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

Deep neural network-based prediction of synthetic dual-energy X-ray fluoroscopic images: a feasibility study

Jiaoyang Wang[1], Kei Ichiji[1], Noriyasu Homma[1], Xiaoyong Zhang[2] and Yoshihiro Takai[3]
[1] Tohoku University, Sendai, Japan [2] National Institute of Technology, Sendai College, Sendai, Japan [3] Southern Tohoku BNCT Research Center, Koriyama, Japan
TOHOKU UNIVERSITY

INTRODUCTION

Tumor localization is a fundamental task to appropriately deliver the treatment beam to the target tumor in radiation therapy (RT). With conventional single-energy kV X-ray fluoroscopy, tumor localization accuracy is often degraded by low target visibility.

Dual-energy (DE) X-ray imaging can enhance the contrast and visibility of the target tumor and improve localization accuracy[1-3]. However, DE X-ray fluoroscopy is not yet clinically available for radiation therapy since the existing DE imaging requires special hardware and extra imaging dose.

The emerging deep-learning (DL) technique can be a solution to emulate DE X-ray imaging in the current RT system since DL has the potential to learn the complex and non-linear relationship in the X-ray images with different energies[4]. i.e., DL possibly predicts DE X-ray images from a single-energy X-ray image without the special hardware and extra imaging dose and improves the tumor localization accuracy.

PURPOSE

To develop a deep-learning technique to predict synthetic dual-energy (DE) X-ray subtraction images from single-energy kV X-ray images for improving tumor localization accuracy during radiation therapy..

METHOD AND MATERIALS

Data preparation Initially 36,000 training datasets of artificial dual-energy X-ray fluoroscopic images of lung area were generated by using XCAT phantom at two energy levels with different imaging angles (θ, ϕ, γ) and regions to simulate possible variations in imaging conditions for RT (Fig 1).

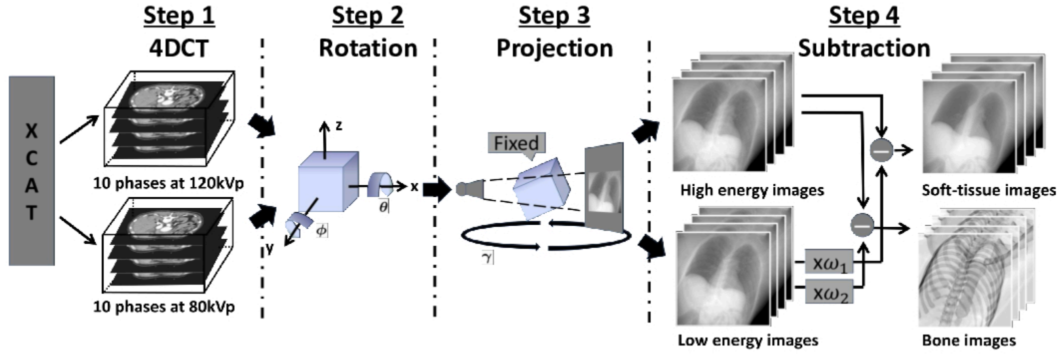

Fig 1. Overview of data preparation.

Building DL models

Recurrent residual convolutional neural networks based on U-Net[5][6] were then trained to respectively predict DE subtraction images of bone and soft tissue from single-energy X-ray images (Fig. 2).

