Am J Nucl Med Mol Imaging 2026;16(1):33-43
www.ajnmmi.us /ISSN:2160-8407/ajnmmi0171210

Original Article

Scatter correction for self-collimating SPECT using a 3D U-Net framework

Yabo Zhao[1,2], Wenyang Jiang[1,2], Hai Hu[1,2], Liang Guo[3,4], Zhenlei Lyu[3], Debin Zhang[3], Li Wang[5], Peng Fan[6], Hui Liu[3], Tianyu Ma[3], Jing Wu[1,2]

[1]*Center for Advanced Quantum Studies, School of Physics and Astronomy, Beijing Normal University, Beijing, China;* [2]*Key Laboratory of Multiscale Spin Physics (Ministry of Education), Beijing Normal University, Beijing, China;* [3]*Department of Engineering Physics, Tsinghua University, Beijing, China;* [4]*Department of Radiology and Biomedical Imaging, Yale University, New Haven, USA;* [5]*Department of Nuclear Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China;* [6]*Beijing Institute of Spacecraft System Engineering, Beijing, China*

Received January 4, 2026; Accepted February 13, 2026; Epub February 15, 2026; Published February 28, 2026

Abstract: Conventional single-photon emission computed tomography (SPECT) relies on mechanical collimators, which impose an inherent trade-off between spatial resolution and sensitivity. A novel cardiac SPECT system that employs a self-collimating design with interleaved mosaic scintillators has been proposed, which markedly enhances sensitivity without compromising resolution. However, the unique self-collimating and closely arranged detector geometry also introduces more complex scatter distribution and increased scatter fractions, making accurate scatter correction essential yet technically challenging. We employed a 3D U-Net framework to directly predict scatter-corrected images from uncorrected images. The network was trained using 36 distinct XCAT phantoms based on GATE simulations, with the true scatter-corrected images (true-SC) precisely obtained from the simulations serving as labels. Quantitative evaluation was performed using another two XCAT phantoms with different contrast levels: a high-contrast phantom (H-Phantom, 10 realizations) and a low-contrast phantom (L-Phantom, 5 realizations). The proposed U-Net approach were compared with two triple energy window (TEW) methods (trapezoidal and triangular). For both contrast levels, the U-Net-based approach achieved higher contrast recovery coefficients, myocardium-to-blood-pool ratios closer to the true-SC, higher contrast-to-noise ratios, and lower relative noise compared to the TEW methods. In addition, the U-Net-based method produced images with higher structural similarity and lower normalized mean square error relative to the true-SC reference, compared with the TEW-corrected images. In conclusion, the proposed 3D U-Net-based scatter correction method provides more accurate scatter estimation and superior quantitative performance for self-collimating SPECT systems than conventional TEW approaches.

Keywords: Deep learning, self-collimation, SPECT, scatter correction, U-Net

Introduction

Single-photon emission computed tomography (SPECT) generates 3D images of in-vivo functional information by detecting gamma photons emitted by radiotracers. It is one of the most widely used nuclear medicine imaging modalities in clinical practice. Quantitative SPECT plays an important clinical role in diagnosis, risk stratification, and therapy monitoring, but its accuracy critically depends on accurate physical corrections. Photon interactions in tissue lead to attenuation and scatter, both of which degrade image quality and increase quantification bias [1]. In quantitative SPECT, it is standard practice to combine attenuation correction (AC) and scatter correction (SC) to enhance quantitative accuracy [2]. The attenuation correction is typically performed using a CT-derived attenuation map [3, 4].

Conventional SC methods can be divided into two main categories [2]. The first category includes model-based approaches [5, 6], such as analytical scatter models and Monte Carlo simulations. This category also includes reconstruction-domain iterative estimators such as the effective source scatter estimation (ESSE) method [2]. The second category includes energy window-based techniques [7-9]. These methods estimate the scatter component by acquiring data in additional energy windows, such as the dual- or triple-energy window methods and their variants. Although energy window-based methods are computationally efficient, they may increase image noise and exhibit residual quantification bias, particularly in low-count imaging cases [10]. Because conventional SC methods are usually based on simplified physical models, data-driven methods are increasingly used to achieve more accurate SC. In recent years, deep learning has attracted substantial attention in medical imaging [11-13]. Following the introduction of the U-Net architecture by Ronneberger et al. in 2015 [14], U-Net and its variants have shown to be highly effective in image segmentation [15], lesion detection [16], image denoising [17, 18] and image reconstruction [19]. Furthermore, researchers have applied these model for physical corrections. Examples include performing joint scatter and attenuation correction in the image domain [20], and using U-Net-based models to predict scatter sinograms in PET [21]. These studies support the feasibility of using deep learn-

https://doi.org/10.62347/ZZHP6340

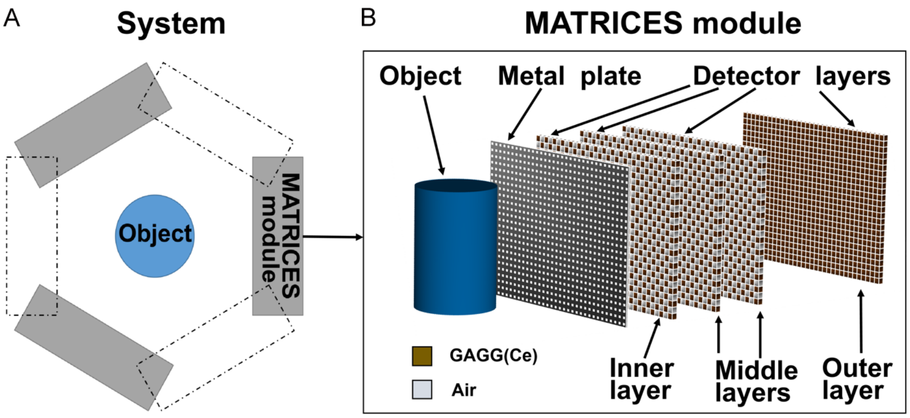

Figure 1. Schematic of the self-collimating SPECT system and the MATRICES detector module. A. Three-module system configuration, where dashed boxes indicate the arrangement of the modules after rotation. B. Internal layout of the MATRICES module, including the metal plate and the four detector layers (inner, two middle, and outer layers).

ing to learn complex scatter correction operators directly from data.

However, the dependence on mechanical collimators poses a fundamental limitation to conventional SPECT performance and quantitative accuracy. This is due to the inherent trade-off between sensitivity and spatial resolution [22-24]. Self-collimating SPECT addresses this limitation by eliminating heavy-metal mechanical collimators and achieving effective photon collimation through specially arranged high-sensitivity detector arrays, representing a new paradigm in nuclear medicine imaging [25, 26]. A representative implementation is the multilayer interleaved mosaic scintillator (MATRICES) architecture, which leverages the spatial arrangement and differential absorption properties of stacked scintillator layers to enable simultaneous photon collimation and detection. Through sequential photon penetration and interaction across the multilayer stack, this design substantially improves system sensitivity without compromising spatial resolution [27, 28].

Self-collimating SPECT systems achieve a higher effective count rate. However, the systems will also detect a larger proportion of scattered photons. Their complex multi-layer detector architecture increases the probability of photon interactions. Consequently, accurate scatter correction becomes especially critical for the imaging performance of self-collimating SPECT. Conventional triple energy window (TEW) correction may systematically over- or under-estimate scatter events, leading to increased noise and quantification bias [29]. To address these challenges, we propose a U-Net-based scatter correction method that aims to improve scatter estimation while suppressing noise and enhancing image contrast [30].

In this study, we compared trapezoidal TEW, triangular TEW, and the proposed U-Net-based method using phantom experiments at two contrast levels. Image quality and quantitative accuracy were evaluated using multiple metrics, demonstrating the applicability and potential advantages of deep learning-based scatter correction for self-collimating SPECT.

Materials and methods

Self-collimating SPECT system setup

We examined a self-collimating SPECT system built with three identical modules. These modules are arranged symmetrically around a cylindrical field of view. They are positioned at 120° intervals from each other. To satisfy angular sampling requirements, the system underwent a single 60° rotation about its central axis. The configuration is shown in **Figure 1**, where dashed boxes indicate the positions of the modules after rotation. Each module adopts the MATRICES architecture, spans 528.0 mm tangentially and 400.0 mm axially, and consists of a 2-mm-thick tungsten plate and four detector layers stacked along the radial direction.

