Generation of PET Attenuation Map for Whole-Body Time-of-Flight ^{18}F-FDG PET/MRI Using a Deep Neural Network Trained with Simultaneously Reconstructed Activity and Attenuation Maps

Donghwi Hwang[1,2], Seung Kwan Kang[1,2], Kyeong Yun Kim[1,2], Seongho Seo[3], Jin Chul Paeng[2,4], Dong Soo Lee[2,4,5], and Jae Sung Lee[1,2,4]

[1]Department of Biomedical Sciences, Seoul National University, Seoul, Korea; [2]Department of Nuclear Medicine, Seoul National University, Seoul, Korea; [3]Department of Neuroscience, College of Medicine, Gachon University, Incheon, Korea; [4]Institute of Radiation Medicine, Medical Research Center, Seoul National University, Seoul, Korea; and [5]Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Suwon, Korea

We propose a new deep learning–based approach to provide more accurate whole-body PET/MRI attenuation correction than is possible with the Dixon-based 4-segment method. We use activity and attenuation maps estimated using the maximum-likelihood reconstruction of activity and attenuation (MLAA) algorithm as inputs to a convolutional neural network (CNN) to learn a CT-derived attenuation map. **Methods:** The whole-body ^{18}F-FDG PET/CT scan data of 100 cancer patients (38 men and 62 women; age, 57.3 ± 14.1 y) were retrospectively used for training and testing the CNN. A modified U-net was trained to predict a CT-derived μ-map (μ-CT) from the MLAA-generated activity distribution (λ-MLAA) and μ-map (μ-MLAA). We used 1.3 million patches derived from 60 patients' data for training the CNN, data of 20 others were used as a validation set to prevent overfitting, and the data of the other 20 were used as a test set for the CNN performance analysis. The attenuation maps generated using the proposed method (μ-CNN), μ-MLAA, and 4-segment method (μ-segment) were compared with the μ-CT, a ground truth. We also compared the voxelwise correlation between the activity images reconstructed using ordered-subset expectation maximization with the μ-maps, and the SUVs of primary and metastatic bone lesions obtained by drawing regions of interest on the activity images. **Results:** The CNN generates less noisy attenuation maps and achieves better bone identification than MLAA. The average Dice similarity coefficient for bone regions between μ-CNN and μ-CT was 0.77, which was significantly higher than that between μ-MLAA and μ-CT (0.36). Also, the CNN result showed the best pixel-by-pixel correlation with the CT-based results and remarkably reduced differences in activity maps in comparison to CT-based attenuation correction. **Conclusion:** The proposed deep neural network produced a more reliable attenuation map for 511-keV photons than the 4-segment method currently used in whole-body PET/MRI studies.

Key Words: PET/MRI; attenuation correction; deep learning; simultaneous reconstruction

J Nucl Med 2019; 60:1183–1189
DOI: 10.2967/jnumed.118.219493

Although PET/MRI is an emerging hybrid imaging modality (*1–4*), the PET attenuation maps (μ-maps) used in whole-body PET/MRI studies have limited accuracy (*5,6*). In brain PET/MRI studies, attenuation correction has advanced greatly through the use of pseudo CT images generated by segmenting ultrashort echo time or zero echo time MR images and by registering an atlas generated from transmission or CT scans to individual data (*7–9*). However, these techniques used in brain PET/MRI have not been successfully applied in whole-body PET/MRI studies. The Dixon sequence–based 4-segment approach used as a standard method in whole-body PET/MRI systems (*10*) underestimates PET activity in bone structures because of the lack of bone information in attenuation maps (*11,12*). Although a model-based approach (*13*) has been suggested to add bone structures to 4-segment maps (*14*), inaccurate registration between PET images and bone models may cause errors in PET activity quantification.

Although algorithms for simultaneous reconstruction of activity and attenuation have been greatly improved by the incorporation of time-of-flight (TOF) information (*15,16*), their accuracy is still far from the clinically relevant level. The maximum-likelihood reconstruction of activity and attenuation (MLAA) algorithm, a simultaneous reconstruction algorithm, has the advantage of providing attenuation maps (*17,18*). MLAA-generated attenuation maps allow the use of image-domain priors to improve the algorithm accuracy and convergence (*19,20*). However, mainly because of the insufficient timing resolution of clinical PET systems, the MLAA suffers from slow convergence, high noise levels in attenuation maps, and crosstalk between activity and attenuation distribution (*21*).

Deep learning–based approaches have been suggested to improve the accuracy of regional PET attenuation correction. In our recent work (*22*), to mitigate the limitations of MLAA in brain PET, deep convolutional neural networks (CNNs) were trained to

Received Aug. 28, 2018; revision accepted Dec. 20, 2018.
For correspondence or reprints contact: Jae Sung Lee, Department of Nuclear Medicine, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea.
E-mail: jaes@snu.ac.kr
Published online Jan. 25, 2019.

learn a true CT-derived attenuation map with the MLAA activity and attenuation maps as their inputs. The CNNs generated less noisy and more uniform attenuation maps than original MLAA, resulting in only 5% errors in activity and binding ratio quantification in the most challenging brain PET cases for simultaneous image reconstruction (dopamine transporter imaging). Another notable application of deep learning for this purpose is CNN-based PET attenuation map generation from the zero echo time and Dixon MRIs (*23*). In pelvis PET/MRI studies, this multiparametric MRI–based approach reduced the PET quantification error in bone lesions by a factor of 4 in comparison to the conventional Dixon sequence–based 4-segment approach. However, this approach requires additional zero echo time MRI acquisition with a relatively long scan time.

