Comparing of two dimensional and three dimensional fully convolutional networks for radiotherapy dose prediction in left-sided breast cancer

Xue Bai¹,², Ze Liu³, Jie Zhang², Shengye Wang², Qing Hou¹, Guoping Shan², Ming Chen² and Binbing Wang²

¹Key Lab of Radiation Physics and Technology, Ministry of Education, Institute of Nuclear Science and Technology, Sichuan University, China
²Department of Radiation Physics, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), China
³School of Electronic Information and Electronical Engineering, Chengdu University, China

Abstract
Fully convolutional networks were developed for predicting optimal dose distributions for patients with left-sided breast cancer and compared the prediction accuracy between two-dimensional and three-dimensional networks. Sixty cases treated with volumetric modulated arc radiotherapy were analyzed. Among them, 50 cases were randomly chosen to conform the training set, and the remaining 10 were to construct the test set. Two U-Net fully convolutional networks predicted the dose distributions, with two-dimensional and three-dimensional convolution kernels, respectively. Computed tomography images, delineated regions of interest, or their combination were considered as input data. The accuracy of predicted results was evaluated against the clinical dose. Most types of input data retrieved a similar dose to the ground truth for organs at risk ($p > 0.05$). Overall, the two-dimensional model had higher performance than the three-dimensional model ($p < 0.05$). Moreover, the two-dimensional region of interest input provided the best prediction results regarding the planning target volume mean percentage difference (2.40 ± 0.18%), heart mean percentage difference (4.28 ± 2.02%), and the gamma index at 80% of the prescription dose are with tolerances of 3 mm and 3% (0.85 ± 0.03), whereas the two-dimensional combined input provided the best prediction regarding ipsilateral lung mean percentage

Corresponding author:
Binbing Wang, Department of Radiation Physics, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou 310022, China.
Email: wangbb@zjcc.org.cn

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).
difference (4.16 ± 1.48%), lung mean percentage difference (2.41 ± 0.95%), spinal cord mean percentage difference (0.67 ± 0.40%), and 80% Dice similarity coefficient (0.94 ± 0.01). Statistically, the two-dimensional combined inputs achieved higher prediction accuracy regarding 80% Dice similarity coefficient than the two-dimensional region of interest input (0.94 ± 0.01 vs 0.92 ± 0.01, p < 0.05). The two-dimensional data model retrieves higher performance than its three-dimensional counterpart for dose prediction, especially when using region of interest and combined inputs.

**Keywords**
Dose prediction, radiotherapy, fully convolutional network, treatment planning, breast cancer

**Introduction**

Intensity-modulated radiotherapy and volumetric modulated arc therapy (VMAT) are widely used to treat malignant tumors. The treatment plan quality, which is a major factor for the curative effect of radiotherapy, is related to the experience and skills of the practitioner and the individual anatomy. As planning based on trial-and-error necessitating a long planning time, the development of techniques such as online adaptive radiotherapy is restricted. Dose prediction before planning is useful for treatment plan quality assurance and subsequent automatic planning. Moreover, accurately predicted doses can reduce the number of interactions between the practitioner and computer and thus the planning time, also improving the plan quality and consistency. Dose prediction has evolved from single-value and dose-volume histogram prediction to three-dimensional (3D) dose distribution prediction. With the increased knowledge of radiobiology, the radiation effects on organs have been unveiled, especially with 3D dose distribution, thus suggesting the higher capabilities of 3D dose prediction over other methods.

A fully convolutional network (FCN) is an end-to-end method that can automatically extract hierarchical features from original images to generate corresponding masks. Its successful application to medical image segmentation has led to its use for dose distribution prediction in head and neck cancer using either two-dimensional (2D) or 3D models. Previous studies have focused on the use of deep learning and the network structure with data from contoured structures which separated from Computed tomography (CT) images as inputs. However, comparisons of different types of input data have not been conducted, despite the dependence of deep learning performance on both the network architecture and input data.

In this paper, we compare the use of 2D and 3D FCN models for predicting dose distribution in left-sided breast cancer. We use either the delineate regions of interest (ROIs), CT images, or their combination with multiple channels to consider the effect of density in dose calculation. The purpose of this study was to determine a superior input data type for the dose prediction FCN model.

