



Original paper

DR-only Carbon-ion radiotherapy treatment planning via deep learning



Xinyang Zhang ^{a,b,c,d}, Pengbo He ^{a,b,c,d}, Yazhou Li ^{a,b,c,d,e}, Xinguo Liu ^{a,b,c,d},
 Yuanyuan Ma ^{a,b,c,d}, Guosheng Shen ^{a,b,c,d}, Zhongying Dai ^{a,b,c,d}, Hui Zhang ^{a,b,c,d},
 Weiqiang Chen ^{a,b,c,d,f}, Qiang Li ^{a,b,c,d,f,*}

^a Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China^b Key Laboratory of Heavy Ion Radiation Biology and Medicine of Chinese Academy of Sciences, Lanzhou 730000, China^c Key Laboratory of Basic Research on Heavy Ion Radiation Application in Medicine, Gansu Province, Lanzhou 730000, China^d University of Chinese Academy of Sciences, Beijing 100049, China^e Gansu Provincial Hospital, Lanzhou 730000, China^f Putian Lanhai Nuclear Medicine Research Center, Putian 351152, China

ARTICLE INFO

ABSTRACT

Keywords:
 Deep learning
 Synthetic DRR
 Predicted CT
 Gamma pass-rate

Purpose: To evaluate the feasibility of patient-specific digital radiography (DR)-only treatment planning for carbon ion radiotherapy in anthropomorphic thorax-and-abdomen phantom and head-and-neck patients.

Methods: The study was conducted on the anthropomorphic phantom and head-and-neck patients. We collected computed tomography (CT) and DR images of the phantom and cone beam CT (CBCT) and DR images of the patients, respectively. Two different deep neural networks were established to correlate the relationships between DR and digitally reconstructed radiograph (DRR) images, as well as DRR and CT images. The similarity between CT and predicted CT images was evaluated by computing the mean absolute error (MAE), root mean square error (RMSE), peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), respectively. Dose calculations on the predicted CT images were compared against the true CT-based dose distributions for carbon-ion radiotherapy treatment planning with intensity-modulated pencil-beam spot scanning. Relative dose differences in the target volumes and organ-at-risks were computed and three-dimensional gamma analyses (3 mm, 3%) were performed.

Results: The average MAE, RMSE, PSNR and SSIM of the framework were 0.007, 0.144, 37.496 and 0.973, respectively. The average relative dose differences between the predicted CT- and CT-based dose distributions at the same carbon-ion irradiation settings for the phantom and the patients were <2% and ≤4%, respectively. The average gamma pass-rates were >98% for the predicted CT-based versus CT-based carbon ion plans of the phantom and the patients.

Conclusion: We have demonstrated the feasibility of a patient-specific DR-only treatment planning workflow for heavy ion radiotherapy by using deep learning approach.

1. Introduction

The increasing dose with penetration makes energetic ions such as protons and heavy ions an ideal option for the radiotherapy of many kinds of solid tumor [1]. Carbon ion beam as one of the heavy ion radiations is commonly used for tumor therapy because it is considered to have the most balanced physical and biological properties [2]. Image-guidance devices are needed to deliver a highly precise dose distribution in carbon ion radiotherapy [3]. Among all the possible image-guidance devices, X-ray digital radiography (DR) is usually used for

patient positioning in carbon-ion radiotherapy to avoid delivering overdosage to the organ at risk (OAR) and the normal tissues surrounding the target volume due to the position uncertainty of the target volume [4]. In recent years, cone beam computed tomography (CBCT) has also been adopted in patient positioning for carbon-ion radiotherapy because it can provide more information of anatomical structures to improve the accuracy of patient positioning [5].

However, although positioning images of patients with image-guided radiotherapy already acquired before treatment, the dose distributions over the tumor target volume in subsequent fractions are still deformed

* Corresponding author at: Institute of Modern Physics, Chinese Academy of Sciences, 509 Nanchang Road, Lanzhou 730000, Gansu Province, China.
 E-mail address: liqiang@impcas.ac.cn (Q. Li).

possibly by changes in anatomy of the patient as a result of the treatment [6]. Thus, carbon-ion adaptive radiotherapy needs to be established to solve the problem of dose distribution deformation caused by changes in anatomy. The conventional approach is to replan the carbon-ion radiotherapy by repeatedly acquiring the planning computed tomography (CT) of the patient, which will increase the number of CT scans and associated patient discomfort, resulting in more radiation dose and treatment cost [7,8].

To alleviate the problems of conventional CT imaging in radiotherapy and other applications, reconstruction methods based on sparse sampling have been extensively investigated using techniques such as compressed sensing [9–12] and maximum a posteriori [13,14]. However, these techniques do not achieve better sparsity and lower patient irradiation without sacrificing image quality. In recent years, artificial intelligence (AI) technology has developed widely among different disciplines [15–18], especially biomedicine [19,20]. Different deep learning methods to generate CT images from magnetic resonance images (MRI) have been developed [21–23]. However, the relatively long imaging time and high cost of MRI make it difficult to deploy in developing countries. In the last few years, deep learning-based approaches have been also proposed to reconstruct CT images using different numbers of digitally reconstructed radiograph (DRR) images [7,24,25]. Due to the low resolution of reconstructed CT images (128*128) and the difference between DRR and DR images [26], these methods are difficult to be applied in clinical radiotherapy. Recently, Montoya et al. reconstructed CT models of patients based on deep learning using two scout views and performed some dose calculations [27]. However, it is prohibitively expensive for most organizations to develop models using a large number of one-to-one scout views and CT images. Moreover, since the scout view and CT images are not acquired at the same time, these images do not correspond actually. Therefore, there are still some problems to be solved in developing models using scout view and CT image pairs.

In this study, two different deep neural networks based on deep learning method were constructed to generate predicted CT images from a single DR image. The aim is to investigate the feasibility of the patient-specific DR-only treatment planning workflow for carbon ion radiotherapy. Dose calculations on the predicted CT images were compared against true CT-based dose distributions for carbon-ion radiotherapy treatment planning with intensity-modulated pencil-beam spot scanning.

2. Materials and methods

2.1. Data

2.1.1. Dataset

This study was carried out with approval from the institutional review board. The requirement for informed consent was waived for this retrospective study. The workflow in this study was developed using an anthropomorphic thorax-and-abdomen phantom and two head-and-neck patients. In the first experiment, the image data of the phantom were used to construct the proposed workflow, including a three-dimensional (3D) CT image dataset (Philips Bello system, 512 × 512 × 100 pixels, 0.67 × 0.67 mm² in-plane resolution, 3 mm slice thickness) and 210 DR images (Rayer Medical Technology Image-guided Positioning system, 1024 × 1024 pixels, 0.40 × 0.40 mm² resolution, 100 kV-64 mA-100 ms). A set of deformation vector fields extracted from a four-dimensional (4D) CT dataset of a thorax-and-abdomen patient were applied to the 3D CT images to synthesize a 4D CT dataset with ten phases of the phantom (The maximum displacement of the diaphragm was 24 mm). The first six phases, as well as the phases seven and eight, were used to generate DRR-CT pairs for training and validating a deep neural network model which will be given below, respectively. DRR-CT samples for testing were generated independently using the last two phases of the 4D CT dataset. With translation, rotation

and deformation applied to the CT images of the different phases, we prepared 1200 DRR-CT samples for the model training and 400 for the validation. The trained model was tested using a completely independent sample of 400 DRR-CT samples. DRR images were obtained under the same imaging parameters as DR images by means of the Plastimatch tool [28], which generates simulated corresponding 2D projection data from 3D CT serial images. To generate CT images directly from clinical DR images, we additionally trained and tested a DR-DRR network, which will be also given below, using paired training dataset (200 samples) and testing dataset (10 samples). Each sample consisted of a DR image of the phantom and a DRR image collected from the dataset of the DRR-CT network.

