People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University Ferhat Abbas Setif 1



Faculty of Sciences, Department of Physics

Master's Thesis

Presented to obtain Master's degree in Medical Physics

Clinical Validation of AI-Based Dose Prediction Models in Radiotherapy

KHARCHI MOHAMED RAMZI

Presented on the 24 of June 2026, in front of the responsible jury:

Kharfi Faycal	Prof. Ferhat Abbas Setif 1 University	Chair
Azizi Hacene	Prof. Ferhat Abbas Setif 1 University	Advisor
Betka Abderraahim	Dr. Ferhat Abbas Setif 1 University	Examiner
Meddas Ahmed Omrane	PhD. Ferhat Abbas Setif 1 University	Guest

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ FERHAT ABBAS SÉTIF 1



FACULTÉ DES SCIENCES, DÉPARTEMENT DE PHYSIQUE

MÉMOIRE

Présenté pour obtenir le titre de Master en Physique Médicale

VALIDATION CLINIQUE DES MODELE DE PREDICITION DE DOSE PAR IA EN RADIOTHERAPIE

KHARCHI MOHAMED RAMZI

Présenté le 24 Juin 2026, devant le jury composé de :

KHARFI FAYCAL	Pr. Université Ferhat Abbas Setif 1	Président
AZIZI HACENE	Pr. Université Ferhat Abbas Setif 1	Encadrant
BETTKA ABDRRAHIM	Dr. Université Ferhat Abbas Setif 1	Examinateur
MEDDAS AHMED OMRANE	PhD. Université Ferhat Abbas Setif 1	Invite

Abstract

the clinical validation of a 3D Attention Residual U-Net dose prediction model, trained on the OpenKBP dataset comprising 340 head and neck cancer patients treated with intensity-modulated radiation therapy to a prescription dose of 70 Gy. The model takes as input a 12-channel three-dimensional tensor combining CT images with binary masks of the planning target volume and organs at risk, and outputs a voxel-wise predicted dose distribution. the objective is To assess the clinical validation of the trained model beyond the training data domain, a dedicated validation protocol is implemented using a CT calibration phantom, designed to replicate human tissue heterogeneities. A reference treatment plan is generated on the phantom using the Monaco treatment planning system, and compare against the dose distribution predicted by the trained deep learning model on the same phantom geometry. The agreement between the predicted and ground-truth dose distributions assessed using the evaluation metrics. The results demonstrate that deep learning model is capable of reproducing the overall spatial dose distribution and dose gradients observed in the groud truth reference plan, with close agreement in DVH curves and dose profiles for the PTV and surrounding regions. However, quantitative discrepancies are identified between the predicted and ground-truth doses, attributable to the domain shift between the anatomically complex training data and the simplified phantom geometry, as well as to differences in dose normalization and calculation algorithms between the training pipeline and the clinical TPS.

Keywords: Deep learning, dose prediction, ground truth, clinical validation, Attention Residual U-Net, ct phantom, Monaco TPS, OpenKBP, DVH.

Résumé

la validation clinique d'un modèle de prédiction de dose de type 3D Attention Residual U-Net, entraîné sur la base de données OpenKBP comprenant 340 patients atteints de cancers de la tête et du cou traités par radiothérapie avec modulation d'intensité à une dose de prescription de 70 Gy. Le modèle reçoit en entrée un tenseur tridimensionnel à 12 canaux, combinant les images CT avec les masques binaires du volume cible prévisionnel et des organes à risque, et produit en sortie une distribution de dose prédite voxel par voxel. L'objectif est d'évaluer la validation clinique du modèle entraîné au-delà du domaine des données d'entraˆınement ; un protocole de validation dédié est mis en œuvre à l'aide d'un fantôme de calibration CT, conçu pour reproduire les hétérogénéités tissulaires humaines. Un plan de traitement de référence est généré sur le fantôme à l'aide du système de planification de traitement Monaco, puis comparé à la distribution de dose prédite par le modèle d'apprentissage profond sur la même géométrie de fantôme. L'accord entre les distributions de dose prédite et de référence est évalué à l'aide des métriques d'évaluation. Les résultats montrent que le modèle d'apprentissage profond est capable de reproduire la distribution spatiale globale de la dose ainsi que les gradients de dose observés dans le plan de référence, avec une bonne concordance des courbes DVH et des profils de dose au niveau du PTV et des régions environnantes. Toutefois, des écarts quantitatifs sont identifiés entre les doses prédites et les doses de référence, attribuables au décalage de domaine entre les données d'entraˆınement anatomiquement complexes et la géométrie simplifiée du fantôme, ainsi qu'aux différences de normalisation de dose et d'algorithmes de calcul entre le pipeline d'entraˆınement et le TPS clinique.