In each MATRICES detector module, the tungsten plate is uniformly perforated with 2 mm × 2 mm square apertures (opening ratio ≈ 4%). The detector assembly is composed of 16 mm × 16 mm × 6 mm detection units arranged in a 33 × 25 array (tangential × axial) within each module. Each unit has many small GAGG(Ce) scintillators. Each one is 4 mm by 4 mm by 6 mm. They are placed with air voxels in between. The first three layers of scintillators are in a mosaic pattern. This pattern lets the layers detect photons and also provides collimation. The fourth layer is different. It is completely filled with GAGG(Ce) material. This full layer catches and absorbs the photons that pass through the layers above it. The distances from the FOV center to the tungsten plate and to the four detector layers are 325, 400, 475, 550, and 825 mm, respectively.

Data simulation and scatter characterization

All projection data in this study were generated using the Geant4 Application for Tomographic Emission (GATE, v9.2) to simulate ^{99m}Tc imaging of the eXtended CArdiaC-Torso phantom (XCAT) [31]. The simulation fully incorporated the key physical interactions of 140 keV gamma photons, including photoelectric effect, Compton scattering, and Rayleigh scattering. We enabled the “process tagging” feature in the simulation to track and record the complete interaction history of each detected photon. Based on this, events recorded by the detector were precisely defined and classified as: true events: referring to photons that were detected without scattering in the imaging object and whose energy deposition fell within the photopeak window (140 keV±20%); scatter events:

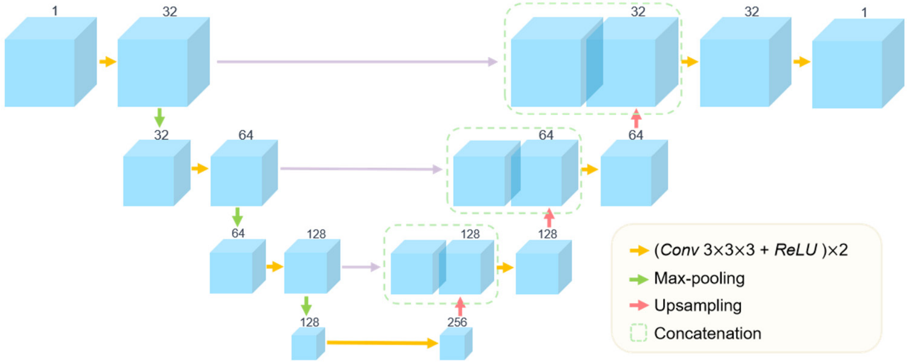

Figure 2. The 3D U-Net architecture.

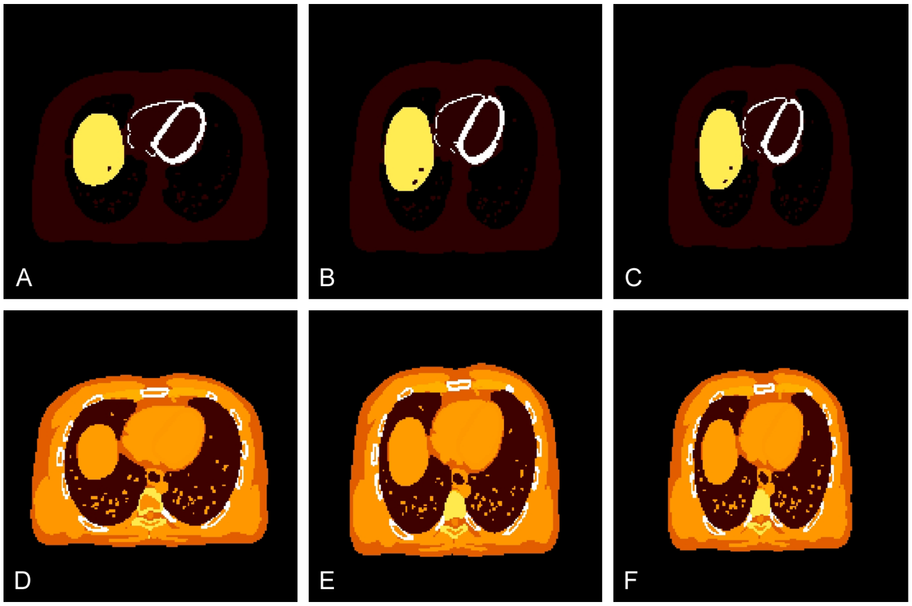

Figure 3. Sample XCAT male slices with a large size (A), a medium size (B) and a small size (C), with an activity ratio of 6:1:0:0.15 for myocardium, liver, lung, and background (including blood). Long axis scale, short axis scale and height scale of phantom for the three different body sizes are 1:1:1, 0.89:1.11:1.00, and 0.78:1.06:1.00, respectively. (D-F) are the attenuation coefficient maps corresponding to (A-C), respectively.

referring to photons that underwent one or more Compton scatterings in the imaging object, with final energy still falling within the photopeak window and being detected. Therefore, we can create two projection files: one contains all events, and the other has scatter events removed. Reconstructing images from these two files gives us the uncorrected image (w/o-SC) and the true scatter-corrected image (true-SC). We can also collect counts from any energy window. This provides the necessary projection data for the TEW scatter correction method.

The self-collimating SPECT system has no mechanical collimator. It shows high sensitivity and uses multi-layer, high-density GAGG(Ce) detectors. Because of this design, scatter effects are more noticeable than in traditional systems. We calculated the average scatter fraction for this system. It reaches 20.90%. Also, photon transport is affected by the multi-layer detector structure. This makes the spatial distribution of scatter events more complex compared to conventional SPECT. The energy spectrum becomes distorted too. For these reasons, scatter correction is very important for imaging with the self-collimating SPECT system.

Network architecture

Figure 2 shows our network. It has a three-level structure with an encoder and a decoder. Skip connections link them. In the encoder part, each level has two blocks. Each block does a 3×3×3 convolution. Then, batch normalization (BN) and a ReLU function follow. After these two blocks, a 2×2×2 max pooling operation is used. Its stride is 2. This step makes the feature maps smaller. It also makes the features deeper. The decoder part is like a mirror of the encoder. It uses a series of up-sampling layers. Each level began with a 2×2×2 transposed convolution operation that doubled the spatial dimensions and halved the feature depth. The upsampled feature maps were concatenated with the corresponding encoder features via skip connections to preserve multi-scale spatial information. The transposed convolution operation was followed by two consecutive 3×3×3 convolutional layers, each with BN and ReLU, to refine the fused features. The input of the U-Net is the reconstructed images without scatter correction while the output is the scatter-corrected images.

Phantom simulation and dataset generation

In this study, the datasets used to train the 3D U-Net and to evaluate the performance of scatter correction methods were generated using GATE-based Monte Carlo simulations of ^{99m}Tc imaging with the XCAT phantom [13]. A total of 36 simulation setups were used, corresponding to all combinations of 6 activity distributions and 6 anatomical types. All the images had a matrix size of 180×180×20 and a voxel size of 2×2×2 mm^3. The anatomical variations included two gender (male and female) and three body sizes (large, medium, and small), as shown in **Figure 3**, thereby yielding 6 different anatomical templates. Each anatomical model was paired with six different activity distributions, defined by uptake ratios listed in **Table 1**.

Table 1. Activity distribution defined by six uptake ratios

Group	Myocardium	Liver	Lung	Background
I	6	1	0	0.15
II	3	1	0	0.07
III	7	1	0	0.08
IV	10	1	0	0.10
V	1	1	0	0.10
VI	0.3	1	0	0.10

We created two additional XCAT phantoms to evaluate the 3D U-Net: a high-contrast phantom (H-Phantom) and a low-contrast phantom (L-Phantom). Their configurations are shown in **Figure 4**. We simulated both phantoms using the GATE software. We did many simulations to make sure our results were reliable. For the H-Phantom, we ran the simulation ten separate times. For the L-Phantom, we ran it five separate times. This repeated simulation process strengthens the reliability of our findings. The activity concentration ratios for the myocardium, liver, lungs, and background (including blood) were set to 10:10:0:1 and 9:9:0:3, respectively, using large-sized anatomical phantoms.