In the second experiment, the image data were collected from two head-and-neck patients to evaluate the clinical performance of the proposed workflow. In order to obtain sufficient DR images to evaluate the workflow method of the previous experiment, two CBCT datasets of the same patient immobilized with a thermoplastic fixation mask were acquired during patient positioning for image-guided radiotherapy in two different fractions. The data of each patient included two sets of CBCT images (Varian Truebeam, full-fan mode, 512 × 512 × 93 pixels, 0.51 × 0.51 mm² resolution, 100 kV-20 mA-20 ms) and 500 DR images (Varian Truebeam, full-fan mode, 1024 × 768 pixels, 0.39 × 0.39 mm² resolution, 100 kV-20 mA-20 ms) were used to reconstruct the CBCT datasets. As shown in Fig. 1, the interval between the two CBCT acquisitions was one week. When building patient-specific models, the datasets should cover as much as possible all clinical scenarios for a given patient's anatomical variation to avoid overfitting [29]. To the extent clinically permissible, the same data augmentation was performed on the CBCT datasets. The deformation vector fields were obtained by co-registering other patients' CBCTs at different fractions. Therefore, although only 2 human subjects were investigated, the datasets used to train and test the deep learning model were abundant for each patient. The first CBCT dataset was used to generate the training (2000 samples) and validation (500 samples) datasets, while the second CBCT dataset was employed to generate the testing dataset (200 samples). Each sample in the dataset consisted of a set of CBCT images and its corresponding DRR image generated by means of the Plastimatch tool. Consistent with the above experiment, we still adopted paired DR-DRR datasets to train (475 samples) and test (25 samples) the DR-DRR network.

2.1.2. Data pre-processing

Data were pre-processed before inputting them into the networks. First, the resolution of DR and DRR images was resized to 512 × 512, corresponding to the size of 3D CT and CBCT images. Moreover, all data samples were normalized, where the pixel or voxel intensity was normalized to the interval [0,1]. In addition, the statistical distribution of the pixel-level intensity values of the DRR images input into the DRR-CT network was normalized to be closer to a standard Gaussian distribution.

2.2. Problem formulation

The purpose of image reconstruction in this study was to predict 3D CT images from a single DR image, which is a cross-dimensional reconstruction task. Suppose $\{(X_{DRi}, X_{DRRi})\}_{i=1}^N$ and $\{(X_{DRRi}, Y_i)\}_{i=1}^N$ denote the training datasets of N DR-DRR and DRR-CT image pairs. Given a DR image X_{DRi} , the goal of the framework is to learn two mapping functions $F_1(\cdot; \theta)$ and $F_2(\cdot; \delta)$ with parameters θ and δ from the training datasets $\{(X_{DRi}, X_{DRRi})\}_{i=1}^N$ and $\{(X_{DRRi}, Y_i)\}_{i=1}^N$, which takes the DR image X_{DRi} as the input, and outputs predicted 3D CT images $\hat{Y} = F_2(F_1(X_{DRi}; \theta); \delta)$.

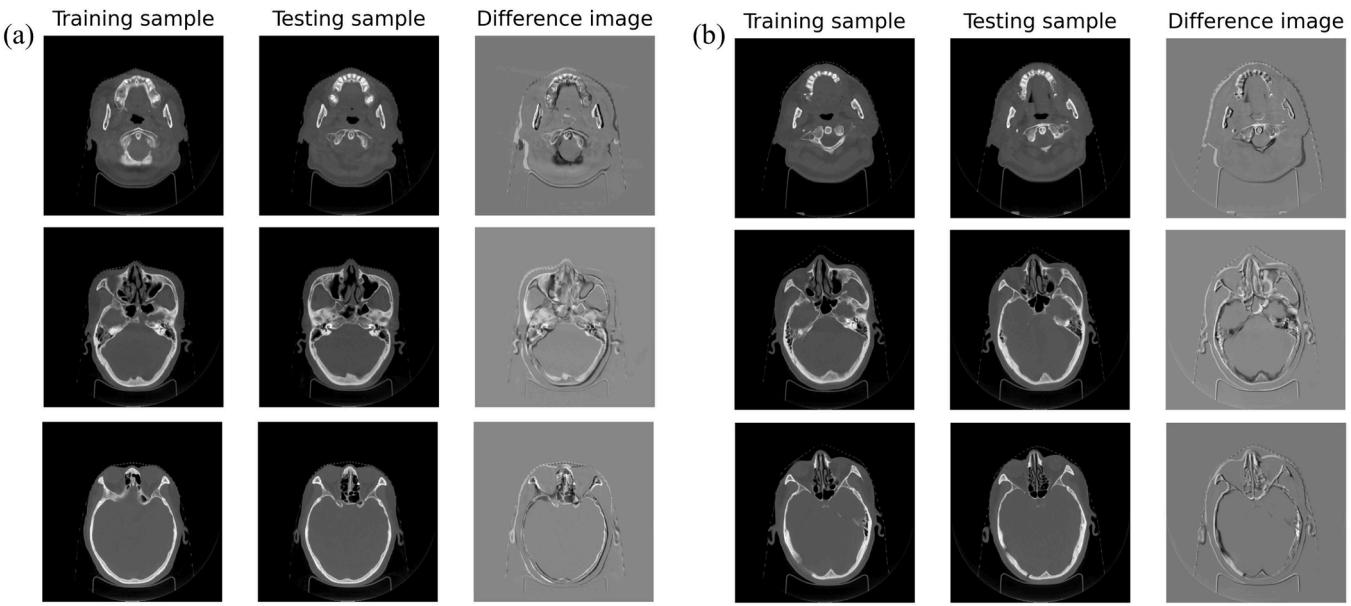


Fig. 1. The transverse views and difference images of the two fractions from the CT images of the patient 1 (a) and patient 2 (b).