**Materials and methods**

**Patient data**

Data from 60 patients with left-sided breast cancer (stage T1M0N0) previously treated using VMAT at the Zhejiang Cancer Hospital (Zhejiang, China) in
2016–2019 were randomly considered for this study. The patients’ volumetric CT data were acquired in freebreath using the Brilliance CT Big Bore system (Philips Healthcare, Best, Netherlands) on the helical scan mode. Each reconstructed CT image had a matrix size of 512×512 and a thickness of 5 mm. Clinical treatment plans were generated using the RayStation treatment planning system (RaySearch Laboratories, Stockholm, Sweden) with 6 MV X-rays on a Trilogy linear accelerator (Varian Medical Systems, Palo Alto, CA, USA). The dose calculation algorithm was RayStation’s collapsed cone convolution algorithm, and the grid size was 3 mm×3 mm×3 mm. Clinical target volumes (CTVs) were delineated by clinical oncologists based on the guidelines developed by the European Society for Radiotherapy and Oncology.24 The CTVs encompasses visible breast tissue and tumor bed. The irradiation area did not include internal breast areas and cervical lymph nodes. The planning target volumes (PTVs) were constructed by adding a 5 mm margin to the CTVs for all photon plans. All PTVs were clipped 5 mm from the skin surface. In addition, the heart, ipsilateral lung, whole lung, and spinal cord were contoured. All plans considered a prescription of 50 Gy in 25 fractions, and the prescription dose covered at 95% of the PTV. The mandatory target volume constraints for the PTVs and organ-at-risk (OAR) dose objectives followed the protocol developed at our institution. The PTV percentage volume above 110% of the prescription dose must not exceed 10%. The dosimetric constraints were \( V_{5} < 55\% \), \( V_{20} < 30\% \), and \( V_{30} < 20\% \) of the left lung volumes, respectively, and 8 Gy to mean heart dose. The VMAT plans were generated using the RayStation treatment planning system with 4° control point spacing. Double partial arc ranging from 290–310° to 115–165° were chosen depending on the anatomy of each patient. The basic characteristics of these contours and plans are listed in Table 1. When meeting all the constraints, the plans were optimized repeatedly to achieve optimal sparing of normal tissues that ensures the minimal OAR doses for each individual plan, thus being considered as expert plans. Two experienced dosimetrists and one senior oncologist at Zhejiang Cancer Hospital reviewed the plans and confirm that all the plans had achieved the best OAR sparing.

### Table 1. Basic characteristics of contours and plans in this study.

| Characteristic                | Training set | Testing set |
|------------------------------|--------------|-------------|
|                              | Mean         | Range       | Mean         | Range       |
| PTV (cm³)                    | 704          | 317–1556    | 692          | 358–1085    |
| Heart volume (cm³)           | 578          | 424–754     | 588          | 439–741     |
| Ipsilateral lung volume (cm³)| 1074         | 639–1752    | 1279         | 863–1983    |
| Lung volume (cm³)            | 2410         | 1470–3801   | 2741         | 1922–4116   |
| Beam start angle (deg)       | 301          | 290–310     | 297          | 290–305     |
| Beam stop angle (deg)        | 134          | 120–165     | 133          | 115–150     |

PTV: planning target volume.
**Data preprocessing**

The adopted FCN were trained using input CT grayscale images, ROI delineations, and the corresponding dose images. The CT images, ROIs, and dose files obtained from the treatment planning system were converted from the DICOM (Digital Imaging and Communications in Medicine) standard to arrays in Python. For the original CT images, the Hounsfield unit values ranging from $-1024$ to $3071$ HU were normalized from 0 to 1. The ROI structures were also converted into array data. Each voxel was assigned to a specific value if the voxel belonged to the ROIs. The ROIs of PTV, ipsilateral lung, heart, whole lung, spinal cord, and body were represented by 1.000, 0.833, 0.667, 0.500, 0.333, and 0.167, respectively, whereas other voxels were represented by 0. The voxels belonging to multiple structures were assigned the maximum value. For all the 60 cases, the maximum planning doses remained below 60 Gy, and hence the dose array was divided by 60 Gy for normalization. Each of the CT, ROI, and dose arrays were interpolated into $64 \times 256 \times 256$ matrices by the cubic method.

We evaluated 2D and 3D prediction models in this study. In the 2D models, predictions were completed slice-by-slice. In the 3D models, each case was represented by one 3D array. There were 50 randomly chosen arrays in the training set and the remaining 10 arrays in the testing set. For each model, the input data were trained in three cases: (a) CT image arrays as a single-channel input; (b) ROI arrays as a single-channel input; and (c) CT image and ROI arrays as a multichannel input. When either the CT images arrays or ROI arrays were used as input data, the channel of the input array was set to 1, and when the two arrays were used as input data, two channels were used to represent the data. Details of the training data are listed in Table 2.

