, à partir de ces disques individuels, nous définissons des zones d'intérêts composées de plusieurs disques (tableau 5.1).

Les DVFs sont calculés avec la méthode pTVreg entre la phase EOE, en tant qu'image mobile, et les autres phases respiratoires de l'acquisition TDM 4D. Nous extrayons tous les vecteurs liés aux points situés dans les zones définies. Les vecteurs sont moyennés pour chaque direction afin d'obtenir un déplacement moyen par zone.

**Influence des couches cachées :** Afin d'étudier l'impact du nombre de couches cachées du MLP sur les résultats, nous proposons aussi d'ajouter à la version présentant les meilleurs résultats des modèles alternatifs avec des nombres différents de couches cachées. Les modèles initiaux seront basés sur trois couches cachées. Nous évaluons par la suite des variations de 1 à 10 couches cachées.

**Métriques :** Nous utilisons les mêmes métriques que dans le chapitre précédent pour évaluer les performances des modèles. Cela inclut l'évaluation globale des images avec la VV, la MAE à l'intérieur des poumons, mais aussi le DSC et l'erreur COM pour les tumeurs. Nous effectuons une évaluation supplémentaire qui consiste à comparer les volumes d'intérêt ITV. Pour cela, nous combinons l'ensemble des segmentations des tumeurs de chaque phase pour simuler l'ITV. Nous utilisons le DSC afin d'évaluer le chevauchement des volumes d'intérêts.

## 5.4 Résultats

### 5.4.1 Effet de chaque disques

Le tableau 5.2 montre les résultats des différents conditionnements pour l'ensemble des phases du cycle respiratoire. Initialement, la VV est de 10,62% et la MAE de 142,75 montrant une différence entre les phases inspiratoires de la phase expiratoire. Après synthèse, nous pouvons observer que les conditionnements avec les disques réduisent les erreurs jusqu'à 3,24% pour la VV (zone 05) et 87,07 pour la MAE (zone 12). L'erreur initiale de chevauchement et distance des COMs des tumeurs est de 0,37 et 7,84 mm respectivement sur toutes les phases respiratoires. La zone 06 obtient le meilleur score de DSC avec 0,58 et la zone 05 obtient la meilleure réduction de l'erreur COM avec 4,33 mm. De plus, les résultats avec les disques de la partie supérieure du corps (1 à 15) présentent des résultats de VV plus proches du conditionnement de référence, environ 3,50%, que ceux de la partie inférieure (16 à 25), environ 4,34%. Les résultats avec la MAE montrent un comportement similaire avec en moyenne une erreur de 88,70 pour la partie supérieure et une erreur de 93,19 pour la partie inférieure. Pour les tumeurs, les deux parties ont des résultats similaires, un score de DSC de 0,55 et une erreur COM de 4,65 pour la partie supérieure et pour la partie inférieure, un DSC de 0,54 et une erreur COM de 4,95.

Le tableau 5.3 présente les résultats uniquement pour la phase EOI. Le conditionnement avec la zone 03 donne les meilleurs résultats de VV et de MAE avec une moyenne

Modèle	VV ↓	MAE ↓	DSC ↑	COM ↓
Erreur initiale	10,62 ± 7,21	142,75 ± 48,2	0,37 ± 0,27	7,84 ± 4,27
Baseline	3,51 ± 3,43	89,82 ± 35,89	0,53 ± 0,25	4,90 ± 2,94
Zone 01	3,74 ± 3,46	90,34 ± 34,79	0,56 ± 0,23	4,65 ± 2,81
Zone 02	3,73 ± 4,00	90,40 ± 37,00	0,56 ± 0,24	4,62 ± 2,96
Zone 03	3,39 ± 3,18	87,74 ± 34,00	0,55 ± 0,25	4,79 ± 3,05
Zone 04	3,62 ± 3,35	88,03 ± 35,13	0,53 ± 0,26	5,04 ± 3,21
Zone 05	3,24 ± 3,21	87,66 ± 35,50	0,58 ± 0,23	4,33 ± 2,75
Zone 06	3,51 ± 3,32	88,79 ± 33,94	0,58 ± 0,25	4,41 ± 3,30
Zone 07	3,68 ± 3,48	88,90 ± 35,47	0,55 ± 0,26	4,85 ± 3,40
Zone 08	3,47 ± 3,42	89,77 ± 34,05	0,56 ± 0,24	4,77 ± 3,22
Zone 09	3,54 ± 3,29	88,71 ± 34,09	0,56 ± 0,24	4,67 ± 3,08
Zone 10	3,43 ± 3,15	88,98 ± 32,46	0,54 ± 0,26	4,69 ± 3,03
Zone 11	3,56 ± 3,79	88,50 ± 37,13	0,56 ± 0,25	4,70 ± 3,26
Zone 12	3,53 ± 3,44	87,07 ± 34,61	0,56 ± 0,25	4,58 ± 3,10
Zone 13	3,30 ± 3,47	88,93 ± 35,83	0,55 ± 0,25	4,67 ± 2,91
Zone 14	3,24 ± 3,31	87,97 ± 34,81	0,56 ± 0,25	4,63 ± 2,96
Zone 15	3,51 ± 3,36	88,74 ± 35,50	0,58 ± 0,26	4,39 ± 3,06
Zone 16	4,53 ± 4,06	94,18 ± 39,83	0,54 ± 0,25	4,96 ± 3,32
Zone 17	3,98 ± 3,85	92,69 ± 38,13	0,56 ± 0,27	4,86 ± 3,61
Zone 18	4,36 ± 4,04	93,42 ± 39,54	0,54 ± 0,26	5,08 ± 3,48
Zone 19	4,47 ± 3,92	93,06 ± 38,49	0,55 ± 0,25	4,72 ± 3,10
Zone 20	3,86 ± 3,80	92,39 ± 37,99	0,56 ± 0,26	4,63 ± 3,37
Zone 21	3,89 ± 3,92	91,63 ± 39,63	0,55 ± 0,25	4,95 ± 3,39
Zone 22	4,57 ± 3,79	93,84 ± 39,75	0,53 ± 0,26	5,08 ± 3,41
Zone 23	4,63 ± 3,94	92,88 ± 40,12	0,53 ± 0,28	5,01 ± 3,56
Zone 24	4,81 ± 4,04	93,13 ± 40,55	0,55 ± 0,27	4,85 ± 3,66
Zone 25	4,29 ± 3,88	94,64 ± 39,49	0,52 ± 0,28	5,39 ± 3,72

TABLE 5.2 – Résultats des 25 zones utilisés indépendamment comme conditionnement sur l'ensemble des phases respiratoires.

de 4,55% et 107,36 respectivement, suivi de la méthode de référence. L'erreur initiale du DSC de 0,14 indique un écart de positionnement important entre la phase EOE et EOI qui est aussi confirmé par une erreur COM importante de 12,06 mm. Après synthèse d'images, le conditionnement basé sur la zone 05 présente les meilleurs résultats avec un DSC atteignant 0,48 et une erreur COM réduite à 5,54 mm. De la même manière que pour le tableau précédent, la partie supérieure génère de meilleurs résultats que la partie inférieure. Pour les métriques globales, la moyenne VV est de 5,20 contre 6,84 et la moyenne MAE est de 110,24 contre 118,93. Pour les métriques locales, le DSC est de 0,44 contre 0,38 et l'erreur COM est de 6,26 mm contre 7,14 mm.

Les résultats présentés dans les tableaux précédents indiquent des résultats similaires à la méthode de référence. Nous observons de meilleurs résultats avec un conditionnement sur les zones supérieures du patient par rapport aux zones inférieures. Parmi l'ensemble des disques que nous avons définis, la zone 05 est apparue comme étant la meilleure, notamment concernant la précision de la position de la tumeur. Pour la suite des expérimentations, nous incluons le disque 05 afin de voir si la composition de plusieurs disques apporte des améliorations.

## 5.4.2 Évaluation par zones d'intérêts

Le tableau 5.4 présente les résultats pour l'ensemble des phases respiratoires avec les zones d'intérêts composées de multiples disques. Le conditionnement avec la zone E permet d'obtenir les meilleurs résultats sur les métriques globales. La zone K présente en moyenne les meilleurs résultats concernant le positionnement des tumeurs avec un score de DSC de 0,62 et une erreur COM de 3,82 mm, suivi de la zone L (DSC = 0,59 ; COM = 4,11 mm).

La figure 5.4 regroupe les boîtes à moustaches avec les métriques centrées sur la tumeur par zones d'intérêts seulement en phase EOI. Nous observons que la zone K présente de faibles améliorations en moyenne avec un score de DSC de 0,52 et une erreur COM de 5,25 mm par rapport à la zone 05 qui a un score de DSC de 0,48 et une erreur COM de 5,54 mm. L'amélioration est plus importante lorsqu'elle est comparée avec les résultats de la méthode de référence qui présente un score de DSC de 0,44 et une erreur COM de 6,14. Pour rappel, les zones K et L représentent la partie supérieure du corps et les zones M et N représentent la partie inférieure. Nous observons des résultats supérieurs avec les zones K et L par rapport aux deux autres zones.

La figure 5.5 présente le chevauchement des ITVs entre les différentes méthodes de

Modèle	VV ↓	MAE ↓	DSC ↑	COM ↓
Erreur initiale	17,10 ± 8,33	184,97 ± 42,44	0,14 ± 0,13	12,06 ± 3,72
Baseline	4,70 ± 4,43	108,15 ± 40,75	0,44 ± 0,26	6,14 ± 3,52
Zone 01	5,66 ± 4,48	113,16 ± 40,97	0,42 ± 0,22	6,64 ± 3,58
Zone 02	5,79 ± 5,43	111,54 ± 45,64	0,46 ± 0,23	6,06 ± 3,33
Zone 03	4,55 ± 4,37	107,36 ± 40,07	0,45 ± 0,24	6,30 ± 3,62
Zone 04	4,87 ± 4,60	108,94 ± 41,59	0,43 ± 0,25	6,30 ± 3,56
Zone 05	5,07 ± 4,61	108,92 ± 42,65	0,48 ± 0,24	5,54 ± 3,04
Zone 06	5,16 ± 4,29	108,70 ± 41,43	0,47 ± 0,26	6,02 ± 4,44
Zone 07	5,51 ± 4,72	110,27 ± 43,25	0,43 ± 0,24	6,66 ± 4,08
Zone 08	5,36 ± 4,53	110,76 ± 41,66	0,45 ± 0,25	6,38 ± 4,27
Zone 09	5,30 ± 4,48	111,20 ± 41,91	0,40 ± 0,25	6,88 ± 3,99
Zone 10	4,91 ± 4,63	110,83 ± 40,04	0,45 ± 0,23	6,10 ± 3,64
Zone 11	5,58 ± 4,96	111,36 ± 43,26	0,43 ± 0,24	6,51 ± 4,00
Zone 12	4,88 ± 4,67	108,82 ± 42,36	0,46 ± 0,26	5,97 ± 3,76
Zone 13	5,10 ± 4,98	111,83 ± 43,03	0,44 ± 0,25	6,28 ± 3,08
Zone 14	4,98 ± 4,70	109,43 ± 41,06	0,45 ± 0,26	6,16 ± 3,35
Zone 15	5,35 ± 4,35	110,53 ± 40,42	0,43 ± 0,26	6,15 ± 3,44
Zone 16	7,12 ± 5,54	121,77 ± 47,68	0,38 ± 0,26	7,28 ± 3,85
Zone 17	6,73 ± 4,81	119,18 ± 44,35	0,41 ± 0,25	6,88 ± 4,11
Zone 18	6,82 ± 5,37	120,33 ± 46,73	0,39 ± 0,22	6,98 ± 3,88
Zone 19	7,10 ± 5,05	118,26 ± 46,23	0,39 ± 0,24	6,72 ± 3,35
Zone 20	6,74 ± 5,11	120,74 ± 43,93	0,39 ± 0,26	7,06 ± 4,13
Zone 21	6,29 ± 5,04	116,24 ± 44,82	0,40 ± 0,24	7,12 ± 4,15
Zone 22	6,55 ± 4,63	115,79 ± 43,81	0,38 ± 0,27	7,07 ± 3,86
Zone 23	6,89 ± 4,98	117,10 ± 45,57	0,34 ± 0,26	7,43 ± 4,09
Zone 24	7,16 ± 5,04	117,96 ± 45,93	0,38 ± 0,26	6,90 ± 4,21
Zone 25	7,00 ± 4,93	121,90 ± 43,27	0,34 ± 0,22	7,98 ± 3,99

TABLE 5.3 – Résultats des 25 zones utilisés indépendamment comme conditionnement pour la phase EOI.

Modèle	VV ↓	MAE ↓	DSC ↑	COM ↓
Erreur initiale	10,62 ± 7,21	142,75 ± 48,20	0,37 ± 0,27	7,84 ± 4,27
Baseline	3,51 ± 3,43	89,82 ± 35,89	0,53 ± 0,25	4,90 ± 2,94
Zone 05	3,24 ± 3,21	87,66 ± 35,50	0,58 ± 0,23	4,33 ± 2,75
Zone A	3,16 ± 3,15	86,40 ± 35,19	0,59 ± 0,25	4,36 ± 3,17
Zone B	3,34 ± 3,30	86,13 ± 35,81	0,58 ± 0,24	4,35 ± 3,11
Zone C	3,80 ± 3,94	89,15 ± 39,34	0,55 ± 0,25	4,73 ± 3,24
Zone D	3,66 ± 3,54	89,90 ± 38,43	0,55 ± 0,25	4,87 ± 3,38
Zone E	3,02 ± 3,15	85,03 ± 35,27	0,58 ± 0,26	4,32 ± 3,03
Zone F	3,57 ± 3,37	89,16 ± 35,61	0,58 ± 0,24	4,39 ± 2,88
Zone G	3,32 ± 3,40	86,99 ± 34,89	0,59 ± 0,25	4,27 ± 3,07
Zone H	3,45 ± 3,65	87,31 ± 37,27	0,56 ± 0,26	4,70 ± 3,25
Zone I	4,42 ± 4,04	93,52 ± 41,17	0,55 ± 0,26	4,80 ± 3,28
Zone J	4,12 ± 4,06	90,62 ± 40,83	0,57 ± 0,25	4,55 ± 3,49
Zone K	3,37 ± 3,38	85,80 ± 34,55	0,62 ± 0,23	3,82 ± 2,77
Zone L	3,32 ± 3,31	85,30 ± 34,72	0,59 ± 0,24	4,11 ± 2,80
Zone M	4,08 ± 4,12	91,40 ± 41,22	0,54 ± 0,26	5,02 ± 3,45
Zone N	3,96 ± 4,16	89,82 ± 41,66	0,54 ± 0,26	4,97 ± 3,57
All	3,41 ± 3,62	87,60 ± 37,65	0,58 ± 0,25	4,45 ± 3,21

TABLE 5.4 – Résultats des zones composées sur l'ensemble des phases respiratoires.