During simulation, both the activity image and attenuation coefficient image of the phantom were voxelized. Energy-windowed projections were derived from ROOT outputs. Total projection counts were ~10^9.

Network training and testing

These 36 XCAT phantoms were simulated to generate projection data and subsequently reconstructed using a self-developed OSEM algorithm incorporating attenuation correction [32]. The system matrix was obtained via Monte Carlo (MC) simulations in GATE using uniformly distributed ^{99m}Tc sources within the FOV. Attenuation factors was generated based on CT-derived attenuation maps using a CUDA-based ray-tracing algorithm. For each phantom, ten reconstructed images were produced by varying the number of OSEM iterations (ranging from 5500 to 10000 in increments of 500). The reconstructed images without scatter correction (w/o SC) served as the input for network training. Since true scatter events can be accurately identified in GATE simulations, images reconstructed with true scatter correction (true SC) were used as the training labels.

For training, we used a 3D-patched strategy with a patch size of 32×32×8. The patches were randomly extracted from the reconstructed images. The initial learning rate and the learning decay rate were set to 0.0001 and 0.996 respectively. The network was trained with the L1 loss function. The number of epochs was 120, with each epoch consisting of 100 batches and a batch size of 16 patches.

The 10000-iteration reconstructed images of H-Phantom and L-Phantom without scatter correction (w/o SC) were adopted as the testing dataset.

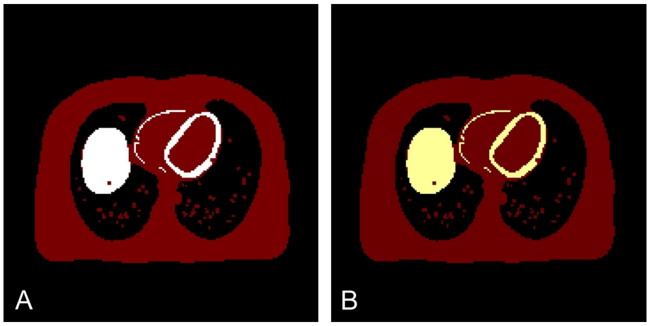

Figure 4. (A) High-contrast phantom (H-Phantom) and (B) low-contrast phantom (L-Phantom) for evaluation. Background display brightness was equalized for visual comparison.

Conventional TEW scatter correction methods

A. Trapezoidal TEW (TR-TEW) scatter correction: As shown in **Figure 5A**, two narrow energy windows for ^{99m}Tc are centered at 112 keV and 168 keV with widths of 10 keV, respectively, while the photopeak window spans 112-168 keV. This window configuration accounts for the 21% energy resolution of the GAGG(Ce) scintillator detectors. It is assumed that all counts within the two narrow windows originate from scattered photons. Accordingly, the scatter count densities at 112 keV and 168 keV are estimated by dividing the detected counts in each narrow window by its width. The number of scattered photons within the photopeak window is then approximated as the area of the trapezoid formed by the estimated scatter count densities at the lower and upper photopeak boundaries. The estimated scatter counts in the photopeak window using the TR-TEW method [33] is given by:

$$S_{pk} = \left[\frac{C_1}{W_1} + \frac{C_2}{W_2}\right] \cdot \frac{W_{pk}}{2} \tag{1}$$

Where S_{pk} denotes the estimated scatter counts in the photopeak window, and C_1 and C_2 are the total counts in the lower and upper narrow windows, respectively. W_1 and W_2 are the widths of the two narrow windows and W_{pk} is the width of the photopeak window.

B. Triangular TEW (TRI-TEW) scatter correction: As shown in **Figure 5B**, this method assumes that the upper narrow energy window contains no scattered photons; that is, all detected counts in this window are considered primary (true) photons. Owing to the assumed symmetry of the energy spectrum about the photopeak, the lower narrow window is expected to contain an equivalent number of primary counts. Under these assumptions, the scatter contribution within the photopeak window using the TRI-TEW method [34] is estimated by:

$$S_{pk} = \left[\frac{C_1}{W_1} - \frac{C_2}{W_2}\right] \cdot \frac{W_{pk}}{2} \tag{2}$$

C. Image reconstruction with TEW scatter corrections: After obtaining the scatter estimates from the TR-TEW

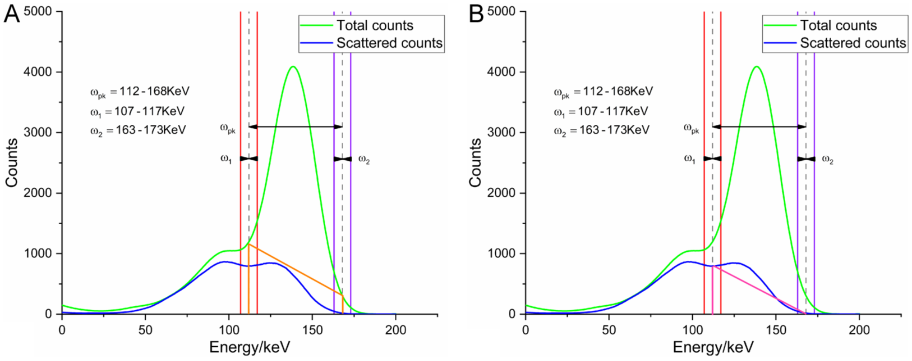

Figure 5. Illustration of the trapezoidal TEW (TR-TEW) scatter correction method (A), and triangular TEW (TRI-TEW) scatter correction (B), both overlaid on the ^{99m}Tc energy spectrum.

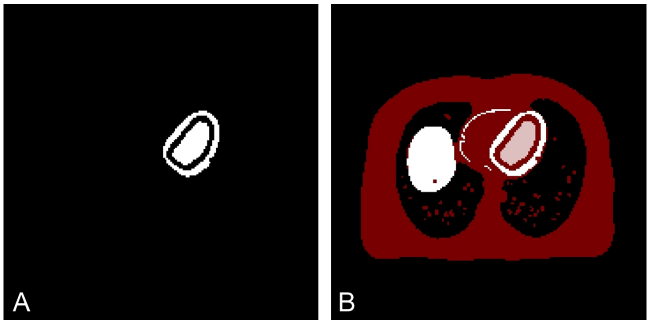

Figure 6. The ROIs were drawn for left ventricular myocardium and blood pool (A), and overlapped with the activity phantom (B).

and TRI-TEW methods, the respective scatter counts were subtracted from the primary window projections to produce two sets of scatter-corrected projection data. Each set was then separately reconstructed using the same OSEM algorithm in MATLAB 2022 (60 subsets; 10,000 iterations).

Performance evaluation

For both H-phantom and L-phantom, we obtained their reconstructed images with U-Net scatter correction (U-Net-SC), along with four conventional OSEM reconstruction images: with attenuation correction but without scatter correction (w/o-SC), with attenuation correction and true scatter correction (true-SC), with attenuation correction and trapezoidal TEW scatter correction (TR-TEW-SC), and with attenuation correction and triangular TEW scatter correction (TRI-TEW-SC).

For quantitative evaluation, we generated the regions of interest (ROIs) for the left ventricular myocardium (LVM, 5062 voxels) and the left ventricular blood pool (LVBP, 4388 voxels) using XCAT. As shown in **Figure 6**, the ROI of blood pool was set to be smaller than the original size to reduce the partial volume effect in the reconstructed images, while the ROI for LVM retained the original size.

Then, we calculated the quantitative metrics for evaluation, including contrast recovery coefficient (CRC), myocardium-to-blood pool ratio (M-B ratio), contrast-to-noise ratio (CNR), relative noise (RN), structural similarity index measure (SSIM), and normalized mean square error (NMSE). The mean and standard deviation of CRC, M-B ratio, CNR, RN, SSIM, and NMSE were calculated respectively.

A. Contrast recovery coefficient (CRC):

$$CRC = \frac{C_{myo}/C_{blood} - 1}{A_{myo}/A_{blood} - 1} \tag{3}$$

C_{myo} and C_{blood} are the mean counts of the LVM and the LVBP ROIs in the reconstructed images; and A_{myo} and A_{blood} are the true activity values of the LVM and the LVBP.