Mot-clés: apprentissage profond, prédiction de dose, doses de référence, validation clinique, Attention Residual U-Net, fantome ct, TPS Monaco, Open KBP, DVH.

ملخص الأطروحة

تتناول هذه الدراسة التحقق السريري من نموذج التعلم العميق ثلاثي الأبعاد **(3D Attention Residual U-Net)** للتنبؤ بالجرعة الإشعاعية، والذي تم تدريبه على قاعدة البيانات المرجعي **(OpenKBP)** التي تضم 340 مريضاً مصاباً بسرطان الرأس والعنق، والذين عولجوا باستخدام تقنية العلاج الإشعاعي معدل الكثافة **(IMRT)** بجرعة موصوفة بلغت 70 غراي.(Gy)

يستقبل هذا النموذج كمدخلات موتراً ثلاثي الأبعاد بـ 12 قناة **(12-channel 3D tensor)**يدمج بين صور التصوير المقطعي المحوسب (CT) والأقنعة الثنائية **(Binary masks)** للحجم المستهدف للتخطيط **(PTV)**والأعضاء المعرضة للخطر **(OARs)**، ليقوم بإخراج التوزيع المكاني للجرعات المتنبأ بها بدقة على مستوى كل عنصر حجمي .**(Voxel-wise)**

يكمن الهدف الأساسي للبحث في تقييم مدى قدرة النموذج المدرب على المصادقة والتحقق السريري خارج نطاق التشريحي لبيانات التدريب؛ ولأجل ذلك، تم تطبيق بروتوكول تحقق مخصص باستخدام مجسم معايرة التصوير المقطعي المحوسب **(CT calibration phantom)**، المصمم خصيصاً لمحاكاة عدم تجانس الأنسجة البشرية. وقد تم حساب خطة علاجية مرجعية (Ground-truth) على هذا المجسم باستخدام نظام تخطيط العلاج "موناكو**(Monaco TPS)** "، ومن ثم مقارنتها بتوزيع الجرعات الإشعاعية التي تنبأ بها نموذج التعلم العميق على هندسة المجسم ذاتها، حيث جرى تقييم مدى التطابق والاتفاق بين التوزيعين باستخدام مقاييس تقييمية هندسية وفيزيائية دقيقة.

أظهرت النتائج أن نموذج التعلم العميق قادر على إعادة إنتاج التوزيع المكاني العام للجرعة الإشعاعية وتدرجات الجرعة **(Dose gradients)**الملاحظة في الخطة المرجعية الحقيقية، مع تسجيل توافق وثيق في منحنيات مدرج الجرعة والحجم **(DVH)**ومقاطع الجرعة الطولية **(Dose profiles)**للحجم المستهدف للتخطيط (PTV) والمناطق المحيطة به.

ومع ذلك، تم رصد تباينات واختلافات كمية بين الجرعات المتنبأ بها والجرعات المرجعية الحقيقية؛ وتُعزى هذه الفروقات أساساً إلى ظاهرة "انزياح النطاق **(Domain shift)** "بين بيانات التدريب البشرية المعقدة تشريحياً وهندسة المجسم البلاستيكي المبسطة، بالإضافة إلى الاختلافات الجوهرية في طرق تسوية الجرعة **(Dose normalization)** وخوارزميات الحساب المعتمدة بين نظام تدريب النموذج ونظام تخطيط العلاج (TPS) السريري.