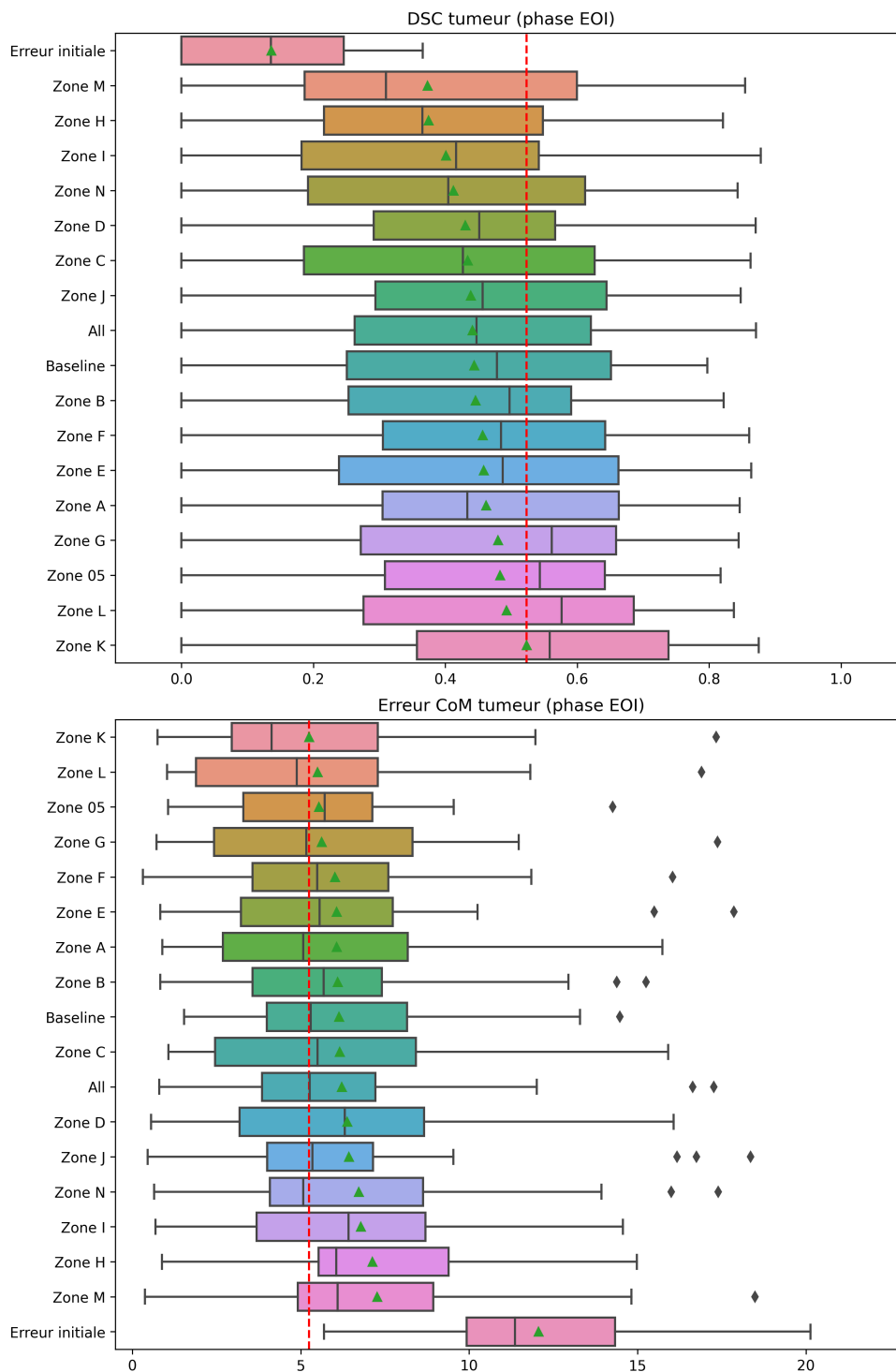


FIGURE 5.4 – Résultats de DSC et de COM pour la phase EOI par conditionnement de zones surfaciques. Les triangles verts indiquent la valeur moyenne de chaque zone. Les pointillés rouges indiquent la meilleure moyenne parmi l'ensemble des zones.

conditionnement et celle de référence. La superposition de l'ITV de référence de la méthode de référence donne un score de DSC de 0,76. Un conditionnement par la zone 05 permet d'obtenir un score similaire de 0,77 tandis qu'avec la zone K, le résultat est légèrement plus élevé avec un DSC de 0,81 en moyenne. Nous observons que les zones M et N performant moins bien, confirmant de meilleurs résultats avec la partie supérieure.

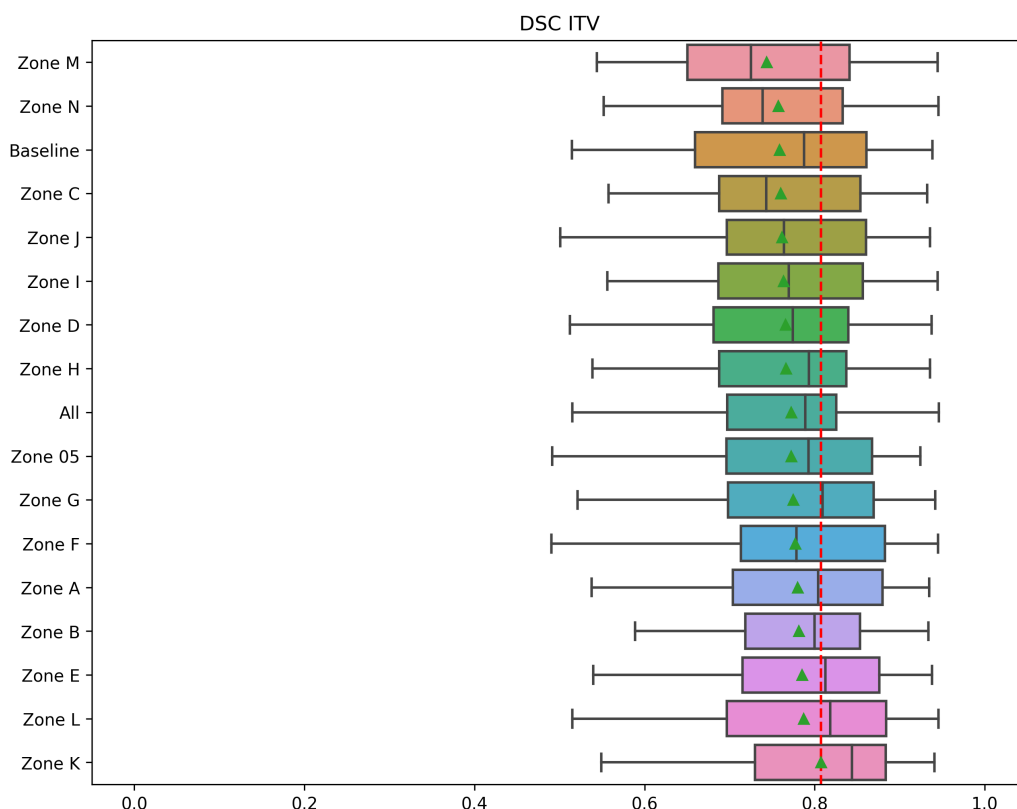


FIGURE 5.5 – Résultats de superposition des ITV avec un conditionnement par zone composée. Les triangles verts indiquent la valeur moyenne de chaque zone. Les pointillés rouges indiquent la meilleure moyenne parmi l'ensemble des zones.

Les résultats indiquent que les zones que nous avons définies permettent d'améliorer faiblement les résultats de manière globale sur l'ensemble de phases respiratoires comparés à la méthode initiale. Lorsque nous nous concentrons sur la phase EOI, les résultats sont sensiblement améliorés, indiquant une difficulté du modèle à générer un mouvement précis de grande amplitude. Le conditionnement par la zone K ou L n'apportent pas de grande amélioration par rapport à un conditionnement avec le disque 05, qui est inclus dans les deux zones. Les diverses analyses montrent des résultats supérieurs avec un conditionnement basé sur la surface supérieure comparés à un conditionnement basé sur

la surface inférieure d'un patient.

### 5.4.3 Nombre de couches cachées pour le MLP

Suite aux résultats précédents, nous considérons la zone K comme étant le meilleur conditionnement. Le tableau 5.5 montre l'effet du nombre de couches cachées sur les résultats de la méthode pour la phase EOI. Pour la VV, la zone K composée de 10 couches obtient la meilleure réduction jusqu'à 4,28%, un gain de 1% par rapport à la version à trois couches utilisées dans la section précédente. Pour la MAE, l'utilisation de neuf couches permet d'atteindre le meilleur résultat avec une MAE de 104,53. Avec les métriques locales de DSC et de COM, les neuf couches présentes à nouveau les meilleures performances avec un DSC de 0,53 et une erreur COM de 5,10 mm.

Modèle	VV ↓	MAE ↓	DSC ↑	COM ↓
Erreur initiale	17,10 ± 8,33	184,97 ± 42,44	0,14 ± 0,13	12,06 ± 3,72
Baseline	4,70 ± 4,43	108,15 ± 40,75	0,44 ± 0,26	6,14 ± 3,52
Zone K1	4,80 ± 4,29	107,06 ± 41,79	0,51 ± 0,26	5,59 ± 4,40
Zone K2	5,03 ± 4,41	108,03 ± 41,21	0,49 ± 0,25	5,57 ± 3,66
Zone K3	5,10 ± 4,63	107,42 ± 42,27	0,52 ± 0,25	5,25 ± 3,84
Zone K4	4,73 ± 4,06	105,72 ± 40,99	0,51 ± 0,27	5,35 ± 4,05
Zone K5	4,72 ± 4,78	106,28 ± 43,04	0,53 ± 0,25	5,31 ± 4,03
Zone K6	6,07 ± 4,74	114,56 ± 44,78	0,44 ± 0,25	6,48 ± 3,93
Zone K7	4,60 ± 4,37	106,67 ± 41,50	0,52 ± 0,24	5,32 ± 3,62
Zone K8	4,81 ± 4,82	107,89 ± 43,51	0,51 ± 0,25	5,59 ± 4,09
Zone K9	4,40 ± 4,15	104,53 ± 40,14	0,53 ± 0,26	5,10 ± 4,07
Zone K10	4,28 ± 5,07	106,86 ± 44,08	0,49 ± 0,24	5,77 ± 3,90

TABLE 5.5 – Résultats de l'influence du nombre de couches cachées dans le MLP pour le conditionnement.

La figure 5.6 montre les scores de DSC entre les ITVs en fonction du nombre de couches cachées du MLP. La méthode initiale présente le score le plus bas avec un DSC de 0,76 et la zone K composée de 3 couches donne le meilleur score de 0,81. Les résultats dans l'ensemble des versions proposées sont similaires et n'indiquent pas d'amélioration ou de dégradation en fonction du nombre de couches utilisées.

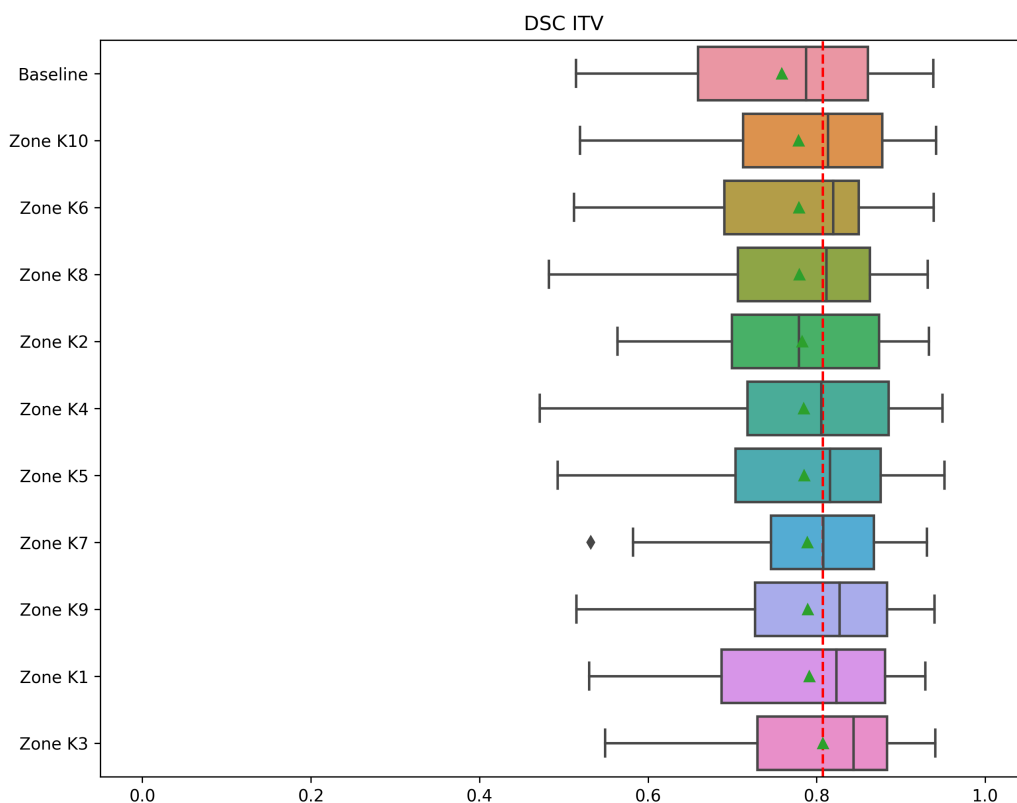


FIGURE 5.6 – Résultats de superposition des ITVs en fonction du nombre de couches cachées du MLP. Les triangles verts indiquent la valeur moyenne de chaque zone. Les pointillés rouges indiquent la meilleure moyenne parmi l’ensemble des zones.

## 5.5 Discussion

Les résultats présentés se basent sur des simulations de mouvements respiratoires superficiels, données qui peuvent être obtenues à l’aide de caméra TOF lors d’acquisition TDM 4D. Ces mouvements simulés ont été utilisés comme un conditionnement plus complexe qu’une valeur scalaire seule. Afin d’évaluer au mieux l’impact de ce conditionnement, nous avons établi un protocole de vérification expérimentale en utilisant différentes zones d’intérêt allant de simples zones uniques à des zones composées. Les résultats montrent que l’utilisation d’un mouvement superficiel ne permet qu’une légère amélioration de la précision du déplacement des tumeurs par rapport à la valeur scalaire utilisée seule.

Ces résultats, peu significatifs, peuvent s’expliquer par différentes limites de notre étude. Tout d’abord, il ne s’agit pas de vraies données superficielles provenant d’une caméra externe, mais d’une simulation faite avec les données disponibles. Le nuage de points obtenu de la phase EOE par notre méthode de conversion ne permet pas de garantir la

bonne correspondance entre chaque point avec les points d'une autre phase respiratoire. De plus, la méthode de recalage retenue ne garantit pas que le déplacement du point obtenu soit exact. Dans le cas d'une utilisation d'une caméra, les points sont les mêmes d'une phase à l'autre, ce qui simplifie le problème et garantit la qualité de l'information extraite. Il serait donc intéressant, à l'avenir, d'approfondir cette approche avec des données appropriées pour confirmer les résultats. Néanmoins, l'impact quasi nul de cette mesure sur la reproduction de la respiration souligne l'importance d'utiliser des méthodes robustes.

Grâce notre évaluation multi-régions, nous avons pu observer que le conditionnement de la partie supérieure du thorax permet d'obtenir de meilleurs résultats que sa partie inférieure ou le haut de l'abdomen. Nous avons aussi évalué l'impact du nombre de couches du MLP et nous avons conclu que le nombre, et donc l'extraction de non linéarité complexes, n'influence que très légèrement les performances du modèle. Quant à l'évaluation sur l'ITV, cela a permis de démontrer que le modèle parvient à obtenir des volumes d'intérêts satisfaisant sur l'ensemble des phases respiratoires malgré des superpositions plus faibles lorsque nous évaluons seulement la phase EOE.

## 5.6 Conclusion

Ce chapitre présente une méthode permettant l'intégration de données respiratoires surfaciques plus complexes qu'une valeur scalaire d'amplitude respiratoire vue dans les chapitres précédents. Ces mesures surfaciques peuvent être par exemple obtenues en pratique clinique à partir de caméra TOF. Ce type d'équipement permet de capturer l'ensemble du mouvement en surface du patient contrairement à une ceinture respiratoire où le mouvement est résumé à une donnée unidimensionnelle. Dans ce travail préliminaire, un MLP apprend à encoder des nuages de points mesurant des mouvements surfaciques simulés pour les injecter dans l'espace latent du générateur, une nouvelle fois en utilisant un mécanisme AdaIN vu aux chapitres précédents. Les premières analyses présentées montrent qu'il est possible d'adapter notre mécanisme d'injection afin d'intégrer des données plus complexes en tant que conditionnement dans le but d'améliorer la précision des mouvements générés. Les résultats expérimentaux obtenus ne permettent pas néanmoins de déterminer un avantage clair après utilisation de cette information plus riche, et une validation expérimentale plus complète avec des données surfaciques réelles est nécessaire.

