Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

CrossMark

PAPER

Deep DoseNet: a deep neural network for accurate dosimetric transformation between different spatial resolutions and/or different dose calculation algorithms for precision radiation therapy

Peng Dong and Lei Xing[1]

Department of Radiation Oncology, Stanford University, Stanford, CA 94305-5847, United States of America
[1] Author to whom any correspondence should be addressed.

E-mail: Lei@stanford.edu

Abstract

The purpose of this work is to introduce a novel deep learning strategy to obtain highly accurate dose plan by transforming from a dose distribution calculated using a low-cost algorithm (or algorithmic settings).

25 168 slices of dose distribution are calculated using Eclipse treatment planning system V15.6 (Varian Medical Systems, Palo Alto, CA) on ten patient CTs whose treatment sites ranging from lung, brain, abdomen and pelvis, with a grid size of $1.25 \times 1.25 \times 1.25$ mm using both anisotropic analytical algorithm (AAA) in 5 mm resolution and Acuros XB algorithm (AXB) in 1.25 mm resolution. The AAA dose slices, and the corresponding down sampled CT slices are combined to form a tensor with a size of $2 \times 64 \times 64$, working as the input to the deep learning-based dose calculation network (deep DoseNet), which outputs the calculated Acuros dose with a size of 256×256. The deep DoseNet (DDN) consists of a feature extraction component and an upscaling part. The DDN converges after ~100 epochs with a learning rate of 10^{-4}, using ADAM.
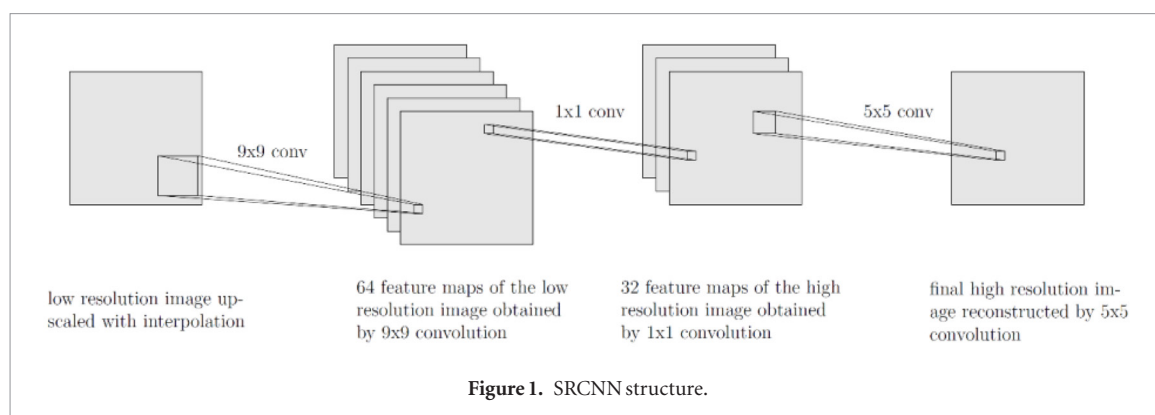
We compared up sampled AAA dose and DDN output with that of AXB. For the evaluation set, the average mean-square-error decreased from 4.7×10^{-4} between AAA and AXB to 7.0×10^{-5} between DDN and AXB, with an average improvement of ~12 times. The average Gamma index passing rate at 3mm3% improved from 76% between AAA and AXB to 91% between DDN and AXB. The average calculation time is less than 1 ms for a single slice on a NVIDIA DGX workstation.

DDN, trained with a large amount of dosimetric data, can be employed as a general-purpose dose calculation acceleration engine across various dose calculation algorithms.

1. Introduction

The success of modern radiotherapy modality, such as stereotactic body radiation therapy (SBRT) (Onishi *et al* 2011), stereotactic radiosurgery (SRS) (Lutz *et al* 1988), and adaptive therapy (Mutic and Dempsey 2014), critically depends on the accuracy and efficiency of dose calculation (Mackie *et al* 1985). Coupled with the increasing utilization of the time intensive dose optimization techniques such as multi criteria optimization (MCO) (Craft *et al* 2012), motion-considered or robust optimization (Chu *et al* 2005) and trajectory optimization (Dong *et al* 2018), the demand for high performance dose calculation increases continuously. In the past few decades, extensive research has been devoted to speed up Monte Carlo calculation (Andreo 1991, Rogers *et al* 1995) by simplifying and approximating electron transport. Others (Failla *et al* 2010) have taken a different route by solving the linear Boltzmann transport equations (LBTE) deterministically, which describes statistically the behaviors of radiation particles during transport through medium without considering self-interaction, providing comparable accuracies to Mote Carlo method with improved calculation speed.

Deep learning (LeCun *et al* 2015), which extracts unique features and learns representations of data by multiple levels of abstraction through multiple processing layers such as convolutional layers, has sparked global interests in recent years and made major impact on several computational tasks (Xing *et al* 2018) such as image

Figure 1. SRCNN structure.

reconstruction, classifications and recognition (Szegedy *et al* 2016), disease detection (Fauw *et al* 2018), segmentation (Ronneberger *et al* 2015), super-resolution imaging (Dong *et al* 2016, Liu *et al* 2018), image style transfer (Gatys *et al* 2016) and Dose prediction (Ma *et al* 2019a, 2019b). Inspired by the success of super-resolution imaging, here we propose a deep learning strategy, named as DDN, for dose calculation and optimization applications (Dong and Xing 2019). The technique is capable of not only up sampling the dose to speed up the calculation, but also transforming the dose calculated in one algorithm to another to enhance the accuracy of dose calculation.

