JULY 12–16 2020 VIRTUAL
JOINT AAPM | COMP MEETING
EASTERN TIME [GMT-4]

A Deep Learning Model to Predict Dose-Volume Histogram of Organs at Risk in Radiotherapy Treatment Plans

Z. Liu [1], X. Chen [1], K. Men [1], J. Yi [1] and J. Dai [1]

1Cancer Hospital, Chinese Academy of Medical Sciences, Beijing, China

INTRODUCTION

There are deep learning models[1-8] including our previous studies[2,8] can predict three-dimensional dose distributions, but none of them can directly predict clinically relevant DVHs for OARs. In this work, we would like to propose a deep learning method to directly predict achievable DVHs for OARs for automation of inverse planning. Based on what we known, this is for the first time to demonstrate a deep learning model to directly predict DVHs for OARs. The model was trained and evaluated with treatment plans of nasopharyngeal cancer cases. The deep learning architecture used is similar to our previous study[8] but with different feature maps in each convolution layers. Compared with UNet, this architecture applies a residual network to perform down-sampling, each down-sampled part then being combined with deconvolution to perform up-sampling. Here, we firstly propose a concept of dose–area histogram as a bridge between a DVH and the deep learning network.

AIM

To develop a deep learning-based model to predict achievable dose–volume histograms (DVHs) of organs at risk (OARs) for automation of inverse planning.

METHOD

- 230 nasopharyngeal cancer patients (170 for training, 60 for testing) were included.
- The model was based on a connected residual deconvolution network (CResDevNet) and compared with UNet as a baseline.
- The input data comprised four channels for each OAR, and the output data is dose-area histogram (DAH) for that OAR:

$$DAH(D) = 1 - \frac{1}{A}\int_0^{D_{max}} \frac{dA(D)}{dD} dD$$

- The model was trained from scratch by correlating anatomical features with DAH of OARs, then accumulating these histograms to obtain DVH for each OAR.
- The DVHs and dose–volume indices (DVIs) for each OAR in the testing dataset were predicted to evaluate the accuracy of the models. Dice similarity coefficients for areas of the DVHs were also evaluated.

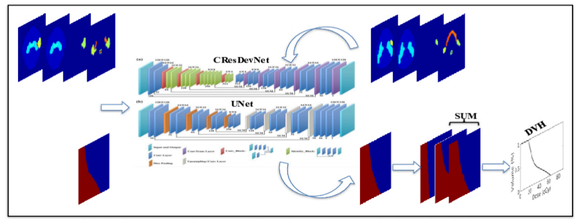

Fig. 1. Overview of the workflow for predicting the DVH for OAR.

RESULTS

- DVHs of 21 OARs in nasopharyngeal cancer were predicted using the models as shown in Fig. 2.
- For each patient, 63 DVIs for all OARs were calculated. Using the 60 patient treatment plans in the testing dataset, 79% and 73% of the DVIs predicted using the CResDevNet and UNet models, respectively, were within 5% of the clinical values. (Table 1)
- The median value of the DVIs' mean absolute errors was 3.2±2.5% and 3.7±2.9% for the CResDevNet and UNet models, respectively. (Table 2)
- The average dice similarity coefficient (DSC) for all OARs was 0.965 using the CResDevNet model and 0.958 using the UNet model. (Table 3)

Table 1. The proportion of DVIs having predicted values within 5% and 10% error.

	Within 5% error (%)		Within 10% error (%)	
	CResDevNet	UNet	CResDevNet	UNet
Brain Stem PRV	93	87	100	100
Brain Stem	86	88	99	98
Spinal Cord PRV	94	94	97	97
Spinal Cord	96	94	99	99
Parotid L	88	67	99	95
Parotid R	82	77	97	97
Thyroid Gland	81	73	96	94
Lens L	98	98	98	98
Lens R	98	100	100	100
Optic Nerve L	58	52	89	78
Optic Nerve R	53	52	86	77
Optic Chiasm	52	40	78	73
Larynx	79	78	95	95
Trachea	87	70	97	95
Mandible L	87	83	98	98
Mandible R	80	76	95	91
TMJ L	58	55	81	79
TMJ R	68	64	87	84
Temporal Lobe L	79	74	100	97
Temporal Lobe R	79	73	97	97
Pituitary	68	42	92	75

Table 3. The DSC of DVHs predicted using CResDevNet and UNet models.