In this study, we investigated the feasibility of the deep learning–based approach to whole-body PET/MRI attenuation correction without use of zero echo time or ultrashort echo time data. Similarly to our previous work on the brain PET studies (*22*), we use MLAA-based activity and attenuation maps as inputs to CNN to learn a CT-derived attenuation map. However, we conducted 3-dimensional (3D) patch-based learning instead of 2-dimensional slice-by-slice mapping because whole-body structures are more complex than head region and 3D learning allows better continuity of image intensity in the axial direction. The CNN was trained and tested using oncologic whole-body ^{18}F-FDG PET/CT scan data. Then, the similarity of the CNN-generated attenuation map and the CT-derived map was evaluated. Also, the attenuation-corrected PET images produced using the conventional 4-segment method and new CNN outcomes were compared with ground truth (CT).

MATERIALS AND METHODS

Subjects and Image Acquisition

The whole-body ^{18}F-FDG PET/CT scan data of 100 cancer patients (38 men and 62 women; age, 57.3 ± 14.1 y) acquired using a Biograph mCT 40 scanner (effective timing resolution, 580 ps; Siemens Healthcare) from March 2017 to May 2017 were retrospectively analyzed. The retrospective use of the scan data and waiver of consent were approved by the Institutional Review Board of our institute. For all patients, PET/CT imaging was performed 60 min after intravenous injection of ^{18}F-FDG (5.18 MBq/kg). The patients' upper bodies from head to upper thigh were covered by a 6- to 8-bed-position emission scan (scan time, 1 min/bed position).

Sinograms of prompt PET counts and correction factors were generated using the e7 toolkit. The CT images were reconstructed in a $512 \times 512 \times 100$ matrix with the voxel size of $1.52 \times 1.52 \times 2.03$ mm and converted into the μ-map for 511-keV photons ($200 \times 200 \times 109$; $4.07 \times 4.07 \times 2.03$ mm). We reconstructed PET datasets using the MLAA with the TOF information (6 iterations and 21 subsets, 5-mm gaussian postprocessing filter). The matrix size of MLAA-reconstructed images was $200 \times 200 \times 109$ ($4.07 \times 4.07 \times 2.03$ mm voxel size) for each bed position.

Simultaneous Activity and Attenuation Reconstruction Algorithm

In each iteration n, a standard maximum-likelihood expectation maximization algorithm updates activity distribution λ for a current attenuation coefficient μ as follows:

$$\forall i : a_i^n = e^{-\sum_j l_{ij}\mu_j^n}$$

$$\forall j : \lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_{it} a_i^n c_{ijt}} \sum_{it} a_i^n c_{ijt} \frac{y_{it}}{\sum_k a_i^n c_{ikt}\lambda_k^n + s_{it}},$$

where l_{ij} is the length of intersection between LOR i and voxel j, and c_{ijt} is the TOF system matrix element between LOR i, TOF bin t, and voxel j. Also, y is the measured emission projection and s is the additive correction factor that contains scatter and random events.

After updating activity distribution, the attenuation coefficient is updated by a maximum-likelihood-for-transmission-tomography algorithm as follows:

$$\forall i : \psi_i^n = a_i^n \sum_{jt} c_{ijt}\lambda_j^{n+1}$$

$$\forall j : \mu_j^{n+1} = \mu_j^n + \frac{\sum_i l_{ij}\frac{\psi_i^n}{\psi_i^n + s_i}\left(\psi_i^n + s_i - y_i\right)}{\sum_i l_{ij}\frac{\left(\psi_i^n\right)^2}{\psi_i^n + s_i}\sum_k l_{ik}}.$$

To resolve the nonunique global scaling problem in MLAA, the boundary constraint was applied during the attenuation image estimation in MLAA, following the original TOF MLAA paper (*17*).

Data Set

The data of 60 patients were used for training the CNNs, that from 20 others as a validation set to prevent overfitting, and the other 20 as a test set (Table 1) for the CNN performance analysis. For the CNN training and testing, activity and attenuation maps derived from MLAA were used as input to the CNN, and an attenuation map converted from CT (μ-CT) as label data (Fig. 1). We generated the attenuation map based on a 4-segment approach (μ-segment) from μ-CT by classifying each patient's body parts into 4 segments using 3 different thresholds applied to μ-CT. The thresholds between air, lung, fat, and water were 0.015, 0.04, and 0.09. The mean μ-value for each segment was then calculated from 60 μ-CT images and assigned to the body segment of each patient.

TABLE 1
Demographic Data of Patients Included in Test Set

Patient	Sex	Height (cm)	Weight (kg)	Diagnosis
1	M	172	70	Thymic cancer
2	F	161	53	Breast cancer
3	F	152	53	Lung cancer
4	F	152	41	Lymphoma
5	F	150	56	Lung cancer
6	F	160	49	Breast cancer
7	M	163	53	Lymphoma
8	M	159	55	Cholangiocarcinoma
9	M	175	76	Breast cancer
10	M	162	76	Oral cancer
11	M	159	54	Lung cancer
12	M	160	40	Osteosarcoma
13	F	164	66	Thyroid cancer
14	M	168	70	Lung cancer
15	F	150	52	Lung cancer
16	F	155	57	Breast cancer
17	F	153	63	Breast cancer
18	F	161	50	Breast cancer
19	M	170	62	Lung cancer
20	M	165	72	Oral cancer