**Prediction model using U-Net**

We used the U-Net FCN to predict dose distributions, as it has been successfully applied in medical image segmentation. U-Net has an upsampling stage with multiple feature channels to propagate context information to higher-resolution layers.20 The implemented U-Net model was based on the network architecture reported by Ronneberger et al.20

### Table 2. Specifications of training data models.

| Model       | Input data                   | Input array shape | Number of training/testing arrays |
|-------------|------------------------------|-------------------|-----------------------------------|
| 2D_CT       | CT image array               | $(256, 256, 1)$   | 3200/640                          |
| 2D_ROI      | ROI structure array          | $(256, 256, 1)$   | 3200/640                          |
| 2D_Combined | CT image and ROI structure arrays | $(256, 256, 2)$   | 3200/640                          |
| 3D_CT       | CT image array               | $(64, 256, 256, 1)$ | 50/10                            |
| 3D_ROI      | ROI structure array          | $(64, 256, 256, 1)$ | 50/10                            |
| 3D_Combined | CT image and ROI structure arrays | $(64, 256, 256, 2)$ | 50/10                            |

CT: computed tomography; ROI: regions of interest; 2D: two dimensional; 3D: three dimensional.
The end-to-end trainable FCN was able to automatically predict dose distributions using the framework shown in Figure 1.

The 2D U-Net comprised 10 layers. The first four layers were composed of two $3 \times 3$ convolutions, each followed by a rectified linear unit (ReLU), and $2 \times 2$ max pooling with a stride of 1 per dimension to reduce the feature size. To capture information from a larger effective receptive field, we used a slightly dilated $3 \times 3$ convolutional kernel with valid padding. The fifth layer was composed of two $3 \times 3$ convolutions. Then, four upsampling layers composed of a concatenate and two $3 \times 3$ convolutions, each followed by a ReLu, were included. Finally, the tenth layer with one convolution retrieved the predicted dose array. The predicted dose arrays were then interpolated to the original resolution using the cubic interpolation method. The activation function of the convolution in the first nine layers was ReLu. Unlike the original architecture, we replaced sigmoid activation in the last layer by a linear activation function. The prediction loss function was defined by the mean squared error (MSE). Hence, the FCN handled regression rather than classification. The initial learning rate was $10^{-3}$ and adjusted by a factor of 0.1 every five epochs if the loss function remained constant. The limit of learning rate was $10^{-6}$.

For the 3D U-Net, operations such as 2D convolution, max pooling, and upsampling were replaced by their 3D counterparts. In addition, the batch size was 16 for the 2D model and 1 for the 3D model due to GPU memory capacity. For all the models, Adam optimization was adopted, and the same hyperparameters and architectures were employed for a fair comparison between input types and models.

**Figure 1.** Framework of dose distribution prediction using U-Net.
**Prediction and evaluation**

A random split of 3200/640 images was used for training and testing the network. The test set was different from the training and validation sets. Ten-fold cross-validation was applied to the training set to assess the performance and stability of each model. For each fold, the training set was divided into 2880 training and 320 validation slices in the 2D model, whereas in the 3D model, the division was 45 training and five validation cases. The model weights that minimize the validation loss were used for external validation. After every fold was trained and the loss function converged, the model weights were used to evaluate the test set.

For dose prediction quality, the mean percentage difference (MPD) was determined for the PTV, heart, whole lung, ipsilateral lung, and spinal cord. This difference is defined for a ROI as follows:

$$\text{MPD(ROI)} = \frac{(1/n) \sum_{i=1}^{n} (|D_{pi} - D_i|)}{\text{prescription dose}} \times 100\% \quad (1)$$

where \(n\) is the number of voxels in the ROI, \(D_{pi}\) and \(D_i\) are the predicted and clinical plan doses for voxel \(i\), respectively.

Furthermore, the Dice similarity coefficient (DSC) was computed for 80% and 100% of the prescription dose area. At a given dose, the DSC is defined as

$$\text{DSC(dose)} = \frac{2(V_p \cap V_m)}{V_p + V_m} \quad (2)$$

where \(V_p\) is the area of predicted dose larger than the given dose and \(V_m\) is the area of clinical plan dose larger than the given dose. Hence, perfect prediction is indicated by a DSC of 1, whereas lower overlap retrieves smaller DSC values.