	DSC		p-value
	CResDevNet	UNet	
Brain Stem PRV	0.970±0.018	0.972±0.017	0.272
Brain Stem	0.963±0.025	0.964±0.028	0.818
Spinal Cord PRV	0.964±0.026	0.961±0.027	0.379
Spinal Cord	0.975±0.025	0.972±0.028	**0.005**
Parotid L	0.971±0.018	0.956±0.026	**0.000**
Parotid R	0.970±0.017	0.962±0.021	**0.009**
Thyroid Gland	0.982±0.011	0.961±0.020	**0.000**
Lens L	0.931±0.074	0.917±0.077	**0.000**
Lens R	0.926±0.050	0.919±0.041	0.225
Optic Nerve L	0.962±0.037	0.950±0.043	**0.004**
Optic Nerve R	0.959±0.027	0.950±0.031	0.075
Optic Chiasm	0.937±0.049	0.935±0.042	0.657
Larynx	0.971±0.018	0.966±0.027	**0.033**
Trachea	0.978±0.012	0.974±0.014	**0.032**
Mandible L	0.984±0.010	0.980±0.012	**0.001**
Mandible R	0.979±0.019	0.976±0.021	**0.043**
TMJ L	0.970±0.016	0.966±0.018	0.222
TMJ R	0.976±0.012	0.972±0.016	**0.041**
Temporal Lobe L	0.965±0.019	0.958±0.026	**0.026**
Temporal Lobe R	0.970±0.020	0.960±0.022	**0.001**
Pituitary	0.968±0.031	0.952±0.035	**0.001**

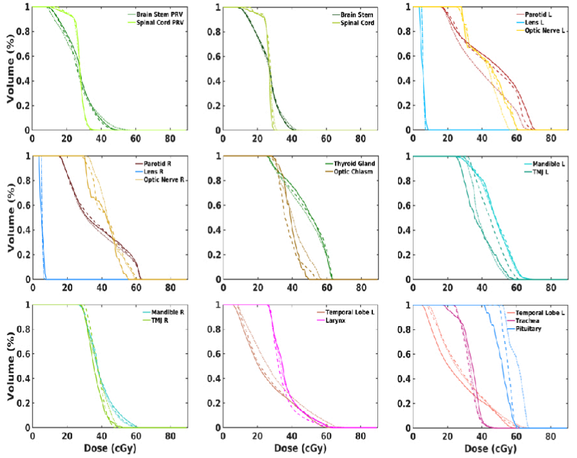

Fig. 2. Example of the "best" similarity between clinical (solid line), CResDevNet (dashed line), and UNet (dotted line) predictions of DVHs for each OAR.

Table 2. Mean absolute error (MAE) of DVIs predicted using CResDevNet and UNet models, together with their statistical comparison for each OAR.

	DVIs	MAE (%)		p-value
		CResDevNet	UNet	
Larynx	V40 (%)	4.5±4.8	4.8±5.6	0.469
	D2%	2.9±2.5	3.4±2.7	0.173
Trachea	Dmean	2.5±1.8	2.5±2.0	0.965
	V40 (%)	2.9±3.4	4.8±5.0	**0.001**
	D2%	3.2±2.7	4.3±3.5	**0.016**
Mandible L	Dmean	1.8±1.5	2.1±1.8	0.223
	V40 (%)	4.1±3.2	3.9±3.2	0.604
	V50 (%)	3.2±2.4	3.8±2.5	**0.032**
	D2%	1.9±1.7	2.7±2.1	**0.003**
Mandible R	Dmean	1.1±1.1	1.4±1.2	**0.022**
	V40 (%)	4.4±3.7	5.1±4.6	0.069
	V50 (%)	4.7±4.6	5.3±5.2	0.089
	D2%	1.5±1.8	2.4±2.6	**0.000**
TMJ L	Dmean	1.4±1.3	1.8±1.5	**0.009**
	V40 (%)	11.0±8.1	11.7±8.7	0.520
	V50 (%)	6.9±6.1	7.4±7.1	0.520
	D2%	3.4±2.6	3.3±3.0	0.836
TMJ R	Dmean	3.2±2.1	3.4±2.5	0.499
	V40 (%)	9.0±7.8	9.2±7.5	0.843
	V50 (%)	5.1±5.1	5.5±5.3	0.541
	D2%	2.9±2.3	3.3±3.0	0.401
Temporal Lobe L	Dmean	2.4±1.6	2.6±2.1	0.581
	V40 (%)	4.1±2.7	4.5±3.8	0.214
	V50 (%)	3.6±2.5	4.0±3.3	0.169
	D2%	3.3±2.5	3.7±2.9	0.255
Temporal Lobe R	Dmean	2.0±1.9	2.6±2.3	**0.005**
	V40 (%)	3.2±2.4	3.8±2.5	**0.038**
	V50 (%)	2.6±2.0	3.9±2.9	**0.000**
	D2%	4.0±3.1	3.9±2.7	0.708
Pituitary	Dmean	2.3±1.7	2.6±1.7	0.196
	D2%	4.4±3.8	7.1±4.9	**0.000**
	Dmean	4.5±5.0	7.1±5.4	**0.001**

