The Open Knowledge-Based Planning (OpenKBP) – An 2020 AAPM Grand Challenge: Sharing Experience and Achievement

Jaehee Chun[1], Sanchit Hira[2], Hao Zhang[3], Justin C. Park[4], Hojin Kim[1], Jin Sung Kim[1*]

[1]Department of Radiation Oncology, Yonsei Cancer Center, Yonsei University College of Medicine, Seoul, South Korea

[2]Department of Robotics, Johns Hopkins University, Baltimore, MD 21218

[3]Google Inc.

[4]Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX 75390

Keywords: Dose Prediction, Deep Learning, Radiation Therapy, IMRT, Knowledge-Based Planning

---

Purpose: Knowledge-based planning (KBP) research is flourishing, but many results are reported using institution-specific datasets and evaluation metrics. While many competing approaches have reported positive results, comparing them is difficult without a large open-source dataset, standardized metrics, and a platform to encourage collaboration, sharing, and benchmarking. The aim of the Open Knowledge-Based Planning (OpenKBP) Challenge is to advance fair and consistent comparisons of dose prediction methods for KBP. Our research group participated in this challenge and competed fiercely for three months. We wanted to share the experiences and the results.

Methods: The treatment information of 340 head and neck cancer patients were provided in this challenge, who were treated using nine equispaced coplanar beams with 6 MV step-and-shoot IMRT. The information includes CT images, structures of 7 OARs and PTVs, and 3D dos distributions. The 200 datasets were provided for training, 40 and 100 datasets for test#1 test#2, respectively. Our deep learning strategy is largely divided into two parts: 3D-model and 2.5D-model. First, 3D-model was developed by Yonsei University and is composed of a conditional generative adversarial network (cGAN) framework based on 3D FC-DenseNet. The 2.5D-model was developed by Johns Hopkins University (JHU) and was called 2.5D because the modified 2D-Unet contains nine depth-directional information in channel dimension. We focused not only on developing advanced deep learning models but also on making informative inputs to make sure that deep learning models are well trained. The inputs contain information of 31 channels, including pseudo beamlines, distance map from PTVs, electron density converted from CT image, and radiological depth information. Each model was trained in a 5-fold manner and the final result was derived from an ensemble of 5 models.

Results: The performance of the model was evaluated by the mean absolute error (MAE, unit: Gy) within the region of interest (ROI). The scores of the models developed by Yonsei University and JHU recorded 2.61 and 2.60 respectively, and we could greatly improve the score to 2.417 after sharing each other's knowledge through collaboration at the end of the competition. We were in the second position on the test#1 phase, with the leaders scoring 2.309, and the third-place group scoring 2.499. Given this performance, we expected a second-place

finish, but it was recorded officially as third place due to some mistakes. We have taken second-place after correcting the mistakes. A total of 230 people participated in the competition, and 50 teams submitted their results at least once.

Conclusion: The dense research during the challenge has led to many advances both technically and intellectually. Through the collaboration, we had a great experience of reaching a score that we couldn't reach alone and felt the importance of ensemble again. Only two teams, including our research group, seemed to be Korean participants this year, and we hope there will be more Korean participants next year.