PhD thesis

Uncertainty quantification in automatic delineation for radiation therapy

I. Context

Recent advances in automatic segmentation using deep learning have led to the emergence of solutions for the delineation of structures in radiotherapy (organs at risk and tumors). These solutions help doctors by providing automatic contouring which can be used as a starting point for delineation. They also enable real-time adaptation of treatment by integrating automatic delineation from the patient's positioning image during treatment, leading to truly adaptive radiotherapy, which is one of the most promising evolutions of radiotherapy. At the end of 2023, ICANS acquired a first commercial solution (ART-Plan Annotate from TheraPanacea) [1, 2] allowing the automatic delineation of organs and tumors, aiming to improve the reproducibility of contours and to help doctors. In 2024, the ICANS technical platform will also see the arrival of the latest generation of radiotherapy system (Fig. 1): Ethos v2 + HyperSight (Varian) [3, 4]. This system includes another commercial solution for auto-contouring, that allows direct online adaptive radiotherapy. In this context, it is crucial for doctors and physicists to know the expected performance of different tools to ensure the validity of delineation and the safety of the treated patients. Furthermore, from

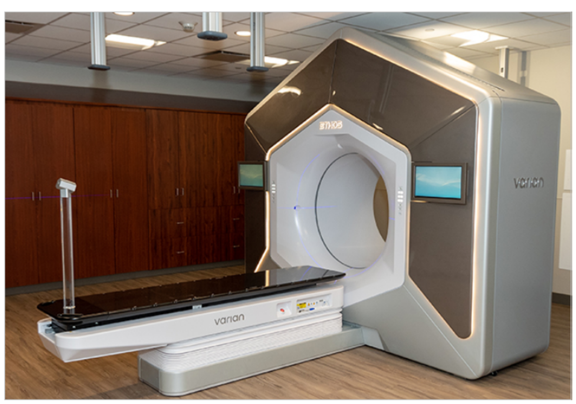

Figure 1: The ETHOS (Varian) medical linear accelerator, that will be installed at ICANS in 2024. This highly sophisticated machine is associated with new AI softwares (including auto-contouring) designed for online adaptive radiotherapy.

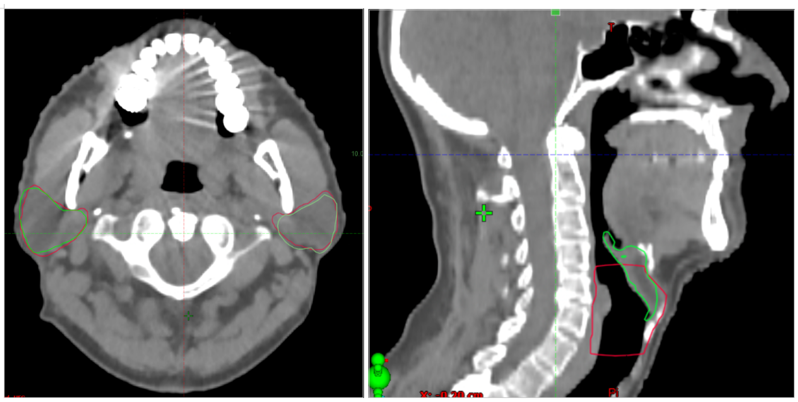

Figure 2: Patient treated for head-and-neck cancer at ICANS, with doctor contours (ground-truth) in green, and automatic contours from ART-Plan (Therapanacea) in red. In this particular case, automatic contours perform well for the parotids, but present a high degree of uncertainty for the larynx. Note that this is only a small sample, given that around 40 organs can be automatically generated by ARTPlan for the head-and-neck region.

a practical standpoint, the ability to predict tool malfunctions would enable the implementation of proactive measures. The objective of the thesis is therefore threefold:

- determine the level of difficulty of automatic segmentation for different tumor locations and different organs at risk.
- determine classes of images or structures for which the algorithm fails to create a segmentation or produces an aberrant segmentation.
- determine the uncertainty associated with the segmentation.

II. Doctors/algorithm comparison [9 months]

In order to identify segmentations subject to error on the part of the delineation software (Fig. 2), the first step will consist in comparing the contours made by doctors to those made by the ART-Plan Annotate algorithm. For this comparison, we can rely on a database of several thousand CT acquisitions including doctors' delineations and covering a large range of the cancers treated. Initially, the comparisons will aim to determine whether differences in performance (through geometrical scores such as but not limited to Dice score) are observed depending on the characteristics of the patients (gender, height, weight, etc.) or the characteristics of the machines used for the production of imagery. The goal will be to identify the organs whose contours must be checked as a priority by doctors. However, this high level understanding of the sources of errors is coarse and cannot provide finer insights into the uncertainty of the delineation.