Cette partie conclut nos contributions liées à l'imagerie du cancer du poumon. Dans le

chapitre suivant, nous verrons comment ces idées développées dans ce contexte applicatif peuvent vraisemblablement être adaptées à un tout autre problème, celui de l'estimation du remplissage vésical pour l'amélioration de la prise en charge des patientes dans le cas du cancer du col de l'utérus.



# APPLICATION EXPLORATOIRE POUR UN AUTRE OBJECTIF : LA VESSIE

## Résumé

Dans ce chapitre, nous envisageons l'application de notre méthode de synthèse dynamique dans un contexte très différent : le remplissage vésical dans la planification du traitement du cancer du col de l'utérus. Le remplissage vésical est un facteur très important de déformation des organes traités et des organes à risque. Il existe ainsi des disparités importantes entre l'anatomie du scanner de planification dosimétrique et celle du jour du traitement. Dans ce contexte, notre objectif est de permettre de produire des images TDMS dynamiques, où la composante supplémentaire serait le niveau de remplissage vésical. Une telle approche de synthèse permettrait d'éviter la nécessité de recourir à des banques d'images de planification à différents niveaux de remplissage vésical, particulièrement contraignantes pour les patientes et nécessairement incorrectes. Après avoir contextualisé le cas spécifique du cancer de l'utérus, nous proposons des ajustements de notre méthode de synthèse respiratoire pour ce contexte particulier. Nous évaluons la performance de la méthode en utilisant un ensemble de données cliniques réduit. Pour combler le manque de données, nous proposons une approche d'augmentation des données basée sur une méthode d'interpolation. Nous concluons sur les résultats de cette application et nous discutons des limitations et des perspectives de la méthode.

## 6.1 Introduction

Le cancer de l'utérus est l'une des formes les plus courantes de cancer touchant les femmes à travers le monde. Il se développe généralement dans les tissus de l'utérus, principalement dans la partie inférieure de l'organe appelée le col de l'utérus. Plusieurs techniques d'imagerie sont utilisées pour examiner les tissus et les structures de l'utérus



afin de détecter toute anomalie liée au cancer [127], [128]. Par exemple, l'IRM peut fournir des informations précieuses sur la taille, la localisation et les caractéristiques des tumeurs. La TDM peut être utilisé pour évaluer la propagation du cancer du col de l'utérus dans les ganglions lymphatiques et d'autres tissus environnants, il est particulièrement utile pour déterminer le stade de la maladie [129].

Les options de traitement pour le cancer de l'utérus dépendent du stade de la maladie et de la propagation des cellules cancéreuses. Les traitements courants comprennent la chirurgie pour enlever les tissus affectés, la radiothérapie pour cibler et détruire les cellules cancéreuses, ainsi que la chimiothérapie pour traiter les cellules cancéreuses à travers le corps. Dans certains cas, une combinaison de ces traitements peut être recommandée pour assurer une approche complète [130].

Lors de la planification du traitement, il est important de prendre en considération l'état de la vessie. En effet, en fonction du niveau de remplissage de cette dernière, le positionnement de l'utérus ainsi que des organes autour peut être différent et influence donc fortement le calcul de dose et le traitement [131]. Afin de prévoir ces changements de position, il est possible d'effectuer des acquisitions d'imagerie médicale en deux temps : une première fois lorsque la vessie est vide et une autre lorsqu'elle est pleine. Cela permet ainsi d'évaluer l'amplitude du déplacement des organes d'intérêts (figure 6.1).

Il est ainsi relativement courant de devoir imposer aux patientes des acquisitions TDMs supplémentaires pour différents niveaux de remplissage de la vessie, pour choisir le plan correspondant au niveau de remplissage vésical le plus proche le jour du traitement. Il est aisé de comprendre que cette situation n'est pas optimale en plus de susciter de l'inconfort chez les patientes. Une méthode permettant de déformer la vessie de manière artificielle semble permettre d'envisager une planification moins lourde. Nous présentons ainsi dans la suite de ce chapitre une preuve de concept relatif à ces enjeux.

Dans ce chapitre, nous souhaitons donc évaluer l'intérêt d'une méthode de synthèse dynamique dans ce contexte de remplissage vésical en se basant essentiellement sur les mêmes fondements méthodologiques que présentés dans les chapitres précédents dans le contexte du cancer du poumon.

## 6.2 Implémentation

Dans un contexte clinique, l'acquisition d'images de vessie vide est plus aisée que celle d'une vessie pleine (confort et sensation de la patiente). Nous décidons de générer les

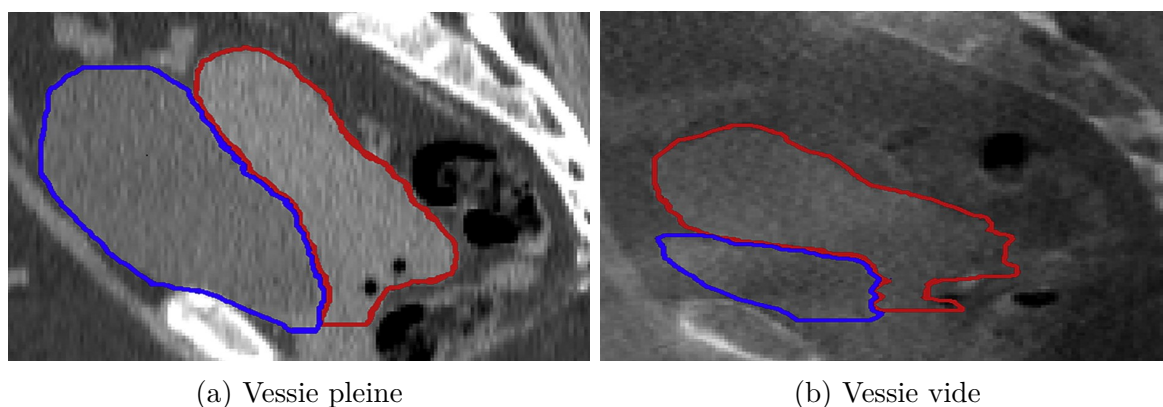


FIGURE 6.1 – Illustration du niveau de remplissage de la vessie. Les contours rouge et bleu montrent le positionnement de l’utérus et de la vessie respectivement.

images de vessies pleines à partir d’image de vessie vide. Nous appliquons la méthode décrite dans le chapitre 4 et nous utilisons en conditionnement la variation de volume de la vessie comme variable scalaire.

### 6.2.1 Jeu de données

Le département de radiothérapie du CHRU de Brest a fourni des acquisitions du pelvis de 34 patientes. Pour chaque patiente, au moins deux acquisitions TDM 3D du bassin sont disponibles avec des niveaux de remplissage différents. Certaines patientes ont des acquisitions supplémentaires où il est possible d’observer la vessie à plusieurs niveaux de remplissage. Les images étant acquises à différents moments de la prise en charge des patients, le positionnement entre les acquisitions d’une même patiente peut varier. Par conséquent, nous avons effectué un recalage rigide entre ces acquisitions afin de superposer le bassin.

Pour réduire la taille des images, nous nous plaçons entre la vessie et l’utérus : nous calculons le centre de masse de la segmentation des deux organes. Puis, nous englobons tous les autres organes qui se trouvent dans un intervalle de 64 pixels dans toutes les directions. Les images sont rognées pour conserver principalement l’utérus et la vessie, donnant des images de tailles  $128 \times 128 \times 128$ . Les segmentations des organes sont fournies par le CHRU de Brest. Pour l’utérus, le CTV est fournie pour tous les cas. Pour la vessie, dans certains cas, seule la segmentation d’un niveau de remplissage est disponible pour la vessie. Afin de combler les cas absents, nous utilisons le modèle TotalSegmentator [38] qui permet de segmenter la vessie. Il est toutefois nécessaire de corriger les résultats obtenus

par la méthode qui n'est pas parfaite.

La vessie n'étant pas très visible ou présentant des valeurs HU très différentes, notamment dû à la présence d'un produit de contraste, un recalage en l'état résulterait à une transformation de mauvaise qualité. Pour aider l'étape de recalage, nous définissons un intervalle entre -150 HU et 150 HU où les organes de la zone sont plus visibles. De plus, pour homogénéiser les vessies entre les acquisitions, nous mettons les vessies en hypersignal (figure 6.2).

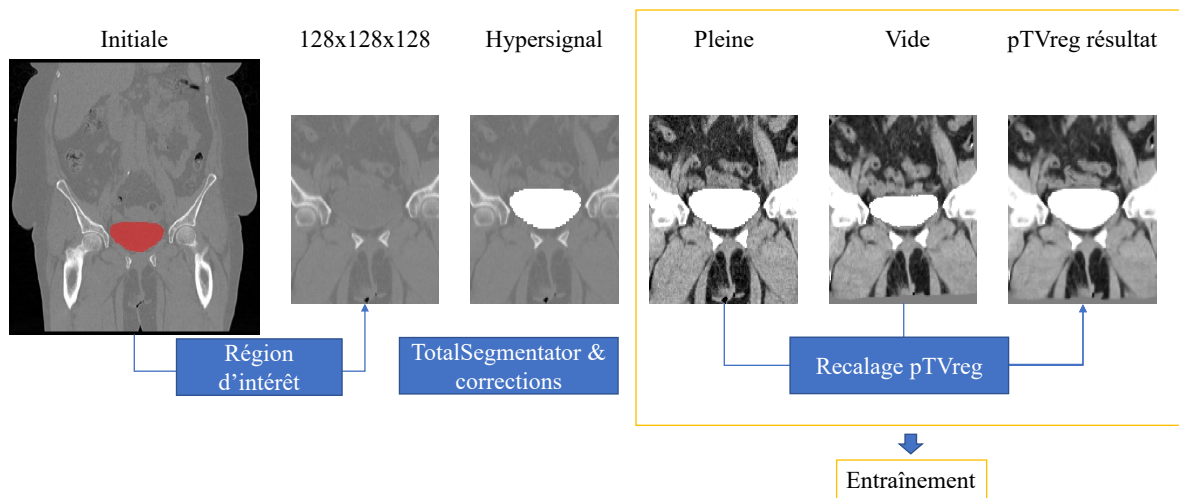


FIGURE 6.2 – Illustration des étapes de prétraitement des données cliniques sur la vessie.

### 6.2.2 Interpolation

Notre objectif est de pouvoir générer différents niveaux de remplissage de la vessie. Toutefois, du fait du faible nombre d'images et surtout l'absence d'images intermédiaires, il est probable que lors de l'entraînement, le réseau n'apprend qu'à générer des images avec une vessie pleine. Pour augmenter le nombre de données d'entraînement, nous proposons une méthode d'interpolation entre les acquisitions afin de simuler des niveaux intermédiaires. À partir des DVFs obtenus entre le recalage des acquisitions, nous pouvons faire varier l'amplitude du déplacement des voxels par un facteur  $f$  et ainsi générer des données supplémentaires. Il est aussi possible d'exagérer la déformation avec un facteur  $f$  supérieur à 1, permettant de simuler des images avec un remplissage encore plus important (figure 6.3). Pour éviter une extrapolation qui entraînerait des déformations non réalistes et donc des organes partiellement ou totalement irréalistes, nous limitons la valeur du facteur  $f$  à 1,5.

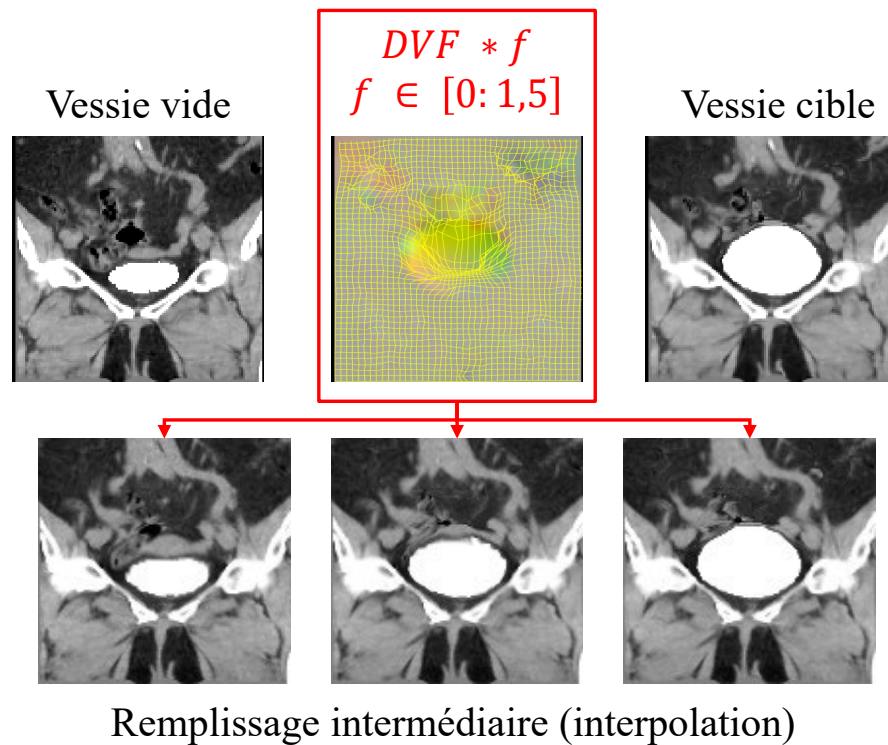


FIGURE 6.3 – Augmentation de données à l’aide d’interpolation.

### 6.2.3 Autres détails d’implémentations

Il s’agit d’un modèle 3D entraîné sur 400 époques en utilisant l’optimisation ADAM [109] ainsi qu’avec une taille de batch de 1. Le jeu de données est séparé en 21 patientes pour l’entraînement et 11 pour la validation. Les DVFs nécessaires à la supervision de l’entraînement du modèle sont obtenus avec la méthode pTVreg [65] qui propose des résultats de bonnes qualités malgré une application différente de celle d’origine (figure 6.2). La valeur  $\alpha$  est calculée à partir de la variation de volume entre les vessies en pourcentage. Nous entraînons notre réseau avec une carte graphique NVIDIA GeForce GTX 1080Ti avec 11 Go de mémoire.

Pour évaluer les performances de notre méthode pour cette application, nous montrons des exemples visuels d’images générées. Nous utilisons des métriques classiques tels que la MAE entre les images générées et les images de références ainsi que le DSC pour mesurer la similarité entre les vessies et les utérus. Nous calculons aussi la VV entre les vessies afin de vérifier si la méthode est capable de générer un remplissage réaliste de la vessie. De plus, nous montrons l’impact de l’ajout de l’augmentation de données basée sur l’interpolation.

### 6.3 Résultats

La figure 6.4 montre des exemples visuels d’images générées par notre méthode à partir d’une vessie vide. Nous pouvons observer une différence initiale importante de la taille de la vessie entre les deux acquisitions de la patiente avec un mouvement arrière de l’utérus dû au remplissage de la vessie (figure 6.4a-6.4b). Notre méthode sans augmentation parvient à générer un remplissage de vessie correct avec aussi un mouvement faible de l’utérus visible sur la coupe sagittale (figure 6.4c). Lorsque nous utilisons l’augmentation de données basée sur l’interpolation, le modèle génère une vessie pleine plus cohérente par rapport à la version sans augmentation (figure 6.4d). Dans les deux cas, il existe bien un mouvement de l’utérus, mais ce dernier est légèrement déformé et mal positionné par rapport à l’image de référence.