2.3. DR-CT framework

2.3.1. DR-DRR network

Previous studies have used CT and DRR images to develop network models due to the difficulty of collecting sufficient DR and corresponding CT image datasets. However, although DRR image is very

similar to DR image, there is still a gap between DRR and DR images. As shown in Fig. 2, here we used an improved cycle-consistent adversarial network (cycle-GAN) that integrates different numbers of residual blocks to learn the mapping between DR and DRR images. Traditional adversarial network (GAN) [30] consists of a generator and a discriminator, which works in opposition to learn specific mappings from two

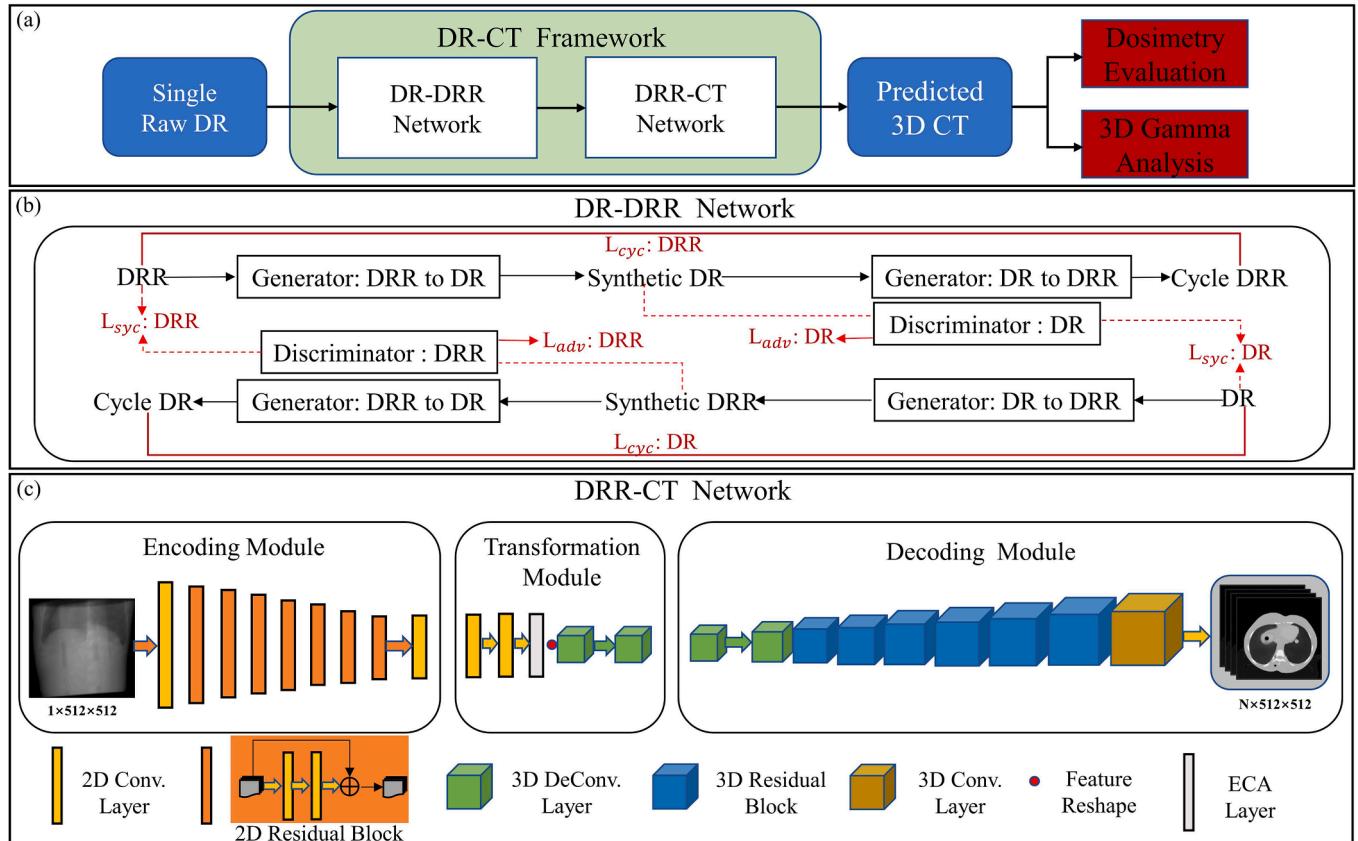


Fig. 2. Schematic flow chart of the proposed method. (a) The DR-only treatment planning workflow in carbon-ion radiotherapy. (b) The architecture of the DR-DRR network. (c) The architecture of the DRR-CT network. L_{cyc} : cycle-consistency loss, L_{syc} : synthetic-consistency loss, L_{adv} : adversarial loss, Conv: convolutional, DeConv: deconvolutional, ECA: efficient channel attention.

datasets. The cycle-GAN network [31] learns the mutual mapping from two datasets, then the bidirectional learning constrains the model while improving the accuracy of the generated images. The generator of the improved cycle-GAN in this study consisted of two down-sampled convolution layers, nine residual blocks, two deconvolution layers and a tanh layer in sequence, thereby learning flexible translations to achieve DR image → Synthetic DRR (sDRR) image and sDRR image → Cycle DR image or vice versa. Similar to the discriminator structure of the original cycle-GAN, the discriminator in this study was to evaluate the authenticity of the generated images and reduce the judgment error of the discriminator network. The loss function was also consistent with that in the original cycle-GAN network.

2.3.2. DRR-CT network

Our purpose was to learn the mapping of 2D to 3D images. To this end, a DRR-CT deep neural network was constructed to generate predicted CT serial images from a single DRR image. The network consisted of an encoding module, a transformation module and a decoding module, as shown in Fig. 2. The encoding network was used to extract 2D features from 2D DRR images and composed of seven 2D convolutional residual blocks [32] with different numbers of convolutional filters to convert the feature map size from $32 \times 512 \times 512$ to $2048 \times 4 \times 4$. To bridge the encoding and decoding networks, the following transformation module composed of 2D convolution layers, efficient channel attention (ECA) layers [33] and 3D deconvolution layers were deployed. This module reshaped the extracted feature maps across dimensions from $2048 \times 4 \times 4$ to $1024 \times 2 \times 4 \times 4$. The role of the ECA layer was to effectively capture cross-channel interaction information while avoiding dimensionality reduction. The decoding network then generated the predicted CT images from these 3D feature maps. It was built upon several 3D deconvolutional residual blocks, which transformed the feature map from $1024 \times 2 \times 4 \times 4$ to $8 \times 256 \times 512 \times 512$. At the end of the decoding network, there was a 3D convolution layer and a 2D convolution layer to generate the predicted CT images that matched the shape of the reconstructed CT images. Finally, the predicted CT images with a size of $N \times 512 \times 512$ were output, where N was the slice number of the target CT images.

2.3.3. Training details

The DR-DRR and DRR-CT networks were implemented using the PyTorch library. The DR-DRR network was trained on a Nvidia A6000 graphics processing unit (GPU) on the training dataset for 380 epochs for the phantom and the patients, respectively, and the remaining parameters were the same as those in the original cycle-GAN network. In the training phase of the DRR-CT network, MSE was defined as the cost function to compare the generated synthetic CT with the target CT images. The Adam optimizer was used to minimize the resulting error with a learning rate of 0.00007. Moreover, we trained the DRR-CT network for 80 epochs using the same GPU. The model was evaluated on the validation set at the end of each training period. If the validation loss remained the same for 6 epochs, the learning rate was reduced by a factor of 2. Finally, the best checkpoint model with the smallest validation loss was saved as the final one in the two experiments. The same network structure and training strategy were used in the two experiments to verify the reliability of the proposed workflow.