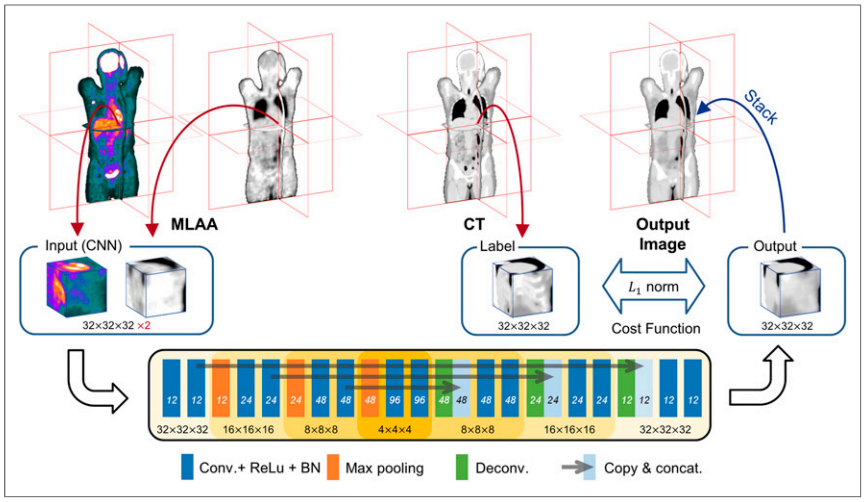

FIGURE 1. Data flow in generation of attenuation (μ) map using proposed deep neural network for whole-body PET/MRI attenuation correction and architecture of deep neural network. Deep neural network was trained to learn ground truth CT-derived μ-map from MLAA-derived activity and attenuation maps. In network architecture, each box represents feature map resulting from operation corresponding to its color. Number of features in each layer is indicated at bottom of box. Dimensions of each feature map are shown below box.

All inputs and labels for the CNN were in the form of $32 \times 32 \times 32$ matrix patches. The intensity of each activity patch was normalized to have the range between 0 and 1. To avoid including meaningless blank patches during the CNN training, the volume patches were used for training only if their centers were included in the body. With this inclusion criterion, approximately 1.3 million patches were used for CNN training.

Network Architecture

The network was designed to predict μ-CT from activity and attenuation maps derived from MLAA. The network consists of a contracting path to capture the context and a symmetric expanding path that enables precise localization (*24*). As shown in Figure 1, the left half of the proposed network (contracting path) includes repeating two 3D convolution layers with rectified linear units and batch normalization, and $2 \times 2 \times 2$ max-pooling layers for down sampling. Similarly, in the right half of the network (expanding path), two 3D deconvolution layers with rectified linear units and batch normalization are repeated. In every layer, the number of feature maps is doubled in the contracting path and reduced by half in the expanding path. In the first layer, a convolution with $3 \times 3 \times 3 \times 2$ kernels is applied to merge 2 input datasets. Then, each convolution and deconvolution layer except for the last one is composed of $3 \times 3 \times 3$ kernels. In the last layer, which provides an output, $1 \times 1 \times 1$ convolution is used for scaling purposes. Symmetric skip connections (copy and concatenation) between the convolution and deconvolution layers are used to achieve fast convergence and attain high frequency of local features (*25*). We implemented the networks using the TensorFlow library (*26*).

Network Training and Loss Function

The L1 norm between output (μ-CNN) and ground truth (μ-CT) was chosen as a cost function to train the network to generate a CT-like attenuation map. The cost function was minimized using adaptive moment estimation with an initial learning rate of 0.001, which was reduced by half every 2 epochs (*27*). In every epoch, we updated the network 9,955 times using a mini-batch of 128 samples. The total number of epochs was 10. The training time was approximately 30 h/epoch when a Ryzen 1700X central processing unit with a GTX 1080 graphics processing unit was used. Supplemental Figure 1 shows that the L1 norm between μ-CNN and μ-CT decreases as the epoch increases, and 10 epochs are sufficient to reach convergence (supplemental materials are available at http://jnm.snmjournals.org).

Image Analysis

The relative performance of CNN-based attenuation correction of whole-body ^{18}F-FDG PET was analyzed by comparing the attenuation and activity maps generated from 20 test sets. The attenuation maps generated using the proposed method, MLAA, and 4-segment method (μ-CNN, μ-MLAA, and μ-segment) were compared with μ-CT, a ground truth. A voxelwise scatterplot was plotted for comparison, and the correlation coefficients were calculated. The similarity of segmented regions from μ-CNN and μ-MLAA attenuation maps was compared using the Dice similarity coefficient (*28,29*). For the segmentation of μ-CNN and μ-MLAA, the sample thresholds used for generating μ-segment were applied (0.015, 0.04, and 0.09 between air, lung, fat, and water).

Attenuation-corrected activity maps were also obtained using the ordered-subset expectation maximization (3 iterations, 21 subsets, 5-mm gaussian postprocessing filter) algorithm with different attenuation maps. We made a voxelwise comparison between the activity maps generated using the proposed method, MLAA, and 4-segment method (λ-CNN, λ-MLAA, and λ-segment) and the activity map obtained using μ-CT (λ-CT). We also

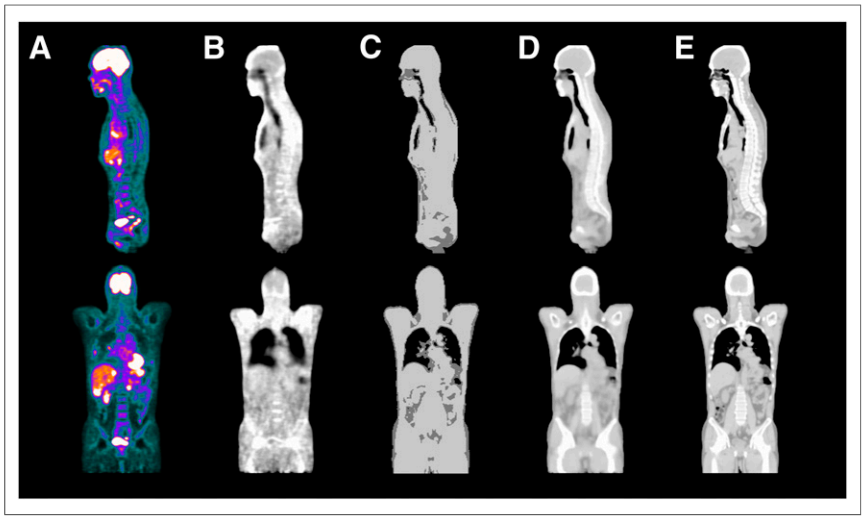

FIGURE 2. Attenuation maps of representative case. (A) λ-MLAA: original MLAA activity. (B) μ-MLAA: original MLAA attenuation map. (C) μ-segment: 4-segment map corresponding to Dixon MRI-based attenuation map. (D) μ-CNN: deep CNN output. (E) μ-CT: ground truth.

