

Amélioration de la robustesse de l'U-Net 3D contre la compression JPEG2000 pour la segmentation des organes pelviens masculins

Karim El Khoury, Martin Fockede, Eliott Brion et Benoît Macq
ICTEAM, UCLouvain, 1348, Louvain-la-Neuve, Belgium

Résumé : La segmentation d'organes est un processus essentiel en imagerie médicale pour la planification et le contrôle des traitements. En cas de grandes déformations, les algorithmes classiques de segmentation basés sur le traitement d'images et les atlas échouent. Dans de telles situations, l'apprentissage profond permet de fournir de meilleures solutions. Cependant, l'entraînement des réseaux d'apprentissage profond nécessite une grande quantité d'images. La disponibilité de ces données d'entraînement nécessite un transfert et un stockage de données importants pour lesquels la compression des images est obligatoire. Cependant, les déformations des données causées par la compression peut influencer les performances de l'apprentissage profond. Dans ce travail, nous proposons d'étudier l'impact de la compression JPEG2000 sur la segmentation U-Net 2D et 3D des organes pelviens masculins. Nous montrons que l'utilisation d'un U-Net 3D finement ajusté permettrait de compresser deux fois plus les scans du patient pour la même performance de segmentation par rapport à un U-Net 2D.

Mots-clés : Segmentation U-Net, JPEG2000, apprentissage profond, compression d'images, imagerie médicale

1 Introduction

Les applications de l'apprentissage profond en imagerie médicale sont diverses et vont de la classification de tissus hépatiques sains [4] à la segmentation multi-organes pour la radiothérapie [1]. Certaines études ont déjà abordé l'impact de la compression d'images sur les tâches d'apprentissage profond. C. Rathgeb et al. [2] observent l'impact de la compression sur la segmentation de la rétine à l'aide d'algorithmes de segmentation traditionnels tels que la transformée de Hough adaptative pondérée, la transformée ellipsopolaire et la transformée de Hough adaptée au contraste. Une étude récente de F. G. Zanjani et al. [5] analyse l'impact de JPEG2000 sur un réseau convolutif profond pour la détection de cancer métastatique dans des images histopathologiques. Leur travail montre qu'un réseau neuronal convolutif entraîné avec des données compressées surpasse le même réseau entraîné avec des données non-compressées pour la classification de données compressées.

2 Travail proposé

Nous proposons d'étudier l'impact de la compression JPEG2000 sur la segmentation U-Net 3D des organes pelviens masculins. Le schéma fonctionnel simplifié de la Figure 1 met en évidence le processus utilisé. Notre travail

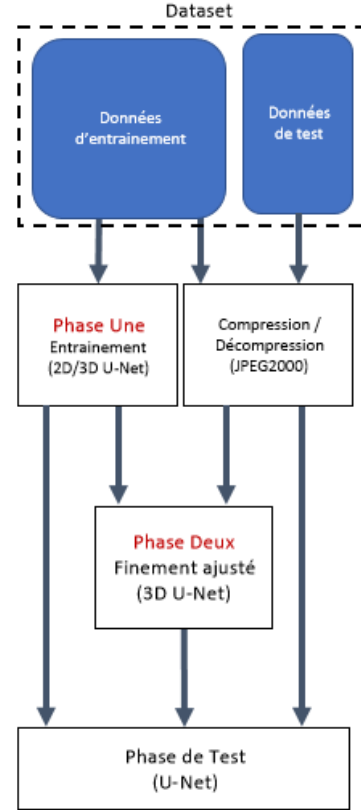


FIGURE 1 – Schéma fonctionnel simplifié du flux de travail proposé, mettant en évidence l'apprentissage du U-Net 3D dans la première phase et son ajustement fin dans la deuxième phase.

se compose de deux phases : la première phase montre la robustesse d'un U-Net 3D par rapport à un U-Net 2D en entraînant les deux réseaux avec des données non-compressées et les testant avec des données compressées. La deuxième phase améliore la première phase en ajustant les poids de l'U-Net 3D avec l'ajout d'images compressées/décompressées à l'ensemble d'entraînement.