2.4. Training and evaluation

2.4.1. Treatment planning

For the phantom and the patients, the gross tumor volume (GTV), clinical target volume (CTV) and OARs delineated by an experienced physician were used in this study. Note that the GTV and CTV of the phantom and the patients were delineated separately, and the GTV and CTV in the phantom were delineated virtually. The planning target volume (PTV) was created by expanding the CTV by 5 mm for the phantom and 3 mm for the patients. The OARs in the phantom included

the heart, the lungs (left and right lungs), the liver and the spinal cord. The brain stem, the lens (left and right lens) and the optical nerves (left and right optical nerves) were considered to be the OARs of the head-and-neck patients.

Carbon-ion radiotherapy treatment planning was performed with intensity-modulated pencil-beam spot scanning in MatRad [34]. In the treatment planning, 2 orthogonal irradiation fields ($0^\circ, 270^\circ$) were used for the phantom while 2 opposite irradiation fields ($90^\circ, 270^\circ$) were adopted for the patients, and the prescribed doses for the target volumes in the phantom and patients were clinically 55 Gy(RBE) and 60 Gy(RBE) in 11 and 15 fractions, respectively.

3D plan optimizations were performed using the planning CT images of the phantom and the CBCT images of the patients, a uniform 2 mm dose grid, a pencil-beam algorithm and the local effects model (LEM IV) to obtain the dose distributions under the pencil-beam spot scanning with scanning point spacings of 3 mm laterally and 2 mm longitudinally as the CT-based treatment plans. Subsequently, the dose distributions on the predicted CT images for the phantom and the patients were recalculated based on the same plan settings and irradiation parameters as those for the CT-based plan optimizations, which were taken as the predicted CT-based treatment plans. Then the difference of the CT- and predicted CT-based dose distributions acquired for the phantom and the patients were compared in dosimetry.

2.4.2. Technical evaluation

To evaluate the performance of the workflow, we deployed the two trained models on the test dataset and analyzed the reconstruction results using four different quantitative evaluation metrics such as mean absolute error (MAE), root mean square error (RMSE), structural similarity (SSIM) [35] and peak signal-to-noise ratio (PSNR). The distance metrics MAE and RMSE were used to estimate the difference between the predicted CT and ground truth CT images. The overall similarity between the predicted CT and ground truth CT images and the quality of the predicted CT images were measured by the structural metrics SSIM and PSNR, respectively.

2.4.3. Dosimetric evaluation

Dose distributions of the CT-based treatment plans were recalculated on the predicted CT images to evaluate the dose acceptability of the predicted CT. Moreover, carbon-ion treatment planning was also conducted on the predicted CT images with the same target volume and organ contours as those in the CT images, as well as the same irradiation settings as those for the CT images. Then the dosimetric metrics of the CT- and predicted CT-based treatment plans were compared under the same planning settings.

Dose differences between the CT- and predicted CT-based treatment plans were compared using dose volume histogram (DVH) metrics and 3D global gamma analysis. DVH metrics included $D_{98\%}$, $D_{50\%}$ and $D_{2\%}$ for the GTV, CTV, and PTV, and $D_{2\%}$ for the OARs. Although there is no established clinical standard for the differences in DVH metrics between CT and predicted CT images in carbon-ion radiotherapy, the standard of heavy ion radiotherapy should be wider than that of photon radiotherapy because the dose calculation accuracy of heavy ions is lower than that of photons. There is only a reference to the clinically acceptable criterion of less than 2% in photon radiotherapy [36] to analyze the differences in DVH metrics between CT and predicted CT images. Gamma pass-rates here were calculated using the CT-based dose as a reference, a consistent distance of 3 mm, and a dose difference of 3% (3 mm, 3%).

Results

In order to qualitatively and quantitatively evaluate the DR-DRR network, test datasets of the phantom and the patients were used for testing. The DRR image corresponding to each sDRR image was generated by CT images according to the imaging parameters of the

corresponding DR image. **Table 1** summarizes the MAE, RMSE, PSNR and SSIM values for the phantom and the patients, respectively. Representative examples during the DR-DRR network test process are shown in **Fig. 3** (column(1)), including the raw DR image, the real DRR image and the sDRR image generated by the DR-DRR network. As we can see from **Table 1**, the quantitative values of the DR-DRR network for the patients were slightly lower than those for the phantom.

The column (2) of **Fig. 3** shows the transverse view of the reconstruction results obtained from single angle-specific sDRR images of the phantom and the patients as the inputs of the DRR-CT network, together with the transverse view of the ground-truth CT images and the difference between the predicted CT and the ground-truth CT images. **Table 2** comprises the average quantitative reconstruction results of the DRR-CT network on the test datasets of the phantom and the patients, and the quantitative reconstruction results of the DRR-CT network on the representative examples when the inputs were DRR and sDRR images, respectively. The results of the DRR-CT network on the entire test dataset, a DRR image from the test dataset and a sDRR image had no significant differences. The differences in the DRR-CT results between the phantom and the patients were probably caused by the differences in the sDRR images generated by the DR-DRR network.

The dose volume histograms (DVH) obtained from the CT and predicted CT images and the relative dose differences of the clinically relevant DVH parameters between the two dose distributions of the carbon-ion treatment plans are shown in **Fig. 4** and **Table 3**, respectively. Similar to the technical performance, the performance of the DVH parameters on the phantom was also higher than that on the patients.

The DVH parameters differed less than 2% for both the target volumes and OARs in the phantom. For patient 1, the $D_{98\%}$ difference for the PTV was 1.75%, and the differences in $D_{98\%}$ for the GTV and CTV did not exceed 4%. Moreover, the $D_{50\%}$ and $D_{2\%}$ differences in all target volumes and the differences in $D_{2\%}$ of the OARs were all less than 1.5%. For patient 2, the $D_{2\%}$ and $D_{98\%}$ differences for the target volumes did not exceed 3.75% and 3%, respectively. Additionally, the differences in $D_{2\%}$ of the OARs and the $D_{50\%}$ differences in all target volumes were both less than 1.5%.

Fig. 5 shows the relative dose difference maps for the phantom and the patients. No significant dose differences were found between the CT-based and predicted CT-based plans for the two cases. In addition, there were almost no differences in the DVH parameters between the CT and predicted CT images of the phantom and the patients under the same carbon-ion treatment planning, except that the $D_{2\%}$ differences of the patient 1's left and right optical nerves were 1.25% and 1.75%, as described in **Table 4**.

The gamma pass-rates (3%, 3 mm) obtained between the CT- and predicted CT-derived dose distributions are shown in **Table 5**. Average gamma pass-rates of the three cases were above 98% for the dose distributions of the carbon ion plans.

4. Discussion

It is an ill-posed problem for reconstructing 3D CT images from 2D DR images because of infinite possibilities of the mapping relationships between 2D and 3D image domains. Therefore, a challenge for the

Table 1

The average quantitative results of the DR-DRR framework on the test datasets of the phantom and the patients using the MAE, RMSE, PSNR and SSIM. All image intensities were normalized between [0,1].