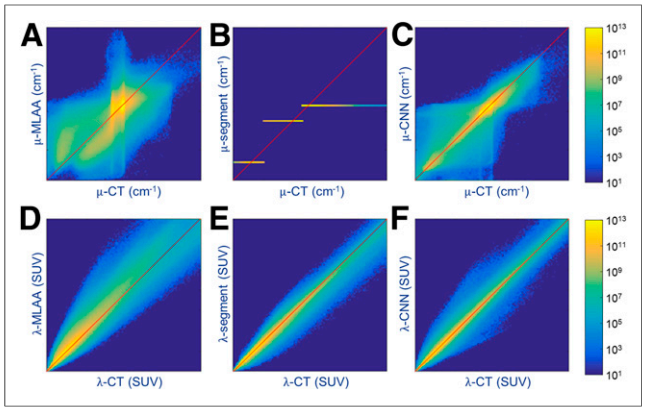

FIGURE 3. (A–C) Correlation between μ-CT and μ-MLAA (A), μ-segment (B), and μ-CNN (C). (D–F) Correlation between λ-CT and λ-MLAA (D), λ-segment (E), and λ-CNN (F). Red lines are identity lines.

compared the SUV_{mean} of primary and metastatic bone lesions (29 lesions from 6 patients) and soft-tissue lesions (11 lesions from 5 patients) obtained by manual drawing of regions of interest on the activity images of the patients who were included in the test set and showed abnormal uptake (Table 1).

The peak signal-to-noise ratio and normalized root-mean-square error values were calculated as additional image quality metrics to quantify the similarity of attenuation and activity maps (*30*).

To understand better what the CNN is really doing, some sanity checks were performed by exploring the CNN output for μ-MLAA input clipped to soft-tissue attenuation values (no higher bone attenuation), λ-MLAA input with high–low flipped activity, and μ-MLAA input globally scaled by 0.95.

RESULTS

Attenuation Maps

The CNN generates less noisy attenuation maps than μ-MLAA and achieves better bone identification than μ-segment and μ-MLAA. Figure 2 shows sagittal and coronal slices of λ-MLAA, μ-MLAA, μ-segment, μ-CNN, and μ-CT in a representative case. Although μ-MLAA shows higher μ-values in bone regions than soft-tissue and lung tissue, the margin between different tissues is not as clear as in the other attenuation maps. Moreover, some soft-tissue regions (e.g., back and shoulder) show improperly high μ-values, and lower spinal bone regions are not well identified (Fig. 2B). In contrast, the large bone structures are more accurately delineated in the μ-CNN, although the small bone structures are not as fine as those obtained with μ-CT (Fig. 2D). Supplemental Figure 2 shows the attenuation maps generated from another representative case that also proves the superiority of the CNN-based approach to the others in terms of the similarity to μ-CT. However, some small air regions in the abdomen are missing in the μ-CNN (Supplemental Fig. 2D), indicating the limited performance of CNN to overcome the high noise-level and low resolution of λ-MLAA and μ-MLAA.

The quantitative analysis on the similarity of attenuation maps confirmed the qualitative comparison results. The μ-CNN achieved the best voxelwise correlation with μ-CT as shown in Figures 3A–3C (the joint histogram of μ-values were plotted in log scale) and Table 2. Also, the μ-CNN yielded the highest peak signal-to-noise ratio and lowest normalized root-mean-square error relative to the μ-CT (Table 3). The average Dice similarity coefficients between μ-CNN and μ-CT were significantly higher than that between μ-MLAA and μ-CT in all regions (Fig. 4). In the regions of interest drawn on the bone lesions, the μ-CNN yielded the smallest error relative to the μ-CT (MLAA, −4.10% ± 10.62%; 4-segment, −4.20% ± 8.23%; CNN, 0.43% ± 6.80%). In the soft-tissue lesions, both 4-segment method and CNN yielded considerably smaller error than MLAA (MLAA, −4.57% ± 6.59%; 4-segment, 0.57% ± 1.74%; CNN, 0.91% ± 3.55%).

Similar to our previous results (*22*), the activity information enhances the CNN performance. Supplemental Figure 3 shows the results of CNNs trained with and without λ-MLAA as input. Using both the λ-MLAA and μ-MLAA as input, we could generate μ-CNN with better image contrast and anatomic detail.

Through the sanity checks, we could verify that the CNN works properly as we intended. As shown in Figure 5, the μ-MLAA input clipped to soft-tissue attenuation values yielded μ-CNN with missing bone regions. Intensity-flipped λ-MLAA input produced irreverently homogeneous μ-CNN, and μ-MLAA input globally scaled by 0.95 led to the underestimation of μ-CNN intensity.

Activity Maps

The accuracy of PET activity quantification was improved by the attenuation correction using μ-CNN. As shown in Figures 3D–3F and Table 2, the λ-CNN more strongly correlates with λ-CT than do the λ-MLAA and λ-segment. In addition, the λ-CNN showed the highest peak signal-to-noise ratio and lowest normalized root-mean-square error relative to λ-CT (Table 3). Figure 6 shows the differences in activity maps (SUV) relative to CT-based in a representative case. Although λ-MLAA (A) shows lower differences in bone regions than λ-segment (B), the differences in soft tissues were higher, particularly in the lungs, heart, liver dome, and bladder. The differences between λ-CNN and λ-CT (C) were smaller and more uniform than the others. The overestimation of activity in the lungs and underestimation in the liver seen in λ-MLAA were considerably reduced in λ-CNN. Supplemental Figure 4 shows attenuation-corrected PET images of a patient with lung lesions using different attenuation maps.