	DVIs	MAE (%)		p-value
		CResDevNet	UNet	
Brain Stem PRV	D2%	2.8±2.2	3.3±2.2	0.196
	Dmean	1.5±1.2	1.7±1.5	0.366
Brain Stem	D2%	3.3±3.0	3.5±2.7	0.551
	Dmean	1.8±1.7	1.7±1.7	0.756
Spinal Cord PRV	D2%	3.2±4.4	2.6±4.1	0.162
	Dmean	1.1±1.0	1.1±0.9	0.982
Spinal Cord	D2%	1.8±2.0	1.9±2.3	0.278
	Dmean	1.2±1.0	1.2±1.2	0.912
Parotid L	V20 (%)	2.8±2.2	3.5±2.5	**0.043**
	V30 (%)	3.4±2.6	5.3±3.8	**0.001**
	V40 (%)	3.0±2.3	4.9±3.5	**0.000**
	V50 (%)	3.0±2.0	4.7±3.4	**0.002**
	Dmean	1.5±1.1	2.8±1.9	**0.000**
Parotid R	V20 (%)	3.3±2.7	2.9±2.2	0.111
	V30 (%)	4.2±3.6	3.8±3.7	0.373
	V40 (%)	3.6±3.0	3.8±3.0	0.641
	V50 (%)	3.0±2.3	4.5±2.9	**0.004**
	Dmean	1.9±1.4	2.2±1.6	0.234
Thyroid Gland	V40 (%)	5.1±4.3	6.2±4.8	0.113
	D2%	1.0±1.2	1.2±1.3	**0.004**
	Dmean	2.1±1.6	3.2±2.2	**0.001**
Lens L	D2%	1.5±2.7	1.7±3.0	**0.010**
	Dmean	1.1±1.7	1.0±1.8	0.368
Lens R	D2%	1.1±1.1	1.3±1.9	0.062
	Dmean	0.8±0.9	0.9±0.6	0.745
Optic Nerve L	D2%	4.8±3.9	7.2±5.9	**0.004**
	Dmean	4.9±4.2	5.5±4.9	0.315
Optic Nerve R	D2%	5.4±3.8	6.7±5.0	0.117
	Dmean	5.0±3.9	5.6±4.4	0.366
Optic Chiasm	D2%	5.9±5.5	6.8±4.6	0.233
	Dmean	6.8±5.1	7.8±5.7	0.149

CONCLUSIONS

A deep learning model was built to directly predict achievable DVHs of OARs, taking a treatment site of nasopharyngeal cancer as an example. The model can automatically extract anatomical features of PTVs, OARs, and their geometric relationship. The predicted DVHs of OARs can be used for automation of inverse planning and quality assessment individual treatment plans.

REFERENCES

1. McIntosh C, Welch M, McNiven A, Jaffray DA, Purdie TG. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. *Phys Med Biol*. 2017;62(15):5926-5944.
2. Chen X, Men K, Li Y, Yi J, Dai J. A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning. *Med Phys*. 2019;46(1):56-64.
3. Fan J, Wang J, Chen Z, Hu C, Zhang Z, Hu W. Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. *Med Phys*. 2019;46(1):370-381.
4. Kearney V, Chan JW, Haaf S, Descovich M, Solberg TD. DoseNet: a volumetric dose prediction algorithm using 3D fully-convolutional neural networks. *Phys Med Biol*. 2018;63(23):235022.
5. Nguyen D, Long T, Jia X, et al. A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. *Sci Rep*. 2019;9(1):1076.
6. Nguyen D, Jia X, Sher D, et al. 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. *Phys Med Biol*. 2019;64(6):065020.
7. Barragán-Montero AM, Nguyen D, Lu W, et al. Three-Dimensional Dose Prediction for Lung IMRT Patients with Deep Neural Networks: Robust Learning from Heterogeneous Beam Configurations. *Med Phys*. 2019;46(8):mp.13597.
8. Liu Z, Fan J, Li M, et al. A deep learning method for prediction of three - dimensional dose distribution of helical tomotherapy. *Med Phys*. 2019;46(5):1972-1983.

ACKNOWLEDGEMENTS

This work was supported by the Beijing Natural Science Foundation (7204295), the National Natural Science Foundation of China [11905295, 11875320, 11975313], the Beijing Hope Run Special Fund of Cancer Foundation of China (LC2018B07, LC2018A14), and the Beijing Municipal Science & Technology Commission (Z181100001918002).

CONTACT INFORMATION

The authors have no conflicts of interest to disclose.