

Segmentation d'images médicales échographiques 3D par deep learning

1- Contexte

La segmentation sémantique d'images médicales 3D consiste à assigner une étiquette à chaque voxel d'un volume d'entrée. L'objectif de cette thèse est de mettre en place des méthodes de segmentation automatique pour les images échographiques 3D (voir Fig. 1). C'est une tâche pour laquelle la mise en place d'outils d'Intelligence Artificielle (IA) revêt des enjeux cruciaux pour l'aide au diagnostic, la planification et le guidage opératoire.



Figure 1 : La thèse consiste à analyser des images échographiques 3D

Co-financement et collaboration. Cette thèse s'inscrit dans le cadre du projet AHEAD (Artificial Intelligence for Health, Physics, Transportation and Defense) porté par le Cnam Paris et financé par l'ANR dans le cadre des contrats doctoraux en Intelligence Artificielle. La thèse est co-financée par l'IRCAD qui est un acteur majeur dans le domaine la chirurgie mini-invasive. Cette thèse s'inscrit dans un projet global de l'IRCAD visant à démocratiser l'usage de l'imagerie échographique dans le cadre du diagnostic automatique et de la chirurgie percutanée augmentée assistés par l'Intelligence Artificielle. L'imagerie échographique est en effet une technique d'imagerie temps-réel, non-invasive et peu onéreuse dont l'utilisation se généralise. Elle permettra de franchir une étape majeure concernant la planification et le guidage intra-opératoire pour la chirurgie percutanée augmentée, qui consiste à positionner des aiguilles avec précision dans des cibles (tumeurs) en prenant soin d'éviter soigneusement les structures fragiles (vascularisation, vésicules biliaires, réseaux biliaire).

Contexte national et international. La segmentation sémantique d'images médicales est largement impactée par le succès actuel de l'IA. Les modèles de deep learning et les réseaux de neurones convolutifs profonds, en particuliers les modèles «Fully Convolutional Networks» [LSD15, CPK+18] et des variantes dédiées aux images médicales comme les U-Net [RFB15] constituent actuellement l'état de l'art pour cette tâche.

2- Objectifs

L'objectif général de la thèse consiste à proposer de nouvelles méthodes pour la segmentation de structures médicales dans les images échographiques, en particulier dans les zones du foie (vésicule biliaire, parenchyme hépatique, vascularisations, réseau biliaire, tumeurs) ou du rein (parenchyme rénal, cortex, tumeurs).

L'application des modèles de deep learning pose des questions lorsqu'on aborde la segmentation de volumes 3D d'images médicales. En effet, les modèles FCN sont rapidement limités par des problématiques d'empreinte mémoire. Par exemple, les images échographiques à traiter dans le cadre de la thèse correspondent à des volumes de l'ordre de 1000x1000x200 voxels, qui ne peuvent pas être analysées naïvement en utilisant des FCN. Les approches classiques pour contrer le problème consistent à effectuer la segmentation en utilisant l'information visuelle uniquement sur chaque coupe 2D ou en effectuant un sous-échantillonnage brutal des images, qui induisent intrinsèquement une perte d'information dans l'analyse globale des données.

L'objectif général de cette thèse consiste à proposer des méthodes de segmentation d'images médicales 3D permettant de prendre en compte explicitement l'aspect 3D pour la segmentation tout en conservant une résolution spatiale importante.

Approches proposées

Axe 1 : Approches de l'état de l'art pour la segmentation sémantique 3D

Le premier axe de la thèse consistera à évaluer des méthodes de références pour la segmentation d'images médicales 3D. Nous explorerons les modèles de référence à base de réseaux FCN 3D, *e.g.* U-Net 3D [LBR+16], V-Net [MNS16]. Nous nous attacherons à calibrer les performances de ces modèles par rapport à des modèles 2D effectuant la segmentation sur chaque coupe, ou des modèles hybrides 2D-3D, *e.g.* [LCQ+18].

Axe 2 : Fusion multi-vues de réseaux 2D de segmentations

Une première piste consistera à effectuer la segmentation d'images dans des coupes 2D prises sous N axes de vues différents, *e.g.* coupes axiales, frontales et sagittales. La motivation de l'approche consiste à effectuer la segmentation 2D de chaque vue en utilisant des FCN avec une très bonne résolution spatiale, et de fusionner les résultats de segmentation de chacun des N réseaux de segmentation pour chaque voxel du volume 3D.