The gamma index was also used to measure dose distribution similarity.\textsuperscript{25,26} The index between a predicted dose-to-voxel \(d_p\) and a clinical plan dose \(d_m\) at point \(r_m\) is defined as follows:

$$\gamma(r_m) = \min \left\{ \frac{|r_p - r_m|^2}{\Delta r^2_M} + \frac{|d_p - d_m|^2}{\Delta d^2_M} \right\} \forall \{r_p\} \quad (3)$$

where \(r_p\) is a search over a neighborhood of voxels in the predicted dose space, \(d_p, \Delta r_M\) is a spatial distance threshold, and \(\Delta d_M\) is a dose difference threshold. The gamma index between two distributions is the percentage of voxels with \(\gamma(r_m) \leq 1\), which is also the percentage of voxels with dose difference below \(\Delta d_M\) to at least one voxel in a space no larger than \(\Delta r_M\) in the predicted dose image. To concentrate on the most crucial area of dose clinically including the PTV and dose falloff at the PTV periphery, the gamma index at 80% of the prescription dose area was determined with tolerances of \(\Delta r_M = 3\) mm, \(d_p = 3\% \text{(33\gamma)}\) and \(\Delta r_M = 5\) mm, \(d_p = 5\% \text{(55\gamma)}\).
**Statistical analyses**

Statistical analyses of prediction results were performed using the Wilcoxon rank test using SPSS v21 (IBM, NY, USA). A $p$-value < 0.05 was considered significant.

**Results**

Figure 2 shows the loss function evolution for the different types of inputs. The ROI and combined inputs have a similar training and validation loss over epochs, while the loss of the CT input is higher in both the 2D and 3D models. In the 2D model, the combined inputs help to achieve a lower validation loss than the ROI inputs. In contrast, in the 3D model, the ROI inputs lead to the lowest validation loss. Overall, the 2D model converges faster and with less overfitting than the 3D model. Still, the 3D model achieves a lower validation loss than the 2D model under ROI inputs, but the 2D outperforms the 3D model under CT and combined inputs.

Table 3 lists the ROI doses including ground truth and predictions. All the models generated PTV $D_{95}$ below the clinical plan. The 2D and 3D models with CT inputs and 3D models with combined inputs retrieved significantly less PTV $D_{\text{mean}}$ than the ground truth. All the models suitably predicted the OAR dose, including the $D_{\text{mean}}$ values for

![Figure 2. Training and validation loss functions for input models.](image-url)
| Model            | PTV $D_{95}$ (Gy) | PTV $D_{\text{mean}}$ (Gy) | PTV $D_{\text{max}}$ (Gy) | Ipsilateral lung $D_{\text{mean}}$ (Gy) |
|------------------|-------------------|----------------------------|---------------------------|----------------------------------|
|                  | Mean ± SD | p-value | Mean ± SD | p-value | Mean ± SD | p-value | Mean ± SD | p-value |
| Ground truth     | 50.00 ± 0.00 | –       | 52.18 ± 0.22 | – | 54.23 ± 1.32 | – | 10.23 ± 1.88 | – |
| 2D_CT            | 39.59 ± 4.34 | 0.043   | 49.99 ± 1.03 | 0.043 | 43.17 ± 2.30 | 0.043 | 10.76 ± 0.30 | 0.500 |
| 2D_ROI           | 48.95 ± 0.52 | 0.043   | 51.62 ± 0.19 | 0.225 | 53.37 ± 1.47 | 0.043 | 10.30 ± 0.46 | 0.686 |
| 2D_Combined      | 48.84 ± 0.52 | 0.043   | 52.20 ± 0.17 | 0.893 | 53.82 ± 0.87 | 0.345 | 10.82 ± 0.58 | 0.345 |
| 3D_CT            | 38.28 ± 10.26 | 0.043  | 43.35 ± 5.17 | 0.043 | 44.57 ± 1.05 | 0.686 | 10.53 ± 1.32 | 0.500 |
| 3D_ROI           | 47.35 ± 1.33 | 0.043   | 51.93 ± 0.52 | 0.5  | 52.65 ± 1.28 | 0.138 | 10.57 ± 0.68 | 0.686 |
| 3D_Combined      | 47.07 ± 1.14 | 0.043   | 51.87 ± 0.22 | 0.043 | 52.87 ± 0.97 | 0.225 | 10.88 ± 0.93 | 0.500 |