2. Methods

2.1. Single dose map super resolution

Image super resolution is widely used to enhance the quality of a low resolution image, and its performance has improved substantially since the original work by Dong *et al* (Dong *et al* 2016), which uses a three-layer deep convolutional neural network (SRCNN) to directly learn an end to end mapping between the low and high resolution images. Kim *et al* (2015) increased the layer numbers of the CNN to 20 and used several training techniques including gradient clipping and skip connection to further boost the accuracy. Ledig *et al* (2016) proposed to use Resblock (He *et al* 2016) and GAN (Radford *et al* 2015) to recover photo realistic textures from heavily down sampled images. Below we briefly summarize SRCNN (figure 1) (Dong *et al* 2016), which is known for its simple and effective structure and forms the basis for more advanced networks and our DDN.

Given a low-resolution image X, SRCNN shown in figure 1 first upscales the image with cubic interpolation and uses it as the input $G(X)$. SRCNN has three parts: feature extraction, nonlinear mapping and reconstruction. Feature extraction uses 9×9 convolution to generate 64 feature maps from the low-resolution image. The second part maps those 64 feature maps nonlinearly to 32 high-resolution patch representations with 1×1 convolution. The last part combines the 32 patch representations with 5×5 convolution to produce the final high-resolution image $F(X)$. The mean square error (MSE) is employed as the loss function where i represents pixel i:

$$Loss = \sum_{i=1}^{n} \| F(x_i) - G(x_i) \|^2$$

and trained the network with 24 800 sub-images which are decomposed from 91 images, achieving the state-of-the-art results when compared with conventional methods (Dong *et al* 2016).

2.2. Analytical anisotropic algorithm (AAA) versus Acuros XB (AXB)

Eclipse treatment planning system V15.6 (Varian Medical Systems, Palo Alto, CA) is used for the dose calculations. The system provides two main dose calculation algorithms: (i) the analytical anisotropic algorithm (AAA) is an analytical calculation algorithm based on the pencil beam convolution superposition technique; and (ii) the grid based the Acuros XB (AXB), which solves the LBTE deterministically. AXB agrees the Monte Carlo method within 2% in low density lung, while the AAA shows a difference of 12% (Failla *et al* 2010). In bone and bone-tissue interface regions, dosimetric inaccuracy as large as 6.4% in AAA calculations has been reported, whereas AXB algorithm markedly improved the situation (Han *et al* 2011). In reality, however, AXB is five time slower than AAA for a single field calculation (Failla *et al* 2010). We use low resolution AAA and high resolution AXB as input and output for supervised training of the proposed DDN model and to showcase the capabilities of DDN. The premise of the DDN is to take advantage of the useful features of both algorithms and provide an computationally efficient and dosimetrically accurate technique for treatment planning.

2.3. Deep DoseNet (DDN)

2.3.1. Data preparation

Ten patient CT image datasets of different disease sites, including brain, thorax (lung and sternum), abdomen (liver and para-aortic) and pelvis, are selected for this study. For each patient, 100 beam dose distributions are computed with different beam iso-center location, field size and gantry angles. Specifically, we set the Y1, Y2, X1, and X2 jaw positions randomly and the beam size from 2 cm to 20 cm, with the iso center of the beam being placed at different locations in different anatomies within the patient body. For each beam configuration, two dose distributions are calculated with the monitor units MUs) fixed to 100 by for a TrueBeam™ LINAC with $10 \times$ flattening filter free (10FFF) photon beam: one with a grid size of $1.25 \times 1.25 \times 1.25$ mm using the AXB algorithm with dose to medium, and the other with a grid size of $5 \times 5 \times 5$ mm using the AAA algorithm. To mitigate potential complication caused by the difference in the coordinates of dose calculation and CT images, we linearly interpolated the AAA (AXB) dose matrix to the CT coordinates at $1.25\,(5) \times 1.25\,(5)$ mm grid size. In this way the dose matrix and the CT matrix match each other voxel to voxel.