Case	MAE	RMSE	PSNR	SSIM
Phantom	0.010	0.027	36.619	0.987
Patient 1	0.019	0.072	30.718	0.970
Patient 2	0.017	0.061	31.667	0.969

Abbreviations = MAE: mean absolute error, RMSE: root mean squared error, PSNR: peak signal-to-noise ratio, SSIM: structural similarity index.

tomographic image reconstruction task is how to choose the best mapping among the many possible solutions [37]. In this study, the feasibility of generating predicted CT images from a single DR image based on the deep learning approach for carbon-ion radiotherapy treatment planning in two diverse subjects of anthropomorphic phantom and patients was evaluated. We proposed a DR-CT framework including a DR-DRR network and a DRR-CT one to generate the predicted CT images from a single DR image. We found acceptable differences between the CT- and predicted CT-based dose distributions for the carbon-ion radiotherapy plans on the image datasets of the phantom and the patients. To our best knowledge, this study is the first to evaluate the feasibility of performing dose calculations for carbon-ion radiotherapy treatment planning on deep learning-based predicted CT images for phantom and patients. Note that although the workflow was developed for carbon-ion radiotherapy, the concept and implementation of this workflow could be fully extended to photon and proton radiotherapy indeed.

Like other studies [7,24,25], DRR-CT data pairs were also used in this study to develop the DRR-CT network because it is extremely difficult to collect enough DR-CT data in the clinic. In addition, the cost and complexity of network development using CT data from a sufficient number of patients are also high. Therefore, the workflow was developed for a specific patient in this study for the sake of simplicity. In the face of the decline in the output results of DRR-CT network caused by the differences between DRR and DR images, we adopted a DR-DRR network to bridge DR and DRR images. The input of the DRR-CT network was changed to the sDRR image generated by the DR-DRR network, which finally realized the generation of predicted CT images from a single raw DR image. To ensure the reliability of the results, we first developed the entire workflow on the phantom, and then validated it using four CBCT datasets from two patients, respectively. The entire phantom and patients' test dataset were used for the qualitative and quantitative testing of the DR-DRR network, and the corresponding generated sDRR images were employed for the subsequent development and validation of the workflow. In the end, in terms of technical and dosimetric evaluation of the workflow, the results of the patients were slightly lower than those of the phantom, but all met or were close to the standards prescribed for photon radiotherapy, confirming the feasibility of the proposed workflow in clinical promotion.

Technical evaluation was used in the DR-DRR network and DRR-CT network of the DR-CT framework. Compared with previous studies [26], we achieved similar results on the DR-DRR network using a smaller number of paired DR-DRR datasets. However, although similar results were obtained, the data and target location of the patients and training epochs used in the two studies were different, which made them difficult to compare. The DR-DRR network performed worse than the phantom on the patients probably because the DR image of the phantom had a sharper image boundary. The DRR-CT network is likewise difficult to compare with some previous studies [7,24,25], because the network structure and resolution of the corresponding predicted CT images in the studies are different. However, our results are reasonable by comparing the method of data augmentation and final results.

In order to evaluate the dosimetric accuracy of the generated predicted CT images, the planning CT images of the phantom and the CBCT images of the patients were used to optimize the dose distributions for the carbon ion plans. For the phantom, clinically acceptable dose differences ($\leq 2\%$) in photon radiotherapy were obtained [36]. Dose differences slightly higher than the standard for photon radiotherapy were obtained on the patients. Thus, our method had acceptable dose differences on the target volumes and OARs between the CT and predicted CT images in the phantom and the patients for carbon-ion treatment planning with intensity-modulated pencil-beam spot scanning. In addition, the results under the same carbon ion plan designed on the CT and predicted CT images showed that there were almost no dose differences in the target volumes and OARs of the phantom and the patients. The gamma pass-rates of the phantom and the patients also showed

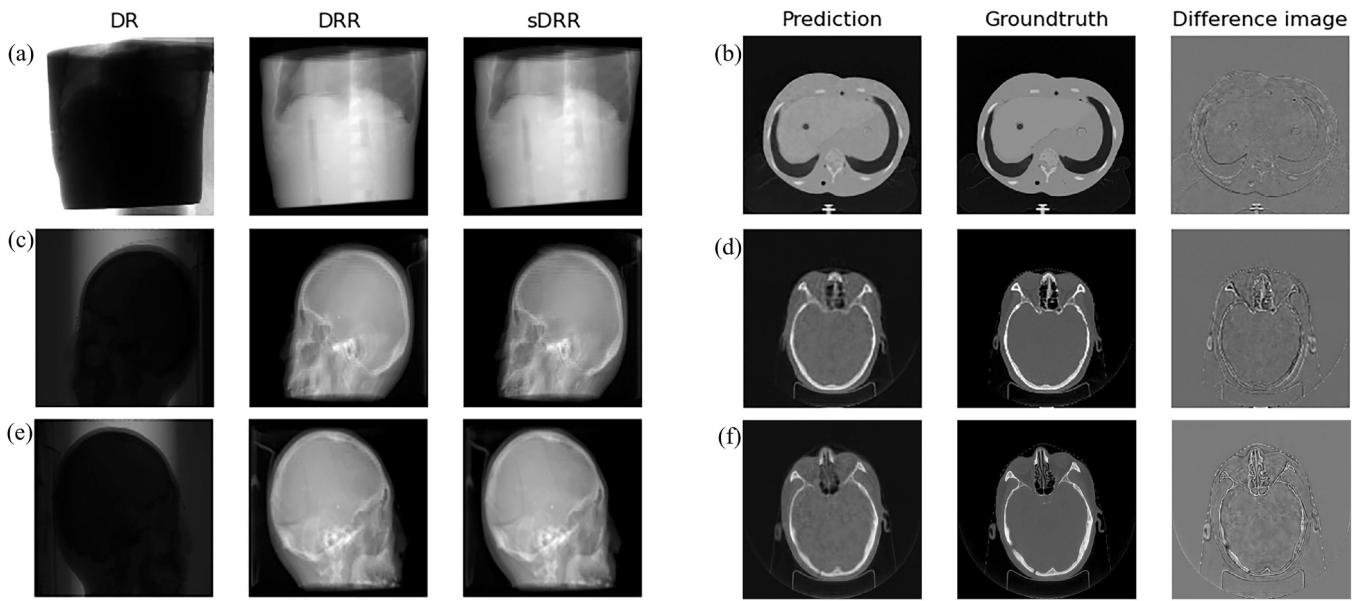


Fig. 3. Column (1) are the representative examples of the comprehensive performance of the DR-DRR network on the phantom (a), patient 1 (c) and patient 2 (e) with DR images in different directions; Column (2) are the transverse views of the representative examples from CT images of the phantom (b), patient 1 (d) and patient 2 (f) reconstructed using sDRR images generated by the DR-DRR network. Ground truth and difference images between the predicted and ground truth ones are also shown.

Table 2

The average quantitative reconstruction results of the DRR-CT network on the phantom and the patients' test datasets, and the quantitative reconstruction results of the DRR-CT network on the representative examples when the inputs were DRR and sDRR images, respectively.