Au niveau de la fusion, nous conditionnerons l'importance de chacune des vues dans la fusion en fonction de l'incertitude décisionnelle des prédictions. Un enjeu clé réside dans la capacité à pouvoir quantifier l'incertitude de manière fiable avec des réseaux de neurones profonds. Nous nous appuyerons sur les travaux récemment menés dans l'équipe académique sur ce point [CTB+19].

Axe 3 : Réseaux réversibles pour la segmentation d'images médicales

Un second axe de la thèse consistera à segmenter directement les volumes 3D de données. Pour surmonter les problèmes d'empreinte mémoire précédemment mentionnés, nous nous appuyerons sur des versions « inversibles » de réseaux convolutifs profonds [GRU+17, JSO18], qui permettent de reconstruire les activations d'une couche à partir des activations de la couche suivante. En particulier, ceci présente l'intérêt de pouvoir recalculer des éléments des tenseurs d'activation « à la volée », et donc de pouvoir traiter des images 3D importantes tout en maintenant une empreinte mémoire maîtrisée.

La thèse consistera d'une part à adapter ces réseaux réversibles à des architectures de segmentation d'images médicales, comme ceci a été récemment proposé [NNK19]. Nous nous intéresserons ici à voir comment l'augmentation des volumes d'entrée peut permettre d'améliorer la qualité de la segmentation, tout en mettant en perspectives ces gains potentiels par rapport au surcoût en termes de temps d'entraînement des réseaux.

Nous nous appuyerons également sur ces réseaux réversibles pour proposer de nouvelles architectures permettant de conserver au maximum l'information spatiales dans l'extraction de représentations, en particulier en étudiant la possibilité de ne pas utiliser d'opérateurs de pooling dans les modèles.

Échéancier

Les premiers mois de la thèse seront consacrés à l'étude bibliographique des différents aspects de l'apprentissage profond pour la segmentation d'images médicales 3D qui sont au cœur du sujet de thèse. L'axe 1 correspondant à la prise en main des outils expérimentaux, des modèles état de l'art sera ensuite abordé. L'axe 2 du programme de recherche sera entamé dès la seconde partie de la première année de thèse. Ce n'est que dans la seconde moitié de la thèse que l'axe 3 sera quant à lui abordé.

3- Candidature

- Master ou école d'ingénieur à dominante informatique ou mathématiques appliquées
- Expérience en machine learning et deep learning, en particulier réseaux convolutifs
- Très bonnes compétences en programmation, avec une expérience sur les bibliothèques de deep learning (tensorflow, pytorch)
- Bonne qualité de synthèse à l'écrit et à l'oral pour la présentation des travaux de recherche. Une expérience d'écriture d'un article serait un plus.

Envoyer un CV et une lettre de motivation à : nicolas.thome@cnam.fr ,
alexandre.hostettler@ircad.fr

Bibliographie

[CPK+18] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(4): 834-848, 2018.

[CTB+19] Charles Corbière, Nicolas Thome, Avner Bar-Hen, Matthieu Cord, Patrick Pérez. Addressing Failure Detection by Learning Model Confidence. *NeurIPS 2019*.

[GRU+17] Gomez, A.N., Ren, M., Urtasun, R., Grosse, R.B. The reversible residual network: Backpropagation without storing activations. In: *Advances in Neural Information Processing Systems 30*, pp. 2214–2224. Curran Associates, Inc. (2017)

[JSO18] Jörn-Henrik Jacobsen, Arnold W.M. Smeulders, Edouard Oyallon. I-RevNet: Deep Invertible Networks, in *ICLR 2018*

[LBR+16] Soeren S. Lienkamp, Thomas Brox, Olaf Ronneberger, Ozgun Cicek, Ahmed Abdulkadir. 3D U-net: Learning dense volumetric segmentation from sparse annotation. *MICCAI*, 2016.

[LCQ+18] Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, Pheng Ann Heng. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes. *IEEE Trans Med Imaging*. 2018 Dec;37(12):2663-2674

[LSD15] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431-3440, 2015.

[MNS16] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565-571, 2016.

[NNK19] Robin Brügger, Christian F. Baumgartner, and Ender Konukoglu. A Partially Reversible U-Net for Memory-Efficient Volumetric Image Segmentation, *MICCAI 2019*.

[RFB15] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of LNCS, pages 234-241. Springer, 2015.