Voxel-based Internal Dosimetry Using Deep Residual Learning

Keon Min Kim[1,2,3], Min Sun Lee[4], Min Seok Suh[5,6], Gi Jeong Cheon[5,6], and Jae Sung Lee[1,2,3,5,6,*]
[1]Interdisciplinary Program in Bioengineering, Seoul National University Graduate School, Seoul 03080, South Korea
[2]Integrated Major in Innovative Medical Science, Seoul National University Graduate School, Seoul 03080, South Korea
[3]Artificial Intelligence Institute, Seoul National University, Seoul 08826, South Korea
[4]Environmental Radioactivity Assessment Team, Nuclear Emergency & Environmental Protection Division, Korea Atomic Energy Research Institute, Daejeon 34057, Korea.
[5]Department of Nuclear Medicine, Seoul National University College of Medicine, Seoul 03080, South Korea
[6]Institute of Radiation Medicine, Medical Research Center, Seoul National University, Seoul 03080, South Korea
* Corresponding author: jaes@snu.ac.kr

Abstract

In this study, we propose a deep learning-based voxel-based dosimetry method in which dose maps acquired using the multiple voxel S-value approach were used for residual learning. 22 SPECT/CT images from seven patients who underwent ^{177}Lu-DOTATATE treatment were used in this study. The dose maps generated from Monte Carlo simulations were used as the reference approach and target images for network training. The multiple VSV approach was used for residual learning and compared with dose maps generated from deep learning. The conventional 3D U-Net network was modified for residual learning. DL-based dosimetry provided a slightly more accurate estimation than the multiple-VSV approach, but the results were not statistically significant. This difference was prominent in the error maps. The multiple VSV approach underestimated doses in the low-dose range, but it accounted for the underestimation when the DL-based approach was applied. In conclusion, the proposed deep learning network is useful for accurate and fast dosimetry after radiation therapy using ^{177}Lu labeled radiopharmaceuticals.

1. Background

^{177}Lu-octreotate (^{177}Lu-DOTATATE) peptide receptor radionuclide therapy is used to treat patients with metastatic neuroendocrine tumors (NETs). Since the peptide receptor has high efficacy for tumor targeting and the therapy has low side effects and good therapeutic efficacy. However, retrospective dosimetry is needed to ensure efficient and accurate treatment.

Several approaches for voxel-based dosimetry using DL has been reported [1-4]. In this study, a U-net-based network for voxel-based dosimetry was proposed to enhance accuracy. In particular, dose maps acquired using the multiple VSV approach were used for residual learning. The proposed network was validated by comparison with MC simulations and the multiple VSV approach.

2. Materials and Methods

2-1. Image acquisition

The 22 sets of SPECT/CT data from 7 patient who underwent ^{177}Lu-DOTATATE therapy at Seoul National University Hospital were used in this study. SPECT/CT images were acquired 4, 24, 48, and 120 h after intravenous injection of ^{177}Lu-DOTATATE.

2-2. MC simulation

The dose maps generated from the MC simulation were used as the reference and target images for network training. Geant4 Application for Emission Tomography (GATE) v.8.2 was used for the simulation. The CT images were resampled to have same voxel sizes as that of the SPECT image. Co-registration of sequential SPECT/CT images was performed to generate time-integrated activity (TIA) maps. The TIA maps were acquired using a voxel-wise trapezoidal sum as:

$$\tilde{A} = \sum_{i=0}^{3} \frac{1}{2}(A_i + A_{i+1})\Delta t_i \ + \int_{t_4}^{\infty} A_4 e^{-\lambda t} dt \qquad \text{Eq. 1}$$

, where A_i is the activity (A_0=0) in each voxel of i-th SPECT images acquired at t_i, $\Delta t\_i = t_{i+1} - t_i$, and λ is the physical decay factor of ^{177}Lu. Patient-specific phantom images derived from the CT image and TIA map for the voxelized source were used as input files for the simulation

2-3. Multiple VSV approach

The multiple VSV approach was used for residual learning and comparison compared with dose maps generated from DL. In this study, the multiple VSV approach with 20 kernels was used for dosimetry. The single-VSV approach has also been used for dosimetry. By contrast, the multiple-VSV approach comprises four steps: (1) generation of VSV kernels using MC simulation, (2) convolution of the TIA map with each VSV kernel, (3) masking each medium-specific dose map with the corresponding medium mask map, and (4) summation of all masked dose maps to generate the final dose map.