On slices close to the beam edge where the dose gradient is high, the AAA dose might be zero, while the AXB dose is still not vanishing due to the resolution difference in the two calculations. The dose mapping is focused on transverse planes with a max AAA plane dose not less than 10% of the max AAA volume dose. In total, we obtained 25 168 CT, AAA and AXB slices with matching coordinates, that is, ~2500 matched slices per patient. The 64×64 AAA dose slices (5×5 mm) combining with the corresponding 256×256 CT slices (1.25×1.25 mm) serve as the input to the DDN (figure 2). The output of the DN model is the inferred AXB dose with a size of 256×256 (1.25×1.25 mm). The CT is linearly scaled from $(-1000\,\text{HU}, 3500\,\text{HU})$ to $(0, 1)$. The AAA beam dose is normalized to the max dose. And the AXB beam dose is linear scaled to keep the max dose ratio between the AAA and AXB the same.
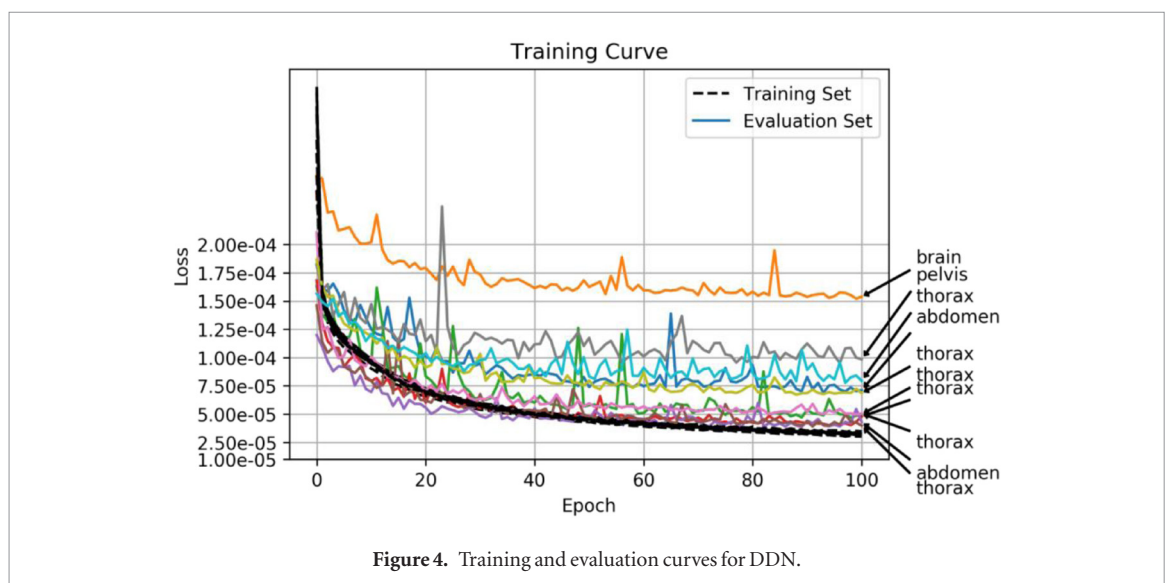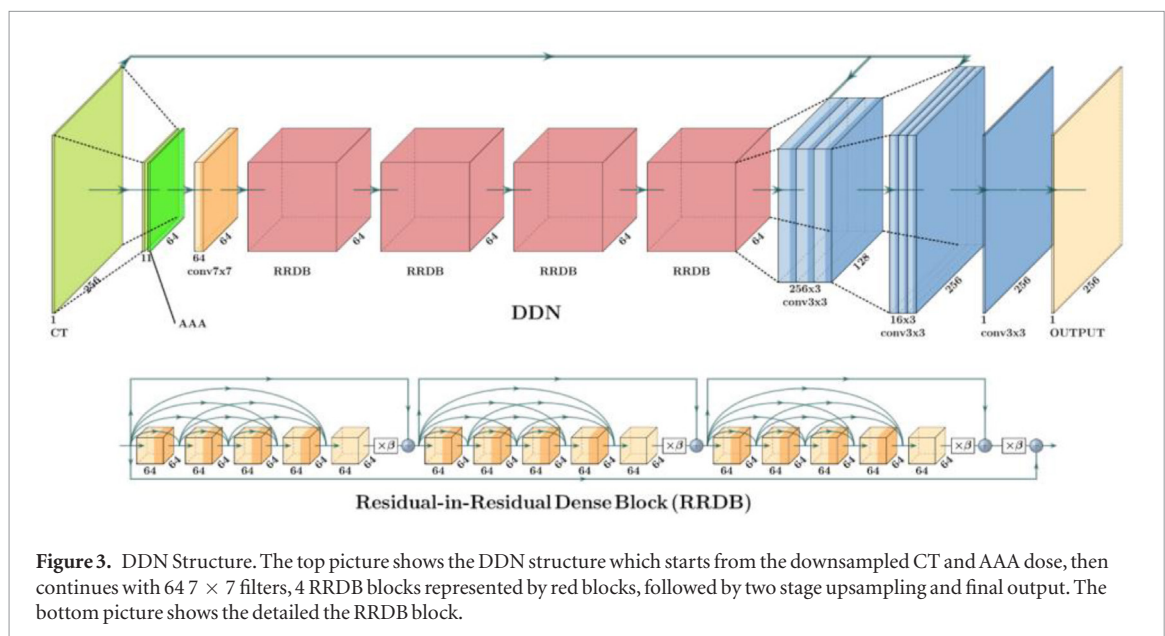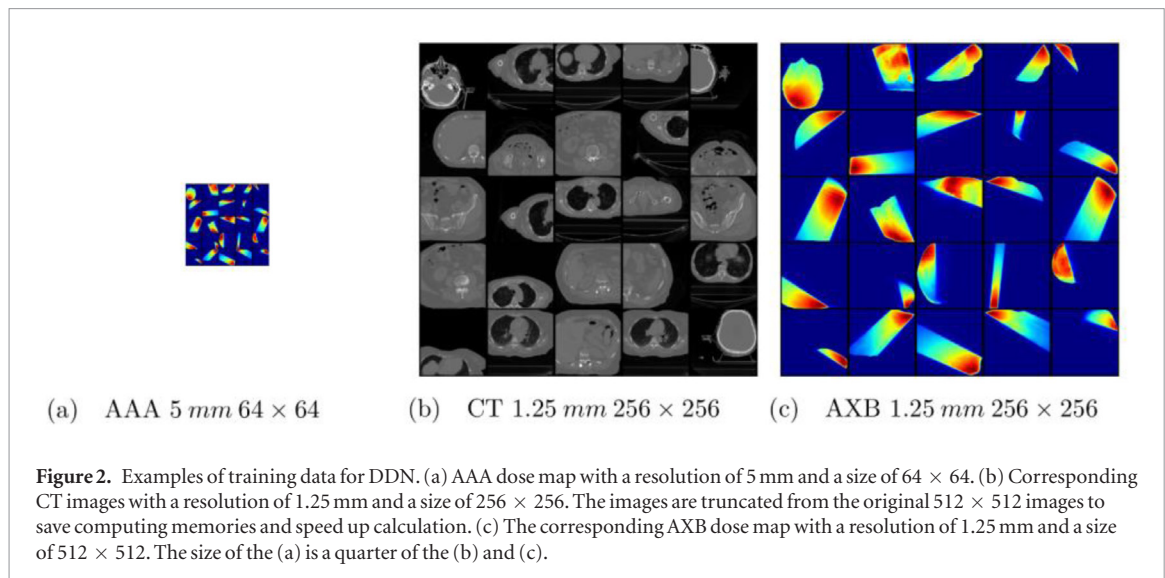
2.3.2. DDN structure