Case	Data	MAE	RMSE	PSNR	SSIM
Phantom	Test dataset	0.006	0.079	38.869	0.977
	DRR	0.005	0.067	40.244	0.984
	sDRR	0.006	0.068	40.077	0.983
Patient 1	Test dataset	0.004	0.132	39.404	0.982
	DRR	0.006	0.173	36.818	0.972
	sDRR	0.007	0.186	36.158	0.968
Patient 2	Test dataset	0.006	0.143	38.339	0.976
	DRR	0.006	0.152	37.624	0.975
	sDRR	0.007	0.178	36.254	0.967

Abbreviations = DRR: digitally reconstructed radiograph, sDRR: synthetic digitally reconstructed radiograph, MAE: mean absolute error, RMSE: root mean squared error, PSNR: peak signal-to-noise ratio, SSIM: structural similarity index.

acceptable values with average gamma pass-rates over 98% for the (3%, 3 mm) criterion. Although there is no established standard for the DVH parameters in heavy ion radiotherapy, it is generally believed that the criterion in heavy ion radiotherapy should be lower than that of photon

radiotherapy. Thus, we think that the results of our workflow on the phantom and the patients are acceptable. At least, the workflow developed in this study can be applied to fast and automatic adaptive radiotherapy, because there is no need to scan patients for acquiring CT images for treatment re-planning.

In addition, as found in the study by Shen et al., the use of two orthogonal images does improve the quality of predicted CT images, but the improvement is limited [25]. Since the accuracy of the obtained CT images is sufficient, the use of a single image can further reduce the additional imaging dose. Furthermore, two complete orthogonal images are sometimes not obtained, considering the patient's body shape and tumor location. All things considered, we decided to use a single DR image to reconstruct 3D CT images for better clinical applicability.

There are some limitations in this study. The first one is the limited number of the datasets. The datasets used in this paper are CT images of the phantom and specific patient. Although the network performed well on the datasets of different fractions of a specific patient, the generalization ability of the network on other patient datasets might be insufficient. Second, although we achieved standard gamma pass rates, some dosimetric results exceeded clinical standards. However, it should be noted that the proposed workflow is only to explore the feasibility of this method. In the future, we will improve the network model in the workflow and use more data from different anatomical regions to verify.

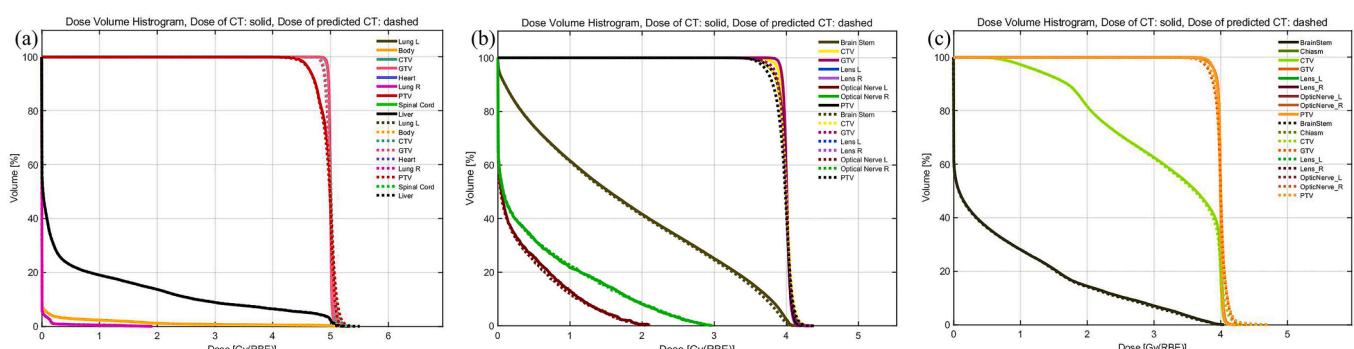


Fig. 4. Dose volume histogram (DVH) between the CT- and predicted CT-based treatment plans of the phantom (a), the patient 1 (b) and patient 2 (c).

Table 3

Clinically relevant relative dose differences derived from the CT- and predicted CT-based dose distributions for the phantom (a), patient 1 and patient 2 (b). (a).

(a)		GTV	CTV	PTV	Lung R	Liver
DVH parameters						
D _{2%}		1.60%	1.60%	1.20%	0	0.20%
D _{50%}		0	0.20%	0	0	0
D _{98%}		1.80%	1.40%	0.40%	0	0

(b)		GTV	CTV	PTV	Brain stem	Optical nerve L	Optical nerve R
DVH parameters	Case						
D _{2%}	Patient 1	1.00%	0.75%	1.00%	1.00%	0.75%	0.75%
	Patient 2	3.25%	3.25%	3.75%	1.00%	0	0
D _{50%}	Patient 1	0.25%	0.25%	0.25%	0.75%	0.25%	0.25%
	Patient 2	0	0.25%	0	0	0	0
D _{98%}	Patient 1	4.00%	3.50%	1.75%	0	0	0
	Patient 2	3.00%	2.75%	2.50%	0	0	0

Abbreviations = CT: computed tomography, GTV: gross tumor volume, CTV: clinical target volume, PTV: planning target volume, D_{x%}: the percent dose covering x% of the target volume.

Note: The OARs with a DVH parameter of 0 are not listed. The parameter difference values are represented as percentages of the respective prescribed doses.