Dans la première phase, nous entraînons les réseaux U-Net 2D et 3D sur des images non-compressées. Nous les testons ensuite sur des images compressées/décompressées à différents taux de compression en utilisant JPEG2000, qui est la norme de compression recommandée pour le format d'image DICOM. En ce qui concerne les architectures U-Net 2D et 3D, nous utilisons les modèles développés dans [3] et [1] respectivement. L'architecture U-Net contient de multiples sous-échantillonnages et sur-échantillonnages et est entièrement convolutive car elle ne possède pas de couche entièrement connectée. Cette

approche multi-échelle permet une segmentation précise grâce à la propagation de l'information multi-résolution. Les réseaux 2D et 3D ont le même nombre d'étapes de sous-échantillonnage et de sur-échantillonnage. Pour chaque niveau, le U-Net 2D possède quatre fois plus de canaux que le U-Net 3D. Cette architecture permet aux deux réseaux d'avoir le même niveau de performance pour les images non comprimées, ce qui est nécessaire pour une comparaison équitable. En ce qui concerne la compression JPEG2000, dans le cas des images 2D, nous compressons des tranches de données 3D pour créer notre ensemble de données de test. Dans le cas des images 3D, nous compressons des tranches de données 3D et reconstruisons ensuite la matrice 3D pour créer notre ensemble de données de test. Dans la deuxième phase, nous cherchons à ajuster finement les poids de l'U-Net 3D pour améliorer sa robustesse. Nous partons du réseau U-Net 3D formé lors de la première phase et nous procédons à un deuxième entraînement en utilisant uniquement des images compressées pour affiner les poids. Ce deuxième entraînement spécialise le réseau pour reconnaître les artefacts créés par la compression. Le réseau est affiné en considérant un taux de compression spécifique et est également testé sur ce même taux.

3 Résultats

Pour nos expériences, nous utilisons des CT-scans (CT) et des CBCT-scans (CBCT) de deux organes de la région pelvienne : le rectum et la vessie. Les données nous sont fournies par le CHU-Charleroi Hôpital André Vésale et le CHU-UCL-Namur, les images CT proviennent de 56 patients masculins et consistent en des coupes 2D individuelles mises dans des matrices 3D de taille 192x192x160 et les images CBCT proviennent de 74 patients masculins et proviennent d'un algorithme de reconstruction 3D pour obtenir directement les images 3D. Les données fournies comprennent des masques des organes sur chaque image qui ont été réalisés par des experts médicaux. En raison de problèmes de convergence et de limitations matérielles, la taille d'entrée de l'U-Net 3D a été limitée à des matrices de 160x160x128 et toutes les images ont été redimensionnées en conséquence. Les images d'entrée ont été compressées en utilisant le format JPEG2000 à des rapports allant de 1 :1 à 128 :1. La taille de l'ensemble d'apprentissage utilisé est composé de 105 images (45 CT et 60 CBCT). L'ensemble de test utilisé est composé de 25 images (11 CT et 14 CBCT). Il convient de noter que pour l'U-Net 2D, ces chiffres sont multipliés par 192 en raison du découpage en tranches le long de l'axe vertical. Pour le réglage fin de l'U-Net 3D, les mêmes images d'entrée ont été compressées/décompressées et utilisées pour le deuxième entraînement. Nous utilisons comme mesure d'évaluation le coefficient de similitude de Dice (DSC) défini dans l'équation (1) :

$$DSC = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad (1)$$

Comme fonction de coût, nous avons utilisé un DSC pondéré des deux organes pour l'U-Net 3D et l'entropie croisée catégorique (ECC) pour l'U-Net 2D. Le choix du ECC a été fait parce que dans l'ensemble de données 2D,

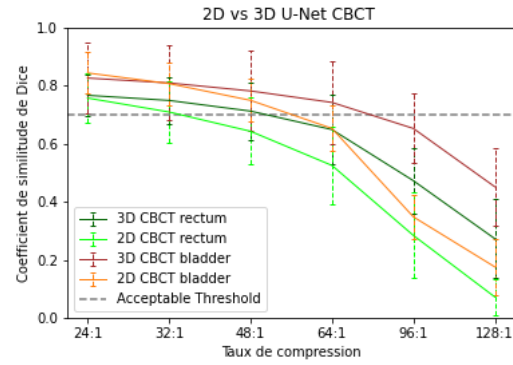


FIGURE 2 – Performances de la segmentation U-Net 2D vs 3D à des taux de compression élevés pour la vessie et le rectum.