2-4. DL approach

The conventional 3D U-net network was modified for residual learning, as illustrated in **Fig. 1**. 3D patch-based supervised learning was performed by considering the size of the network, which took CT and TIA patch images with a size of 64 × 64 × 64 as inputs, and was trained to yield a dose map. Each encoding layer comprised two 3 × 3 × 3 convolution layers and a 2 × 2 × 2 max-pooling layer. The decoding layer comprised one 3 × 3 × 3 convolution layer, a concatenating path, and two 3 × 3 × 3 convolution layers. Each convolution layer was followed by batch normalization and rectified linear unit (ReLU) activation function. The number of feature maps in the first layer was set to 16 via empirical fine-tuning. It was then doubled as it passed through the encoding layer, and reduced by half as it passed through the decoding layer. After passing through all the

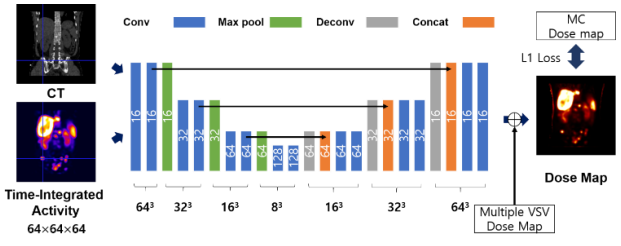

Fig. 1. Modified 3D U-net structure for residual learning

layers, the feature maps were contracted to one image by a 1 × 1 × 1 convolution, and the image was then summated with a multiple VSV dose map for the residual learning. The multiple VSV dose map using 20 VSV kernels was used.

The network was trained and tested using a cross-validation strategy with 7-fold, the same as the number of patients. Therefore, all 22 SPECT/CT data set were used for the training and validation of the network. Each SPECT and CT dataset was split into 125 patches of size 64 × 64 × 64, forming one dataset for training and testing. The L1 loss between the dose map from the MC simulation and the network was estimated for training. An adaptive moment optimizer was used to minimize loss.

3. Results

The absorbed doses of organs estimated using different approaches are summarized in **Table 1**. DL-based dosimetry provided a slightly more accurate estimation than the multiple VSV approach but was not statistically significant. The single VSV approach especially yielded high error for bone marrow and tumor regions where the medium properties of the regions were different than that of water.

Table 1. Absorbed Dose (AD, Gy) Estimated using 4 different approaches (Mean ± Standard Deviation) and Mean Absolute Error (MAE, %) compared to Monte Carlo simulation

AD(Gy) &MAE(%)		Kidneys	Bone marrow	Liver	Spleen	Tumors
MC	Dose	6.41 ± 1.34	1.76 ± 1.61	10.00 ± 7.29	6.25 ± 3.04	15.28 ± 12.63
Single VSV	Dose	6.75 ± 1.40	1.99 ± 1.82	10.69 ± 7.80	6.62 ± 3.32	16.81 ± 13.36
	MAE	5.45	11.84	6.82	9.30	13.62
Multiple VSV	Dose	6.36 ± 1.32	1.78 ± 1.64	9.76 ± 7.05	6.18 ± 2.96	15.15 ± 12.10
	MAE	1.56	1.21	2.79	2.28	4.66
DL	Dose	6.46 ± 1.36	1.77 ± 1.62	10.02 ± 7.28	6.26 ± 3.05	15.54 ± 12.64
	MAE	1.04	0.81	0.54	1.34	3.18

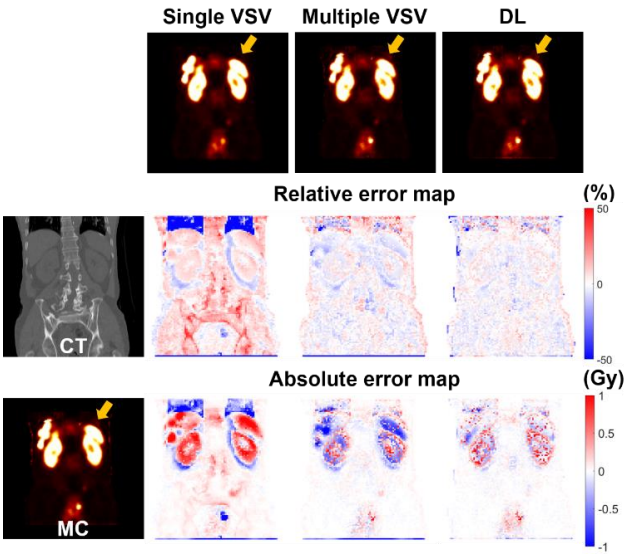

Fig. 2. Relative and absolute error maps of single VSV, multiple VSV, and DL-based approaches compared to Monte Carlo simulation.