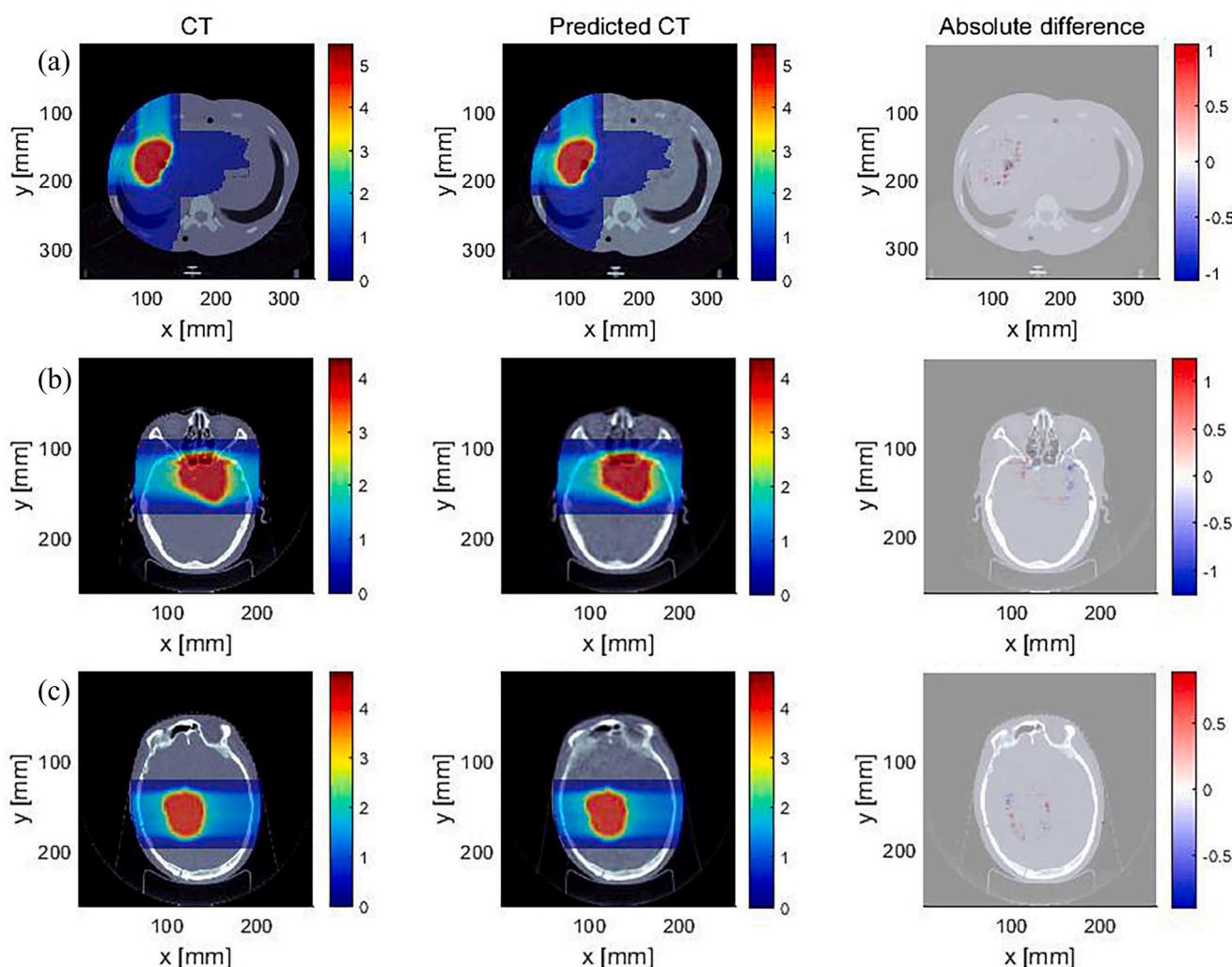


Fig. 5. Dose maps optimized on the CT and re-calculated on the predicted CT for the dose distributions of the phantom (a), patient 1 (b) and patient 2 (c). The last column displays the dose difference maps (CT versus predicted CT) overlaid on the planning CT image.

Finally, an inherent limitation of the presented workflow is the dose calculations for the patients were performed on the CBCT images, which are not usually applicable in the clinic. The main reason is that there were not enough DR images corresponding to the patient's planning CT

for developing the DR-DRR network in the DR-CT framework. Moreover, the formulation of treatment plans and dose calculations on the CBCT images was only to verify the reliability of the results of the workflow on the phantom, and it does not mean that CBCT images are used in the

Table 4

Clinically relevant relative dose differences between the CT and predicted CT images of the phantom (a), patient 1 and patient 2 (b) under the same carbon-ion treatment planning.

(a)						
DVH parameters	GTV	CTV	PTV	Lung R	Liver	
D _{2%}	0	0	0.20%	0	0	
D _{50%}	0.20%	0	0	0	0	
D _{98%}	0.40%	0	0.20%	0	0	

(b)								
DVH parameters	Case	GTV	CTV	PTV	Brain stem	Optical nerve L	Optical nerve R	
D _{2%}	Patient 1	0.25%	0.50%	0	0	1.25%	1.75%	
	Patient 2	0.50%	0.25%	0.25%	0.25%	0	0	
D _{50%}	Patient 1	0.25%	0.25%	0.25%	1.50%	0.25%	0.50%	
	Patient 2	0	0	0	0.25%	0	0	
D _{98%}	Patient 1	0.50%	0.25%	0.50%	0	0	0	
	Patient 2	0.25%	0.50%	0.50%	0	0	0	

Abbreviations = CT: computed tomography, GTV: gross tumor volume, CTV: clinical target volume, PTV: planning target volume, D_{x%}: the percent dose covering x% of the target volume.

Note: The OARs with a DVH parameter of 0 are not listed. The parameter difference values are represented as percentages of the respective prescribed doses.

Table 5

Gamma pass-rates using a 3% and 3 mm criterion for the dose distributions optimized on the CT images and re-calculated on the predicted CT images of the phantom and the patients.

Case	GTV	CTV	PTV
Phantom (%)	99.49	99.13	98.44
Patient 1 (%)	99.05	98.56	98.84
Patient 2 (%)	97.41	98.50	98.36

Abbreviations = GTV: gross tumor volume, CTV: clinical target volume, PTV: planning target volume.

clinic instead of planning CT for carbon-ion treatment planning. In the future, the planning CT of patients will be used to further verify and evaluate the entire workflow definitely. For example, cascading other networks later in the workflow enables the conversion of predicted CBCT to planning CT images [38,39].

5. Conclusions

In this study, a deep learning-based workflow was demonstrated to be clinically feasible to perform patient-specific DR-only carbon-ion radiotherapy treatment planning for head-and-neck patients on the generated predicted CT images even using a single unprocessed raw DR image. The data of different tumor sites can be used to build different patient-specific models according to the deep learning approach used to generate predicted CT images, reducing additional radiation dose to patients and improve the treatment efficiency in carbon-ion radiotherapy. Additionally, it is highly possible to adopt time-series DR images derived from fluoroscopy during a breathing cycle to generate 4D CT images of patients with the continuous improvement of the DR-CT framework. Practically, the proposed workflow provides a highly potential solution for realizing adaptive radiotherapy without CT rescan for carbon ions, and even photons and protons.