B. Myocardium-to-blood pool ratio (M-B ratio): The M-B ratio is calculated as the mean counts within LVM ROI divided by the mean counts within LVBP ROI:

$$M - B \text{ ratio} = \frac{C_{myo}}{C_{blood}} \tag{4}$$

C. Contrast-to-noise ratio (CNR):

$$CNR = (c_{myo} - c_{blood})/\delta_{blood} \tag{5}$$

δ_{blood} is the standard deviation of the LVBP ROI in the reconstructed images.

D. Relative noise (RN): The RN is defined as the ratio of the standard deviation of the LVBP ROI to mean counts within the LVBP ROI:

$$RN = \delta_{blood}/c_{blood} \tag{6}$$

E. Structural similarity index measure (SSIM):

$$SSIM = \frac{(2\mu_x\mu_y + c1)(2\sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x^2 + \sigma_y^2 + c2)} \tag{7}$$

μ_x and μ_y represent the mean values of the evaluated image and the reference image reconstructed with true-SC, respectively; σ_x and σ_y denote the standard deviations of the evaluated image and the reference image, respectively; and σ_{xy} indicates the cross-covariance between the evaluated and the reference images. c1 and c2 are two small constants introduced to prevent division by zero or excessively small values.

F. Normalized mean square error (NMSE):

$$NMSE = \frac{\| I^S - I^{REF} \|^2}{\| I^{REF} \|^2} \tag{8}$$

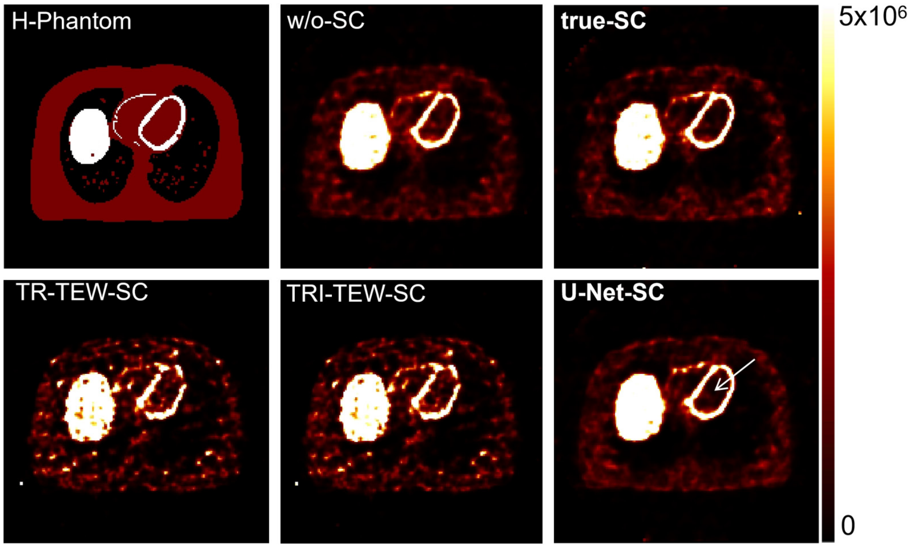

Figure 7. Reconstructed images with different scatter correction methods for H-Phantom.

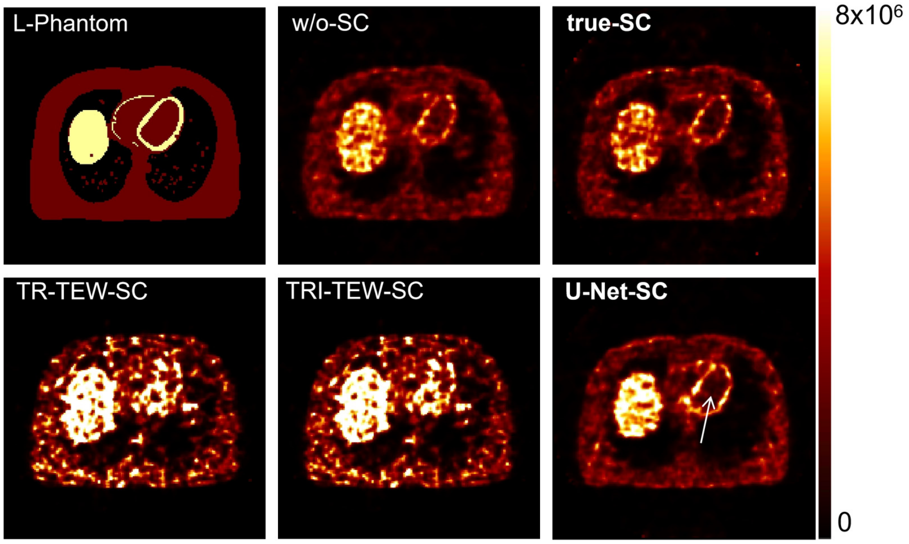

Figure 8. Reconstructed images with different scatter correction methods for L-Phantom.

I^S denotes the evaluated image, and I^{REF} represents the reference image.

Results

Reconstruction images with different scatter correction methods

Figures 7 and **8** show the reconstructed images with different scatter correction methods for the H-Phantom and the L-Phantom. Comparing the true-SC image with the w/o-SC image reveals that when scattered photon events can be accurately removed, the contrast between the myocardium and blood pool is enhanced, leading to improved image quality. We computed the scatter fractions (SF) of two TEW SC methods (TR-TEW and TRI-TEW) and the true-SC. The true-SC SF is 20.90%. The TR-TEW method gives an SF estimate of 30.51%. This value is higher than the true one. So, TR-TEW overestimates the scatter part. The TRI-TEW method gives an SF estimate of 16.87%. This value is lower than the true one. So, TRI-TEW underestimates the scatter part. Both window-based methods have clear limits when we compare them to true-SC. TR-TEW and TRI-TEW can improve image contrast to some extent. But their results are still much lower than the theoretical values. This likewise produces poorer image quality. The U-Net-SC method provides a more effective scatter correction. The images it reconstructs are closer to true-SC. It performs especially well in myocardium-to-blood-pool contrast. For the L-Phantom, which has a lower tracer uptake ratio, U-Net-SC can still reconstruct high-quality images. This method not only corrects scatter accurately but also reduces image noise at the same time. It shows clear advantages in the self-collimating SPECT system.

Figures 9 and **10** display the reconstructed cardiac images and their intensity profiles for both the H-Phantom and the L-Phantom. We compared different scatter correction methods. In these results, the U-Net-SC method consistently produces images with less noise in the blood pool area. The U-Net-SC images are closer to the true-SC reference images. The TEW method images are not as close. This shows that the U-Net method gives much more accurate quantitative results.

Quantitative evaluation results with different scatter correction methods

Table 2 shows the quantitative results for the H-Phantom images. We measured CRC, M-B ratio, CNR, and RN. With true-SC, the average CRC went from 0.50 to 1.00. This means accurate scatter correction greatly improves image contrast in our system. U-Net-SC gave a CRC of 0.99. This value is nearly the same as the true-SC result. It is higher than the results from the two TEW methods. We also looked at the M-B ratio. With true-SC, the ratio went from 5.45 to 9.97. The TEW methods did not reach as high. TR-TEW-SC had a value of 5.94. TRI-TEW-SC had a value of 6.06. The U-Net-SC method achieved a ratio of