TABLE 2
Summary of Voxelwise Correlation of Attenuation (μ) and Activity (λ) Relative to Ground Truth (μ-CT and λ-CT) for 20 Subjects Used in Test Set

	MLAA			4-Segment			CNN		
Parameter	Slope	Intercept	R^2	Slope	Intercept	R^2	Slope	Intercept	R^2
μ	0.73 (0.06)	0.02 (0.01)	0.56 (0.05)	0.73 (0.03)	0.02 (0.00)	0.78 (0.05)	0.86 (0.05)	0.01 (0.00)	0.87 (0.04)
λ	0.97 (0.06)	0.09 (0.04)	0.97 (0.02)	0.94 (0.03)	0.03 (0.02)	1.00 (0.00)	0.99 (0.02)	0.01 (0.02)	1.00 (0.00)

Data are mean (SD).

TABLE 3
Peak Signal-to-Noise Ratio and Normalized Root-Mean-Square Error Relative to Ground Truth (μ-CT and λ-CT) for 20 Subjects Used in Test Set

Parameter	MLAA		4-Segment		CNN	
	Peak signal-to-noise ratio	Normalized root-mean-square error	Peak signal-to-noise ratio	Normalized root-mean-square error	Peak signal-to-noise ratio	Normalized root-mean-square error
μ	23.23 (1.30)	0.070 (0.011)	26.66 (0.99)	0.047 (0.006)	29.12 (1.90)	0.036 (0.008)
λ	44.29 (4.89)	0.007 (0.004)	49.92 (5.38)	0.004 (0.003)	60.41 (6.79)	0.001 (0.001)

Data are mean (SD).

The same trend was observed in the regional SUV_{mean} quantification in bone lesions (Fig. 7) and soft lesions. The relative differences in bone lesions and soft-tissue lesions are plotted in Supplemental Figure 5. Although the 4-segment map-based correction ($y = 0.947x - 0.043$, $R^2 = 0.964$) improved the SUV correlation with CT-based attenuation correction in comparison to the original MLAA ($y = 1.020x - 0.062$, $R^2 = 0.848$), the CNN ($y = 1.000x - 0.017$, $R^2 = 0.992$) outperformed the 4-segment map-based correction. However, in 6 vertebrae regions of 29 bone lesions, MLAA (2.71% ± 4.61%) and CNN (−2.22% ± 1.77%) were more accurate than 4-segment map-based correction (−9.40% ± 5.16%). In soft lesions, CNN (1.31% ± 3.35%) showed smaller error than MLAA (8.78% ± 7.41%) and 4-segment map-based correction (−2.90% ± 1.22%). Also, the CNN yielded the lowest voxelwise average error relative to the CT-based attenuation correction (λ-MLAA, 12.82% ± 2.45%; λ-segment, 5.61% ± 0.68%; λ-CNN, 2.05% ± 1.51%).

Computation Time

The MLAA took approximately 2 times longer than ordered-subset expectation maximization, but the CNN inference did not require long computation time (<30% of ordered-subset expectation maximization reconstruction for whole-body PET).

DISCUSSION

Attenuation correction is an essential procedure in the generation of PET images with quantitatively accurate regional activity information. In PET/CT systems, the Hounsfield unit in CT images is converted into the linear attenuation coefficient for 511-keV annihilation photons based on their bilinear relationship to generate patient-specific PET attenuation maps (μ-CT). However, attenuation map generation in PET/MRI is not so straightforward because the MR signal is not directly related to photon attenuation (*6*). The Dixon MRI based on the chemical shift difference between water and fat provides noiseless 4-segment maps for PET attenuation correction (*10*). However, these 4-segment maps cannot accurately account for photon attenuation by bone tissue (*11,12,31*). In addition, the inter- and intrapatient variability of μ-values is ignored in this method (*11,32*). Moreover, the bone model–based approach cannot handle heterogeneous bone attenuation coefficients even in cases with accurate bone registration. Although ultrashort echo time MRI sequence-based conversion of *R2** to the CT Hounsfield unit was proposed, its application is limited to neurologic studies (*29,33*). Despite recent advances in the accuracy of TOF measurement in PET, limitations of simultaneous activity and attenuation reconstruction still exist in real clinical cases. The deep CNN trained in this study to learn μ-CT from MLAA activity and attenuation maps handled well the information provided by this physically relevant but imperfect attenuation correction method.

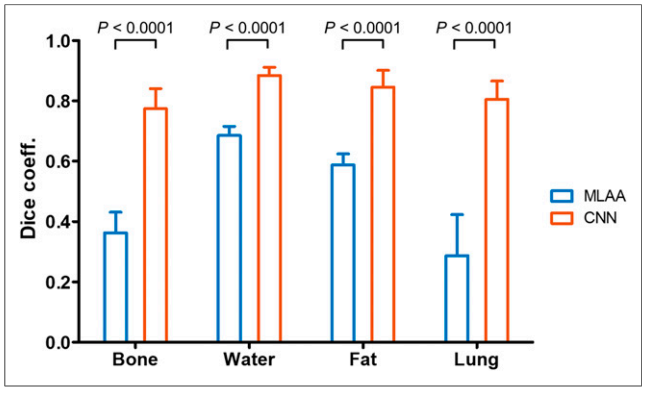

FIGURE 4. Statistical analysis on Dice similarity coefficients between μ-CNN and μ-CT and those between μ-MLAA and μ-CT.