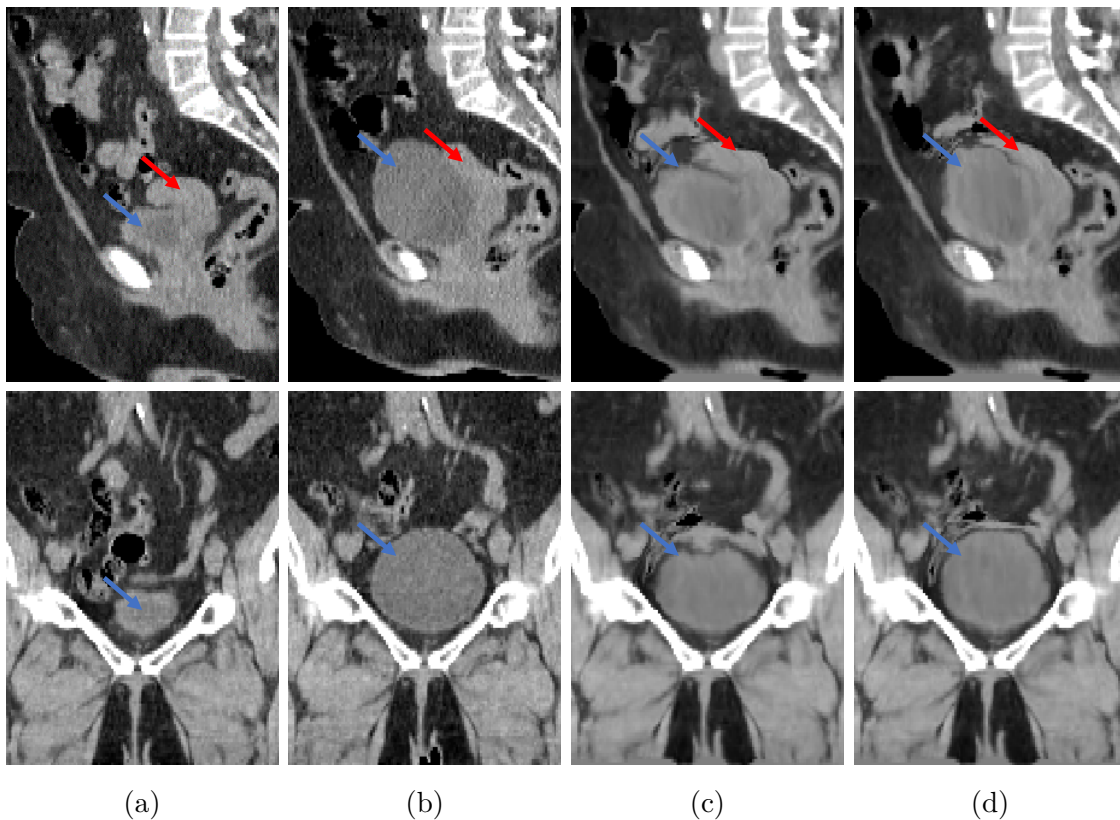


FIGURE 6.4 – Exemple de génération de vessies pleines à partir de vessie vide. Les flèches bleues indiquent la vessie et les flèches rouges indiquent l’utérus. (a) Vessie vide; (b) Vessie pleine; (c) Image générée sans augmentation; (d) Image générée avec augmentation.

Le tableau 6.1 donne les résultats pour les deux modèles utilisés, avec ou sans augmentation de données basée sur l’interpolation. Nous pouvons observer qu’initialement, il

existe une différence importante entre les volumes d'une vessie vide et pleine (en moyenne : 199,78%). Après synthèse, la VV est considérablement réduite à 21,07% pour le modèle sans augmentation par interpolation et jusqu'à 14,45% avec augmentation. Pour la MAE, une réduction est aussi observable avec une meilleure similarité avec la méthode avec augmentation, l'erreur initiale passe de 75,14 à 43,43 après synthèse. La moyenne initiale du DSC est de 0,55. Cela indique une faible similarité entre les phases en termes de forme et de localisation de la vessie. La moyenne augmente à 0,76 sans augmentation et à 0,81 avec augmentation. La synthèse a conduit à une amélioration notable de la similarité et l'augmentation semble avoir contribué à une similarité encore meilleure. Le DSC moyen initial était de 0,79 indiquant un bon chevauchement initial entre les contours de l'utérus. Après la synthèse, avec ou sans augmentation, le DSC moyen est resté relativement stable. Cela suggère que la synthèse n'ait pas eu d'effet significatif sur le chevauchement des contours de l'utérus.

Les résultats du tableau montrent que la méthode de synthèse proposée a réussi à réduire considérablement la variabilité des volumes de vessie, à améliorer la précision de localisation et à augmenter la similarité entre les niveaux de remplissage, en particulier lorsque l'augmentation est utilisée. Cela suggère que la méthode est efficace pour générer un remplissage de vessie cohérent et précis à partir de données de vessies vides. Pour l'utérus, le chevauchement initial est déjà correct, et la synthèse n'entraîne pas de changements significatifs à cet égard.

		Moyenne
VV	Initial	199,78 ± 180,31
	Après synthèse sans augmentation	21,07 ± 20,59
	Après synthèse avec augmentation	14,45 ± 13,08
MAE	Initial	75,14 ± 8,27
	Après synthèse sans augmentation	57,43 ± 10,40
	Après synthèse avec augmentation	43,43 ± 6,03
DSC Vessie	Initial	0,55 ± 0,18
	Après synthèse sans augmentation	0,76 ± 0,15
	Après synthèse avec augmentation	0,81 ± 0,09
DSC Utérus	Initial	0,79 ± 0,15
	Après synthèse sans augmentation	0,78 ± 0,07
	Après synthèse avec augmentation	0,80 ± 0,10

TABLE 6.1 – Résultats quantitatifs sur les données de vessies.

## 6.4 Discussion

Dans ce chapitre, nous avons exploré la possibilité d’appliquer notre méthode, initialement conçue pour générer des mouvements respiratoires chez les patients, à un contexte différent, celui du remplissage de la vessie. L’approche se sert des DVFs afin de déformer les organes autour pour positionner les organes en fonction du volume de la vessie. Les résultats préliminaires ont permis de montrer la possibilité de générer différents niveaux de remplissage.

Nos premiers résultats préliminaires suggèrent que nous pouvons générer différents niveaux de remplissage à partir d’une acquisition TDM d’une vessie vide. La vessie synthétique parvient à atteindre un niveau de remplissage similaire à celui d’une acquisition cible. Cependant, les déformations principalement concentrées sur la vessie induisent une déformation irréaliste de l’utérus et montrent un positionnement insatisfaisant. Cela pourrait s’expliquer en partie par un manque de données qui ne permet pas de représenter l’ensemble des types d’utérus et des différents niveaux de remplissage de la vessie. Cela semble suggéré par le fait que l’utilisation de l’augmentation de données basée sur l’interpolation permet d’améliorer sensiblement les résultats. Pour continuer à améliorer les résultats, il serait également intéressant d’explorer l’utilisation d’un deuxième réseau réduit aux déplacements de l’utérus, ce qui permettrait d’obtenir des déformations et un positionnement plus réaliste.

## 6.5 Conclusion

Nous avons étendu notre approche de synthèse indirecte pour générer différents niveaux de remplissage vésical dans le contexte du cancer de l’utérus dans le cadre d’une preuve de concept. Afin d’aider le modèle à générer des niveaux intermédiaires de remplissage de la vessie, une augmentation de données basée sur de l’interpolation est proposée. La méthode génère des images avec différents niveaux de remplissage vésical à partir d’un niveau défini par une image TDM 3D et contrôlé par un conditionnement basé sur le volume de la vessie.

Nous avons ainsi montré que notre méthode de synthèse peut être étendue à d’autres cas pratiques en radiothérapie, moyennant des adaptations spécifiques à chaque application. Dans cet exemple, il a été nécessaire de remplacer la valeur scalaire basée sur la surface du patient par une mesure de variation de volume entre les vessies. Dans des

travaux futurs envisagés, nous pensons qu'il serait également pertinent d'utiliser une méthodologie de synthèse similaire pour prédire l'évolution temporelle d'une tumeur (croissance ou régression), en utilisant un conditionnement basé sur une mesure temporelle, par exemple.





# CONCLUSION GÉNÉRALE ET PERSPECTIVES

---

Par sa capacité à mesurer l'amplitude de la dynamique respiratoire, l'imagerie TDM 4D joue un rôle essentiel dans la planification du traitement du cancer du poumon par radiothérapie. Elle permet de visualiser le mouvement interne, offrant une meilleure compréhension de la mobilité et de la forme des tumeurs et des organes au cours de la respiration. Cependant, son utilisation expose le patient à un rayonnement considérable, potentiellement préjudiciable. De plus, la fiabilité des marges TDM 4D dépend de la fidélité du mouvement respiratoire du patient entre le scanner de planification et le jour du traitement. Plusieurs études (dont la nôtre) ont montré que ces mesures de mouvement internes, notamment au niveau de la tumeur et des OARs, sont faiblement reproductibles entre deux acquisitions TDM 4D. Ces incertitudes nous ont ainsi conduits à envisager des pistes alternatives pour prendre en compte ces marges, par des méthodes de synthèse d'image artificielle pour réduire l'utilisation d'acquisitions TDM 4D réelles.

Dans une première exploration du domaine, il nous est apparu pertinent de tirer profit de méthodologies GANs éprouvées, largement majoritaires en synthèse médicale, pour établir une preuve de concept de la capacité de ce type de modèles à générer des dynamiques respiratoires. C'est ce que nous avons obtenu par une modification d'une architecture I2I directe de type pix2pix, où une image de phase respiratoire est directement produite par le modèle. Elle nous a permis d'identifier un mécanisme d'injection d'amplitude respiratoire scalaire par normalisation d'instance (AdaIN) dans le goulot d'étranglement du réseau, mécanisme préalablement utilisé dans la littérature en transfert de style. L'utilisation de ce mécanisme permet de contrôler la génération de la phase respiratoire cible avec un conditionnement secondaire basé sur une information respiratoire du patient. Ce type d'information peut être obtenu en routine clinique avec, par exemple, une ceinture respiratoire.

Néanmoins, cette première approche directe s'est révélée produire des images de qualité insuffisante. Cela nous a amenés en deuxième intention à déformer la phase réelle initiale vers la phase générée par des méthodes de recalage conventionnelles (la phase synthétique

---

servant alors d’oracle), dans le but de préserver des attributs importants de l’image de phase initiale et de paramétrer explicitement les déformations induites. Des premiers résultats encourageants ont été obtenus, bien que soulignant des problèmes fondamentaux liés à ce type d’approches de synthèse directe d’image par GAN.

Nous avons été ainsi amenés à envisager une approche indirecte basée sur la synthèse de DVF (et non d’image) pour répondre à ces limitations. Le cas de l’apprentissage de DVF est généralement abordé du point de vue du recalage non supervisé, où les réseaux prennent en entrée des paires d’images de phase pour apprendre une transformation de manière implicite en minimisant une fonction de coût composée d’un terme de similarité et d’un terme de régularité de champ (par exemple  $\ell_2$ ). Toutefois, malgré des progrès certains ces dernières années dans le domaine de recalage pulmonaire, les méthodes basées DL n’ont pas encore atteint des niveaux de performance garantissant leur efficacité, avec notamment une tendance nette à engendrer des champs trop lisses du fait de cette contrainte de régularité explicite. Nous avons ainsi choisi d’entraîner de manière supervisée notre méthode en tirant profit de DVF vérité terrain issu de méthodes de recalage de l’état de l’art, en l’occurrence une approche conventionnelle très performante non basée sur des techniques d’apprentissage [65]. Cette stratégie permet de découpler le problème du recalage de celui de la synthèse. En outre, compte tenu des incertitudes sur ces champs, une contribution importante de notre travail a consisté à proposer conjointement un apprentissage adversaire à la fois sur l’image déformée (de manière analogue aux méthodes de recalage non supervisées) mais aussi sur les cartes d’amplitude des champs DVFs. Nous montrons que c’est cette contribution qui permet d’atteindre des hauts niveaux de réalisme, autorisant des irrégularités plus représentatives du mécanisme respiratoire interne, qui ne pourraient être pris en compte par une contrainte de régularité explicite.

Un autre point clé de nos travaux concerne les critères de validation. Les méthodes de l’état de l’art utilisent préférentiellement des métriques globales sur les images générées pour évaluer leurs performances. Cependant, dans le cadre d’une utilisation clinique, l’évaluation de la qualité visuelle globale de l’image n’est pas suffisante, et nous avons montré qu’il est impératif d’analyser le mouvement à une échelle plus locale afin de considérer les variations de positions des tumeurs et des OARs. Nos analyses mettent en avant ce point critique en évaluant différentes structures anatomiques. D’autre part, mentionné précédemment, nous comparons également les images générées par notre approche à des répétitions d’acquisitions TDM 4D réelles et montrons que les images générées présentent des variabilités de l’ordre de celles mesurées entre acquisitions réelles.

---

Ces nouveaux critères de validation proposés ont pour but de mieux tenir en compte les objectifs à atteindre, afin de faciliter la compréhension des possibilités et limitations de ces approches vers l’acceptabilité clinique. Dans le futur, des expérimentations supplémentaires comme la comparaison de la carte de dose obtenue à partir d’acquisitions réelles TDM 4D et d’images générées seront également nécessaires pour confirmer l’intérêt de la méthode dans un contexte de planification dosimétrique en radiothérapie.

À la suite de ce travail de validation, nous avons également questionné notre méthode de conditionnement. En dépit de résultats satisfaisants, nous associons de manière bijective la respiration et une seule valeur scalaire, ce qui est une simplification qui nous semble excessive et sous-optimale. Pour mieux capturer l’ensemble des informations respiratoires de surface, de nouveaux outils se sont développés avec des appareils externes fournissant une information plus riches tels que des caméras TOF. Ces caméras ont la particularité de produire des nuages de points de la surface et de donner des informations sur le mouvement en surface grâce à des mesures de profondeur. Compte tenu de l’émergence de ces outils en routine clinique, nous avons exploré l’utilisation de conditionnements surfaciques plus complexes sur la base de nuage de points dérivé de l’image TDM 4D, dans le but d’étudier la capacité de notre méthode à intégrer ce type de données. Toutefois, ces explorations préliminaires montrent que l’utilisation de conditionnement plus complexe n’améliore pas significativement les résultats par rapport à un conditionnement scalaire. Cependant, des améliorations modestes sont observables localement, notamment pour les zones surfaciques situées dans la partie supérieure du corps. Une prochaine étape pour valider ces résultats de manière plus déterminante serait d’utiliser des données réelles de mouvement surfacique plutôt que des signaux dérivés de l’image.

Notre approche de synthèse par coût conjointement adversaire et supervisé peut vraisemblablement s’étendre à d’autres applications médicales, comme nous l’avons montré dans le chapitre 6. Nous avons adapté notre modèle pour simuler différents états de remplissage vésical dans un contexte de cancer du col de l’utérus, le positionnement des organes périphériques étant grandement influencé par ce paramètre. Bien que prometteuse, la méthodologie nécessite encore des développements pour permettre d’accroître le réalisme des images synthétisées, ainsi qu’une validation dosimétrique. Des modifications simples du conditionnement et de l’architecture, ainsi que l’apport d’une stratégie d’augmentation originale, ont néanmoins permis de montrer des pistes d’amélioration vraisemblables. D’autres aspects sont à envisager, comme l’utilisation d’un réseau spécialisé pour traiter le déplacement de l’utérus afin d’améliorer la prise en compte des mouvements in-

---

ternes. Quoi qu'il en soit, cette application distante du cadre initialement proposé permet de laisser envisager de nombreux autres cas applicatifs de ce type de modèles génératifs dynamiques. Par exemple, une piste non explorée serait la visualisation temps réel du traitement en couplant ces modèles à de l'imagerie cone beam.

Dans l'ensemble, ces expérimentations montrent la puissance et la flexibilité des développements en synthèse proposés. D'autres pistes d'amélioration et d'extension du modèle pourraient consister à fournir davantage d'informations, afin de les rendre, par exemple, patient spécifique en ajoutant des données cliniques (âge, tabagisme, etc.). De plus, le découplage proposé entre le choix de la méthode de recalage et de synthèse permet de substituer les champs DVF de référence en suivant l'évolution des performances des méthodes de recalage. Il est aussi possible d'envisager d'autres pistes d'améliorations au niveau architectural du modèle, en faisant le choix d'autres architectures génératives telles que les transformers ou les modèles de diffusion en plein essor à l'heure de la finalisation de ce manuscrit [132]-[134]. Des architectures hybrides qui combinent ces approches pour proposer des méthodes plus performantes en tirant parti des avantages de chacune semblent aussi opportunes [135]-[137]. Il serait ainsi intéressant d'étudier l'impact d'architectures alternatives tout en conservant l'aspect du conditionnement par injection pour la génération d'images médicales dynamiques.