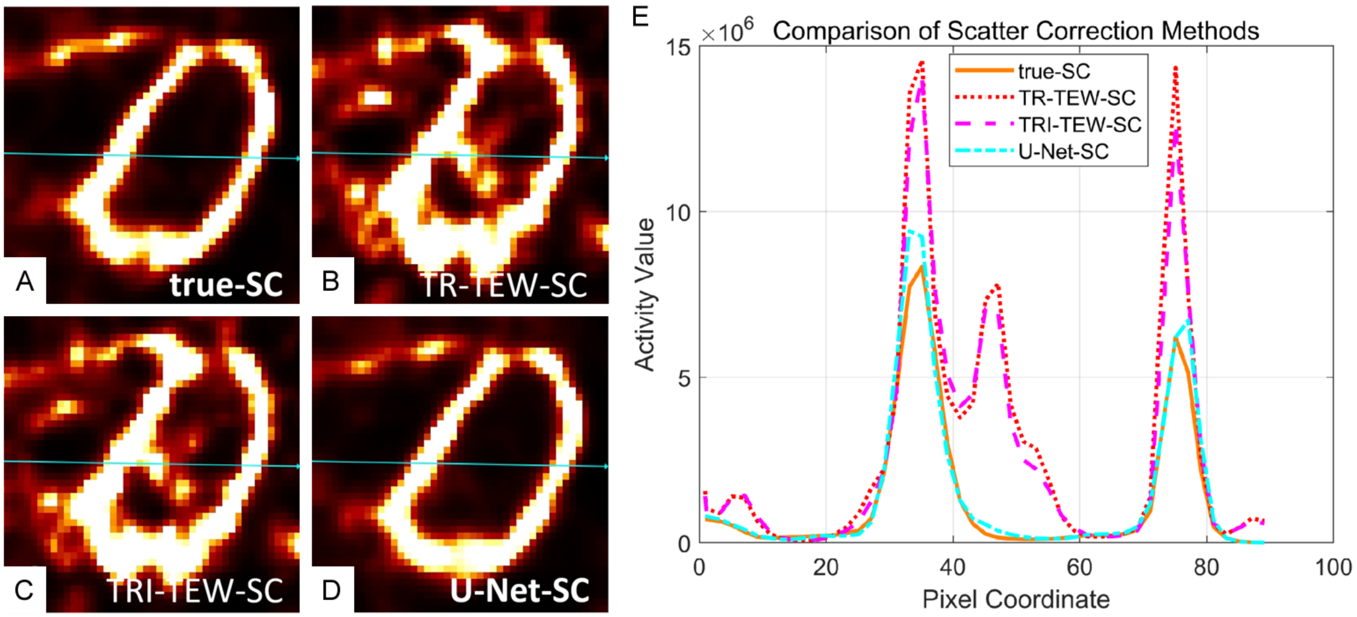

Figure 9. Reconstructed images of the heart region (A-D) and the corresponding profiles (E) obtained using different scatter-correction methods in the H-Phantom.

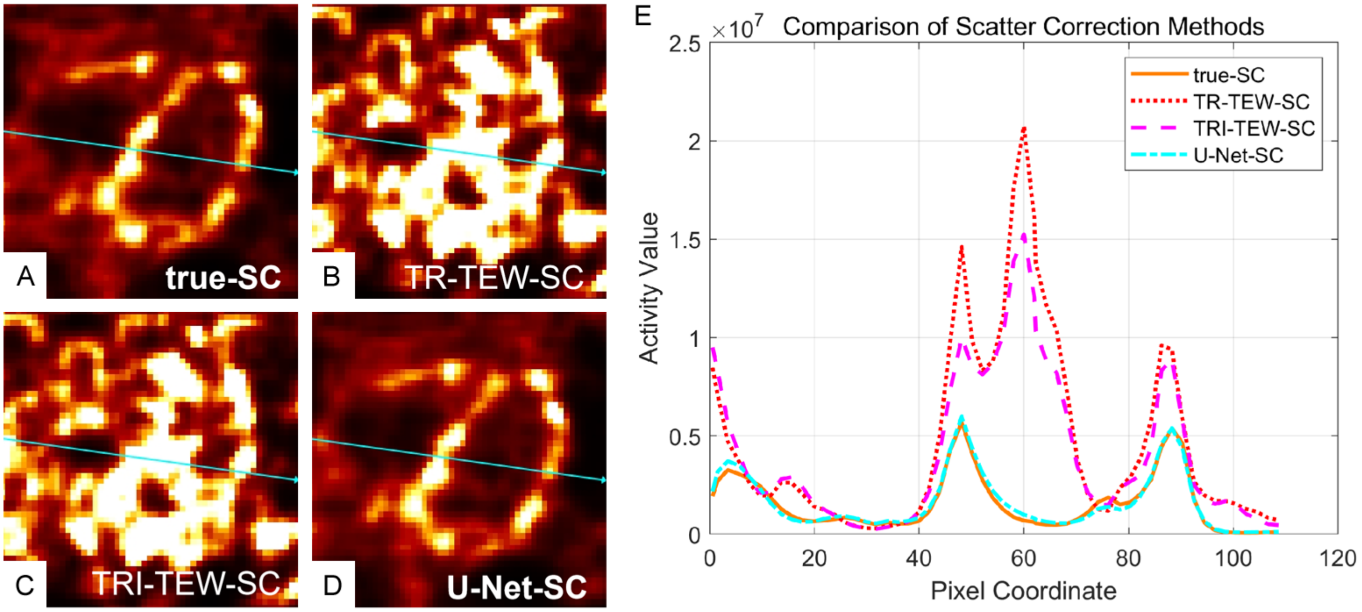

Figure 10. Reconstructed images of the heart region (A-D) and the corresponding profiles (E) obtained using different scatter-correction methods in the L-Phantom.

9.88. This is close to the ideal value from true-SC. For the CNR, the value was 9.37 before any correction. After applying true-SC, the CNR became 13.45. U-Net-SC achieved a similar CNR of 13.65. This shows good contrast with less noise. In the blood pool, RN increased from 0.48 to 0.67 with true-SC and to 0.65 with U-Net-SC. Both methods add more noise than no correction. But the TEW methods performed worse than U-Net-SC in controlling noise.

Table 3 shows the quantitative results for the L-Phantom images. With true-SC, the CRC was 0.76. Before correction (w/o-SC), it was 0.45. U-Net-SC gave a CRC of 0.74. This is close to the true-SC result. The TEW methods did not perform as well. TR-TEW-SC had a CRC of 0.44, and TRI-TEW-SC had 0.49. The M-B ratio improved from 1.91 to 2.52 with true-SC; U-Net-SC achieved 2.48, again near the ideal level, whereas the TEW methods yielded only 1.87-1.89. For CNR, true-SC raised the value from 2.49 to 3.14, and U-Net-SC attained a comparable 3.09, far above the TEW results (1.19 and 1.37). In terms of RN, both true-SC and U-Net-SC gave 0.48 (up from 0.36 without SC), while the TEW methods introduced substantially higher noise (0.74 and 0.65). Overall, U-Net-SC delivers

Table 2. Contrast recovery coefficients (CRC), Myocardium-to-blood pool (M-B ratio), contrast-to-noise ratios (CNR) and relative noise (RN) for each reconstructed image of H-Phantom

H-Phantom	w/o-SC	true-SC	TR-TEW-SC	TRI-TEW-SC	U-Net-SC
CRC	0.50±0.01	1.00±0.03	0.55±0.02	0.56±0.02	0.99±0.03
M-B ratio	5.45±0.11	9.97±0.23	5.94±0.14	6.06±0.14	9.88±0.31
CNR	9.37±0.43	13.45±0.77	5.34±0.20	6.13±0.18	13.65±0.87
RN	0.48±0.02	0.67±0.03	0.93±0.02	0.83±0.02	0.65±0.03

The data in the tables are presented as the mean ± standard deviation from 10 sets of reconstruction results.

Table 3. Contrast recovery coefficients (CRC), Myocardium-to-blood pool (M-B ratio), contrast-to-noise ratios (CNR) and relative noise (RN) for each reconstructed image of L-Phantom

L-Phantom	w/o-SC	true-SC	TR-TEW-SC	TRI-TEW-SC	U-Net-SC
CRC	0.45±0.01	0.76±0.01	0.44±0.01	0.49±0.01	0.74±0.02
M-B ratio	1.91±0.02	2.52±0.03	1.87±0.03	1.89±0.03	2.48±0.03
CNR	2.49±0.06	3.14±0.07	1.19±0.03	1.37±0.03	3.09±0.09
RN	0.36±0.01	0.48±0.01	0.74±0.02	0.65±0.26	0.48±0.01

The data in the tables are presented as the mean ± standard deviation from 5 sets of reconstruction results.