Funding

This work was jointly supported by the National Natural Science Foundation of China (Grant Nos. 11875299, U1532264, 12005271 and 12005273), the West Light Foundation of Chinese Academy of Sciences (Grant No. E023222Y) and the Youth Innovation Promotion Association of Chinese Academy of Sciences (Grant No. 2019405).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Kraft G. Tumor therapy with heavy charged particles. *Prog Part Nucl Phys* 2000;45: S473–544. [https://doi.org/10.1016/S0146-6410\(00\)00112-5](https://doi.org/10.1016/S0146-6410(00)00112-5).
- [2] Tsujii H, Kamada T, Shirai T, Noda K, Tsujii H, Karasawa K, editors. Carbon-Ion Radiotherapy. Tokyo: Springer Japan; 2014.
- [3] Rath AK, Sahoo N, editors. Particle Radiotherapy. New Delhi: Springer India; 2016.
- [4] Mori S, Shirai T, Takei Y, Furukawa T, Inanaga T, Matsuzaki Y, et al. Patient handling system for carbon ion beam scanning therapy. *J Appl Clin Med Phys* 2012;13(6):226–40.
- [5] Li Y, Kubota Y, Tashiro M, Ohno T. Value of Three-Dimensional Imaging Systems for Image-Guided Carbon Ion Radiotherapy. *Cancers* 2019;11(3):297.
- [6] Nagano Ai, Minohara S, Kato S, Kiyohara H, Ando K. Adaptive radiotherapy based on the daily regression of a tumor in carbon-ion beam irradiation. *Phys Med Biol* 2012;57(24):8343–56.
- [7] Ying XD, Guo H, Ma K, Wu J, Weng ZX, Zheng YF. X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks. *Proc Cvpr Ieee* 2019; 10611–20. <https://doi.org/10.1109/Cvpr.2019.01087>.
- [8] Karlsson M, Karlsson MG, Nyholm T, Amies C, Zackrisson B. Dedicated Magnetic Resonance Imaging in the Radiotherapy Clinic. *Int J Radiat Oncol* 2009;74:644–51. <https://doi.org/10.1016/j.ijrobp.2009.01.065>.
- [9] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* 2007;58:1182–95. <https://doi.org/10.1002/mrm.21391>.
- [10] Sidky EY, Pan XC. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys Med Biol* 2008;53: 4777–807. <https://doi.org/10.1088/0031-9155/53/17/021>.
- [11] Chen GH, Tang J, Leng SH. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys* 2008;35:660–3. <https://doi.org/10.1118/1.2836423>.
- [12] Yu HY, Wang G. Compressed sensing based interior tomography. *Phys Med Biol* 2009;54:2791–805. <https://doi.org/10.1088/0031-9155/54/9/014>.
- [13] Fessler JA, Rogers WL. Spatial resolution properties of penalized-likelihood image reconstruction: Space-invariant tomographs. *Ieee T Image Process* 1996;5: 1346–58. <https://doi.org/10.1109/83.535846>.
- [14] Ji SH, Xue Y, Carin L. Bayesian compressive sensing. *Ieee T Signal Proces* 2008;56: 2346–56. <https://doi.org/10.1109/Tsp.2007.914345>.
- [15] Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature* 2018;555:487–92. <https://doi.org/10.1038/nature25988>.
- [16] Schmidhuber J. Deep learning in neural networks: An overview. *Neural Networks* 2015;61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
- [17] Chen CY, Seff A, Kornhauser A, DeepDriving XJX. Learning Affordance for Direct Perception in Autonomous Driving. *Ieee I Conf Comp Vis* 2015;2722–30. <https://doi.org/10.1109/Iccv.2015.312>.
- [18] Ibragimov B, Toesca D, Chang D, Yuan YX, Koong A, Xing L. Development of deep neural network for individualized hepatobiliary toxicity prediction after liver SBRT. *Med Phys* 2018;45:4763–74. <https://doi.org/10.1002/mp.13122>.

- [19] Mori S, Hirai R, Sakata Y. Simulated four-dimensional CT for markerless tumor tracking using a deep learning network with multi-task learning. *Phys Med* 2020; 80:151–8. <https://doi.org/10.1016/j.ejmp.2020.10.023>.
- [20] Douglass MJJ, Keal JA. DeepWL: Robust EPID based Winston-Lutz analysis using deep learning, synthetic image generation and optical path-tracing. *Phys Med* 2021;89:306–16. <https://doi.org/10.1016/j.ejmp.2021.08.012>.
- [21] Florkow MC, Guerreiro F, Zijlstra F, Seravalli E, Janssens GO, Maduro JH, et al. Deep learning-enabled MRI-only photon and proton therapy treatment planning for paediatric abdominal tumours. *Radiother Oncol* 2020;153:220–7.
- [22] Bahrami A, Karimian A, Arabi H. Comparison of different deep learning architectures for synthetic CT generation from MR images. *Phys Med* 2021;90: 99–107. <https://doi.org/10.1016/j.ejmp.2021.09.006>.
- [23] Boulanger M, Nunes J-C, Chourak H, Largent A, Tahri S, Acosta O, et al. Deep learning methods to generate synthetic CT from MRI in radiotherapy: A literature review. *Phys Med* 2021;89:265–81.
- [24] Almeida DF, Astudillo P, Vandermeulen D. Three-dimensional image volumes from two-dimensional digitally reconstructed radiographs: A deep learning approach in lower limb CT scans. *Med Phys* 2021;48:2448–57. <https://doi.org/10.1002/mp.14835>.
- [25] Shen LY, Zhao W, Xing L. Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning. *Nat Biomed Eng* 2019;3:880–8. <https://doi.org/10.1038/s41551-019-0466-4>.
- [26] Dhont J, Verellen D, Mollaert I, Vanreusel V, Vandemeulebroucke J. RealDRR - Rendering of realistic digitally reconstructed radiographs using locally trained image-to-image translation. *Radiother Oncol* 2020;153:213–9. <https://doi.org/10.1016/j.radonc.2020.10.004>.
- [27] Montoya JC, Zhang CZ, Li YS, Li K, Chen GH. Reconstruction of three-dimensional tomographic patient models for radiation dose modulation in CT from two scout views using deep learning. *Med Phys* 2022;49:901–16. <https://doi.org/10.1002/mp.15414>.
- [28] Zaffino P, Raudaschl P, Fritscher K, Sharp GC, Spadea MF. Technical Note: PLASTIMATCH MABS, an open source tool for automatic image segmentation. *Med Phys* 2016;43:5155–60. <https://doi.org/10.1118/1.4961121>.
- [29] Zhao W, Shen L, Han B, Yang Y, Cheng K, Toesca DAS, et al. Markerless Pancreatic Tumor Target Localization Enabled By Deep Learning. *Int J Radiat Oncol* 2019;105 (2):432–9.
- [30] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets *Adv Neur In* 2014;27:2672–80.
- [31] Zhu JY, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *Ieee International Conference on Computer Vision (Iccv) 2017;2017:2242–51*. <https://doi.org/10.1109/Iccv.2017.244>.
- [32] He KM, Zhang XY, Ren SQ, Sun J. Deep Residual Learning for Image Recognition. *Ieee Conference on Computer Vision and Pattern Recognition (Cvpr) 2016;2016: 770–8*. <https://doi.org/10.1109/Cvpr.2016.90>.
- [33] Wang Q, Wu B, Zhu P, Li P, Hu Q. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)2020*.
- [34] Bangert M, Wieser H, Stadler A, Jaekel O. MatRad - An Open-Source Multi-Modality Treatment Planning Toolkit for Educational Purposes. *Med Phys* 2017;44: 3067–8.
- [35] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *Ieee T Image Process* 2004;13:600–12. <https://doi.org/10.1109/TIP.2003.819861>.
- [36] Korsholm ME, Waring LW, Edmund JM. A criterion for the reliable use of MRI-only radiotherapy. *Radiat Oncol* 2014;9. <https://doi.org/10.1186/1748-717x-9-16>.
- [37] Belthangady C, Royer LA. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat Methods* 2019;16:1215–25. <https://doi.org/10.1038/s41592-019-0458-z>.
- [38] Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X, et al. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. *Med Phys* 2019;46(9):3998–4009.
- [39] Chen LY, Liang X, Shen CY, Jiang S, Wang J. Synthetic CT generation from CBCT images via deep learning. *Med Phys* 2020;47:1115–25. <https://doi.org/10.1002/mp.13978>.