Les développements proposés dans cette thèse semblent prometteurs pour la génération d'images dynamiques dans différents contextes. Nous sommes ainsi convaincus que les travaux menés, par leur caractère générique, ont une portée qui dépasse le cadre applicatif de la thèse.

Pour résumer, ces travaux s'inscrivent dans le contexte de la médecine personnalisée et des jumeaux numériques, où l'individu est au centre du modèle, et s'attachent à apporter une composante dynamique à cet effort, souvent négligée. La nécessité d'incorporer des informations précises et personnelles du patient dans les modèles, en complément de la base d'entraînement, est essentielle pour adapter la prise en charge du patient et améliorer les traitements personnalisés. Nous espérons que ces travaux contribuent ainsi à l'évolution des liens entre numérique et médecine vers de nouveaux protocoles thérapeutiques, et par conséquent des soins de meilleure qualité.

## Code et exemples visuels

**Partage de code :** Pour faciliter la diffusion de la méthode proposée, nous avons mis à disposition le code utilisé dans le chapitre 4 sur un dépôt GitHub : <https://github.com/cyiheng/Dynagan>

Ce dépôt comprend des instructions pour utiliser les poids du modèle entraîné dans le chapitre 4. Il fournit également aux utilisateurs les clés nécessaires pour qu'ils puissent entraîner leur propre modèle avec leurs propres données. À noter que les données que nous avons utilisées ne seront pas disponibles, conformément aux réglementations sur la protection des données personnelles. Des informations concernant l'implémentation graphique présentée dans la section suivante y sont aussi incluses.

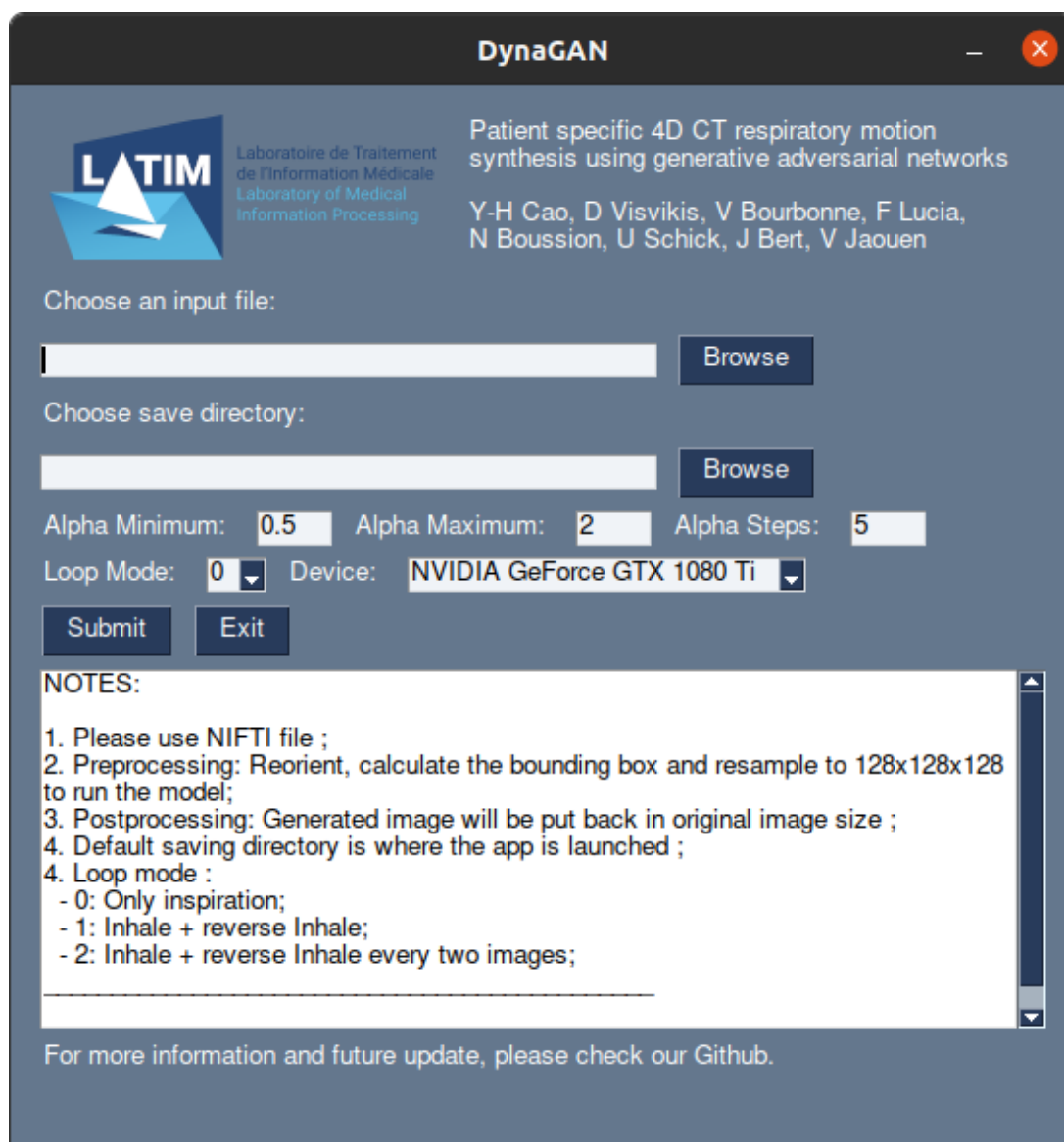
**Exemples visuels :** Plusieurs exemples visuels sont disponibles sous forme d'un jeu. Il s'agit d'un questionnaire dans lequel le but est de deviner si une animation est extraite d'une acquisition TDM 4D réelle ou d'une acquisition générée par notre modèle. Le questionnaire est disponible via le QR code ci-dessous et également via le lien suivant : <https://forms.gle/9LPMTNKak4gRnohm9>



QR code vers le questionnaire

## Implémentation graphique

Nous avons développé une interface graphique pour faciliter l'utilisation de la génération d'images à un public plus élargi sans connaissance de programmation. Cette interface utilise les poids du modèle entraîné dans le chapitre 4. À partir d'une image 3D de poumon de format Nifti, l'application peut générer un mouvement respiratoire de l'image choisie. L'image utilisée sera automatiquement pré-traitée par l'application pour correspondre au format requis pour le bon fonctionnement du modèle.



Interface graphique de l'application proposée

---

L'amplitude du mouvement respiratoire est contrôlée de manière linéaire à l'aide des paramètres *Alpha minimum*, *Alpha maximum* et *Alpha steps*. *Alpha minimum* définit la première phase initialisée par l'application, tandis que *alpha maximum* correspond à la phase maximale du mouvement respiratoire. *Alpha steps* permet de définir le nombre de phases intermédiaires entre ces deux paramètres.

Le paramètre *Loop* détermine comment l'image TDM 4D sera construite :

- 0 : De la phase source jusqu'à la phase *alpha maximum* - [0,1,...,*alpha maximum*]
- 1 : De la phase source à la phase *alpha maximum*, puis réutilisation des phases dans l'ordre inversé - [0,1,...,*alpha maximum*,...,2,1]
- 2 : Identique à 1, mais avec un pas de 2, évitant de réutiliser une même image deux fois - [0,2,4,...,*alpha maximum*,...,5,3,1]

Il est possible d'utiliser l'application sans GPU mais le temps de traitement total sera plus long. Les temps indiqués dans le tableau sont obtenus à partir des configurations suivantes pour la génération de 5 phases respiratoires :

- un GPU NVIDIA GeForce GTX 1080Ti GPU avec 11Go de mémoire ( $\approx$  5Go sont utilisés pour générer les images)
- un CPU Intel I7-3930M @ 3.2GHz

Appareil	Taille d'image	Prétraitement	Génération	Sauvegarde	Total
GPU	256 x 256 x 96	$\approx$ 10sec	$\approx$ 5sec	$\approx$ 10sec	$\approx$ 25sec
	512 x 512 x 117	$\approx$ 20sec	$\approx$ 10sec	$\approx$ 30sec	$\approx$ 70sec
CPU	256 x 256 x 96	$\approx$ 120sec	$\approx$ 20sec	$\approx$ 10sec	$\approx$ 160sec
	512 x 512 x 117	$\approx$ 120sec	$\approx$ 40sec	$\approx$ 30sec	$\approx$ 190sec

Détails du temps de génération d'une image dynamique

L'application est disponible sous :

- Windows : <https://ubocloud.univ-brest.fr/s/tqjfEe39Q3J8qyD>
- Linux : <https://ubocloud.univ-brest.fr/s/aL8yffHNjWYSNtm>



---

## Communications

### Revue internationale

- **Y. Cao**, V. Bourbonne, F. Lucia, U. Schick, J. Bert, V. Jaouen, D. Visvikis. *CT respiratory motion synthesis using joint supervised and adversarial learning. Physics in Medicine & Biology*, 2023 (en cours de révision)

### Conférences internationales

- **Y. Cao**, V. Jaouen, V. Bourbonne, N. Boussion, U. Schick, J. Bert, D. Visvikis. *Image and volume conditioning for respiratory motion synthesis using GANs. IEEE nuclear science symposium and medical imaging conference*, 2021, au Japon (virtuel) – Oral – Sélectionné pour le *best student paper award* (Top 8)
- **Y. Cao**, V. Jaouen, V. Bourbonne, F. Lucia, N. Boussion, U. Schick, J. Bert, D. Visvikis. *Patient-specific 4DCT respiratory motion synthesis using tumor-aware GANs. IEEE nuclear science symposium and medical imaging conference*, 2022, à Milan – Poster

### Conférences nationales

- **Y. Cao**, V. Jaouen, V. Bourbonne, N. Boussion, U. Schick, J. Bert, D. Visvikis. Génération d'image dynamique par un réseau antagoniste génératif doublement conditionnel. Colloque ORASIS, journées francophones des jeunes chercheurs en vision par ordinateur, 2021, à Toulouse – Oral
- **Y. Cao**, V. Jaouen, V. Bourbonne, N. Boussion, U. Schick, J. Bert, D. Visvikis. *Patient-specific 4DCT respiratory motion synthesis using generative adversarial networks. Cancéropôle grand ouest, Adaptation of the tumour and its ecosystem to radiotherapies*, 2021, à Vannes – Oral
- **Y. Cao**, V. Jaouen, V. Bourbonne, N. Boussion, U. Schick, J. Bert, D. Visvikis. Synthèse du mouvement respiratoire par GAN doublement conditionnel en imagerie 4DCT. Colloque de recherche en imagerie et technologies pour la santé, 2022, à Brest – Oral
- **Y. Cao**, V. Jaouen, V. Bourbonne, F. Lucia, N. Boussion, U. Schick, J. Bert, D. Visvikis. Synthèse d'image dynamique à l'aide d'image statique et d'une valeur scalaire basée sur l'amplitude respiratoire. Colloque français d'intelligence artificielle

- 
- en imagerie biomédicale, 2023, à Paris – Poster
- **Y. Cao**, V. Jaouen, V. Bourbonne, F. Lucia, N. BouSSION, U. Schick, J. Bert, D. Visvikis. Synthèse d'image dynamique à l'aide d'image statique et d'une valeur scalaire basée sur l'amplitude respiratoire. Séminaire de la recherche de l'Institut Mines-Télécom Atlantique, 2023, à Brest – Poster – Prix du meilleur poster



# BIBLIOGRAPHIE

---

- [1] J. R. van SÖRNSEN DE KOSTE, S. SENAN, C. E. KLEYNEN, B. J. SLOTMAN et F. J. LAGERWAARD, « Renal mobility during uncoached quiet respiration : An analysis of 4DCT scans », en, *International Journal of Radiation Oncology\*Biolog\*Physics*, t. 64, 3, p. 799-803, mars 2006, ISSN : 03603016. DOI : 10.1016/j.ijrobp.2005.09.012. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S0360301605026039>.
- [2] D. SHEN, G. WU et H.-I. SUK, « Deep Learning in Medical Image Analysis », en, 2017.
- [3] S. KAZEMINIA, C. BAUR, A. KUIJPER et al., « GANs for medical image analysis », en, *Artificial Intelligence in Medicine*, t. 109, p. 101-938, sept. 2020, ISSN : 09333657. DOI : 10.1016/j.artmed.2020.101938. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S0933365719311510>.
- [4] P. ISOLA, J.-Y. ZHU, T. ZHOU et A. A. EFROS, « Image-to-Image Translation with Conditional Adversarial Networks », en, *arXiv :1611.07004 [cs]*, nov. 2018, arXiv : 1611.07004. adresse : <http://arxiv.org/abs/1611.07004>.
- [5] H. SUNG, J. FERLAY, R. L. SIEGEL et al., « Global Cancer Statistics 2020 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries », en, *CA : A Cancer Journal for Clinicians*, t. 71, 3, p. 209-249, 2021, ISSN : 1542-4863. DOI : 10.3322/caac.21660. adresse : <https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21660>.
- [6] G. DELANEY, S. JACOB, C. FEATHERSTONE et M. BARTON, « The role of radiotherapy in cancer treatment : estimating optimal utilization from a review of evidence-based clinical guidelines », eng, *Cancer*, t. 104, 6, p. 1129-1137, sept. 2005, ISSN : 0008-543X. DOI : 10.1002/cncr.21324.
- [7] D. JONES, « ICRU Report 50-Prescribing, Recording and Reporting Photon Beam Therapy », en, *Medical Physics*, t. 21, 6, p. 833-834, juin 1994, ISSN : 00942405. DOI : 10.1118/1.597396. adresse : <http://doi.wiley.com/10.1118/1.597396>.

- 
- [8] Z. LI, « Prescribing, Recording, and Reporting Proton-Beam Therapy », en, *International Journal of Radiation Oncology\*Biology\*Physics*, t. 73, 5, p. 1602, avr. 2009, ISSN : 03603016. DOI : 10.1016/j.ijrobp.2008.10.084. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S0360301608039059>.
- [9] S. SARUDIS, A. KARLSSON HAUER, J. NYMAN et A. BÄCK, « Systematic evaluation of lung tumor motion using four-dimensional computed tomography », en, *Acta Oncologica*, t. 56, 4, p. 525-530, avr. 2017, ISSN : 0284-186X, 1651-226X. DOI : 10.1080/0284186X.2016.1274049. adresse : <https://www.tandfonline.com/doi/full/10.1080/0284186X.2016.1274049>.
- [10] A. REHAILIA-BLANCHARD, S. DE OLIVEIRA DUARTE, M. BAURY et al., « Utilisation de la scanographie quadridimensionnelle : principaux aspects techniques et intérêts cliniques », fr, *Cancer/Radiothérapie*, t. 23, 4, p. 334-341, juill. 2019, ISSN : 12783218. DOI : 10.1016/j.canrad.2018.07.143. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S1278321819300757>.
- [11] D. PANETTA, « Computed Tomography », en, in *Comprehensive Biomedical Physics*, Elsevier, 2014, p. 65-88, ISBN : 978-0-444-53633-4. DOI : 10.1016/B978-0-444-53632-7.00203-3. adresse : <https://linkinghub.elsevier.com/retrieve/pii/B9780444536327002033>.
- [12] L. A. SHEPP et Y. VARDI, « Maximum Likelihood Reconstruction for Emission Tomography », *IEEE Transactions on Medical Imaging*, t. 1, 2, p. 113-122, 1982. DOI : 10.1109/TMI.1982.4307558.
- [13] I. ELBAKRI et J. FESSLER, « Statistical image reconstruction for polyenergetic X-ray computed tomography », *IEEE Transactions on Medical Imaging*, t. 21, 2, p. 89-99, 2002. DOI : 10.1109/42.993128.
- [14] C. HEINZ, M. REINER, C. BELKA, F. WALTER et M. SÖHN, « Technical evaluation of different respiratory monitoring systems used for 4D CT acquisition under free breathing », en, *Journal of Applied Clinical Medical Physics*, t. 16, 2, p. 334-349, mars 2015, ISSN : 15269914. DOI : 10.1120/jacmp.v16i2.4917. adresse : <http://doi.wiley.com/10.1120/jacmp.v16i2.4917>.
- [15] C. SHI, X. TANG et M. CHAN, « Evaluation of the new respiratory gating system », en, *Precision Radiation Oncology*, t. 1, 4, p. 127-133, déc. 2017, ISSN : 2398-7324, 2398-7324. DOI : 10.1002/pro6.34. adresse : <https://onlinelibrary.wiley.com/doi/10.1002/pro6.34>.