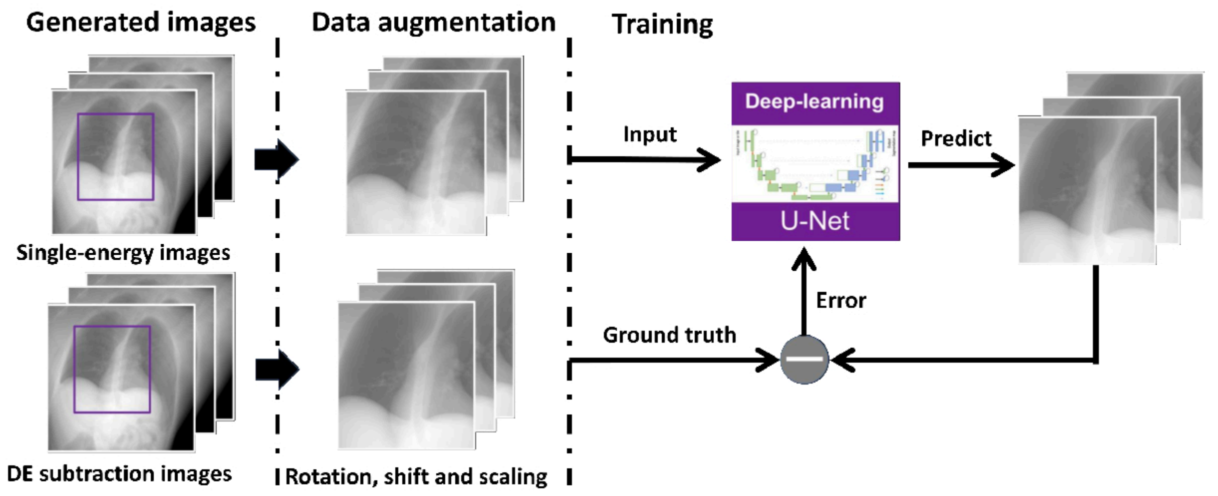

Fig 2. Schematic of model building process.

Evaluation

The trained DL models were tested on 5,000 test datasets, and the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) values were calculated for the predicted DE subtraction images. For further RT application, we also evaluated tumor tracking performance using the predicted soft tissue images.

Detail on training and testing datasets for predicting DE subtraction images

- Training datasets: 36,000 images of three sets of X-ray images. Imaging angles: $(\theta, \phi) = (0°, 0°), (1°, 1°), \dots, (35°, 35°)$, γ is from 3.6° to 360°
- Validation datasets: 9,000 images of three sets of X-ray images. Imaging angles: (θ, ϕ) are in random angle combinations, γ is from 3.6° to 360°
- Testing datasets: 5,000 images of three sets of X-ray images. Imaging angles: (θ, ϕ) are in random angle combinations, γ is from 3.6° to 360°

Detail on tracking performance evaluation

- Performance improvement of target localization was evaluated by template matching based on normalized cross-correlation (NCC) on predicted soft-tissue images.

RESULT AND DISCUSSION

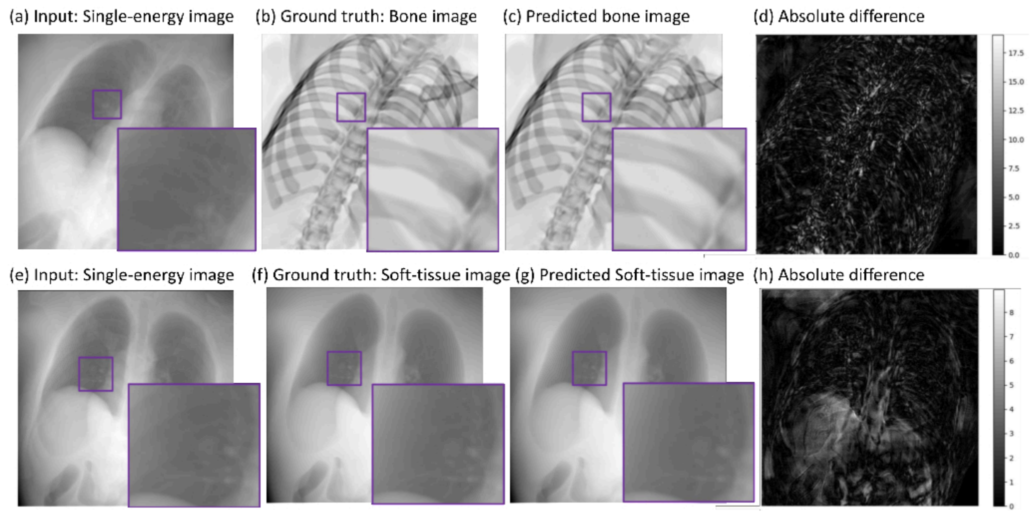

Fig 3. Examples of predicted DE subtraction.

** Imaging angles of input image (a) and (e) are ($\theta = 25°, \phi = 21°, \gamma = 86.4°$) and ($\theta = 20°, \phi = 32°, \gamma = 111.6°$) respectively.*

Table 1. Evaluation results of the predicted DE subtraction images.