| Model  | PTV (%) | Ipsilateral lung (%) | Heart (%) | Lung (%) | Spinal cord (%) |
|--------|---------|----------------------|-----------|----------|-----------------|
| 2D_ROI | 2.40 ± 0.18 | 4.28 ± 2.02 | 2.32 ± 0.62 | 2.55 ± 1.38 | 0.75 ± 0.37 |
| 2D_Combined | 2.49 ± 0.24 | 4.16 ± 1.48 | 2.39 ± 0.63 | 2.41 ± 0.95 | 0.67 ± 0.40 |
| p-value | 0.138   | 0.893 | 0.225 | 0.893 | 0.225 |

CT: computed tomography; ROI: regions of interest; PTV: planning target volume; SD, standard deviation; 2D: two dimensional; 3D: three dimensional.
ipsilateral lung, heart, and lung, as well as $D_{\text{max}}$ for the PTV and spinal cord, as no significant difference was found from the ground truth.

Figure 3 shows the prediction errors on all the OARs and PTV from the test set. Using 3D CT inputs provides the worst prediction results, especially in PTV and the surrounding high-dose area. In addition, the 2D models have an overall lower prediction error than the 3D models.

Using the 2D model either with ROI or combined inputs leads to better prediction accuracy than the other model and inputs. Table 4 lists the prediction accuracy of the 2D_ROI and 2D_Combine models. The 2D ROI inputs retrieved the best prediction regarding PTV and heart MPDs and $33\gamma$, and the 2D combined inputs retrieved the best prediction regarding ipsilateral lung, lung, and spinal cord MPDs as well as 80% DSC. Moreover, the two models provide the same results for 100% DSC and $55\gamma$. Statistically, the 2D combined inputs are more accurate for prediction regarding 80% DSC than the 2D ROI inputs.

An open-source implementation of U-Net based on the Keras deep learning library was used in this study. Training on 50 cases can take 2–3 h for 100 iterations on an
NVIDIA GeForce GTX 1080 GPU with 11 GB memory. However, the execution time during testing is much faster. On the 10 cases from the test set, the processing time for each volume was 6–22 s per stage depending on the size of the candidate regions.

**Discussion**

We evaluated six types of inputs for left-sided breast cancer. Although CT images without contouring can retrieve patient-specific dose distributions for radiotherapy, using ROI delineations with a unique label for PTV to serve as input improves the prediction accuracy. Moreover, combining CT images and delineated ROIs can further improve accuracy.

For CT image, ROI, and combined inputs, we tested 2D and 3D input models. It can be expected that prediction using 2D models can cause more errors, particularly in slices at the superior and inferior PTV edges, compared to 3D models.\(^{15,16}\) However, we found that the proposed FCN with 2D inputs retrieved an overall higher performance \((p < 0.05)\). Figure 4 shows an example of the error between predicted dose and ground truth in the last slice of superior and inferior PTV edges. All the models predicted higher doses outside the PTV and lower doses inside it, especially along the PTV boundary. Hence, the FCN models should be improved to predict high-dose gradients in the boundary. The poorer 3D model performance in this study may be caused by the limited batch size. For 2D models, we tested different batch sizes (2, 4, 8, 16, 32, and 64), with 16 providing the best performance. However, the batch size was set to 1 in the 3D models due to GPU memory limitations, establishing a compromise between

### Table 4

| Model          | Heart \(D_{\text{mean}}\) (Gy) Mean ± SD | p-value | Lung \(D_{\text{mean}}\) (Gy) Mean ± SD | p-value | Spinal cord \(D_{\text{max}}\) (Gy) Mean ± SD | p-value |
|----------------|-----------------------------------------|---------|----------------------------------------|---------|----------------------------------------------|---------|
| Ground truth   | 5.37 ± 0.62                             | –       | 5.43 ± 1.18                            | –       | 2.37 ± 1.27                                 | –       |
| 2D_CT          | 6.50 ± 0.98                             | 0.138   | 5.58 ± 0.20                            | 0.138   | 2.57 ± 0.36                                 | 0.138   |
| 2D_ROI         | 5.75 ± 0.44                             | 0.5     | 5.35 ± 0.15                            | 0.686   | 1.85 ± 0.11                                 | 0.686   |
| 2D_Combined    | 6.02 ± 0.28                             | 0.138   | 5.49 ± 0.26                            | 0.500   | 1.90 ± 0.55                                 | 0.686   |
| 3D_CT          | 6.32 ± 0.83                             | 0.225   | 5.61 ± 0.77                            | 0.500   | 1.48 ± 0.18                                 | 0.345   |
| 3D_ROI         | 6.06 ± 0.56                             | 0.345   | 5.51 ± 0.30                            | 0.893   | 1.53 ± 0.47                                 | 0.345   |
| 3D_Combined    | 6.26 ± 0.75                             | 0.225   | 5.66 ± 0.39                            | 0.686   | 4.08 ± 3.21                                 | 0.225   |