- 
- [16] S. NAZIR, J. BERT, H. FAYAD et D. VISVIKIS, « Technical note : Surface imaging for real-time patient positioning in external radiation therapy », en, *Medical Physics*, t. 48, 12, p. 8037-8044, déc. 2021, ISSN : 0094-2405, 2473-4209. DOI : 10.1002/mp.15300. adresse : <https://onlinelibrary.wiley.com/doi/10.1002/mp.15300>.
- [17] A. F. ABDELNOUR, S. A. NEHMEH, T. PAN et al., « Phase and amplitude binning for 4D-CT imaging », en, *Physics in Medicine and Biology*, t. 52, 12, p. 3515-3529, juill. 2007, ISSN : 0031-9155, 1361-6560. DOI : 10.1088/0031-9155/52/12/012. adresse : <https://iopscience.iop.org/article/10.1088/0031-9155/52/12/012>.
- [18] K. E. ROSENZWEIG, J. HANLEY, D. MAH et al., « The deep inspiration breath-hold technique in the treatment of inoperable non-small-cell lung cancer », en, *International Journal of Radiation Oncology\*Biophysics*, t. 48, 1, p. 81-87, août 2000, ISSN : 03603016. DOI : 10.1016/S0360-3016(00)00583-6. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S0360301600005836>.
- [19] P. HUBBARD, J. CALLAHAN, J. CRAMB, R. BUDD et T. KRON, « Audit of radiation dose delivered in time-resolved four-dimensional computed tomography in a radiotherapy department : Dose in 4D CT for radiotherapy planning », en, *Journal of Medical Imaging and Radiation Oncology*, t. 59, 3, p. 346-352, juin 2015, ISSN : 17549477. DOI : 10.1111/1754-9485.12284. adresse : <https://onlinelibrary.wiley.com/doi/10.1111/1754-9485.12284>.
- [20] G. FERRETTI et A. JANKOWSKI, « Doses d'exposition des examens radiologiques thoraciques », fr, *Revue des Maladies Respiratoires*, t. 27, 6, p. 644-650, juin 2010, ISSN : 07618425. DOI : 10.1016/j.rmr.2010.03.024. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S0761842510002603>.
- [21] Y. G. van der GELD, F. J. LAGERWAARD, J. R. van SÖRNSEN DE KOSTE, J. P. CUIJPERS, B. J. SLOTMAN et S. SENAN, « Reproducibility of target volumes generated using uncoached 4-dimensional CT scans for peripheral lung cancer », *Radiation Oncology*, t. 1, 1, p. 43, nov. 2006, ISSN : 1748-717X. DOI : 10.1186/1748-717X-1-43. adresse : <https://doi.org/10.1186/1748-717X-1-43>.
- [22] M. GUCKENBERGER, J. WILBERT, J. MEYER, K. BAIER, A. RICHTER et M. FLENTJE, « Is a single respiratory correlated 4D-CT study sufficient for evaluation

- 
- of breathing motion? », *International Journal of Radiation Oncology\* Biology\* Physics*, t. 67, 5, p. 1352-1359, 2007.
- [23] P. TRÉMOLIÈRES, A. GONZALEZ-MOYA, A. PAUMIER et al., « Lung stereotactic body radiation therapy : personalized PTV margins according to tumor location and number of four-dimensional CT scans », *Radiation Oncology*, t. 17, 1, p. 5, jan. 2022, ISSN : 1748-717X. DOI : 10.1186/s13014-021-01973-5. adresse : <https://doi.org/10.1186/s13014-021-01973-5>.
- [24] F. KHAN, G. BELL, J. ANTONY et al., « The Use of 4DCT to Reduce Lung Dose : A Dosimetric Analysis », en, *Medical Dosimetry*, t. 34, 4, p. 273-278, déc. 2009, ISSN : 0958-3947. DOI : 10.1016/j.meddos.2008.11.005. adresse : <https://www.sciencedirect.com/science/article/pii/S0958394708001684>.
- [25] D. SARRUT, V. BOLDEA, S. MIGUET et C. GINESTET, « Simulation of four-dimensional CT images from deformable registration between inhale and exhale breath-hold CT scans : Simulation of 4D CT images », en, *Medical Physics*, t. 33, 3, p. 605-617, fév. 2006, ISSN : 00942405. DOI : 10.1118/1.2161409. adresse : <http://doi.wiley.com/10.1118/1.2161409>.
- [26] H. J. FAYAD, F. LAMARE, C. C. LE REST, V. BETTINARDI et D. VISVIKIS, « Generation of 4-Dimensional CT Images Based on 4-Dimensional PET-Derived Motion Fields », en, *Journal of Nuclear Medicine*, t. 54, 4, p. 631-638, avr. 2013, ISSN : 0161-5505, 2159-662X. DOI : 10.2967/jnumed.112.110809. adresse : <http://jnm.snmjournals.org/lookup/doi/10.2967/jnumed.112.110809>.
- [27] A. DUETSCHLER, G. BAUMAN, O. BIERI et al., « Synthetic 4DCT(MRI) lung phantom generation for 4D radiotherapy and image guidance investigations », en, *Medical Physics*, mp.15591, mars 2022, ISSN : 0094-2405, 2473-4209. DOI : 10.1002/mp.15591. adresse : <https://onlinelibrary.wiley.com/doi/10.1002/mp.15591>.
- [28] X. HUANG, M.-Y. LIU, S. BELONGIE et J. KAUTZ, « Multimodal Unsupervised Image-to-Image Translation », en, *arXiv :1804.04732 [cs, stat]*, août 2018, arXiv : 1804.04732. adresse : <http://arxiv.org/abs/1804.04732>.
- [29] J.-Y. ZHU, R. ZHANG, D. PATHAK et al., « Toward Multimodal Image-to-Image Translation », en, *arXiv :1711.11586 [cs, stat]*, oct. 2018, arXiv : 1711.11586. adresse : <http://arxiv.org/abs/1711.11586>.

- 
- [30] J.-Y. ZHU, T. PARK, P. ISOLA et A. A. EFROS, « Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks », en, *arXiv :1703.10593 [cs]*, août 2020, arXiv : 1703.10593. adresse : <http://arxiv.org/abs/1703.10593>.
- [31] Y.-C. CHEN, X. XU, Z. TIAN et J. JIA, « Homomorphic Latent Space Interpolation for Unpaired Image-To-Image Translation », en, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA : IEEE, juin 2019, p. 2403-2411, ISBN : 978-1-72813-293-8. DOI : 10.1109/CVPR.2019.00251. adresse : <https://ieeexplore.ieee.org/document/8954444/>.
- [32] Q. MAO, H.-Y. TSENG, H.-Y. LEE, J.-B. HUANG, S. MA et M.-H. YANG, « Continuous and Diverse Image-to-Image Translation via Signed Attribute Vectors », en, arXiv, rapp. tech. arXiv :2011.01215, avr. 2021, arXiv :2011.01215 [cs] type : article. adresse : <http://arxiv.org/abs/2011.01215>.
- [33] F. PIZZATI, P. CERRI et R. de CHARETTE, « CoMoGAN : continuous model-guided image-to-image translation », en, *arXiv :2103.06879 [cs]*, avr. 2021, arXiv : 2103.06879. adresse : <http://arxiv.org/abs/2103.06879>.
- [34] O. RONNEBERGER, P. FISCHER et T. BROX, « U-Net : Convolutional Networks for Biomedical Image Segmentation », en, *arXiv :1505.04597 [cs]*, mai 2015, arXiv : 1505.04597. adresse : <http://arxiv.org/abs/1505.04597>.
- [35] I. J. GOODFELLOW, J. POUGET-ABADIE, M. MIRZA et al., « Generative Adversarial Networks », en, *arXiv :1406.2661 [cs, stat]*, juin 2014, arXiv : 1406.2661. adresse : <http://arxiv.org/abs/1406.2661>.
- [36] M. MIRZA et S. OSINDERO, « Conditional Generative Adversarial Nets », en, *arXiv :1411.1784 [cs, stat]*, nov. 2014, arXiv : 1411.1784. adresse : <http://arxiv.org/abs/1411.1784>.
- [37] E. HUYNH, A. HOSNY, C. GUTHIER et al., « Artificial intelligence in radiation oncology », en, *Nature Reviews Clinical Oncology*, t. 17, 12, p. 771-781, déc. 2020, ISSN : 1759-4774, 1759-4782. DOI : 10.1038/s41571-020-0417-8. adresse : <http://www.nature.com/articles/s41571-020-0417-8>.
- [38] J. WASSERTHAL, H.-C. BREIT, M. T. MEYER et al., « TotalSegmentator : Robust Segmentation of 104 Anatomic Structures in CT Images », *Radiology : Artificial Intelligence*, e230024, juill. 2023, Publisher : Radiological Society of North America.



- 
- DOI : 10.1148/ryai.230024. adresse : <https://pubs.rsna.org/doi/10.1148/ryai.230024>.
- [39] F. ISENSEE, P. F. JAEGER, S. A. A. KOHL, J. PETERSEN et K. H. MAIER-HEIN, « nnU-Net : a self-configuring method for deep learning-based biomedical image segmentation », en, *Nature Methods*, t. 18, 2, p. 203-211, fév. 2021, Number : 2 Publisher : Nature Publishing Group, ISSN : 1548-7105. DOI : 10.1038/s41592-020-01008-z. adresse : <https://www.nature.com/articles/s41592-020-01008-z>.
- [40] M. D. CIRILLO, D. ABRAMIAN et A. EKLUND, « Vox2Vox : 3D-GAN for Brain Tumour Segmentation », en, *arXiv :2003.13653 [cs, eess]*, nov. 2020, arXiv : 2003.13653. adresse : <http://arxiv.org/abs/2003.13653>.
- [41] S. JAIN, S. INDORA et D. K. ATAL, « Lung nodule segmentation using Salp Shuffled Shepherd Optimization Algorithm-based Generative Adversarial Network », en, *Computers in Biology and Medicine*, t. 137, p. 104811, oct. 2021, ISSN : 00104825. DOI : 10.1016/j.compbimed.2021.104811. adresse : <https://linkinghub.elsevier.com/retrieve/pii/S0010482521006053>.
- [42] M. MASPERO, M. H. F. SAVENIJE, A. M. DINKLA et al., « Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy », en, *Physics in Medicine & Biology*, t. 63, 18, p. 185001, sept. 2018, ISSN : 1361-6560. DOI : 10.1088/1361-6560/aada6d. adresse : <https://iopscience.iop.org/article/10.1088/1361-6560/aada6d>.
- [43] D. L. HILL, P. G. BATCHELOR, M. HOLDEN et D. J. HAWKES, « Medical image registration », eng, *Physics in Medicine and Biology*, t. 46, 3, R1-45, mars 2001, ISSN : 0031-9155. DOI : 10.1088/0031-9155/46/3/201.
- [44] W. R. CRUM, T. HARTKENS et D. L. G. HILL, « Non-rigid image registration : theory and practice », *The British Journal of Radiology*, t. 77, suppl\_2, S140-S153, déc. 2004, Publisher : The British Institute of Radiology, ISSN : 0007-1285. DOI : 10.1259/bjr/25329214. adresse : <https://www.birpublications.org/doi/abs/10.1259/bjr/25329214>.
- [45] J. SARVAIYA, S. PATNAIK et S. BOMBAYWALA, « Image Registration by Template Matching Using Normalized Cross-Correlation », in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, déc.

- 
- 2009, p. 819-822. DOI : 10.1109/ACT.2009.207. adresse : <https://ieeexplore.ieee.org/abstract/document/5375779>.
- [46] F. MAES, D. VANDERMEULEN et P. SUETENS, « Medical image registration using mutual information », *Proceedings of the IEEE*, t. 91, 10, p. 1699-1722, oct. 2003, Conference Name : Proceedings of the IEEE, ISSN : 1558-2256. DOI : 10.1109/JPROC.2003.817864. adresse : <https://ieeexplore.ieee.org/abstract/document/1232201>.
- [47] J. B. MAINTZ et M. A. VIERGEVER, « A survey of medical image registration », eng, *Medical Image Analysis*, t. 2, 1, p. 1-36, mars 1998, ISSN : 1361-8415. DOI : 10.1016/s1361-8415(01)80026-8.
- [48] A. SOTIRAS, C. DAVATZIKOS et N. PARAGIOS, « Deformable Medical Image Registration : A Survey », en, *IEEE Transactions on Medical Imaging*, t. 32, 7, p. 1153-1190, juill. 2013, ISSN : 0278-0062, 1558-254X. DOI : 10.1109/TMI.2013.2265603. adresse : <http://ieeexplore.ieee.org/document/6522524/>.
- [49] X. CHEN, A. DIAZ-PINTO, N. RAVIKUMAR et A. FRANGI, « Deep learning in medical image registration », en, *Progress in Biomedical Engineering*, déc. 2020, ISSN : 2516-1091. DOI : 10.1088/2516-1091/abd37c. adresse : <https://iopscience.iop.org/article/10.1088/2516-1091/abd37c>.
- [50] R. BAJCSY et S. KOVAČIČ, « Multiresolution elastic matching », *Computer Vision, Graphics, and Image Processing*, t. 46, 1, p. 1-21, avr. 1989, ISSN : 0734-189X. DOI : 10.1016/S0734-189X(89)80014-3. adresse : <https://www.sciencedirect.com/science/article/pii/S0734189X89800143>.
- [51] C. DAVATZIKOS, « Spatial Transformation and Registration of Brain Images Using Elastically Deformable Models », *Computer Vision and Image Understanding*, t. 66, 2, p. 207-222, mai 1997, ISSN : 1077-3142. DOI : 10.1006/cviu.1997.0605. adresse : <https://www.sciencedirect.com/science/article/pii/S1077314297906056>.
- [52] D. SHEN et C. DAVATZIKOS, « HAMMER : hierarchical attribute matching mechanism for elastic registration », *IEEE Transactions on Medical Imaging*, t. 21, 11, p. 1421-1439, nov. 2002, Conference Name : IEEE Transactions on Medical Imaging, ISSN : 1558-254X. DOI : 10.1109/TMI.2002.803111. adresse : <https://ieeexplore.ieee.org/document/1175091>.

- 
- [53] J. ASHBURNER et K. J. FRISTON, « Voxel-Based Morphometry—The Methods », *NeuroImage*, t. 11, 6, p. 805-821, juin 2000, ISSN : 1053-8119. DOI : 10.1006/nimg.2000.0582. adresse : <https://www.sciencedirect.com/science/article/pii/S1053811900905822>.
- [54] D. RUECKERT, L. SONODA, C. HAYES, D. HILL, M. LEACH et D. HAWKES, « Non-rigid registration using free-form deformations : application to breast MR images », *IEEE Transactions on Medical Imaging*, t. 18, 8, p. 712-721, août 1999, Conference Name : IEEE Transactions on Medical Imaging, ISSN : 1558-254X. DOI : 10.1109/42.796284. adresse : <https://ieeexplore.ieee.org/document/796284>.
- [55] J. -. THIRION, « Image matching as a diffusion process : an analogy with Maxwell's demons », *Medical Image Analysis*, t. 2, 3, p. 243-260, sept. 1998, ISSN : 1361-8415. DOI : 10.1016/S1361-8415(98)80022-4. adresse : <https://www.sciencedirect.com/science/article/pii/S1361841598800224>.
- [56] M. F. BEG, M. I. MILLER, A. TROUVÉ et L. YOUNES, « Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms », en, *International Journal of Computer Vision*, t. 61, 2, p. 139-157, fév. 2005, ISSN : 1573-1405. DOI : 10.1023/B:VISI.0000043755.93987.aa. adresse : <https://doi.org/10.1023/B:VISI.0000043755.93987.aa>.
- [57] M. ZHANG, R. LIAO, A. V. DALCA et al., « Frequency Diffeomorphisms for Efficient Image Registration », *Information processing in medical imaging : proceedings of the ... conference*, t. 10265, p. 559-570, juin 2017, ISSN : 1011-2499. DOI : 10.1007/978-3-319-59050-9\_44. adresse : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5788203/>.
- [58] T. VERCAUTEREN, X. PENNEC, A. PERCHANT et N. AYACHE, « Diffeomorphic demons : Efficient non-parametric image registration », *NeuroImage, Mathematics in Brain Imaging*, t. 45, 1, *Supplement 1*, S61-S72, mars 2009, ISSN : 1053-8119. DOI : 10.1016/j.neuroimage.2008.10.040. adresse : <https://www.sciencedirect.com/science/article/pii/S1053811908011683>.
- [59] T. D. ZUK, M. S. ATKINS et K. S. BOOTH, « Approaches to registration using 3D surfaces », in *Medical Imaging 1994 : Image Processing*, t. 2167, SPIE, mai 1994, p. 176-187. DOI : 10.1117/12.175052. adresse : <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/2167/0000/Approaches-to-registration-using-3D-surfaces/10.1117/12.175052.full>.