Our DDN (figure 3) model adapts from RESNET, DENSENET (Wang *et al* 2018) and SRCNN. It has a horizontal hourglass shape and can be roughly split into three parts: downscale, feature extraction, and upscale. In the beginning, the 256×256 CT slice is linear interpolated to 64×64 and combined with the AAA slice of the same resolution. $64\ 7 \times 7$ convolution filters take the combined AAA and CT as input and feed the output to four residuals in the residual dense blocks (RRDB) (Wang *et al* 2018). The RRDB block is very good at ensuring the input can flow freely to the output in very deep network structures. A RRDB block possesses three parts, each has five $64\ 3 \times 3$ filters where the output from each filter is sent to each of the following 3×3 filters as additional input. The original input is also sent directly to the final stage to be summed up with an output scaled by $\beta = 0.2$ (Wang *et al* 2018). That is, the final output of the RRDB is a summation of the original input and the final output scaled by β. In the upscaling part, the images are firstly resized to 64×64 with linear interpolation and combined with a 64×64 CT slice. This then passes through three consecutive $256\ 3 \times 3$ convolution filters (256×3, 3×3). The output is resized to 256×256 and combined with a CT slice in the original resolution (1.25 mm). It is followed by convolution filters with a size of (16×3, 3×3). The images go through a final single 3×3 convolution layer and exit as a single slice with a resolution of 256×256, which is then added with the linearly interpolated 256×256 AAA. This sum is our final output. We use leaky ReLu activation for all convolution layers except the last layer with ReLu activation, which forces all values in the output to be positive.
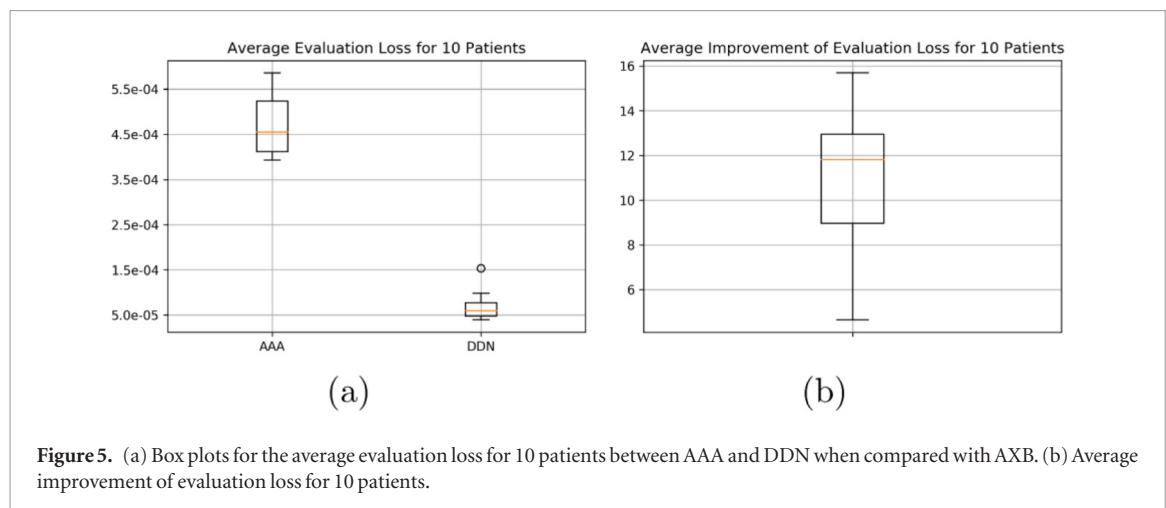
2.3.3. Training calculation

The performance of the DDN model is evaluated extensively. For an evaluation patient, we used the datasets of the rest 9 as the training data and compare the model output with the known dose distribution (i.e. the ground truth) of the patient quantitatively. The evaluation data were unseen during the model training process, thus the trained DDN model is not influenced by the evaluation data. Altogether, 10 DDN models were obtained. At the start of a model training process, we reinitialized the model parameters with He norm (He *et al* 2015). The same training parameters were employed. MSE between the pixels of the training data set was used as the loss function. The optimization uses ADAM (Le *et al* 2011) algorithm with a learning rate of 10^{-4} and takes ~100 Epochs of training with a batch of 8. The learning rate of 10^{-4} is selected to ensure fast converging while keeping the gradient explosion from happening. The computation was done on a Nvidia DGX system. We use the framework of TensorFlow 1.12 (Abadi *et al* 2015). Each epoch takes ~20 min.