Contents

Acknowledgments i

Abstract ii

Résumé iii

List of Figures vii

List of Tables ix

List of Abbreviations x

Introduction 1

1 FUNDAMENTALS OF EXTERNAL BEAM RADIOTHERAPY 3

1.1 INTRODUCTION 3

1.2 External Beam Radiation Therapy Definition 3

1.3 External Beam Radiatherapy Workflow 4

1.3.1 Virtual Simulation 4

1.3.2 Contouring/Segmentation 5

1.3.3 Treatment Planning 7

1.3.4 Treatment Delivery 8

1.4 External Radiotherapy Techniques 8

1.4.1 Three-Dimensional Conformal Radiotherapy (3D-CRT) 9

1.4.2 Intensity modulated radiation therapy (IMRT) 9

1.4.3 Volumetric modulated arc therapy (VMAT) 11

1.4.4 Image guided radiation therapy (IGRT) 12

1.4.5 Adaptive Radiotherapy (ART) 13

1.5 Conclusion 14

2 Evolution of Dose Calculation: From Conventional TPS to Artificial Intelligence 15

2.1 Introduction 15

2.2 TPS: Definition and Historical Evolution 15
2.3 Dose Calculation Algorithm Definition 16
2.4 Dose Calculation Algorithms 17
2.4.1 Pencil Beam (PB) Algorithm 17
2.4.2 Anisotropic Analytical Algorithm (AAA) 17
2.4.3 Collapsed Cone Convolution (CCC) Algorithm 18
2.4.4 Monte Carlo (MC) Algorithm 18
2.5 Algorithm Comparison: Performance,Advantages and Clinical Evidence 19
2.6 Input Data and Output of Dose Calculation Process 20
2.6.1 Input Data 20
2.6.2 Output 3D Dose Distribution 22
2.7 Advanced Dose Calculation Approach: Artificial Intelligence 22
2.7.1 Artificial Intelligence and Deep Learning 22
2.7.2 Motivation for AI Integration in Radiotherapy 23
2.7.3 Volumetric Dose Prediction via Deep Learning 23
2.7.4 State of the Art in Deep Learning Dose Prediction (2022 2025) 24
2.8 Comparative Assessment: Conventional TPS versus AI-Based Dose Prediction 25
2.9 Problem Statement 25
2.9.1 Conclusion 26

3 MATERIALS & METHODS 27
3.1 Context 27
3.2 Internship Location Description 27
3.3 Material Resource and Clinical Tools 30
3.3.1 Clinical Dataset (Open KBP) 30
3.3.2 3D Attention Residual U-Net Architecture(Attention Res U-Net) 31
3.3.3 Clinical Dose Calculation Engine (Monaco) 32
3.3.4 Validation Phantom 33
3.4 Methodological Work Flow and Validation Protocol 34
3.4.1 3D Attention Res U-Net Implementation 34
3.4.2 TPS Based Dose Calculation 34
3.4.3 Phantom Based Dose Prediction 35
3.5 Evaluation Metrics 36
3.5.1 Mean Absolute Error (MAE) 36
3.5.2 Root Mean Square Error (RMSE) 36
3.5.3 Homogeneity Index (HI) 36
3.5.4 Paddick Conformity Index (CI) 37
3.5.5 Volumetric and DVH Criteria 37
3.5.6 $V_{95\%}$ (Target Prescribed Volume) 39

3.5.7 Dose profile Analysis 39
3.6 Conclusion 40

4 EXPERIMENTS, RESULTS & DISCUSSION **41**
4.1 Introduction 41
4.2 Qualitative Results 41
4.3 Quantitative Results 43
4.3.1 Dose-Volume Histogram Analysis 43
4.3.2 Spatial Dose Profile Analysis 44
4.3.3 Dose Distribution Inside the Body (Voxel Histogram) 45
4.3.4 Discussion and Justification of Quantitative Discrepancies 45
4.4 Conclusion 47

CONCLUSION & PERSPECTIVES **48**

BIBLIOGRAPHY **49**
Bibliography 49