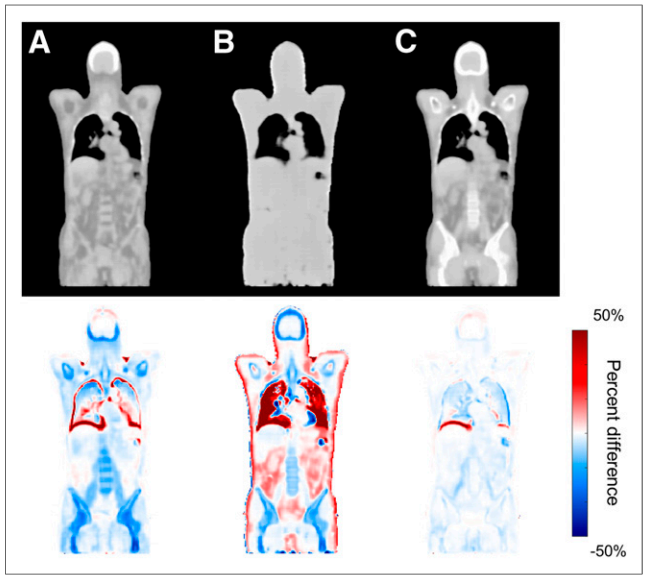

FIGURE 5. Attenuation maps generated in sanity checks to understand how CNN works. (A–C, top) CNN outcomes for μ-MLAA input clipped to soft-tissue attenuation values (no higher bone attenuation) (A), λ-MLAA input with high–low flipped activity (B), μ-MLAA input globally scaled by 0.95 (C). (A–C, bottom) Differences from original CNN output.

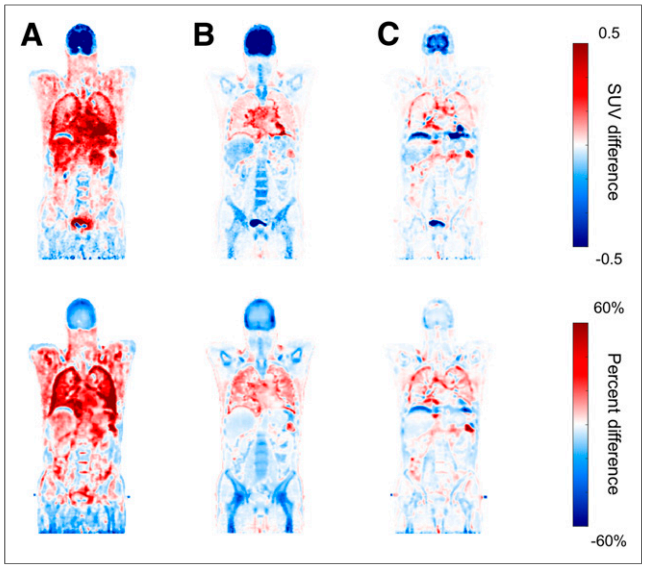

FIGURE 6. Differences of SUV between λ-CT and λ-MLAA (A), λ-segment (B), and λ-CNN (C). SUV difference = λ-method − λ-CT; percent difference = (λ-method − λ-CT)/λ-CT×100%; method = MLAA, segment, or CNN.

The convolution kernels of CNN, which determine how the MLAA images are merged, evolve during the network training based on the given training set. Therefore, the CNN approach does not need additional image segmentation, tissue probability prior, control parameters, and so forth (*34–36*), which are required in other approaches that have been previously proposed to combine MLAA and Dixon attenuation correction methods (*37,38*). For example, the MRI-guided MLAA algorithm (*37,39*), which imposes MR spatial and CT statistical constraints on the MLAA estimation of attenuation maps using a constrained gaussian mixture model and a Markov random field smoothness prior, needs to use a coregistered bone probability map. Another joint estimation algorithm that uses Dixon MRI as a prior requires uncertain MR-region segmentation and emphatically determined prior weights for this segment (*38*).

The difference between PET activities corrected for attenuation using proposed CNN-based method and CT was relatively high in lung boundary and liver dome (Fig. 6C). Incorrect activity measurement in lung and upper liver due to the respiratory motion–induced mismatch between PET and CT data is a well-known artifact in CT-based PET attenuation correction. In the MLAA, the activity and attenuation are simultaneously estimated only from the emission PET data without use of any transmission data. Therefore, the MLAA activity and attenuation maps would be free from or less vulnerable to the position mismatch artifact. We used CT as the ground truth for training the CNN. However, major data used for network training came from other regions rather than lung and upper liver because we performed patch-based learning. In fact, the λ-MLAA, μ-MLAA, and μ-CNN show well-matched boundaries of liver dome (Figs. 2A, 2B, and 2C) although the liver dome in μ-CT is elevated (Fig. 2E). Accordingly, the CNN-based approach that derived an attenuation map from the MLAA outputs would also be less influenced by the respiratory motion than CT-based attenuation correction, resulting in the activity difference shown in Figure 6C.

A limitation of this study is that we trained and validated the CNN only for whole-body ^{18}F-FDG PET scans. It is uncertain yet whether the trained deep network for simultaneously reconstructed activity and attenuation maps from ^{18}F-FDG will work for other types of PET radiotracers. Further investigation is required to answer this clinically important question. Even if different network parameters are required for each type of radiotracer, the network parameters derived from ^{18}F-FDG PET in this study could serve as initial values assigned before fine parameter tuning for each radiotracer. This technique, called transfer learning, allows faster convergence in training with a limited training set (*40*). The transfer learning will also be useful for training the CNN when it is applied to the PET systems with different timing resolutions. Another direction to advance in future investigations would be the combination of MLAA outputs and Dixon MR images as the input to the CNN. The detailed anatomic information and tissue characterization provided by the Dixon MR images would be useful for improving the CNN performance. It should also be noted that the performance of MLAA is sensitive to accurate scatter estimation and TOF timing offset calibration (*41*). Thus also, the CNN MLAA performance would be influenced by the residual uncertainties in scatter estimation and TOF timing offset calibration.

