3D Dose Predictions and Plan Quality Assessment in MRI Guided Online Plan Adaptation Using Artificial Neural Network Models

Deshan Yang[1], Allan Thomas[2], Yabo Fu[3], H Michael Gach[1], Harold Li[1]

[1]Department of Radiation Oncology, School of Medicine, Washington University in St. Louis

[2]Department of Radiology, MD Anderson Cancer Center

[3]Department of Radiation Oncology, School of Medicine, Emory University

JULY 12–16 2020 VIRTUAL JOINT AAPM | COMP MEETING

EASTERN TIME [GMT-4]

INTRODUCTION

MRIGART (MR Image-Guided Online Adaptive RT) is very useful for treating abdominal cancer by 1) adapting the treatment plan, 2) real-time MRI 2D-CINE imaging, and 3) gating by tracking tumor position on 2D-CINE.

How to evaluate each adapted plan's quality is an unsolved and very important question. The online plan adaptation workflow does not allow sufficient time to manually adjust plan optimization settings and seek optimal plan quality for each individual plan. It is important to quickly test the quality of the adapted plan against the previous knowledge of high-quality, off-line plans.

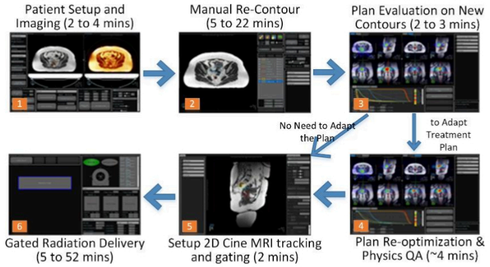

Fig. 1: MRIGART workflow and the average time by each step.

AIM

ANN models were developed to predict 3D dose distributions, enabling the evaluation of online adapted plan quality to better inform adaptive decision-making in MRgRT. Given the contoured structures GTV, PTV and OARs (stomach, duodenum, small bowel, large bowel, spinal cord, liver, kidneys, etc.), to directly predict the 3D dose distribution in the GTV, then compute the DVH metrics V95 and D95.

METHOD

Over 300 treatment plans from 53 abdominal cancer patients undergoing linac-based MRgRT were analyzed. ANN models were developed to predict per voxel dose inside the GTV using input variables related to patient anatomy and target/OAR relationships. The models were designed to be simple (two nodes in a single hidden layer) in order to avoid overfitting. Beam related variables such as beam angles or fluence are not included as input parameters to enable 3D dose prediction using only target and OAR geometry for guiding treatment planning. Five inferior plans identified by the models were manually re-planned to confirm if plan quality could be improved.

METHOD & RESULTS

Rationale and top-level study design considerations:

The shallow ANN model (Fig. 2) is a great choice for plan quality evaluation:

1. It does not require a large amount of training datasets. This is very important to avoid over-fitting the limited training datasets.
2. It is computationally efficient, does not require a GPU, and can be easily integrated into current clinical tools.
3. With the novel data preprocessing procedure, high-level relationships of the OARs and tumor target are effectively presented by the variables input into the ANN models. The overall prediction accuracy is greater than complex deep-learning models (which requires a much larger amount of training datasets).

Key contributions:

1. The idea of using a simple and shallow ANN model for 3D dose prediction.
2. The data pre-processing procedure to extract geometrical information per voxel, and high-dimensional relationships between the OARs and GTV/CTV. See the list of model input parameters in Table 1.
3. Using both per-voxel information and high-dimensional information into ANN model prediction.
4. All input parameters are derived from target and OAR geometries. Beam related variables such as beam angles or fluence are not included. In this way, the ANN prediction model can be used before the actual beam plan is devised and therefore could be useful to guide treatment planning.

Data preparation:

The total 310 plans (50 simulation, 260 adapted) were separated into training and testing groups as follows: the 53 patients were split into ten groups of five and one group of three, then each group was cycled through as the test group. For each iteration of the cross validation, all plans from the test group patients made up the testing data while all the plans from the other ten groups served as training data. The cross validation was done on a patient-by-patient basis instead of a plan-by-plan basis to avoid testing the models on any plan coming from a patient whose plans were already used for model training.

Model training:

Our model training process is shown in Fig. 3. The trained model was applied to detect the inferior plans, which were then removed from the training datasets before the model was retrained and thus refined.

Performance evaluation:

The trained model was applied to predict 3D dose in the tumor target on all plans from the 5 patients that were not used for training in each cross-validation iteration. Dose prediction errors were computed and analyzed. Dosimetric metrics V95, V100 and D95 were computed on the predicted 3D dose distribution and compared to the ground truth values. This sequence was repeated until all 53 patients cycled through the test group and all results were collected.