**CT:** computed tomography; **ROI:** regions of interest; **PTV:** planning target volume; **2D:** two dimensional; **DSC:** Dice similarity coefficient; **MPD:** mean percentage dose.
model accuracy and computing power. Hence, more memory allowing larger batch sizes may improve the performance of 3D over 2D models. Nevertheless, using the reported parameters, the 2D models outperformed their 3D counterparts, even in the superior and inferior PTV edges. We believe that the reason for this phenomenon is that when the plans were designed with a coplanar field, the geometry of 2D ROI has the greatest correlation with the dose distribution. The 3D input may not provide more valuable information, but added noise that interfered the prediction model. Figure 5 shows examples of dose color wash and dose-volume histogram for a patient from the test set. Visually, most 3D dose prediction models generate similar doses to the ground truth, with 2D and 3D CT input models having the lower dose coverage in PTV. All the models predicted comparable doses to OARs. The heart exhibited higher doses in all the models compared to the ground truth, which may represent scarce representative data in the training set or unsuitability of the models using the current FCN architecture.

Figure 4. Example of error between predicted dose and ground truth in the last slices of superior (top) and inferior (bottom) planning target volume (PTV) edges. Yellow represents predicted dose higher than the ground truth, whereas green represents predicted dose lower than the ground truth.
In this study, we predicted accurate volumetric doses for breast cancer patients using an FCN. McIntosh and Purdie used contextual atlas regression forests to predict the dose distribution for whole breast tangent intensity-modulated radiotherapy plans. Their study reported the \( \gamma \) index at 80% of the prescription dose with a tolerance of 5% and 5 mm of 0.79 \( \pm 0.08 \), and the DSC by thresholding the distribution at 80% of the prescription dose was 0.86 \( \pm 0.05 \). Bai et al. predicted dose distribution for left breast cancer using deformable image registration. Their study reported \( \gamma \) index at 80% of the prescription dose with a tolerance of 5% and 5 mm of 0.84 \( \pm 0.10 \), and the DSC by thresholding the distribution at 80% of the prescription dose was 0.87 \( \pm 0.05 \). Although the different patient datasets and treatment protocols prevent direct comparison with our study, we demonstrated the feasibility of accurately predicting dose distributions for breast cancer treatment. Unlike studies on head-and-neck and prostate cancer with identical beam arrangements, the beam angles for breast cancer are patient-specific, and we verified that the FCN approach is robust to slight changes in beam angles.

Various limitations of this study remain to be addressed. First, we considered data from only 60 patients, which may be insufficient for deep learning. The FCN model

![Figure 5. Example of predicted dose distribution generated with the different input types.](image-url)
can become more robust and accurate with larger training sets. Second, we only used left-sided breast cancer data. Still, radiotherapy is used in many sites for cancer treatment, and the extension of our results to other sites, such as head, neck, and prostate, remains to be investigated. Third, we focused on the influence of input data for dose prediction rather than on the network structure, thus using the same hyperparameters for every input type. Adjusting the hyperparameters according to the input type may improve accuracy.

In future work, we will include more patient data to confirm the generality of our findings. We also plan to expand the study to several clinical conditions, including different treatment sites and tumor sizes. Moreover, the influence of the prediction results from different inputs should be investigated for different FCN architectures.

**Conclusions**

We propose FCN-based prediction of radiation dose for breast cancer and find that 2D models generally outperform 3D models under our experimental conditions. Although the radiotherapy dose distributions can be predicted directly from CT images without delineated ROIs, using ROIs or combinations of the inputs in 2D models retrieves the highest prediction accuracy of doses for left-sided breast cancer patients. In fact, CT images are complemented with ROI information for improved prediction performance.

**Acknowledgments**

The authors sincerely thank all study participants.

**Declaration of conflicting interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by the National Natural Science Foundation of China (12005190), the Zhejiang Province Key Research and Development Program (2019C03003), the Zhejiang Medical and Health Discipline Platform Project (2018ZD014), and the Zhejiang Basic Public Welfare Research Program (LSY19H180002).