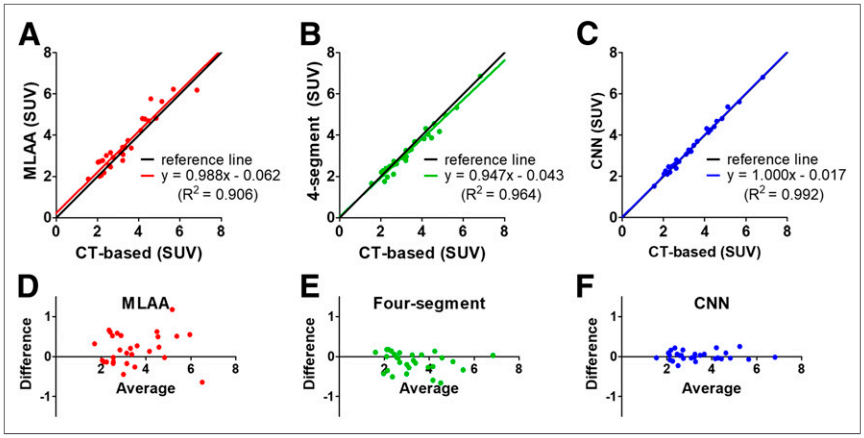

FIGURE 7. Comparison of SUV_{mean} measured in bone lesions (29 lesions in 10 patients in test set). Scattered plots between λ-CT and λ-MLAA (A), λ-segment (B), and λ-CNN (C) and corresponding Bland–Altman plots (D–F).

CONCLUSION

We have developed a deep neural network that successfully processes the information obtained from simultaneous activity and attenuation reconstruction to produce a more reliable attenuation map for 511-keV photons in comparison to the conventional 4-segment method. We also verified its feasibility using a whole-body ^{18}F-FDG PET dataset with TOF information. Accordingly, the proposed method has potential to replace the current 4-segment–based attenuation correction in whole-body FDG PET/MRI in which bones are poorly identified in whole-body PET/MRI studies and the local MRI signal loss produced by metallic implants results in the considerable error in image segmentation. Also, we expect that

the accuracy of this new method will improve as the TOF and machine learning technologies advance further.

DISCLOSURE

This work was supported by grants from the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Science and ICT (NRF-2014M3C7034000, NRF-2016R1A2B3014645, and NRF-2017M3C7A1044367). The funding source had no involvement in the study design, collection, analysis, or interpretation. No other potential conflict of interest relevant to this article was reported.