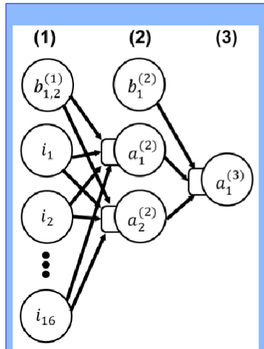

Fig. 2: The designed ANN model, which has only 3 layers and 21 total nodes.

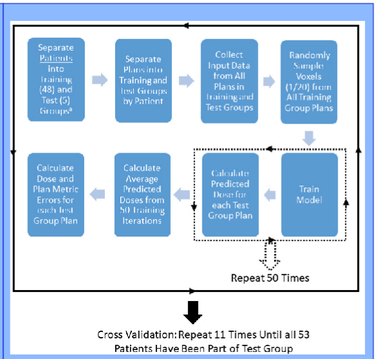

Fig. 3: Process flowchart for ANN model training and testing.

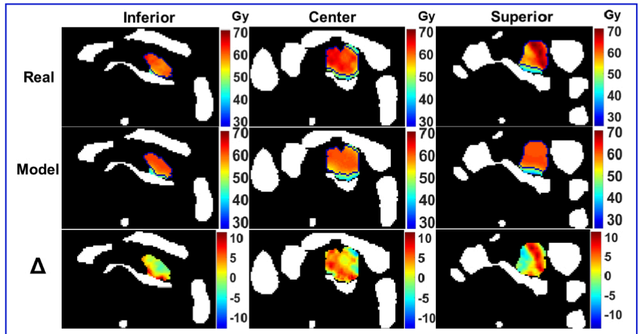

Fig. 4: Three axial slices (inferior, center, and superior of GTV) of planned dose, model predicted dose, and ΔDose of a representative patient plan (SBRT: Rx = 50 Gy, OARCRIT constraint = 36 Gy). All OARs (stomach, duodenum, small bowel, large bowel, and spinal cord) are shown in white. The prescription isodose lines are shown in blue for the planned and predicted dose views. Dose prediction errors = 0.7 ± 3.9 Gy, absolute errors = 3.3 + 2.3 Gy. The plan metrics (planned/predicted) are V95 = 90.4%/88.5%, V100 = 86.8%/85.1%, D95 = 42.6Gy/41.4 Gy, Dmean = 57.2Gy/56.6 Gy.

RESULTS

The dose prediction error and the absolute error were 0.1 ± 3.4 Gy (0.1 ± 6.2%) and 3.5 ± 2.4 Gy (6.4 ± 4.3%), respectively. Plan metric prediction errors were -0.1 ± 1.5%, -0.5 ± 2.1%, -0.9 ± 2.2 Gy, and 0.1 ± 2.7 Gy for V95, V100, D95, and Dmean, respectively. Plan metric prediction absolute errors were 1.1 ± 1.1%, 1.5 ± 1.5%, 1.9 ± 1.4 Gy, and 2.2 ± 1.6 Gy. Approximately 10% of the plans studied were clearly identified by the prediction models as inferior quality plans needing further optimization and refinement. Manual replanning of 5 of such inferior plans increased GTV V95% by 5% on average.

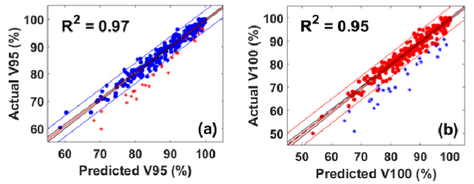

Fig. 5: Results for (a) V95, (b) V100 predictions calculated from the model's 3D dose predictions. The 45° dashed lines in each plot represent where the predicted and actual values are equal. The linear fit line and coefficient of determination (R^2) are included. The outside boundary lines represent the 95% prediction interval. Plans identified as inferior during model training and refinement are labeled with *.

CONCLUSIONS

The developed ANN models can accurately predict 3D dose distributions and overall plan quality metrics for the adapted plans. The models are useful to identify inferior plans and recommend adjustments in planning optimization.

ACKNOWLEDGEMENTS

AHRQ grant R01-HS022888, NIBIB grant R03-EB028427.

REFERENCES

1 **M. Allan Thomas, Yabo Fu, Deshan Yang**. Development and evaluation of machine learning models for voxel dose predictions in online adaptive magnetic resonance guided radiation therapy. *JACMP 2020; Apr.;: 1-10*

2 Bin Cai, **Deshan Yang***, et al, A Practical Implementation of Physics Quality Assurance for Photon Adaptive Radiotherapy, *Zeitschrift fuer Medizinische Physik, Z Med Phys. 2018 Aug;28(3):211-223*

CONTACT INFORMATION

Dr. Deshan Yang. yangdeshan@wustl.edu