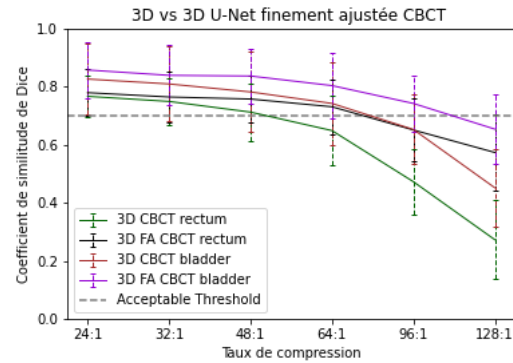


FIGURE 3 – Performance de segmentation de l'U-Net 3D avant et après l'ajustement fin de ses poids.

les deux organes ne sont pas toujours visibles, et le DSC échoue dans ce cas. Les résultats de la comparaison des U-Net 2D et 3D sont présentés dans la Figure 2. Nous pouvons clairement voir que, pour les deux organes, l'U-Net 3D est nettement plus performant que l'U-Net 2D à des taux de compression élevés. Si nous fixons un seuil DSC acceptable de 0,7, pour la vessie, nous pouvons aller jusqu'à un taux de compression de 48 :1 et 64 :1 pour la segmentation U-Net 2D et 3D respectivement. Pour le rectum, nous pouvons aller jusqu'à un taux de compression de 32 :1 et 48 :1 pour la segmentation U-Net 2D et 3D respectivement. Un autre résultat que nous observons est que le rectum est plus difficile à segmenter que la vessie. Cela est principalement dû à la petite taille et à la forme spéciale du rectum. Les résultats après le réglage fin des poids du U-Net 3D sont présentés dans la Figure 3. Nous observons que l'ajustement fin étend la robustesse du U-Net 3D à des taux de compression plus élevés. Si nous gardons le même seuil DSC de 0,7, pour la vessie, nous pouvons aller jusqu'à un taux de compression de 64 :1 et 96 :1 pour la segmentation U-Net 3D et U-Net 3D ajustée respectivement. Pour le rectum, nous pouvons atteindre un taux de compression de 48 :1 et 64 :1 pour la segmentation U-Net 3D et la segmentation U-Net 3D finement ajustée respectivement. Par conséquent, pour un même DSC fixé à 0,7, nous montrons que nous sommes capables de compresser nos données deux fois plus entre l'U-Net 2D et l'U-Net 3D finement ajusté pour la vessie (de 48 :1 à 96 :1) et le rectum (de 32 :1 à 64 :1).

4 Conclusion et perspectives

Dans cet article, nous avons étudié l'impact de la compression JPEG2000 sur la segmentation U-Net 2D et 3D des organes pelviens masculins. Nous avons montré que la segmentation U-Net 3D est plus robuste à la compression que la segmentation U-Net 2D. De plus, nous avons montré qu'après un réglage fin de l'U-Net 3D, nous pouvons passer à un taux de compression deux fois plus élevé que l'U-Net 2D pour une même performance de segmentation. Intuitivement, un processus 3D s'exécutant sur un ensemble d'images 2D englobant un bruit de compression 2D, qui est de faible corrélation d'une tranche à l'autre, compensera naturellement les artefacts avec la troisième dimension. Dans le cadre de nos travaux futurs, nous envisageons d'élaborer la modélisation mathématique de ce phénomène en détail.

Références

- [1] Jean Léger, Eliott Brion, Paul Desbordes, Christophe De Vleeschouwer, John A. Lee, and Benoit Macq. Cross-domain data augmentation for deep-learning-based male pelvic organ segmentation in cone beam ct. *Applied Sciences*, 10(3) :1154, 2020. [doi :10.3390/app10031154].
- [2] Christian Rathgeb, Andreas Uhl, and Peter Wild. Effects of severe image compression on iris segmentation performance. *IEEE International Joint Conference on Biometrics*, 2014.
- [3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net : Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, page 234–241, 2015. [doi :10.1007/978-3-319-24574-4_28].
- [4] K. Steinhofel, C.F. Dewey, D. Janssens, and B. Macq. Classification of compressed dicom liver tissue images. *Proceedings of the IEEE-EMBS Special Topic Conference on Molecular, Cellular and Tissue Engineering*.
- [5] Farhad Ghazvinian Zanjani, Svitlana Zinger, Bastian Piepers, Saeed Mahmoudpour, and Peter Schelkens. Impact of jpeg 2000 compression on deep convolutional neural networks for metastatic cancer detection in histopathological images. *Journal of Medical Imaging*, 6(02) :1, 2019. [doi :10.1117/1.jmi.6.2.027501].