Table 4. Structural similarity index measure (SSIM) and normalized mean square error (NMSE) for each reconstructed image of H-Phantom and L-Phantom

H&L	w/o-SC	TR-TEW-SC	TRI-TEW-SC	U-Net-SC
SSIM$^{(H)}$	0.66±0.01	0.15±0.01	0.16±0.01	0.65±0.01
NMSE$^{(H)}$	0.25±0.17	0.59±0.14	0.55±0.15	0.23±0.18
SSIM$^{(L)}$	0.69±0.01	0.13±0.01	0.14±0.01	0.66±0.01
NMSE$^{(L)}$	0.13±0.05	0.77±0.02	0.66±0.03	0.10±0.05

Each reconstructed image with different scatter correction methods was evaluated using the true-SC reconstruction as the reference.

performance close to true-SC across all metrics and is markedly superior to the conventional TEW methods in this low-contrast setting, which is consistent with the H-Phantom results.

The SSIM and NMSE of reconstructed images of H-Phantom and L-Phantom were shown in **Table 4**. After applying TR-TEW-SC and TRI-TEW-SC, the SSIM of images was much lower than w/o-SC and the NMSE was much higher, indicating that both TEW-based methods introduced considerable noise and degraded structural fidelity. Our U-Net-SC method produced better images. For the H-Phantom, the SSIM was 0.65. The NMSE was 0.23. The U-Net-SC image looks more like the true-SC reference image. The images from the TEW methods do not match as well. We saw the same for the L-Phantom. We checked the SSIM and NMSE for U-Net-SC and both TEW methods. In every case, U-Net-SC had a higher SSIM number and a lower NMSE number.

Discussion

In this work, we studied scatter correction methods for a novel self-collimating SPECT system with a staggered mosaic scintillator structure. This system provides both high sensitivity and high spatial resolution. Monte Carlo simulations of the XCAT phantoms were performed using GATE to evaluate image quality and quantitative accuracy. The TEW scatter correction methods provided some improvement in quantitative results, but the improvement in overall image quality was limited. One reason is that the energy spectrum varies across detector layers, while the TEW-based methods were applied using the total energy spectrum. In addition, TEW-based methods introduce increased image noise. In contrast, the proposed U-Net scatter correction method significantly improved quantitative accuracy and reduced image noise. These results indicate that the deep-learning-based approach is more effective for this self-collimating SPECT system.

To mitigate the partial volume effect, reduced ROI sizes were used for quantification in this study. To better separate the effects of scatter correction and ROI reduction on quantification performance, we conducted an additional analysis using the original (non-reduced) ROI sizes. The corresponding quantification results for both the H-Phantom and the L-Phantom were summarized in the Supplementary Materials. As shown in Table S1 for the H-Phantom using the original ROI sizes, accurate scatter correction increased the CRC from 0.30 (w/o-SC) to 0.49 (true-SC). The proposed U-Net-SC achieved a comparable CRC of 0.48, outperforming the conventional TEW-based methods. Because the original ROI sizes were used in this analysis, the CRC improvement from 0.30 to 0.48 reflects the contribution of the scatter correction method alone. We then compared these results with those reported in **Table 2**, where reduced ROI sizes were used. Without scatter correction (w/o-SC), reducing the ROI size alone increased the CRC for the H-Phantom from 0.30 to 0.50, indicating that this improvement is attributable solely to mitigation of partial volume effects. Overall, these results suggest that the quantitative improvements from scatter correction and ROI optimization are of comparable magnitude. Under the optimized ROI selection, the proposed U-Net-SC further increased the CRC from 0.50 to 0.99. As shown in Table S2, the L-Phantom exhibits the same quantitative trends as the H-Phantom. In future work, explicit partial volume correction methods [35, 36] could be incorporated to further reduce partial volume effects and improve quantitative accuracy.

We also observed that the overall image quality and quantitative accuracy achieved with the H-Phantom were superior to those of the L-Phantom. It has two main reasons. First, the higher background activity in the L-Phantom leads to more pronounced scattering from the blood pool to the myocardial. Second, a lower myocardium-to-background ratio in the L-Phantom generally requires more reconstruction iterations to achieve clearer contrast. In this study, all results were reconstructed using the same 10,000 iterations. For comparison, we also reconstructed the L-Phantom with 20,000 iterations using true-SC, which showed that the CRC increased from 0.76 to 0.93 and the M-B ratio increased from 2.52 to 2.85. However, this improvement came at the cost of a noticeable increase in RN, from 0.48 to 0.65. This shows that increasing the number of reconstruction iterations improves contrast but also increases noise. For a fair comparison between the two phantoms, we used the same iteration number for image reconstruction in this study. To further improve low-contrast image quality, future work may focus on optimizing the U-Net framework. This can include increasing low-count and low-contrast samples in the training data and adding data with different noise levels. In addition, attention mechanisms can be introduced to help the network focus on relevant regions and suppress background noise under low signal-to-noise conditions.

To evaluate the model's generalization ability to anatomical size and gender variations, we performed an additional experiment on the current U-Net model using a female, small-sized phantom (W-H-Phantom). The results are provided in the Supplementary Materials. This phantom represents an extreme case of anatomical variation in terms of gender and size within the training dataset. The activity distribution of the W-H-Phantom was kept the same as that of the H-Phantom, as shown in Figure S1, and was used for the GATE Monte Carlo simulation. As shown in Figure S2, the U-Net-SC reconstructed result is close to the true-SC result and shows better performance than the conventional TEW methods. The quantitative results are summarized in Table S3. For the W-H-Phantom, the U-Net-SC achieves a CRC of 1.09, which is close to the true-SC value of 1.15 and higher than the TEW result (0.77). Noise-related metrics show similar trends to those observed in previous experiments. These results suggest that the proposed model is able to handle anatomical size and gender variations within the range covered by the training data. The current framework appears to be suitable for common adult anatomical variations. If future studies require more extreme anatomical cases, they can be addressed by expanding the training dataset.

We also tried training the network using only the first four activity distribution groups from **Table 1**. With this limited training data, the U-Net-SC results for the L-Phantom were not good enough. The main reason is that the myocardium-to-blood pool activity ratio in the L-Phantom is very different from the ratios covered in our smaller training set. These findings emphasize the necessity of close agreement between training and testing activity distributions. It is anticipated that the performance of this U-Net network can be further enhanced by expanding both the diversity and volume of the training data.

For clinical generalizability, the diversity of the training data is more important than the number of samples. A robust dataset should span a wide range of activity distributions, from low-dose and low-contrast cases to standard-dose cases, and to high-contrast cases. These distributions should be guided by physiological knowledge and clinical imaging protocols. To support this point, we trained the 3D U-Net using dataset groups 1, 3, 5, and 6 in **Table 1**, which cover a broader range of activity ratios. With this training set, the model achieved CRC values of 0.98 for the H-Phantom and 0.74 for the L-Phantom, which are close to the values reported in **Tables 2** and **3**, respectively. This confirms that coverage of activity distribution is more critical than simply increasing the dataset size.

Computational efficiency is indeed an important factor for clinical application. In our experiments, using a workstation equipped with an NVIDIA GeForce RTX 3090 GPU, the trained U-Net model requires approximately 2 seconds to process one reconstructed image. In comparison, conventional TEW scatter correction combined with OSEM reconstruction (10,000 iterations) takes about 1 hour. Therefore, the additional time introduced by the U-Net scatter correction is negligible, accounting for approximately 0.06% of the total reconstruction time. Once the reconstructed image is available, the U-Net can perform scatter correction almost in real time. In addition, TEW-based methods require projection data from multiple energy windows, whereas the U-Net-based approach only requires data from the main energy window. This further reduces data handling complexity. Combined with the high-sensitivity characteristics of the self-collimating SPECT system, the proposed method is computationally practical for quantitative SPECT imaging.

Although the proposed method was designed for single-radionuclide ^{99m}Tc imaging, its generalizability allows flexible extension to other single- or dual-radionuclide imaging scenarios, such as ^{99m}Tc/^{123}I or ^{201}Tl/^{99m}Tc. In practice, this can be achieved by using a dual-channel U-Net with two input channels and two output channels. Ideal single-radionuclide images can be generated separately and used as reference labels. Training data are then created by simulating images that include both scatter and inter-isotope crosstalk. The uncorrected reconstructed images of the two radionuclides are used as network inputs, and the network outputs the corrected activity distributions for both radionuclides simultaneously. Compared with the single-radionuclide case, the main added complexity of dual-radionuclide imaging lies in training data generation. It requires carefully designed datasets that cover different activity ratios and crosstalk levels between the two isotopes. In contrast, the network architecture and the prediction process remain similar to those used in the single-isotope case.