3. Results

Figure 4 shows the loss values of the training sets (black dashed curves) and the validation sets (colored curves) as a function epoch. It is seen that the training error starts to converge around the 100th epoch and reaches a minimum average loss value of 2.6×10^{-5}, compared with an average loss at 4.7×10^{-4} before the training. The value of loss function for the evaluation set shows a large variation among patients with different treatment sites,

Figure 2. Examples of training data for DDN. (a) AAA dose map with a resolution of 5 mm and a size of 64 × 64. (b) Corresponding CT images with a resolution of 1.25 mm and a size of 256 × 256. The images are truncated from the original 512 × 512 images to save computing memories and speed up calculation. (c) The corresponding AXB dose map with a resolution of 1.25 mm and a size of 512 × 512. The size of the (a) is a quarter of the (b) and (c).

Figure 3. DDN Structure. The top picture shows the DDN structure which starts from the downsampled CT and AAA dose, then continues with 64 7 × 7 filters, 4 RRDB blocks represented by red blocks, followed by two stage upsampling and final output. The bottom picture shows the detailed the RRDB block.

Figure 4. Training and evaluation curves for DDN.

Figure 5. (a) Box plots for the average evaluation loss for 10 patients between AAA and DDN when compared with AXB. (b) Average improvement of evaluation loss for 10 patients.

with the worst evaluation loss occurring for the brain case at 1.5×10^{-4} and the best evaluation loss occurring for one of the thoracic case at 4.0×10^{-4}. As shown in the box plot of figure 5(a)), the average evaluation loss is 4.5×10^{-4} for linearly upsampled AAA and 5.2×10^{-4} for DDN. The average improvement of evaluation loss is ~12 times when all cases are considered.

We linearly interpolated the AAA dose distribution of 5 mm pixel size to 1.25 mm resolution. For each of the 10 models, the up-sampled dose distribution was compared with the AXB calculation with 1.25 mm resolution using 2D Gamma analysis for each slice. The same was performed for the DDN-derived dose distribution. We found that the average Gamma index passing rate at 3mm3% improved from 76% between the upsampled AAA and AXB to 91% between the DDN and AXB, from 66% to 86% at 2mm2%, and from 40% to 70% at 1mm1%. For the brain case, the 3mm3% passing rate improved from 69.9% to 89.0%, and for the pelvis case, 77.7% to 91.9%. The average calculation time of DDN was less than 1 milliseconds for a single slice. Figure 6(a) shows the boxplot summarizing the above statements and figure 6(b) shows the violin plot of the gamma passing rate for the three different gamma indices.
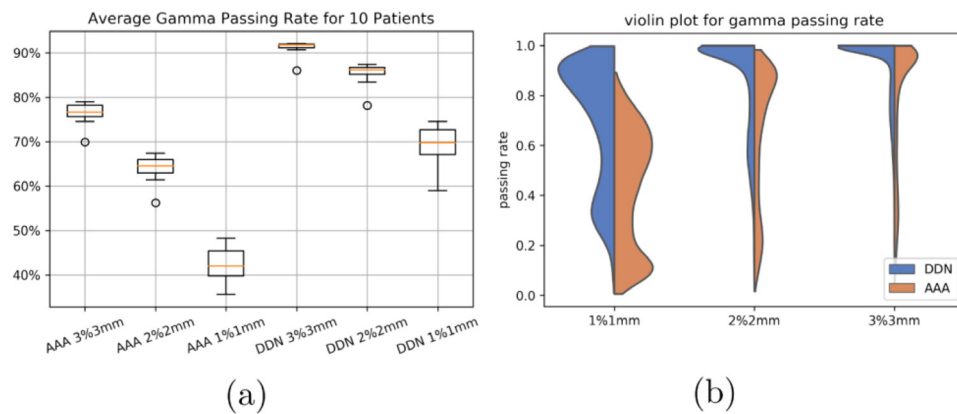
It is also intriguing that, as can be seen from figure 7, the DDN inference yields sharper beam edges as compared with that of AAA. The results are more consistent with the AXB calculations. Moreover, the DDN also reconciles the tissue density heterogeneities along the beam path and adjusts the dose accordingly. In general, a better agreement between the AXB and DDN calculations in the low-density lung regions and in the high-density bones is clearly demonstrated in figure 7.