- 
- [60] G. CHRISTENSEN, A. KANE, J. MARSH et M. VANNIER, « Synthesis of an individualized cranial atlas with dysmorphic shape », in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, juin 1996, p. 309-318. DOI : 10.1109/MMBIA.1996.534083. adresse : <https://ieeexplore.ieee.org/abstract/document/534083>.
- [61] D. N. LEVIN, C. A. PELIZZARI, G. T. CHEN, C. T. CHEN et M. D. COOPER, « Retrospective geometric correlation of MR, CT, and PET images. », *Radiology*, t. 169, 3, p. 817-823, déc. 1988, Publisher : Radiological Society of North America, ISSN : 0033-8419. DOI : 10.1148/radiology.169.3.3263666. adresse : <https://pubs.rsna.org/doi/10.1148/radiology.169.3.3263666>.
- [62] W. M. WELLS, P. VIOLA, H. ATSUMI, S. NAKAJIMA et R. KIKINIS, « Multi-modal volume registration by maximization of mutual information », *Medical Image Analysis*, t. 1, 1, p. 35-51, mars 1996, ISSN : 1361-8415. DOI : 10.1016/S1361-8415(01)80004-9. adresse : <https://www.sciencedirect.com/science/article/pii/S1361841501800049>.
- [63] P. VIOLA, N. SCHRAUDOLPH et T. J. SEJNOWSKI, « Empirical Entropy Manipulation for Real-World Problems », in *Advances in Neural Information Processing Systems*, D. TOURETZKY, M. C. MOZER et M. HASSELMO, éd., t. 8, MIT Press, 1995. adresse : [https://proceedings.neurips.cc/paper\\_files/paper/1995/file/537d9b6c927223c796cac288cced29df-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/1995/file/537d9b6c927223c796cac288cced29df-Paper.pdf).
- [64] H. JI, Y. LI, E. DONG et al., « A non-rigid image registration method based on multi-level B-spline and L2-regularization », en, *Signal, Image and Video Processing*, t. 12, 6, p. 1217-1225, sept. 2018, ISSN : 1863-1711. DOI : 10.1007/s11760-018-1274-0. adresse : <https://doi.org/10.1007/s11760-018-1274-0>.
- [65] V. VISHNEVSKIY, T. GASS, G. SZEKELY, C. TANNER et O. GOKSEL, « Isotropic Total Variation Regularization of Displacements in Parametric Image Registration », *IEEE Transactions on Medical Imaging*, t. 36, 2, p. 385-395, fév. 2017, Conference Name : IEEE Transactions on Medical Imaging, ISSN : 1558-254X. DOI : 10.1109/TMI.2016.2610583.
- [66] S. SOMAYAJULA, A. A. JOSHI et R. M. LEAHY, « Non-Rigid Image Registration Using Gaussian Mixture Models », *Biomedical image registration : second international workshop, WBIR 2003, Philadelphia, PA, USA, June 23-24, 2003 : revised papers / James C. Gee, J.B. Antoine Maintz, Michael W. Vannier (eds.)*.

- 
- International Workshop on Biomedical Imag.*, t. 7359, p. 286-295, 2012. DOI : 10.1007/978-3-642-31340-0\_30. adresse : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4702048/>.
- [67] D. F. PACE, S. R. AYLWARD et M. NIETHAMMER, « A Locally Adaptive Regularization Based on Anisotropic Diffusion for Deformable Image Registration of Sliding Organs », en, *IEEE Transactions on Medical Imaging*, t. 32, 11, p. 2114-2126, nov. 2013, ISSN : 0278-0062, 1558-254X. DOI : 10.1109/TMI.2013.2274777. adresse : <http://ieeexplore.ieee.org/document/6568964/>.
- [68] A. HERING, L. HANSEN, T. C. W. MOK et al., *Learn2Reg : comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning*, en, arXiv :2112.04489 [cs, eess], oct. 2022. adresse : <http://arxiv.org/abs/2112.04489>.
- [69] M. JADERBERG, K. SIMONYAN, A. ZISSERMAN et K. KAVUKCUOGLU, « Spatial Transformer Networks », *CoRR*, t. abs/1506.02025, 2015. arXiv : 1506.02025. adresse : <http://arxiv.org/abs/1506.02025>.
- [70] G. BALAKRISHNAN, A. ZHAO, M. R. SABUNCU, J. GUTTAG et A. V. DALCA, « VoxelMorph : A Learning Framework for Deformable Medical Image Registration », en, *IEEE Transactions on Medical Imaging*, t. 38, 8, p. 1788-1800, août 2019, arXiv : 1809.05231, ISSN : 0278-0062, 1558-254X. DOI : 10.1109/TMI.2019.2897538. adresse : <http://arxiv.org/abs/1809.05231>.
- [71] G. BALAKRISHNAN, A. ZHAO, M. R. SABUNCU, J. GUTTAG et A. V. DALCA, « An Unsupervised Learning Model for Deformable Medical Image Registration », en, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, arXiv :1802.02604 [cs], juin 2018, p. 9252-9260. DOI : 10.1109/CVPR.2018.00964. adresse : <http://arxiv.org/abs/1802.02604>.
- [72] F. MILLETARI, N. NAVAB et S.-A. AHMADI, *V-Net : Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation*, en, arXiv :1606.04797 [cs], juin 2016. adresse : <http://arxiv.org/abs/1606.04797>.
- [73] Y. FU, Y. LEI, T. WANG et al., « LungRegNet : an unsupervised deformable image registration method for 4D-CT lung », *Medical physics*, t. 47, 4, p. 1763-1774, avr. 2020, ISSN : 0094-2405. DOI : 10.1002/mp.14065. adresse : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7165051/>.

- 
- [74] M. MODAT, G. R. RIDGWAY, Z. A. TAYLOR et al., « Fast free-form deformation using graphics processing units », eng, *Computer Methods and Programs in Biomedicine*, t. 98, 3, p. 278-284, juin 2010, ISSN : 1872-7565. DOI : 10.1016/j.cmpb.2009.09.002.
- [75] J. A. SHACKLEFORD, N. KANDASAMY et G. C. SHARP, « On developing B-spline registration algorithms for multi-core processors », en, *Physics in Medicine & Biology*, t. 55, 21, p. 6329, oct. 2010, ISSN : 0031-9155. DOI : 10.1088/0031-9155/55/21/001. adresse : <https://dx.doi.org/10.1088/0031-9155/55/21/001>.
- [76] R. WERNER, A. SCHMIDT-RICHBERG, H. HANDELS et J. EHRHARDT, « Estimation of lung motion fields in 4D CT data by variational non-linear intensity-based registration : A comparison and evaluation study », en, *Physics in Medicine & Biology*, t. 59, 15, p. 4247, juill. 2014, Publisher : IOP Publishing, ISSN : 0031-9155. DOI : 10.1088/0031-9155/59/15/4247. adresse : <https://dx.doi.org/10.1088/0031-9155/59/15/4247>.
- [77] T. SENTKER, F. MADESTA et R. WERNER, « GDL-FIRE4D : Deep Learning-Based Fast 4D CT Image Registration », en, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, A. F. FRANGI, J. A. SCHNABEL, C. DAVATZIKOS, C. ALBEROLA-LÓPEZ et G. FICHTINGER, éd., t. 11070, Series Title : Lecture Notes in Computer Science, Cham : Springer International Publishing, 2018, p. 765-773. DOI : 10.1007/978-3-030-00928-1\_86. adresse : [http://link.springer.com/10.1007/978-3-030-00928-1\\_86](http://link.springer.com/10.1007/978-3-030-00928-1_86).
- [78] R. CASTILLO, E. CASTILLO, R. GUERRA et al., « A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets », eng, *Physics in Medicine and Biology*, t. 54, 7, p. 1849-1870, avr. 2009, ISSN : 0031-9155. DOI : 10.1088/0031-9155/54/7/001.
- [79] J. FAN, X. CAO, P.-T. YAP et D. SHEN, « BIRNet : Brain image registration using dual-supervised fully convolutional networks », en, *Medical Image Analysis*, t. 54, p. 193-206, mai 2019, ISSN : 1361-8415. DOI : 10.1016/j.media.2019.03.006. adresse : <https://www.sciencedirect.com/science/article/pii/S1361841519300283>.
- [80] B. B. AVANTS, N. J. TUSTISON, G. SONG, P. A. COOK, A. KLEIN et J. C. GEE, « A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration », *NeuroImage*, t. 54, 3, p. 2033-2044, fév. 2011, ISSN : 1053-8119.

- 
- DOI : 10.1016/j.neuroimage.2010.09.025. adresse : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3065962/>.
- [81] M. LORENZI, N. AYACHE, G. B. FRISONI et X. PENNEC, « LCC-Demons : A robust and accurate symmetric diffeomorphic registration algorithm », en, *NeuroImage*, t. 81, p. 470-483, nov. 2013, ISSN : 1053-8119. DOI : 10.1016/j.neuroimage.2013.04.114. adresse : <https://www.sciencedirect.com/science/article/pii/S1053811913004825>.
- [82] X. TENG, Y. CHEN, Y. ZHANG et L. REN, « Respiratory deformation registration in 4D-CT/cone beam CT using deep learning », *Quantitative Imaging in Medicine and Surgery*, t. 11, 2, p. 737-748, fév. 2021, ISSN : 2223-4292. DOI : 10.21037/qims-19-1058. adresse : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7779910/>.
- [83] T.-C. WANG, M.-Y. LIU, J.-Y. ZHU et al., « Video-to-Video Synthesis », in *Advances in Neural Information Processing Systems*, S. BENGIO, H. WALLACH, H. LAROCHELLE, K. GRAUMAN, N. CESA-BIANCHI et R. GARNETT, éd., t. 31, Curran Associates, Inc., 2018. adresse : [https://proceedings.neurips.cc/paper\\_files/paper/2018/file/d86ea612dec96096c5e0fcc8dd42ab6d-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2018/file/d86ea612dec96096c5e0fcc8dd42ab6d-Paper.pdf).
- [84] J. THIES, M. ZOLLHÖFER, M. STAMMINGER, C. THEOBALT et M. NIESSNER, *Face2Face : Real-time Face Capture and Reenactment of RGB Videos*, en, arXiv :2007.14808 [cs], juill. 2020. adresse : <http://arxiv.org/abs/2007.14808>.
- [85] A. SIAROHIN, S. LATHUILÈRE, S. TULYAKOV, E. RICCI et N. SEBE, « First Order Motion Model for Image Animation », en, *arXiv :2003.00196 [cs]*, oct. 2020, arXiv : 2003.00196. adresse : <http://arxiv.org/abs/2003.00196>.
- [86] C. CHAN, S. GINOSAR, T. ZHOU et A. A. EFROS, « Everybody Dance Now », en, p. 10, 2019.
- [87] C. FINN, I. GOODFELLOW et S. LEVINE, *Unsupervised Learning for Physical Interaction through Video Prediction*, en, arXiv :1605.07157 [cs], oct. 2016. adresse : <http://arxiv.org/abs/1605.07157>.
- [88] A. X. LEE, R. ZHANG, F. EBERT, P. ABBEEL, C. FINN et S. LEVINE, *Stochastic Adversarial Video Prediction*, en, arXiv :1804.01523 [cs], avr. 2018. adresse : <http://arxiv.org/abs/1804.01523>.

- 
- [89] J. WALKER, K. MARINO, A. GUPTA et M. HEBERT, *The Pose Knows : Video Forecasting by Generating Pose Futures*, en, arXiv :1705.00053 [cs], avr. 2017. adresse : <http://arxiv.org/abs/1705.00053>.
- [90] L. CHEN, Z. LI, R. K. MADDOX, Z. DUAN et C. XU, *Lip Movements Generation at a Glance*, en, arXiv :1803.10404 [cs], mai 2018. adresse : <http://arxiv.org/abs/1803.10404>.
- [91] X. REN, H. LI, Z. HUANG et Q. CHEN, *Music-oriented Dance Video Synthesis with Pose Perceptual Loss*, en, arXiv :1912.06606 [cs, eess], déc. 2019. adresse : <http://arxiv.org/abs/1912.06606>.
- [92] J. P. FERREIRA, T. M. COUTINHO, T. L. GOMES et al., « Learning to dance : A graph convolutional adversarial network to generate realistic dance motions from audio », en, *Computers & Graphics*, t. 94, p. 11-21, fév. 2021, arXiv : 2011.12999, ISSN : 00978493. DOI : 10.1016/j.cag.2020.09.009. adresse : <http://arxiv.org/abs/2011.12999>.
- [93] K. VOUGIOUKAS, S. PETRIDIS et M. PANTIC, *End-to-End Speech-Driven Facial Animation with Temporal GANs*, en, arXiv :1805.09313 [cs, eess], juill. 2018. adresse : <http://arxiv.org/abs/1805.09313>.
- [94] V. JAOUEN et D. VISVIKIS, « 4D respiratory motion synchronized image synthesis from static CT images using GANs », en, p. 3, 2019.
- [95] S. MORI, R. HIRAI et Y. SAKATA, « Simulated four-dimensional CT for marker-less tumor tracking using a deep learning network with multi-task learning », en, *Physica Medica*, t. 80, p. 151-158, déc. 2020, ISSN : 1120-1797. DOI : 10.1016/j.ejmp.2020.10.023. adresse : <https://www.sciencedirect.com/science/article/pii/S1120179720302660>.
- [96] Y. CHANG, Z. JIANG, W. P. SEGARS et al., « A generative adversarial network (GAN)-based technique for synthesizing realistic respiratory motion in the extended cardiac-torso (XCAT) phantoms », en, *Physics in Medicine & Biology*, t. 66, 11, p. 115018, mai 2021, Publisher : IOP Publishing, ISSN : 0031-9155. DOI : 10.1088/1361-6560/ac01b4. adresse : <https://doi.org/10.1088/1361-6560/ac01b4>.