Fig. 2 shows the dose maps generated using each approach and corresponding relative and absolute error maps from SPECT/CT images of a 74-year-old male patient diagnosed with rectal NET. The dose maps generated by the single VSV approach showed large errors in the regions. On the other hand, the multiple VSV and DL-based approaches provided similar dose maps compared to the MC simulation as shown in the difference maps. However, the errors in voxel level were minimal at the lung-liver interface regions, kidneys, spleen, and tumors in the liver when the DL-based approach was used.

Fig. 3 shows dose-line profiles of each approach acquired using

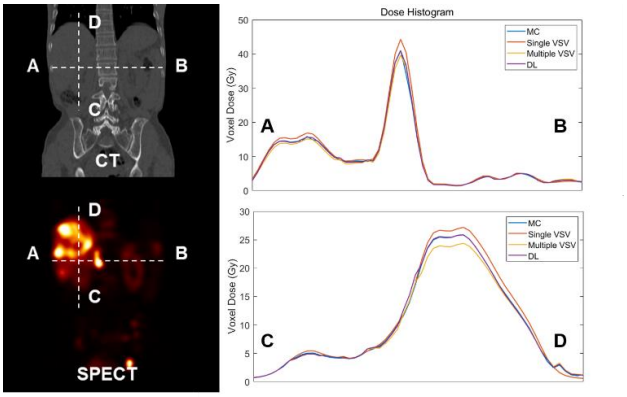

Fig. 3. The dose-line profile of MC, single VSV, multiple VSV, and DL-based approaches.

SPECT/CT images of a 63-year-old male patient diagnosed with rectal NET. As shown in the dose-line profile along the horizontal direction, the DL-based approach was almost identical to MC simulation, especially in the bone metastasis region.

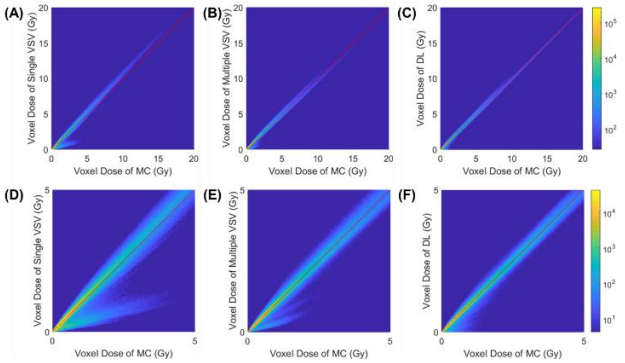

Fig. 4. The voxel-wise correlation between the MC and (A, D) single VSV approach, (B, E) the multiple VSV approach, and (C, F) the DL-based approach.

The voxel-wise correlations between MC and 3 different approaches were analyzed to observe the accuracy of each model at the voxel level. As shown in Fig. **4-A** and **D**, the single VSV approach overestimated the doses in the high-dose range (>5 Gy), and underestimated the doses in the low-dose range (< 5 Gy). The multiple VSV approach underestimated doses in the low-dose range, as shown in **Fig. 6-B** and **E** but it accounted for the underestimation when the DL-based approach was applied, as shown in **Fig. 6-C** and **F**. Furthermore, the DL-based approach had a slightly narrower plotting shape than that in the multiple VSV approach.

4. References

[1] Lee MS, Hwang D, Kim JH, Lee JS. Sci Rep. 2019;9(1):10308.

[2] Gotz TI, Schmidkonz C, Chen S, Al-Baddai S, Kuwert T, Phys Med Biol. 2020;65(3):035007.

[3] Akhavanallaf A, Shiri I, Arabi H, Zaidi H. Eur J Nucl Med Mol Imaging. 2021;48(3):670-82.

[4] Li Z, Fessler JA, Mikell JK, Wilderman SJ, Dewaraja YK. Med Phys. 2022;49(2):1216-30.