**Ethics approval**

This study was carried out in accordance with the declaration of Helsinki and was approved by the Ethics Committee of Zhejiang cancer hospital with the following number: IRB-2020-2.

**ORCID iDs**

Xue Bai https://orcid.org/0000-0002-9056-0882
Jie Zhang https://orcid.org/0000-0002-0817-0464
References

1. Staffurth J. A review of the clinical evidence for intensity-modulated radiotherapy. *Clin Oncol (R Coll Radiol)* 2010; 22: 643–657.
2. Bedford JL. Treatment planning for volumetric modulated arc therapy. *Med Phys* 2009; 36: 5128–5138.
3. Karl OJMP. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys* 2008; 35: 310–317.
4. Nelms BE, Robinson G, Markham J, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. *Pract Radiat Oncol* 2012; 2: 296–305.
5. Zhu X, Ge Y, Li T, et al. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Med Phys* 2011; 38: 719–726.
6. Hunt MA, Jackson A, Narayana A, et al. Geometric factors influencing dosimetric sparing of the parotid glands using IMRT. *Int J Radiat Oncol Biol Phys* 2006; 66: 296–304.
7. Wu BB, Ricchetti F, Sanguineti G, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys* 2009; 36: 5497–5505.
8. Reddy NMS, Chang NH, Lange CS, et al. Prostate and seminal vesicle volume based consideration of prostate cancer patients for treatment with 3D-conformal or intensity-modulated radiation therapy. *Med Phys* 2010; 37: 3791–3801.
9. Moore KL, Brame RS, Low DA, et al. Experience-based quality control of clinical intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys* 2011; 81: 545–551.
10. Petit SF, Wu B, Kazhdan M, et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. *Radiother Oncol* 2012; 102: 38–44.
11. Appenzoller LM, Michalski JM, Thorstad WL, et al. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Med Phys* 2012; 39: 7446–7461.
12. Shiraishi S and Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. *Med Phys* 2016; 43: 378.
13. McIntosh C and Purdie TG. Contextual atlas regression forests: multiple-atlas-based automated dose prediction in radiation therapy. *IEEE Trans Med Imaging* 2016; 35: 1000–1012.
14. Bai X, Shan G, Chen M, et al. Approach and assessment of automated stereotactic radiotherapy planning for early stage non-small-cell lung cancer. *Biomed Eng Online* 2019; 18: 101.
15. Chen X, Men K, Li Y, et al. A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning. *Med Phys* 2019; 46: 56–64.
16. 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: 065020.
17. Dan N, Long T, Jia X, et al. A feasibility study for predicting dose distributions from contours using deep learning on prostate IMRT patients. https://arxiv.org/abs/1709.09233, 2018.
18. Chao M, Wei J, Narayanasamy G, et al. Three-dimensional cluster formation and structure in heterogeneous dose distribution of intensity modulated radiation therapy. *Radiother Oncol* 2018; 127: 197–205.
19. Shelhamer E, Long J and Darrell T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2017 Apr; 39(4): 640–651. doi: 10.1109/TPAMI.2016.2572683. Epub 2016 May 24. PMID: 27244717.
20. Ronneberger O, Fischer P and Brox T. U-Net: convolutional networks for biomedical image segmentation. https://arxiv.org/abs/1505.04597v1, 2015.
21. Thong W, Kadoury S, Piché N, et al. Convolutional networks for kidney segmentation in contrast-enhanced CT scans. *Comput Methods Biomech Biomed Eng Imaging Vis* 2016; 6: 277–282.

22. Christ PF, Elshaer MEA, Ettlinger F, et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016: 415–423.

23. Men K, Dai J and Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. *Med Phys* 2017; 44: 6377–6389.

24. Offersen BV, Boersma LJ, Kirkove C, et al. ESTRO Consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. *Radiother Oncol* 2015; 114: 3–10.

25. Low DA, Harms WB, Mutic S, et al. A technique for the quantitative evaluation of dose distributions. *Med Phys* 1998; 25: 656–661.

26. Low DA and Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Med Phys* 2003; 30: 2455–2464.

27. Chollet F. Keras: deep learning library for Theano and TensorFlow, https://github.com/fchollet/keras (2015).

28. Bai X, Wang B, Wang S, et al. Radiotherapy dose distribution prediction for breast cancer using deformable image registration. *Biomed Eng Online* 2020; 19: 39.