On average, it takes ~5 h to calculate the 1.25 mm resolution AXB dose for 100 radiation therapy beams on our clinical 7 xenon servers with 6–8 cores each. The same calculations using AAA algorithm takes about 3 min when a pixel size of 5 mm is selected. The proposed DDN method presents orders of magnitude improvement in computational efficiency with only limited (and controllable) compromise in the accuracy of resultant dose distribution. Indeed, the DDN super-resolution calculation for the same 100 beams would take only several seconds. The longer AXB calculation time is partly due to the inefficiency of the AXB algorithm to calculate the field dose rather than the plan dose. The plan dose calculation treats multiple fields in a plan as a whole and it is not possible to get the dose contribution from each field separately, while the field dose option calculates each field in a plan independently. Ideally, we should compare with the performance of the AXB algorithm on a GPU, however, we do not have access to one of the GPU systems. There are published data showing a speed up of 4.4 times for a GPU AXB implementation, which can have memory problems for field areas of 300 cm² or below 0.15 cm grid resolution (Aland *et al* 2019).
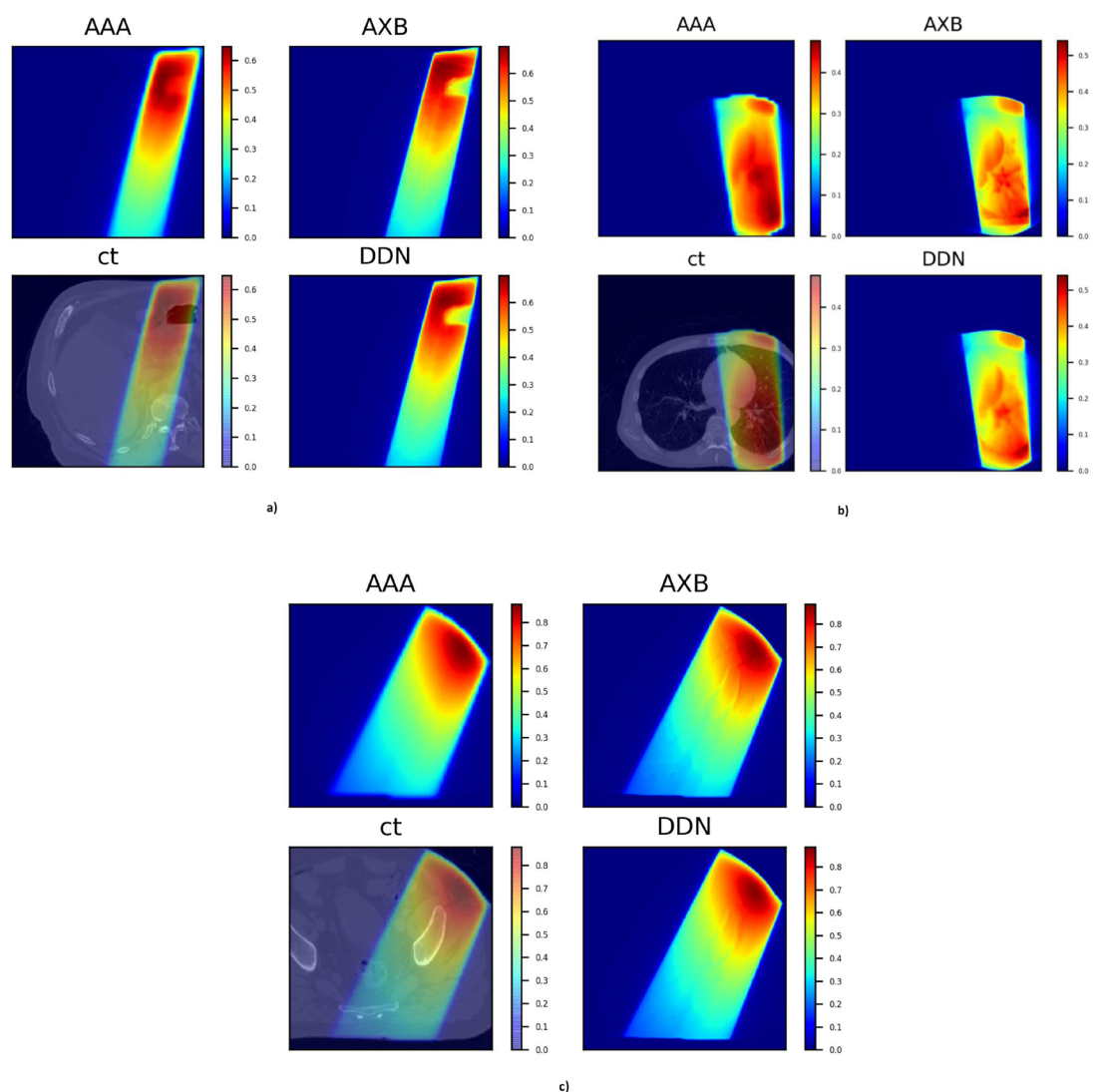
4. Discussions

The problem of single image super resolution is generally ill posed (Yang *et al* 2010), with multiple solutions for the low-resolution image. Thus, the approach cannot be directly applied to dose super resolution. Effective incorporation of CT information into the DDN helps to alleviate the problem and facilitate the dose transformation calculation. In DDN, the CT image and low-resolution dose distribution are fed into the neural network simultaneously. Note that the CT information is reintroduced at each layer of the network, forcing the network to learn from CT when forming feature maps. The examples shown in figure 7 suggest that the DDN picks up the missing density information from the CT images. Generally, a high-performance image super resolution network structure would work better for the DDN with inclusion of the CT information.