Imaging angle* (θ and ϕ)	PSNR [dB]**	SSIM***
(9°,28°)	30.1438	0.9612
(17°,5°)	32.5120	0.9695
(20°,32°)	32.2932	0.9685
(25°,21°)	36.1355	0.9775
(27°,8°)	29.3343	0.9555
(9°,28°)	40.1683	0.9933
(17°,5°)	42.1068	0.9931
(20°,32°)	42.1361	0.9935
(25°,21°)	44.0892	0.9936
(27°,8°)	39.8684	0.9922

** Gantry angle γ: 3.6 ° to 360 ° at interval of 3.6 °*
*** The maximum value of PSNR is about 50[dB]*
**** SSIM = 1 means they have same structure*

- The predicted dual-energy X-ray images were visually similar to the ground truth images
- They had almost the same structure and sharpness of edge for bones, and rib bones in soft-tissue images were effectively suppressed.

- Most of PSNRs >30 dB and SSIM>0.95.
- The quality of the predicted bone and soft tissue images were reasonably high.

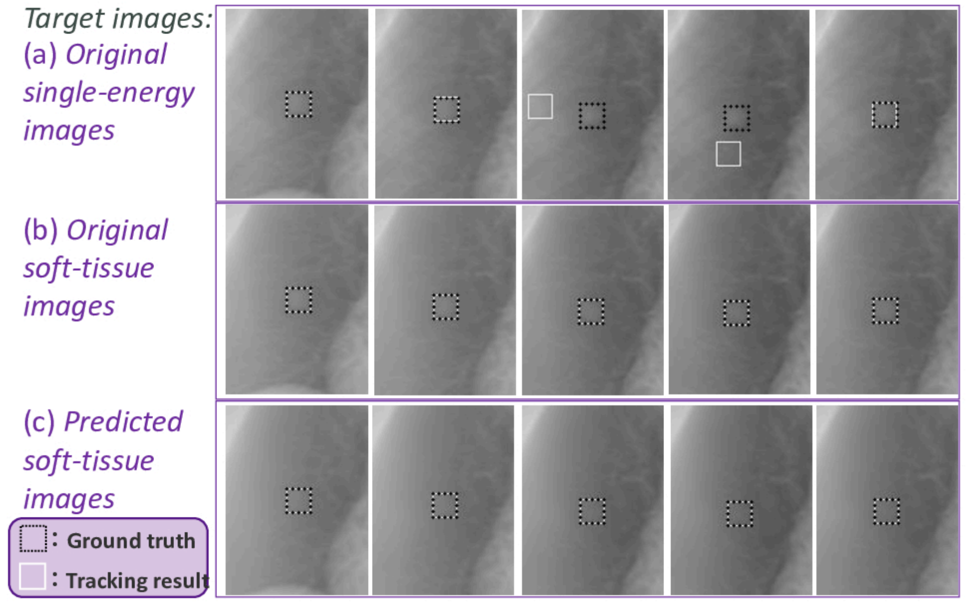

Fig 4. Examples on target tumor tracking.

Table 2. Evaluation results of the predicted DE subtraction images.

	Mean tracking error (pixel)
(a)	12.3859 ± 18.8658
(b)	0.5196 ± 0.5879
(c)	0.6396 ± 0.5713

- Fig. 4 and Table 2 shows the tracking errors at $\theta = 0°, \phi = 3°, \gamma = 285°$.
- A major improvement can be found for the predicted soft-tissue images as tracking target images.

CONCLUSION

The proposed method has successfully demonstrated the feasibility of DE X-ray image prediction from single-energy image, and potentially assist target tumor localization for more precise radiation therapy without special hardware.

Contact wang@rii.med.tohoku.ac.jp (Jiaoyang Wang)

REFERENCES

[1] **D Jennifer et al**. Radiotherapy and Oncol., 117(3): 487-490. (2015)
[2] **N.A Shkumat et al**. Med. Phys., 35: 629-632. (2008)
[3] **M Haytmyradov et al**. Med. Phys., 46: 3235-3244. (2019)
[4] **D Lee et al**. Phys Med Biol. 2019 May 31;64(11):115017. (2019)
[5] **O Ronneberger et al**.. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. (2015)
[6] **M. Z Alom et al**. arXiv:1802.06955v5 [cs.CV]. (2018)