**Author biographies**

**Xue Bai** received the a B.S. degree in physics from the Zhejiang University, in 2008, and an M.S. degree in nuclear energy and nuclear technology engineering from the Sichuan University, in 2017. From 2008 to 2013, she was an engineer assistant with the department of radiation physics in Zhejiang Cancer Hospital (Cancer Hospital of the University of Chinese Academy of Sciences). Since 2014, she has been an engineer with in the same department. She is the author of more than 10 articles and more than 5 inventions. Her research interest includes the dosimetry study in radiotherapy, medical physics, the deep learning used in radiotherapy, and automatic planning.

**Ze Liu** received a Ph.D. degree in nuclear energy and nuclear technology engineering from the Sichuan University, in 2019. Since graduation in 2019, he was a teacher with the school of electronic information and electronic engineering in at Chengdu University. He is the author of more than 8 articles and more than 3 inventions. His research interest includes surface modification, design, characterization, and preparation of thin solid films.

**Jie Zhang** received the a B.S. degree in communication engineering and a Ph.D. degree in electronic science and technology from Nanjing University. Since 2018, she has been a radiation physicist with the radiation physics department, Zhejiang Cancer Hospital (Cancer Hospital of the University of Chinese Academy of Sciences). She is the author of more than 10 articles and 2 inventions. Her research interests include tumor motion tracking and the medical application of artificial intelligence.

**Shengye Wang** received the a B.S degree in Clinical Medicine in from Wenzhou Medical University in 2005 and M.D. in Oncology from Zhejiang University in 2018. From
2005 to 2021, She has been a radiation oncologist in the department of radiation oncology in Zhejiang Cancer Hospital. She is the author of more than 20 articles. Her research domain includes the heart dosimetry study in left-sided breast cancer, medical dosimetry, radiation of head & neck tumors, and thoracic tumors.

Qing Hou received his Ph.D. in physics from the Institute of Physics, Academia Sinica of China in 1990. He is a professor of the Institute of Nuclear Science and Technology, Sichuan University. His interest is on in the fundamental research of radiation–material interactions, and the applications of nuclear technology.

Guoping Shan received the a B.S. degree in Technical Physics from the University of South China, in 1997 and an M.S. degree in Biomedical Engineering from the Tsinghua University, in 2015. From 1997 to 2012, he was a senior staff radiation oncology physicist with the department of radiotherapy in Taizhou Hospital. From 2008 to 2009, he also was a radiation oncology physics resident with the department of radiation oncology in Fox Chase Cancer Center, Philadelphia, PA. From 2013 to 2014, he worked as the chief physicist in Mali National Radiotherapy Center, Bamako, Mali in Africa. Since 2012, he has been the director of the medical physics department in Zhejiang Cancer Hospital (Cancer Hospital of the University of Chinese Academy of Sciences). He has published more than 30 articles. His main research direction is medical physics. His awards and honors include Mali “bee” National Medal (Mali Africa 2014) and the Best Medical Workers (Zhejiang Province 2015).

Ming Chen received his bachelor’s degree in Medical College of Soochow University, the a master’s degree in Shanghai Medical College of Fudan University, and an M.D. degree in Wuhan University. From 2004 to 2005, he was a postdoc at the University of Michigan. He has been a professor at Sun Yat-Sen University, Zhejiang University, and Soochow University. From 2012 to 2021, he is with Zhejiang Cancer Hospital (Cancer Hospital of Chinese Academy of Sciences), and from 2019 to 2021, he was the executive vice president of the hospital. Since 2021, he is the director of the Medical Physics Department at Sun Yat-Sen University Cancer Center. He is the Standing Committee Member and General Secretary of the Chinese Society of Radiation Oncology, Chinese Medical Association, the Standing Committee Member of the Society of Radiation Therapy, China Anti-Cancer Association, and the Vice Chairman of the Experts Committee on Radiation Therapy, Chinese Society of Clinical Oncology. He has been a principal investigator of over 20 national or ministry projects, and editors of 9 academic journals and 11 professional associations. He has published over 100 papers and was licensed with 11 patents. His research interest is radiation oncology.

Binbing Wang is a professor of medical physics at Zhejiang Cancer Hospital. He has participated as a principal investigator for more than 5 clinical research studies. He has published more than 20 journal articles and scientific abstracts. He has a personal and professional interest in deep learning and medical physics.