- 
- [97] W. P. SEGARS, G. STURGEON, S. MENDONCA, J. GRIMES et B. M. W. TSUI, « 4D XCAT phantom for multimodality imaging research : 4D XCAT phantom for multimodality imaging research », en, *Medical Physics*, t. 37, 9, p. 4902-4915, août 2010, ISSN : 00942405. DOI : 10.1118/1.3480985. adresse : <http://doi.wiley.com/10.1118/1.3480985>.
- [98] L. V. ROMAGUERA, T. MEZHERITSKY, R. MANSOUR, J.-F. CARRIER et S. KADOURY, « Probabilistic 4D predictive model from in-room surrogates using conditional generative networks for image-guided radiotherapy », en, *Medical Image Analysis*, t. 74, p. 102250, déc. 2021, ISSN : 1361-8415. DOI : 10.1016/j.media.2021.102250. adresse : <https://www.sciencedirect.com/science/article/pii/S1361841521002954>.
- [99] Y. HUANG, Z. DONG, H. WU, C. LI, H. LIU et Y. ZHANG, « Deep learning-based synthetization of real-time in-treatment 4D images using surface motion and pretreatment images : A proof-of-concept study », en, *Medical Physics*, t. n/a, n/a, 2022, \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mp.15858>, ISSN : 2473-4209. DOI : 10.1002/mp.15858. adresse : <https://onlinelibrary.wiley.com/doi/abs/10.1002/mp.15858>.
- [100] Y. SANG et D. RUAN, « A conditional registration network for continuous 4D respiratory motion synthesis », en, *Medical Physics*, t. n/a, n/a, 2023, \_eprint : <https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/mp.16226>, ISSN : 2473-4209. DOI : 10.1002/mp.16226. adresse : <http://onlinelibrary.wiley.com/doi/abs/10.1002/mp.16226>.
- [101] D. LEE, E. YORKE, M. ZAREPISHEH, S. NADEEM et Y. HU, « RMSim : controlled respiratory motion simulation on static patient scans », en, *Physics in Medicine & Biology*, 2023, ISSN : 0031-9155. DOI : 10.1088/1361-6560/acb484. adresse : <http://iopscience.iop.org/article/10.1088/1361-6560/acb484>.
- [102] Y. CHOI, Y. UH, J. YOO et J.-W. HA, « StarGAN v2 : Diverse Image Synthesis for Multiple Domains », en, *arXiv :1912.01865 [cs]*, avr. 2020, arXiv : 1912.01865. adresse : <http://arxiv.org/abs/1912.01865>.
- [103] T. KARRAS, S. LAINE, M. AITTALA, J. HELLSTEN, J. LEHTINEN et T. AILA, « Analyzing and Improving the Image Quality of StyleGAN », en, arXiv, rapp. tech. arXiv :1912.04958, mars 2020, arXiv :1912.04958 [cs, eess, stat] type : article. adresse : <http://arxiv.org/abs/1912.04958>.

- 
- [104] G. YANG, N. FEI, M. DING, G. LIU, Z. LU et T. XIANG, « L2M-GAN : Learning to Manipulate Latent Space Semantics for Facial Attribute Editing », en, in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA : IEEE, juin 2021, p. 2950-2959, ISBN : 978-1-66544-509-2. DOI : 10.1109/CVPR46437.2021.00297. adresse : <https://ieeexplore.ieee.org/document/9578611/>.
- [105] P.-W. WU, Y.-J. LIN, C.-H. CHANG, E. Y. CHANG et S.-W. LIAO, « RelGAN : Multi-Domain Image-to-Image Translation via Relative Attributes », en, *arXiv :1908.07269 [cs]*, août 2019, arXiv : 1908.07269. adresse : <http://arxiv.org/abs/1908.07269>.
- [106] X. HUANG et S. BELONGIE, « Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization », en, *arXiv :1703.06868 [cs]*, juill. 2017, arXiv : 1703.06868. adresse : <http://arxiv.org/abs/1703.06868>.
- [107] STARING et al., « Pulmonary image registration with elastix using a standard intensity-based algorithm », *Medical Image Analysis for the Clinic : A Grand Challenge*, p. 73-79, 2010.
- [108] KLEIN et al., « Adaptive stochastic gradient descent optimisation for image registration », *International journal of computer vision*, t. 81, 3, p. 227, 2009.
- [109] D. P. KINGMA et J. BA, « Adam : A Method for Stochastic Optimization », en, *arXiv :1412.6980 [cs]*, jan. 2017, arXiv : 1412.6980. adresse : <http://arxiv.org/abs/1412.6980>.
- [110] J. HOFMANNINGER, F. PRAYER, J. PAN, S. RÖHRICH, H. PROSCH et G. LANGS, « Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem », en, *European Radiology Experimental*, t. 4, 1, p. 50, déc. 2020, ISSN : 2509-9280. DOI : 10.1186/s41747-020-00173-2. adresse : <https://eurradiolexp.springeropen.com/articles/10.1186/s41747-020-00173-2>.
- [111] M. ANTONELLI, A. REINKE, S. BAKAS et al., « The Medical Segmentation Decathlon », en, *Nature Communications*, t. 13, 1, p. 4128, juill. 2022, Number : 1 Publisher : Nature Publishing Group, ISSN : 2041-1723. DOI : 10.1038/s41467-022-30695-9. adresse : <https://www.nature.com/articles/s41467-022-30695-9>.

- 
- [112] H. XIAO, R. NI, S. ZHI et al., « A dual-supervised deformation estimation model (DDEM) for constructing ultra-quality 4D-MRI based on a commercial low-quality 4D-MRI for liver cancer radiation therapy », eng, *Medical Physics*, t. 49, 5, p. 3159-3170, mai 2022, ISSN : 2473-4209. DOI : 10.1002/mp.15542.
- [113] G. D. HUGO, E. WEISS, W. C. SLEEMAN et al., *Data from 4D Lung Imaging of NSCLC Patients*, Version Number : 2 Type : dataset, 2016. DOI : 10.7937/K9/TCIA.2016.ELN8YGLE. adresse : <https://wiki.cancerimagingarchive.net/x/1oNEAQ>.
- [114] A. REINKE, M. D. TIZABI, C. H. SUDRE et al., « Common Limitations of Image Processing Metrics : A Picture Story », en, arXiv, rapp. tech. arXiv :2104.05642, juill. 2022, arXiv :2104.05642 [cs, eess] type : article. adresse : <http://arxiv.org/abs/2104.05642>.
- [115] R. FINNEGAN, J. DOWLING, E.-S. KOH et al., « Feasibility of multi-atlas cardiac segmentation from thoracic planning CT in a probabilistic framework », en, *Physics in Medicine & Biology*, t. 64, 8, p. 085006, avr. 2019, Publisher : IOP Publishing, ISSN : 0031-9155. DOI : 10.1088/1361-6560/ab0ea6. adresse : <https://doi.org/10.1088/1361-6560/ab0ea6>.
- [116] Y. JI, H. BAI, J. YANG et al., « AMOS : A Large-Scale Abdominal Multi-Organ Benchmark for Versatile Medical Image Segmentation », en, arXiv, rapp. tech. arXiv :2206.08023, sept. 2022, arXiv :2206.08023 [cs, eess] type : article. adresse : <http://arxiv.org/abs/2206.08023>.
- [117] H. FAYAD, T. PAN, J. FRANÇOIS CLEMENT et D. VISVIKIS, « Technical Note : Correlation of respiratory motion between external patient surface and internal anatomical landmarks : Correlation between external patient surface and internal landmarks », en, *Medical Physics*, t. 38, 6Part1, p. 3157-3164, mai 2011, ISSN : 00942405. DOI : 10.1118/1.3589131. adresse : <http://doi.wiley.com/10.1118/1.3589131>.
- [118] X. ZHOU, T. HARA, H. FUJITA, R. YOKOYAMA, T. KIRYU et H. HOSHI, « Automated segmentations of skin, soft-tissue, and skeleton, from torso CT images », J. M. FITZPATRICK et M. SONKA, éd., San Diego, CA, mai 2004, p. 1634. DOI : 10.1117/12.534843. adresse : <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.534843>.

- 
- [119] R. GROENENDIJK, S. KARAOGLU, T. GEVERS et T. MENSINK, « On the Benefit of Adversarial Training for Monocular Depth Estimation », en, *Computer Vision and Image Understanding*, t. 190, p. 102848, jan. 2020, arXiv :1910.13340 [cs, eess], ISSN : 10773142. DOI : 10.1016/j.cviu.2019.102848. adresse : <http://arxiv.org/abs/1910.13340>.
- [120] A. NGUYEN et B. LE, « 3D point cloud segmentation : A survey », en, in *2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, Manila, Philippines : IEEE, nov. 2013, p. 225-230. DOI : 10.1109/RAM.2013.6758588. adresse : <http://ieeexplore.ieee.org/document/6758588/>.
- [121] S. A. BELLO, S. YU, C. WANG, J. M. ADAM et J. LI, « Review : Deep Learning on 3D Point Clouds », en, *Remote Sensing*, t. 12, 11, p. 1729, mai 2020, ISSN : 2072-4292. DOI : 10.3390/rs12111729. adresse : <https://www.mdpi.com/2072-4292/12/11/1729>.
- [122] X. HUANG, G. MEI, J. ZHANG et R. ABBAS, « A comprehensive survey on point cloud registration », en, arXiv, rapp. tech. arXiv :2103.02690, mars 2021, arXiv :2103.02690 [cs] type : article. adresse : <http://arxiv.org/abs/2103.02690>.
- [123] T. WENTZ, J.-F. CLÉMENT, H. FAYAD et D. VISVIKIS, « Time-of-Flight camera for patient motion monitoring in medical applications », in *RITS 2011 : Recherche en Imagerie et Technologies pour la Santé*, 2011.
- [124] T. WENTZ, H. FAYAD, J. BERT et al., « Accuracy of dynamic patient surface monitoring using a time-of-flight camera and B-spline modeling for respiratory motion characterization », en, *Physics in Medicine and Biology*, t. 57, 13, p. 4175-4193, juill. 2012, ISSN : 0031-9155, 1361-6560. DOI : 10.1088/0031-9155/57/13/4175. adresse : <https://iopscience.iop.org/article/10.1088/0031-9155/57/13/4175>.
- [125] T. KARRAS, M. AITTALA, S. LAINE et al., « Alias-Free Generative Adversarial Networks », en, arXiv, rapp. tech. arXiv :2106.12423, oct. 2021, arXiv :2106.12423 [cs, stat] type : article. adresse : <http://arxiv.org/abs/2106.12423>.
- [126] M. GAO, N. RUAN, J. SHI et W. ZHOU, « Deep Neural Network for 3D Shape Classification Based on Mesh Feature », en, *Sensors*, t. 22, 18, p. 7040, jan. 2022, Number : 18 Publisher : Multidisciplinary Digital Publishing Institute, ISSN : 1424-8220. DOI : 10.3390/s22187040. adresse : <https://www.mdpi.com/1424-8220/22/18/7040>.

- 
- [127] A. S. FELIX et L. A. BRINTON, « Cancer Progress and Priorities : Uterine Cancer », *Cancer Epidemiology, Biomarkers & Prevention*, t. 27, 9, p. 985-994, sept. 2018, ISSN : 1055-9965. DOI : 10.1158/1055-9965.EPI-18-0264. adresse : <https://doi.org/10.1158/1055-9965.EPI-18-0264>.
- [128] S. WHETSTONE, W. BURKE, S. S. SHETH et al., « Health Disparities in Uterine Cancer », *Obstetrics and Gynecology*, t. 139, 4, p. 645-659, avr. 2022, ISSN : 0029-7844. DOI : 10.1097/AOG.0000000000004710. adresse : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8936152/>.
- [129] O. AKIN, S. MIRONOV, N. PANDIT-TASKAR et L. E. HANN, « Imaging of Uterine Cancer », English, *Radiologic Clinics*, t. 45, 1, p. 167-182, jan. 2007, Publisher : Elsevier, ISSN : 0033-8389, 1557-8275. DOI : 10.1016/j.rcl.2006.10.009. adresse : [https://www.radiologic.theclinics.com/article/S0033-8389\(06\)00103-5/fulltext](https://www.radiologic.theclinics.com/article/S0033-8389(06)00103-5/fulltext).
- [130] C. E. DESANTIS, C. C. LIN, A. B. MARIOTTO et al., « Cancer treatment and survivorship statistics, 2014 », en, *CA : A Cancer Journal for Clinicians*, t. 64, 4, p. 252-271, 2014, \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.3322/caac.21235>, ISSN : 1542-4863. DOI : 10.3322/caac.21235. adresse : <https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21235>.
- [131] R. JADON, C. A. PEMBROKE, C. L. HANNA et al., « A Systematic Review of Organ Motion and Image-guided Strategies in External Beam Radiotherapy for Cervical Cancer », *Clinical Oncology*, t. 26, 4, p. 185-196, avr. 2014, ISSN : 0936-6555. DOI : 10.1016/j.clon.2013.11.031. adresse : <https://www.sciencedirect.com/science/article/pii/S0936655513005396>.
- [132] P. DHARIWAL et A. NICHOL, « Diffusion Models Beat GANs on Image Synthesis », en, 2021.
- [133] K. HE, C. GAN, Z. LI et al., « Transformers in medical image analysis », *Intelligent Medicine*, t. 3, 1, p. 59-78, fév. 2023, ISSN : 2667-1026. DOI : 10.1016/j.imed.2022.07.002. adresse : <https://www.sciencedirect.com/science/article/pii/S2667102622000717>.
- [134] C. LI, Y. QI, Q. ZENG et L. LU, « Comparison of Image Generation methods based on Diffusion Models », in *2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, mai 2023, p. 1-4. DOI : 10.1109/CVIDL58838.2023.10167336.

- 
- [135] D. A. HUDSON et C. L. ZITNICK, « Generative Adversarial Transformers », en, 2021.
- [136] S. R. DUBEY et S. K. SINGH, « Transformer-based Generative Adversarial Networks in Computer Vision : A Comprehensive Survey », en, arXiv, rapp. tech. arXiv :2302.08641, fév. 2023, arXiv :2302.08641 [cs, eess] type : article. adresse : <http://arxiv.org/abs/2302.08641>.
- [137] Z. WANG, H. ZHENG, P. HE, W. CHEN et M. ZHOU, « Diffusion-GAN : Training GANs with Diffusion », en, arXiv, rapp. tech. arXiv :2206.02262, août 2023, arXiv :2206.02262 [cs, stat] type : article. adresse : <http://arxiv.org/abs/2206.02262>.







**Titre :** Apprentissage génératif pour la synthèse d'images médicales dynamiques 4D

**Mot clés :** Synthèse d'image, imagerie dynamique, mouvement respiratoire, radiothérapie

**Résumé :** La tomodensitométrie quadridimensionnelle (TDM 4D) consiste à reconstruire une acquisition en plusieurs phases afin de suivre les mouvements des organes internes et des tumeurs. Elle est couramment utilisée dans la planification des traitements de radiothérapie pour le cancer du poumon, mais elle expose les patients à des doses de radiation plus élevées, allant jusqu'à six fois supérieures à celles d'une tomodensitométrie tridimensionnelle (TDM 3D) classique. Les méthodes d'apprentissage automatique profond, issues du domaine de la vision par ordinateur, suscitent un vif intérêt au sein de la communauté de l'imagerie médicale. Parmi ces approches, les modèles génératifs se distinguent en leur capacité à générer

des images synthétiques reproduisant fidèlement l'apparence et les caractéristiques statistiques d'images obtenues à partir de systèmes réels. Dans cette thèse, nous explorons l'utilisation d'un modèle génératif pour la génération d'image dynamique. Nous proposons un modèle capable de générer un mouvement respiratoire spécifique à un patient à partir d'une image TDM 3D de diagnostic et de données respiratoires. L'objectif est de permettre aux radiologues d'effectuer la délimitation des volumes cibles et des organes à risque, ainsi que le calcul de la dose sur des images dynamiques de synthèses. Cette méthode éviterait ainsi la nécessité de réaliser une acquisition TDM 4D, réduisant ainsi l'exposition du patient aux radiations.

**Title:** Dynamic generative modelling for medical applications

**Keywords:** Image synthesis, dynamic imaging, respiratory motion, radiotherapy

**Abstract:** Four-dimensional computed tomography (4DCT) involves reconstructing an acquisition in multiple phases to track the movements of internal organs and tumors. It is used routinely for radiotherapy treatment planning of lung cancer, but it exposes patients to higher radiation doses, up to six times greater than those of a conventional three-dimensional computed tomography (3DCT). Deep learning methods from the field of computer vision are gaining significant interest within the medical imaging community. Among these approaches, generative models stand out due to their ability to generate synthetic

images that faithfully replicate the appearance and statistical characteristics of images acquired from real systems. In this thesis, we explore the use of a generative model for dynamic image generation. We propose a model capable of generating patient-specific respiratory motion from a diagnostic 3DCT image and respiratory data. The goal is to enable radiologists to delineate target volumes and organs at risk, as well as perform dose calculations on these dynamic synthetic images. This method would reduce the need for a 4DCT acquisition, thereby reducing the patient's radiation exposure.