REFERENCES

1. Yoon HS, Ko GB, Kwon SI, et al. Initial results of simultaneous PET/MRI experiments with an MRI-compatible silicon photomultiplier PET scanner. *J Nucl Med.* 2012;53:608–614.
2. Levin CS, Maramraju SH, Khalighi MM, Deller TW, Delso G, Jansen F. Design features and mutual compatibility studies of the time-of-flight PET capable GE SIGNA PET/MR system. *IEEE Trans Med Imaging.* 2016;35:1907–1914.
3. Ko GB, Yoon HS, Kim KY, et al. Simultaneous multiparametric PET/MRI with silicon photomultiplier PET and ultra-high-field MRI for small-animal imaging. *J Nucl Med.* 2016;57:1309–1315.
4. Judenhofer MS, Wehrl HF, Newport DF, et al. Simultaneous PET-MRI: a new approach for functional and morphological imaging. *Nat Med.* 2008;14:459–465.
5. Yoo HJ, Lee JS, Lee JM. Integrated whole body MR/PET: where are we? *Korean J Radiol.* 2015;16:32–49.
6. Vandenberghe S, Marsden PK. PET-MRI: a review of challenges and solutions in the development of integrated multimodality imaging. *Phys Med Biol.* 2015;60:R115–R154.
7. Yang J, Jian Y, Jenkins N, et al. Quantitative evaluation of atlas-based attenuation correction for brain PET in an integrated time-of-flight PET/MR imaging system. *Radiology.* 2017;284:169–179.
8. Sekine T, Ter EV, Warnock G, et al. Clinical evaluation of zero-echo-time attenuation correction for brain ^{18}F-FDG PET/MRI: comparison with atlas attenuation correction. *J Nucl Med.* 2016;57:1927–1932.
9. Keereman V, Fierens Y, Broux T, De Deene Y, Lonneux M, Vandenberghe S. MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences. *J Nucl Med.* 2010;51:812–818.
10. Martinez-Möller A, Souvatzoglou M, Delso G, et al. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. *J Nucl Med.* 2009;50:520–526.
11. Kim JH, Lee JS, Song I-C, Lee DS. Comparison of segmentation-based attenuation correction methods for PET/MRI: evaluation of bone and liver standardized uptake value with oncologic PET/CT data. *J Nucl Med.* 2012;53:1878–1882.
12. Keller SH, Holm S, Hansen AE, et al. Image artifacts from MR-based attenuation correction in clinical, whole-body PET/MRI. *MAGMA.* 2013;26:173–181.
13. Paulus DH, Quick HH, Geppert C, et al. Whole-body PET/MR imaging: quantitative evaluation of a novel model-based MR attenuation correction method including bone. *J Nucl Med.* 2015;56:1061–1066.
14. Rausch I, Quick HH, Cal-Gonzalez J, Sattler B, Boellaard R, Beyer T. Technical and instrumentational foundations of PET/MRI. *Eur J Radiol.* 2017;94:A3–A13.
15. Defrise M, Rezaei A, Nuyts J. Time-of-flight PET data determine the attenuation sinogram up to a constant. *Phys Med Biol.* 2012;57:885–899.
16. Berker Y, Li Y. Attenuation correction in emission tomography using the emission data: a review. *Med Phys.* 2016;43:807–832.
17. Rezaei A, Defrise M, Bal G, et al. Simultaneous reconstruction of activity and attenuation in time-of-flight PET. *IEEE Trans Med Imaging.* 2012;31:2224–2233.
18. Salomon A, Goedicke A, Schweizer B, Aach T, Schulz V. Simultaneous reconstruction of activity and attenuation for PET/MR. *IEEE Trans Med Imaging.* 2011;30:804–813.
19. Boellaard R, Hofman M, Hoekstra O, Lammertsma A. Accurate PET/MR quantification using time of flight MLAA image reconstruction. *Mol Imaging Biol.* 2014;16:469–477.
20. Chun SY, Kim KY, Lee JS, Fessier JA. Joint estimation of activity distribution and attenuation map for TOF-PET using alternating direction method of multiplier. *Proc IEEE Int Symp Biomed Imaging*. 2016;1:86–89.
21. Son JW, Kim KY, Yoon HS, et al. Proof-of-concept prototype time-of-flight PET system based on high-quantum-efficiency multi-anode PMTs. *Med Phys.* 2017;44:5314–5324.
22. Hwang D, Kim KY, Kang SK, et al. Improving accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. *J Nucl Med.* 2018;59:1624–1629.
23. Leynes AP, Yang J, Wiesinger F, et al. Direct pseudoCT generation for pelvis PET/MRI attenuation correction using deep convolutional neural networks with multi-parametric MRI: zero echo-time and Dixon deep pseudoCT (ZeDD-CT). *J Nucl Med.* 2018;59:852–858.
24. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham, Switzerland: Springer: 2015;9351:234–241.
25. Mao X, Shen C, Yang Y-B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Proceedings of Advances in Neural Information Processing Systems. *NIPS*. 2016;1:1–9.
26. Abadi M, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv.org website. https://arxiv.org/abs/1603.04467. Published 2016. Accessed April 2, 2019.
27. Kingma D, Ba J. Adam: a method for stochastic optimization. arXiv.org website. https://arxiv.org/abs/1412.6980. Published 2014. Accessed April 2, 2019.
28. An HJ, Seo S, Kang H, et al. MRI-based attenuation correction for PET/MRI using multiphase level-set method. *J Nucl Med.* 2016;57:587–593.
29. Juttukonda MR, Mersereau BG, Chen Y, et al. MR-based attenuation correction for PET/MRI neurological studies with continuous-valued attenuation coefficients for bone through a conversion from R2* to CT-Hounsfield units. *Neuroimage.* 2015;112:160–168.
30. Park J, Hwang D, Kim KY, Kang SK, Kim YK, Lee JS. Computed tomography super-resolution using deep convolutional neural network. *Phys Med Biol.* 2018;63:145011.
31. Hofmann M, Pichler B, Schölkopf B, Beyer T. Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques. *Eur J Nucl Med Mol Imaging.* 2009;36(suppl):S93–S104.
32. Marshall HR, Prato FS, Deans L, Théberge J, Thompson RT, Stodilka RZ. Variable lung density consideration in attenuation correction of whole-body PET/MRI. *J Nucl Med.* 2012;53:977–984.
33. Ladefoged CN, Benoit D, Law I, et al. Region specific optimization of continuous linear attenuation coefficients based on UTE (RESOLUTE): application to PET/MR brain imaging. *Phys Med Biol.* 2015;60:8047–8065.
34. Park C, Took CC, Seong J-K. Machine learning in biomedical engineering. *Biomed Eng Lett.* 2018;8:1–3.
35. Mansour RF. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed Eng Lett.* 2017;8:41–57.
36. Choi H. Deep learning in nuclear medicine and molecular imaging: current perspectives and future directions. *Nucl Med Mol Imaging.* 2018;52:109–118.
37. Mehranian A, Zaidi H. Joint estimation of activity and attenuation in whole-body TOF PET/MRI using constrained gaussian mixture models. *IEEE Trans Med Imaging.* 2015;34:1808–1821.
38. Ahn S, Cheng L, Shanbhag DD, et al. Joint estimation of activity and attenuation for PET using pragmatic MR-based prior: application to clinical TOF PET/MR whole-body data for FDG and non-FDG tracers. *Phys Med Biol.* 2018;63:045006.
39. Mehranian A, Zaidi H. Clinical assessment of emission-and segmentation-based MR-guided attenuation correction in whole-body time-of-flight PET/MR imaging. *J Nucl Med.* 2015;56:877–883.
40. Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging.* 2016;35:1299–1312.
41. Nuyts J, Rezaei A, Defrise M. The validation problem of joint emission/transmission reconstruction from TOF-PET projections. *IEEE Trans Radiat Plasma Med Sci.* 2018;2:273–278.

The Journal of NUCLEAR MEDICINE

Generation of PET Attenuation Map for Whole-Body Time-of-Flight ^{18}F-FDG PET/MRI Using a Deep Neural Network Trained with Simultaneously Reconstructed Activity and Attenuation Maps

Donghwi Hwang, Seung Kwan Kang, Kyeong Yun Kim, Seongho Seo, Jin Chul Paeng, Dong Soo Lee and Jae Sung Lee

J Nucl Med. 2019;60:1183-1189.
Published online: January 25, 2019.
Doi: 10.2967/jnumed.118.219493

The Journal of Nuclear Medicine is published monthly.
SNMMI | Society of Nuclear Medicine and Molecular Imaging
1850 Samuel Morse Drive, Reston, VA 20190.
(Print ISSN: 0161-5505, Online ISSN: 2159-662X)

SNMMI SOCIETY OF NUCLEAR MEDICINE AND MOLECULAR IMAGING