We noticed that the evaluation losses for brain and pelvis cases are slightly worse than other sites, presumably due to less abundant training datasets with similar anatomies. By including more training data with more
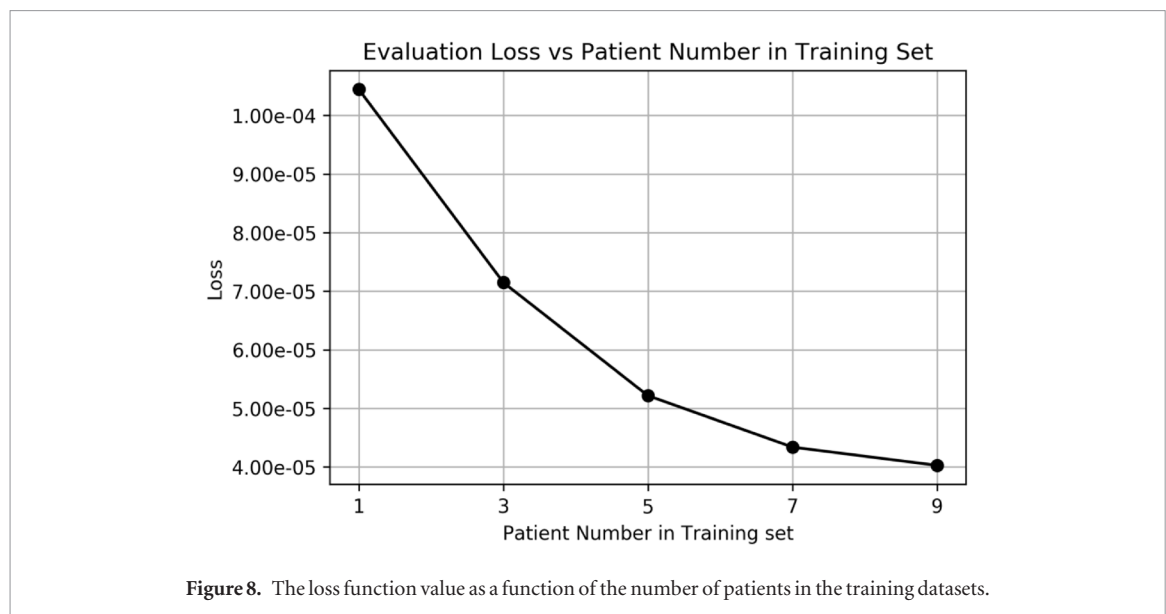
Figure 6. Gamma passing rate of the AAA and DDN methods as compared with that of the AXB at 1%1mm, 2%2mm, and 3%3mm. DDN method achieves markedly improved results.

Figure 7. Three typical dose distributions in the validation study obtained using the DDN. For visualization/comparison, the AAA dose overlaid on CT, AAA and AXB doses are also displayed.

anatomical scenarios included, the performance of DDN is expected to improve. As shown in figure 8, the evaluation loss for a lung case improves steadily as more patients are included in the training dataset.

Our study here is focused on beams with an energy of 10FFF. Compared to lower energy beams, larger discrepancy has been reported between the doses calculated by AAA and AXB (Huang *et al* 2015). In reality, it is required to train a separate model for each energy. However, the procedure presented in the above is quite

Figure 8. The loss function value as a function of the number of patients in the training datasets.

broad and generally applicable to any other beam energy or even other modality such as proton therapy and brachytherapy. A study of using super resolution dose approach for proton therapy is still in progress and will be reported in the future.

While the speed of current dose calculation algorithms is improving steadily over the years, the development seems to be outpaced by the increasing demands for higher spatial resolution for SBRT/SRS and treatment of small lesions such as oligomets and for the emerging plan optimization techniques such as MCO and robust optimization. Our study has clearly shown that significant acceleration of dose calculation is achievable by up-sampling the dose with deep learning techniques. The data-driven DDN provides an efficient tool for obtaining accurate dose distribution with high spatial resolution without paying the widely recognized overhead associated with the high-performance dose computing. Finally, we note that the DDN is not yet a full-fledged dose calculation engine. Future directions of improvement might include inferring from 3D CT and AAA data directly, expanding the training dataset to include more CTs, including more beams with different energies and machines, and replacing AAA dose with a simple ray tracing to further speed up the total dose calculation.

5. Conclusion

We have proposed a novel deep learning strategy to obtain highly accurate dose plan by transforming from a dose distribution calculated using a low-cost algorithm or algorithmic setting. Although the feasibility of the approach is demonstrated by using AAA and AXB, the method is quite general and can easily be applied to other algorithms, such as Monte Carlo based calculations. It is important to emphasize that the speed of the DDN does not depend on the algorithm it mimics, only the resolution. This makes it an ideal acceleration add-on for any dose calculation engine to speed up the treatment planning process. Dose super resolution represents a fertile area of research and development in medical physics to harness the enormous potential of deep learning.

