Deep Learning CT segmentation for dosimetry in postoperative endometrial carcinoma treatment

A. Casamitjana, R. Sala-Llonch, R. Tudela, A. Andrés, S. Orio, J.R. Casas, C. Candela-Juan, A. Herreros, A. Niñerola.

[1]Universitat Politècnica de Catalunya, Spain;

[2]Department of Biomedicine, Faculty of Medicine, University of Barcelona, Spain.

[3]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)

[4] Institute of Medical Physics section. The Radiotherapy Oncology Service. Hospital Clínic Barcelona, Barcelona, Spain

[5]Nuclear Medicine Service. Hospital Clínic Barcelona, Barcelona, Spain

1. Introduction

Prophylactic vaginal brachytherapy (VBT) is a common treatment after tumor resection in endometrial cancer. It involves the placement of an applicator inside the vagina that is designed to efficiently adapt to each individual's anatomy and delivers the required radiation dose. In current clinical settings, the specialist delineates several organs from Computed Tomography (CT) images, including the Clinical Target Volume (CTV) which is later used for treatment planning and to optimize the dose distribution in the vaginal cuff for each patient.

The main aim of this study was to assess the viability of automatic Deep Learning (DL) algorithms for segmenting the CTV from pelvic CT images.

2. Methods

We collected a dataset of 220 CT images acquired at the Radiotherapy Department of the Hospital Clínic de Barcelona, together with their corresponding labels obtained manually. The labels included the CTV (Vaginal cuff), and the adjacent organs at risk, including the bladder, rectum, sigma, and bowel. After quality control, 208 images were retained.

We implemented two network architectures using the MONAI framework [1]: V-Net [2] and UNETR [3]. The dataset was divided into *training*, *validation*, and *test* sets (70%, 20%, and 10% of the data respectively). We used Dice loss and Adam algorithms as loss function and optimizer, respectively. We employed an early stopping strategy on the validation set to avoid overfitting. Depending on the experiment, the networks were trained during 1200 or 2400 epochs. The models were evaluated on the test set using geometric and dosimetric metrics. After obtaining the predicted segmentation for each test image, we performed a post-processing clustering to remove possible outliers appearing away from the CTV. We then assessed the quality of the segmentations using the Dice Coefficient (DICE) and the 95% Hausdorff Distance (HD).

We used the VBT planning software (Oncentra Brachy ELEKTA) treatment planning system to compute the absorbed radioactivity dose for the predicted CTV segmentations and compare them against manual delineations. We used 2 dose-volume histogram parameters, namely the $D_{90\%}$ and the D_{2cc}, that measure the minimum dose received by the 90% of the target volume and the dose received by the more exposed 2 cm^3 from the organs at risk, respectively [4].

Finally, we evaluated the effect of data augmentation (DA) strategies during training. These consisted of the application of filters and deformations to the original images.

3. Results

Out of the 208 images, 146 were used for training, 41 for validation, and 21 for testing. We evaluated the 95% HD and the DICE for the V-Net and UNETR first with no DA, using 1200 epochs. Results are summarized in *Table 1.*

	HD (mean, *SD*)	DICE (mean, SD)
V-Net	(8.2, 3.5) mm	(0.78, 0.07)
UNETR	(8.7, 3.3) mm	(0.79, 0.07)

Tabla 1. *Table 1. Results of the main algorithms tested.*

Increasing the number of epochs to 2400 did not increase the performance of the algorithm. Using image filters (IF) was beneficial for V-Net, and spatial transformations (ST) slightly increased the scores for UNETR.

Finally, we compared the volume-dose parameters obtained from manual and automatic segmentations and there were no statistical differences (all Wilcoxon *signed-rank test, $p>0.05$)* between manual and automatic labels, indicating that they might be interchangeable in a clinical setting.

4. Conclusions

We proved the applicability of DL for CTV segmentation in postoperative endometrial carcinoma, using two different model networks, UNETR and V-Net, and with DA. Both nets showed good HD and DICE scores compared to manual labels and no differences were identified in terms of dose-volume histogram metrics.

References

[1] The MONAI Consortium. (2020). Project MONAI. Zenodo. https://doi.org/10.5281/zenodo.4323059

[2] Milletari, F. et al. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. arXiv (Cornell University).

[3] Hatamizadeh, A. et al. (2021). UNETR: Transformers for 3D Medical Image Segmentation. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2103.10504

[4] Rovirosa. A et al. Braquiterapia 3D guiada por la imagen , (EdikaMed S.L., Barcelona 2016, 1st ed.)

Short Abstract (250): Prophylactic vaginal brachytherapy (VBT) is a common treatment after tumor resection in endometrial cancer that requires individual delineation of the Clinical Target Volume (CTV). The main aim of this study was to assess the viability of automatic Deep Learning (DL) algorithms for segmenting the CTV from pelvic CT images. We collected a dataset of 220 CT images, labelled manually.

We implemented and trained V-Net and UNETR networks and we assessed the performance in a cross-validated setting, by dividing the into *training*, *validation*, and *test* sets. We assessed the quality of the segmentations using the Dice Coefficient (DICE) and the 95% Hausdorff Distance (HD) and using dose-volume histogram parameters. We also evaluated data augmentation (DA) strategies.

Both algorithms gave HD values between 8.2-8.7 and DICE of 0.78-0.79. There were no statistical differences in the volume-dose parameters between automatic and manual labels. DA slightly improved the performance of the algorithms. We proved the applicability of DL for CTV segmentation in postoperative endometrial carcinoma, using two different model networks, UNETR and V-Net, and with DA.