III. Determination of uncertainties linked to contouring [21 months]

The study will then focus on the possibility of predicting delineation errors from the raw image and from the generated contours. Several approaches have been proposed to develop systems for detecting delineation errors [5, 6, 7, 8, 9] downstream of the segmentation. However, as ART-Plan Annotate is a proprietary software, the possibilities for studying the uncertainties associated with predictions are limited compared to the possibilities offered by an open-source tool. Two types of approaches are possible, an approach studying the model directly or a proxy approach by studying a model trained with similar inputs and outputs (and potentially guided by the predictions of the proprietary model).

The direct study of the uncertainties in a non-open model (unknown architecture and weights) can only rely on modifications of the inputs. In this context, our plan is to create variations of the images by injecting noise or eliminating pixels to study their impact on the produced contours. However, such an approach does not allow the construction of uncertainties whose coverage can be guaranteed.

The indirect study of uncertainties involves the use of a model close to the one used. Although limited information is available about the architecture of the network used for Annotate, a natural choice is the nn-UNet model [10], whose performance in recent international competitions positions it as a strong competitor to the commercial software being studied. By providing access to prediction scores during segmentation, it will be possible to explore the feasibility of using conformal predictions [11, 12] to assess the model's confidence in its predictions.

Working environment

The student will be a member of ICANS and the IMAGeS team (`http://images.icube.unistra.fr/`) in the ICube laboratory in Illkirch. The PhD thesis will start in october 2024.

Supervisors: Sylvain Faisan (`faisan@unistra.fr`), Philippe Meyer (`p.meyer@icans.eu`), Xavier Coubez (`x.coubez@icans.eu`).

Profile of the candidate

- Last year of Master studies in the following fields: computer science, applied mathematics and machine learning.
- Good programming skills (the coding language will be Python).
- Interest for image processing and medical applications.

Application

Send a CV and a short description of your motivation, as well as the transcript of marks for the past 2 years to Sylvain Faisan (`faisan@unistra.fr`), Philippe Meyer (`p.meyer@icans.eu`) and Xavier Coubez (`x.coubez@icans.eu`).

References

[1] Therapanacea. *Annotate*. URL: `https://www.therapanacea.eu/our-products/annotate/`. (accessed: 18.12.2023).

[2] U.S. Food & Drug Administration. *510(k) Premarket Notification (of ART-Plan)*. URL: `https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/pmn.cfm?ID=K230023`. (accessed: 18.12.2023).

[3] Varian. *Ethos*. URL: `https://www.varian.com/products/adaptive-therapy/ethos`. (accessed: 18.12.2023).

[4] U.S. Food & Drug Administration. *510(k) Premarket Notification (of Ethos)*. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/pmn.cfm?ID=K192377. (accessed: 18.12.2023).

[5] Altman et al. "A framework for automated contour quality assurance in radiation therapy including adaptive techniques". In: *Physics in Medicine and Biology* (2015). DOI: 10.1088/0031-9155/60/13/5199.

[6] Hui et al. "Quality assurance tool for organ at risk delineation in radiation therapy using a parametric statistical approach". In: *Medical Physics* (2018). DOI: 10.1002/mp.12835.

[7] Rhee et al. "Automatic detection of contouring errors using convolutional neural networks". In: *Medical Physics* (2019). DOI: 10.1002/mp.13814.

[8] Mody et al. "Improving error detection in deep learning based radiotherapy autocontouring using bayesian uncertainty". In: *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging. UNSURE 2022* (2022). DOI: 10.1007/978-3-031-16749-2_7.

[9] Sandfort et al. "Use of Variational Autoencoders with Unsupervised Learning to Detect Incorrect Organ Segmentations at CT". In: *Radiology: Artificial Intelligence* (2021). DOI: 10.1148/ryai.2021200218.

[10] Isensee et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: *Nature Methods* (2020). DOI: 10.1038/s41592-020-01008-z.

[11] Vovk et al. *Algorithmic Learning in a Random World*. Springer, 2005.

[12] Balasubramanian et al. *Conformal Prediction for Reliable Machine Learning*. Elsevier Inc., 2014. ISBN: 978-0-12-398537-8.