Conclusion

In this GATE simulation study of a novel self-collimating SPECT system, we investigated the effects of photon scatter on image quality and proposed a deep learning-based scatter correction method. We demonstrate that photon scatter substantially degrades image quality and quantitative accuracy, and effective scatter correction is essential for this system. Compared to conventional trapezoidal and triangular triple-energy-window methods, the proposed U-Net scatter correction method consistently achieves superior performance across all evaluated metrics in terms of contrast recovery, myocardium-to-blood-pool ratios, noise characteristics, and structural fidelity that closely approached the ideal true-scatter-corrected reference. These findings indicate that deep learning-based scatter correction can more accurately model complex scatter behavior than conventional energy-window-based approaches, and offers a practical and effective solution to enhance quantitative reliability in self-collimating SPECT imaging.

Acknowledgements

This study was supported by Beijing Natural Science Foundation (Z220010), National Natural Science Foundation of China (No. 12305379, No. 12205166), and CIRP Open Fund of Radiation Protection Laboratories (TNMPP-2024-No.1).

Disclosure of conflict of interest

None.

Address correspondence to: Jing Wu, Center for Advanced Quantum Studies, School of Physics and Astronomy, Key Laboratory of Multiscale Spin Physics (Ministry of Education), Beijing Normal University, Beijing, China. E-mail: jingwu@bnu.edu.cn; Tianyu Ma and Hui Liu, Department of Engineering Physics, Tsinghua University, Beijing, China. E-mail: maty@tsinghua.edu.cn (TYM); liuhui2020@tsinghua.edu.cn (HL)

References

[1] Niu X, Yang Y, Jin M, Wernick MN and King MA. Effects of motion, attenuation, and scatter corrections on gated cardiac SPECT reconstruction. Med Phys 2011; 38: 6571-6584.

[2] Hutton BF, Buvat I and Beekman FJ. Review and current status of SPECT scatter correction. Phys Med Biol 2011; 56: R85-112.

[3] Jacobs F, Sundermann E, De Sutter B, Christiaens M and Lemahieu I. A fast algorithm to calculate the exact radiological path through a pixel or voxel space. J Comput Inf Technol 1998; 6: 89-94.

[4] Berker Y and Li Y. Attenuation correction in emission tomography using the emission data--A review. Med Phys 2016; 43: 807-832.

[5] Fan P, Hutton BF, Holstensson M, Ljungberg M, Pretorius PH, Prasad R, Ma T, Liu Y, Wang S, Thorn SL, Stacy MR, Sinusas AJ and Liu C. Scatter and crosstalk corrections for (99m)Tc/(123)I dual-radionuclide imaging using a CZT SPECT system with pinhole collimators. Med Phys 2015; 42: 6895-6911.

[6] Velo AF, Fan P, Xie H, Chen X, Boutagy N, Feher A, Sinusas AJ, Ljungberg M and Liu C. (99m)Tc/(123)I dual-radionuclide correction for self-scatter, down-scatter, and tailing effect for a CZT SPECT with varying tracer distributions. IEEE Trans Radiat Plasma Med Sci 2023; 7: 839-850.

[7] Yang JT, Yamamoto K, Sadato N, Tsuchida T, Takahashi N, Hayashi N, Yonekura Y and Ishii Y. Clinical value of triple-energy window scatter correction in simultaneous dual-isotope single-photon emission tomography with 123I-BMIPP and 201Tl. Eur J Nucl Med 1997; 24: 1099-1106.

[8] Noori-Asl M, Sadremomtaz A and Bitarafan-Rajabi A. Evaluation of three scatter correction methods based on estimation of photopeak scatter spectrum in SPECT imaging: a simulation study. Phys Med 2014; 30: 947-953.

[9] Michael K, Frangos S, Iakovou I, Lontos A, Demosthenous G and Parpottas Y. The impact of dual and triple energy window scatter correction on I-123 postsurgical thyroid SPECT/CT imaging using a phantom with small sizes of thyroid remnants. Life (Basel) 2024; 14: 113.

[10] Robinson AP, Tipping J, Cullen DM and Hamilton D. The influence of triple energy window scatter correction on activity quantification for (177)Lu molecular radiotherapy. Phys Med Biol 2016; 61: 5107-5127.

[11] Liu H, Wu J, Miller EJ, Liu C, Liu Y and Liu YH. Diagnostic accuracy of stress-only myocardial perfusion SPECT improved by deep learning. Eur J Nucl Med Mol Imaging 2021; 48: 2793-2800.

[12] Chen X, Zhou B, Xie H, Shi L, Liu H, Holler W, Lin M, Liu YH, Miller EJ, Sinusas AJ and Liu C. Direct and indirect strategies of deep-learning-based attenuation correction for general purpose and dedicated cardiac SPECT. Eur J Nucl Med Mol Imaging 2022; 49: 3046-3060.

[13] Kim H, Li Z, Son J, Fessler JA, Dewaraja YK and Chun SY. Physics-guided deep scatter estimation by weak supervision for quantitative SPECT. IEEE Trans Med Imaging 2023; 42: 2961-2973.

[14] Ronneberger O, Fischer P and Brox T. U-Net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015; 2015: 234-241.

[15] Lian S, Luo Z, Zhong Z, Lin X, Su S and Li S. Attention guided U-Net for accurate iris segmentation. J Visual Commun Image Represent 2018; 56: 296-304.

[16] Chen W, Liu B, Peng S, Sun J and Qiao X. S3D-UNet: separable 3D U-Net for brain tumor segmentation. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2019. pp. 358-368.

[17] Liu H, Wu J, Lu W, Onofrey JA, Liu YH and Liu C. Noise reduction with cross-tracer and cross-protocol deep transfer learning for low-dose PET. Phys Med Biol 2020; 65: 185006.

[18] Liu H, Yousefi H, Mirian N, Lin M, Menard D, Gregory M, Aboian M, Boustani A, Chen MK, Saperstein L, Pucar D, Kulon M and Liu C. PET image denoising using a deep-