Acknowledgment

References

Abadi M *et al* 2016 TensorFlow: Large-scale machine learning on heterogeneous systems (arXiv:1603.04467)
Aland T, Walsh A, Jones M, Piccini A and Devlin A 2019 Accuracy and efficiency of graphics processing unit (GPU) based Acuros XB dose calculation within the varian eclipse treatment planning system *Med. Dosim.* **44** 219–25
Andreo P 1991 Monte Carlo techniques in medical radiation physics *Phys. Med. Biol.* **36** 861–920
Chu M, Zinchenko Y, Henderson S G and Sharpe M B 2005 Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty *Phys. Med. Biol.* **50** 5463–77
Craft D L, Hong T S, Shih H A and Bortfeld T R 2012 Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy *Int. J. Radiat. Oncol.* **82** e83–90
Dong C, Loy C C, He K and Tang X 2016 Image super-resolution using deep convolutional networks *IEEE Trans. Pattern Anal. Mach. Intell.* **38** 295–307

Dong P and Xing L 2019 DoseNet: a deep neural network for accurate dosimetric transformation between different spatial resolutions and/ or different dose calculation algorithms for precision radiation therapy *Med. Phys.* **46** E473

Dong P, Liu H and Xing L 2018 Monte Carlo tree search-based non-coplanar trajectory design for station parameter optimized radiation therapy (SPORT) *Phys. Med. Biol.* **63** 135014

Failla G A, Wareing T, Archambault Y and Thompson S 2010 *Acuros XB Advanced Dose Calculation for the Eclipse Treatment Planning System* (Palo Alto, CA: Varian Medical Systems)

Fauw J D *et al* 2018 Clinically applicable deep learning for diagnosis and referral in retinal disease *Nat. Med.* **24** 1342

Gatys L A, Ecker A S and Bethge M 2016 Image style transfer using convolutional neural networks *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* pp 2414–23

Han T, Mikell J K, Salehpour M and Mourtada F 2011 Dosimetric comparison of Acuros XB deterministic radiation transport method with Monte Carlo and model-based convolution methods in heterogeneous media *Med. Phys.* **38** 2651–64

He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *IEEE Conf. on Computer Vision and Pattern Recognition (Las Vegas, NV, 27–30 June 2016)* pp 770–8

He K, Zhang X, Ren S and Sun J 2015 Delving deep into rectifiers: surpassing human-level performance on imagenet classification *Proc. IEEE Int. Conf. on Computer Vision (Santiago, 7–13 December 2015)* pp 1026–34

Huang B, Wu L, Lin P and Chen C 2015 Dose calculation of Acuros XB and anisotropic analytical algorithm in lung stereotactic body radiotherapy treatment with flattening filter free beams and the potential role of calculation grid size *Radiat. Oncol.* **10** 53

Kim J, Lee J K and Lee K M 2016 Accurate image super-resolution using very deep convolutional networks *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, 27–30 June 2016)* pp 1646–54

Le Q V, Ngiam J, Coates A, Lahiri A, Prochnow B and Ng A Y 2011 On optimization methods for deep learning *Proc. of the 28th Int. Conf. on Machine Learning* pp 265–72

LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44

Ledig C *et al* 2016 Photo-realistic single image super-resolution using a generative adversarial network (arXiv:160904802 Cs Stat)

Liu H, Xu J, Wu Y, Guo Q, Ibragimov B and Xing L 2018 Learning deconvolutional deep neural network for high resolution medical image reconstruction *Inf. Sci.* **468** 142–54

Lutz W, Winston K R and Maleki N 1988 A system for stereotactic radiosurgery with a linear accelerator *Int. J. Radiat. Oncol.* **14** 373–81

Ma M, Buyyounouski M K, Vasudevan V, Xing L and Yang Y 2019a Dose distribution prediction in isodose feature-preserving voxelization domain using deep convolutional neural network *Med. Phys.* **46** 2978–87

Ma M, Kovalchuk N, Buyyounouski M K, Xing L and Yang Y 2019b Dosimetric features-driven machine learning model for DVH prediction in VMAT treatment planning *Med. Phys.* **46** 857–67

Mackie T R, Scrimger J W and Battista J J 1985 A convolution method of calculating dose for 15 MV x rays *Med. Phys.* **12** 188–96

Mutic S and Dempsey J F 2014 The ViewRay system: magnetic resonance–guided and controlled radiotherapy *Semin. Radiat. Oncol.* **24** 196–9

Onishi H *et al* 2011 Stereotactic body radiotherapy (SBRT) for operable stage I non-small-cell lung cancer: can SBRT be comparable to surgery? *Int. J. Radiat. Oncol.* **81** 1352–8

Radford A, Metz L and Chintala S 2015 Unsupervised representation learning with deep convolutional generative adversarial networks (arXiv:151106434 Cs)

Rogers D W O, Faddegon B A, Ding G X, Ma C-M, We J and Mackie T R 1995 BEAM: a Monte Carlo code to simulate radiotherapy treatment units *Med. Phys.* **22** 503–24

Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* pp 234–41

Szegedy C, Ioffe S, Vanhoucke V and Alemi A 2017 Inception-v4, inception-ResNet and the impact of residual connections on learning *31st AAAI Conf. on Artificial Intelligence (AAAI'17)* pp 4278–84

Wang X *et al* 2018 Esrgan: enhanced super-resolution generative adversarial networks *Proc. of the European Conf. on Computer Vision (ECCV)*

Xing L, Krupinski E A and Cai J 2018 Artificial intelligence will soon change the landscape of medical physics research and practice *Med. Phys.* **45** 1791–3

Yang J, Wright J, Huang T S and Ma Y 2010 Image super-resolution via sparse representation *IEEE Trans. Image Process.* **19** 2861–73