HOKKAIDO UNIVERSITY

Title	Study on Quick Prediction of Dose Volume Statistics in Proton Beam Therapy using Deep Learning [an abstract of dissertation and a summary of dissertation review]
Author(s)	JAMPA-NGERN, SIRA
Degree Grantor	北海道大学
Degree Name	博士(医理工学)
Dissertation Number	甲第15034号
Issue Date	2022-03-24
Doc URL	https://hdl.handle.net/2115/86132
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	doctoral thesis
File Information	Sira_Jampa-ngern_abstract.pdf, 論文内容の要旨

学位論文内容の要旨

（Abstract of thesis）

博士の専攻分野の名称 博士（医理工学）　　　氏 名　シラ　ジャンパングン

（Degree conferred: Doctor of Philosophy）　　（Name of recipient: SIRA JAMPA-NGERN）

学 位 論 文 題 名

（Title of thesis）

Study on Quick Prediction of Dose Volume Statistics in Proton Beam Therapy using Deep Learning

(深層学習を用いた陽子線治療における線量体積統計量の迅速予測に関する研究)

[Background and Objective]

Proton beam therapy (PBT) is effective for cancer treatment due to advantageous physical characteristics which allow minimizing unwanted doses and limit complicating outcomes from radiation exposure. However, the treatment costs of PBT are much higher than with intensity modulated X-ray therapy (IMXT) and stereotactic irradiation (STI), and in many countries, access to PBT systems is very limited. For this reason, the patient selection for PBT is important from the aspect of social healthcare economics. In the patient selection for proton beam therapy (PBT), predictions of the liver D_{mean} with 3-dimensional radiation treatment planning (3DRTP) are time consuming. In evaluations of X-ray therapy (XRT) use, an automatic 3DRTP system is developed using deep learning. However, to develop models for PBT, it will be more difficult to collect a sufficient number of data sets than for XRT studies because of the small number of patients who have been (and are being) treated with PBT, and deep learning prediction models generally require large and tumor-specific databases. With a limited number of patient data sets, we developed a simple dose prediction tool (SDP) using deep learning and a novel contour-based data augmentation (CDA) approach and assessed its usability. In this study, the development is presented in 2 chapters, the first chapter covers the development of a prototype of SDP tool to predict liver D_{mean} for PBT using CNN and contour-based data augmentation using our in-house programs. The second chapter is about improving the accuracy of the SDP tool by applying the architecture of pre-trained models.

Chapter 1: The development of simple dose prediction tool (SDP) tool

[Materials and Methods]

The liver D_{mean} prediction model was trained with 7 actual patients with liver cancer, which were split into 5 patients for training and 2 patients for validation. Data augmentation was performed by artificially embedding 199 contours of virtual clinical target volumes (CTV) into CT images for each patient. The feature to train the model is 2-dimensional profiling (2DP), which was the extracted information of the CTV and liver. The data sets of features are labeled with liver D_{mean} for 6 different treatment plans using two-dimensional calculations assuming all tissue densities as 1.0.

[Results]

The test of the validated model was performed using 10 unlabeled CT data sets of actual patients as test data set. Contouring of CTV and 3 OARs, including the skin surface, liver, and spinal cord, was required as input. The mean relative error and regression coefficient between the planned and predicted D_{mean} were 0.1637 and 0.9455, respectively. The result for the model trained without CDA were compared to investigate whether CDA was effective to improve the prediction model. The MRE between the planned and predicted liver D_{mean}, which was trained without CDA is 0.3211. The comparison shows that there was a statistically significant difference between the MRE of the model

trained with and without CDA. The mean time required for the inference of the liver D_{mean} of the 6 different treatment plans for a patient using the SDP was 4.47±0.09 seconds in 10 actual patients in the test data sets.

[Discussion]

We have generated a model for the prediction of the liver D_{mean} in 6 different treatment plans from the contouring of only the CTV and liver. The time required to predict the D_{mean} by the SDP was short enough to be of use clinically. The time to do the contouring of the CTV and 3OARs will depend on the tool being used in a clinic. The accuracy of our liver D_{mean} prediction model is inferior to those in other publications. However, our approach requires only 2D contours of the liver and CTV as input, which requires fewer resources and is cost-effective to directly predict liver D_{mean} compared to those that need precise 3D volumes and 3D dose distributions as the input.

Chapter 2: Attempts to improve the SDP

[Materials and Methods]

The architecture of the prediction model in the SDP tool was modified to improve the accuracy of prediction results for the liver D_{mean}. The feature to train the prediction model in this chapter was changed to 1) 3-dimension profiling (3DP), and 2) bitmap contour image of CTV and OARs for pre-trained model-based structures of prediction models. The pre-trained model structures used in this study are image classification model structures in the Pytorch package including Alexnet, VGG-net, GoogLeNet, Inception net V3, Resnet, Wide Resnet, Densenet, and Shufflenet V2.

[Results]

In the model trained by 3DP as the input, The MRE between the predicted and the planned liver D_{mean} is 0.4611 and for the β it is 1.1070. There was no improvement by using the 3DP instead of the 2DP in chapter 1. In pre-trained model-based architecture models, the most accurate predicted liver D_{mean} is obtained with the GoogLenet-based architecture. The MRE and β between predicted and planned liver D_{mean} are 0.0797 and 0.9747, respectively. The MRE between the planned and predicted liver D_{mean} is significantly reduced when compare to the 2DP-based models in chapter 1. However, the training time for a GoogLenet-based prediction model is approximate 3.5 times longer than for a model trained by 2DP.

[Discussion]

Compared to the model trained by 2DP in chapter 1, most pre-trained model-based structure of prediction models have better predictive accuracy for the liver D_{mean} prediction, all except the Alexnet-based model. The GoogLenet-based prediction model has the highest accuracy in liver D_{mean} predictions of 10 patients in the test data set. In this study, the GoogLenet-based model can predict liver D_{mean} with the highest accuracy with shorter training time than other pre-trained model-based liver D_{mean} prediction models but requiring longer times than 2DP-based prediction models. The result of the MRE of various pre-trained model-based structures, such as VGGnet, Inception net V3, Resnet, Wide Resnet, Densenet, and Shufflenet V2, is not significantly different. It is possible that the key aspects of improving the accuracy are not the model architecture but other particulars, for example data preprocessing, augmentation techniques, or model hyper-parameter optimization.

[Conclusion]

The SDP tool is cost-effective and usable for clinical gross estimates of liver D_{mean}. A physician only needs to do contouring before using the SDP. In the future, the accuracy of the SDP should be improved further if we need the accuracy of liver D_{mean} to be compatible with 3DRTP.