learning method for extremely obese patients. IEEE Trans Radiat Plasma Med Sci 2022; 6: 766-770.
[19] Liu H, Zhang Y, Lyu Z, Cheng L, Gao L, Wu J and Liu Y. Investigation of a deep learning-based reconstruction approach utilizing dual-view projection for myocardial perfusion SPECT imaging. Am J Nucl Med Mol Imaging 2025; 15: 15-27.
[20] Yang J, Park D, Gullberg GT and Seo Y. Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain (18)F-FDG PET. Phys Med Biol 2019; 64: 075019.
[21] Laurent B, Bousse A, Merlin T, Nekolla S and Visvikis D. PET scatter estimation using deep learning U-Net architecture. Phys Med Biol 2023; 68: 065004.
[22] Moore SC, Kouris K and Cullum I. Collimator design for single photon emission tomography. Eur J Nucl Med 1992; 19: 138-150.
[23] Van Audenhaege K, Van Holen R, Vandenberghe S, Vanhove C, Metzler SD and Moore SC. Review of SPECT collimator selection, optimization, and fabrication for clinical and preclinical imaging. Med Phys 2015; 42: 4796-4813.
[24] Wu J and Liu C. Recent advances in cardiac SPECT instrumentation and imaging methods. Phys Med Biol 2019; 64: 06TR01.
[25] Ma T, Wei Q, Lyu Z, Zhang D, Zhang H, Wang R, Dong J, Liu Y, Yao R and He ZX. Self-collimating SPECT with multi-layer interspaced mosaic detectors. IEEE Trans Med Imaging 2021; 40: 2152-2169.
[26] Zhang D, Lyu Z, Liu Y, He ZX, Yao R and Ma T. Characterization and assessment of projection probability density function and enhanced sampling in self-collimation SPECT. IEEE Trans Med Imaging 2023; 42: 2787-2801.
[27] Zhang D, Cheng L, Hu Y, Lyu Z, Wang L, Fang W, Yao R and Ma T. Design optimization and simulated performance evaluation of a self-collimated cardiac SPECT with simultaneous high sensitivity and resolution. Med Phys 2025; 52: 3823-3839.
[28] Wang R, Zhang D, Hu Y, Lyu Z and Ma T. High-sensitivity cardiac SPECT system design with collimator-less interspaced mosaic-patterned scintillators. Front Med (Lausanne) 2023; 10: 1145351.
[29] Zhao Y, Guo L, Liu H, Lyu Z, Zhang D, Liu Y, Ma T and Wu J. Comparison of three triple-energy-window scatter correction approaches for self-collimating SPECT. J Nucl Med 2024; 65: 241929.
[30] Zhao Y, Liu H, Guo L, Jiang W, Lyu Z, Zhang D, Ma T and Wu J. Deep learning-based scatter correction for self-collimating SPECT. 2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD) 2024; 1-1.
[31] Kayal G, Chauvin M, Vergara-Gil A, Clayton N, Ferrer L, Moalosi T, Knoll P, Struelens L and Bardiès M. Generation of clinical 177Lu SPECT/CT images based on Monte Carlo simulation with GATE. Phys Med 2021; 85: 24-31.
[32] Wu J, Guo L, Liu H, Lyu Z, Zhao Y, Zhang D, He ZX, Liu Y and Ma T. Accurate attenuation correction for self-collimating SPECT imaging. 2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD) 2023; 1-1.
[33] Ogawa K, Harata Y, Ichihara T, Kubo A and Hashimoto S. A practical method for position-dependent Compton-scatter correction in single photon emission CT. IEEE Trans Med Imaging 1991; 10: 408-412.
[34] Buvat I, Rodriguez-Villafuerte M, Todd-Pokropek A, Benali H and Di Paola R. Comparative assessment of nine scatter correction methods based on spectral analysis using Monte Carlo simulations. J Nucl Med 1995; 36: 1476-1488.
[35] Wu J, Liu H, Hashemi Zonouz T, Sandoval VM, Mohy-Ud-Din H, Lampert RJ, Sinusas AJ, Liu C and Liu YH. A blind deconvolution method incorporated with anatomical-based filtering for partial volume correction: validations with (123)I-mIBG cardiac SPECT/CT. Med Phys 2017; 44: 6435-6446.
[36] Qin J, Cheng L, Jiang N, Gao L, Huang S, Zhang D, Zhao R, Zhang J, Liu H, Liu Y and Wu J. Deconvolution-based partial volume correction for dynamic cardiac PET imaging of (18)F-FDG in rats using a time-variant point spread function. 2025 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD) 2025; 1-2.

Table S1. Contrast recovery coefficients (CRC), Myocardium-to-blood pool (M-B ratio), contrast-to-noise ratios (CNR) and relative noise (RN) for each reconstructed image of H-Phantom, quantified by using the original sizes of the left ventricular myocardium and left ventricular blood pool as ROIs

H-Phantom	w/o-SC	true-SC	TR-TEW-SC	TRI-TEW-SC	U-Net-SC
CRC	0.30±0.01	0.49±0.01	0.33±0.01	0.34±0.01	0.48±0.01
M-B ratio	3.74±0.04	5.43±0.07	3.98±0.06	4.05±0.06	5.35±0.08
CNR	4.13±0.07	4.72±0.08	3.17±0.08	3.48±0.09	4.65±0.09
RN	0.66±0.01	0.94±0.02	0.94±0.02	0.88±0.02	0.94±0.02

The data in the tables are presented as the mean ± standard deviation from 10 sets of reconstruction results.

Table S2. Contrast recovery coefficients (CRC), Myocardium-to-blood pool (M-B ratio), contrast-to-noise ratios (CNR) and relative noise (RN) for each reconstructed image of L-Phantom, quantified by using the original sizes of the left ventricular myocardium and left ventricular blood pool as ROIs

L-Phantom	w/o-SC	true-SC	TR-TEW-SC	TRI-TEW-SC	U-Net-SC
CRC	0.33±0.01	0.50±0.01	0.32±0.01	0.33±0.01	0.48±0.01
M-B ratio	1.65±0.01	2.01±0.01	1.63±0.02	1.65±0.02	1.96±0.02
CNR	1.62±0.03	1.87±0.04	0.88±0.02	1.02±0.03	1.76±0.03
RN	0.40±0.01	0.54±0.01	0.71±0.01	0.64±0.01	0.55±0.01

The data in the tables are presented as the mean ± standard deviation from 5 sets of reconstruction results.

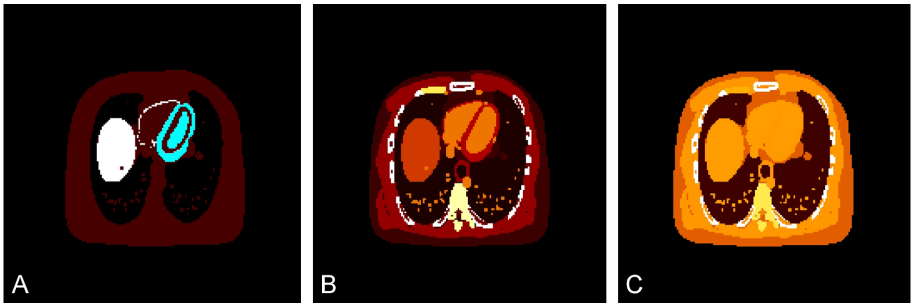

Figure S1. W-H-Phantom. A. Small-sized XCAT female phantom slices. The activity ratio was set at 10:10:0:1 for myocardium, liver, lungs, and background (including blood). ROIs for the LVM and LVBP are outlined. B. Tissue distribution image. C. Linear attenuation coefficient image.

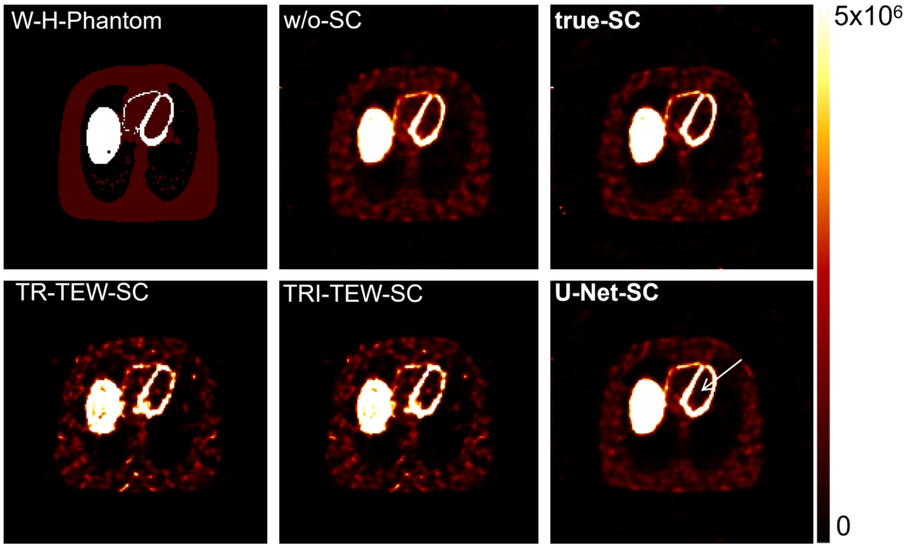

Figure S2. Reconstructed images with different scatter correction methods for W-H-Phantom.

Table S3. Contrast recovery coefficients (CRC), Myocardium-to-blood pool (M-B ratio), contrast-to-noise ratios (CNR) and relative noise (RN) for each reconstructed image of W-H-Phantom

W-H-Phantom	w/o-SC	true-SC	TR-TEW-SC	TRI-TEW-SC	U-Net-SC
CRC	0.56	1.15	0.77	0.77	1.09
M-B ratio	6.00	11.34	7.91	7.97	10.77
CNR	9.75	16.03	6.18	6.98	13.95
RN	0.51	0.65	1.12	1.00	0.70