Am J Nucl Med Mol Imaging 2025;15(1):15-27
www.ajnmmi.us /ISSN:2160-8407/ajnmmi0162985

Original Article

Investigation of a deep learning-based reconstruction approach utilizing dual-view projection for myocardial perfusion SPECT imaging

Hui Liu[1,2], Yajing Zhang[3], Zhenlei Lyu[1], Li Cheng[4], Lilei Gao[4], Jing Wu[5,6], Yaqiang Liu[1,2]

[1]Department of Engineering Physics, Tsinghua University, Beijing, China; [2]Key Laboratory of Particle and Radiation Imaging (Tsinghua University), Ministry of Education, Beijing, China; [3]Department of Nuclear Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China; [4]Chengdu Novel Medical Equipment Ltd., Chengdu, Sichuan, China; [5]Center for Advanced Quantum Studies, School of Physics and Astronomy, Beijing Normal University, Beijing, China; [6]Key Laboratory of Multiscale Spin Physics (Ministry of Education), Beijing Normal University, Beijing, China

Received December 30, 2024; Accepted February 23, 2025; Epub February 25, 2025; Published February 28, 2025

Abstract: Single-photon emission computed tomography (SPECT) is widely used in myocardial perfusion imaging (MPI) in clinic. However, conventional dual-head SPECT scanners require lengthy scanning times and gantry rotation, which limits the application of SPECT MPI. In this work, we proposed a deep learning-based approach to reconstruct dual-view projections, aiming to reduce acquisition time and enable non-rotational imaging for MPI based on conventional dual-head SPECT scanners. U-Net was adopted for the dual-view projection reconstruction. Initially, 2D U-Nets were used to evaluate various data organization schemes for dual-view projection as input, including paved projection, interleaved projection, and stacked projection, with and without an attenuation map. Subsequently, we developed 3D U-Nets using the optimal data organization scheme as input to further enhance reconstruction performance. The dataset consisted of a total of 116 SPECT/CT scans with ^{99m}Tc-tetrofosmin tracer acquired on a GE NM/CT 640 scanner. Reconstruction performance was assessed using quantitative metrices and absolute percentage errors, while the reconstruction images from the full-view projection were used as reference images. The 2D U-Nets provided reasonable transverse view images but exhibited slight axial discontinuity compared to the reference images, regardless of the data organization schemes. Incorporating the attenuation map reduced this axial discontinuity. Quantitatively, the 2D U-Net trained using both stacked projection and attenuation map achieved the best performance, with a normalized mean absolute error of 0.6%±0.3% and a structural similarity index measure (SSIM) of 0.93±0.04. The 3D U-Net further improved the performance with less axial discontinuity and a higher SSIM of 0.94±0.03. The localized absolute percentage errors were 1.8±16.8% and -2.0±6.3% in the left ventricular (LV) cavity and myocardium, respectively. We developed a deep learning-based image reconstruction approach for dual-view projection from a conventional SPECT scanner. The 3D U-Net, trained with the stacked projection with an attenuation map is effective for non-rotational imaging and could benefit dynamic myocardium perfusion imaging.

Keywords: Deep learning, image reconstruction, dual view projection, myocardial perfusion imaging

Introduction

Myocardial perfusion imaging (MPI) with single photon emission computed tomography (SPECT) has been widely used for patients with known or suspected coronary artery disease (CAD) [1, 2]. Despite the commercialization of several advanced cardiac-dedicated scanners, conventional dual-head scanners equipped with parallel-hole collimators remain the most commonly used for clinical SPECT MPI [3-5]. However, due to the imaging principles of conventional SPECT, gantry rotation is required to acquire sufficient projections for accurate image reconstruction with minimized artifacts [6]. This results in prolonged long acquisition times, which is typically around 20 minutes for a 180°-scanning orbit with 30 steps in clinical setting [7]. Such long acquisition times not only reduce the patient comfort but also degrade the image quality that is potentially caused by involuntary body movements [8, 9]. Additionally, the long acquisition time for SPECT imaging provides little benefit for myocardial blood flow quantification, which has demonstrated supplementary prognostic values in CAD assessment [10-13]. Moreover, the efficiency of the scanner is constrained by the acquisition time [14]. Therefore, reducing acquisition time is crucial for enhancing the clinical utility of SPECT MPI [15].

In practice, there are two approaches to shortening the SPECT acquisition time: reducing the scanning time for each projection and decreasing the number of projections [16]. However, the former approach results in elevated noise levels, similar to those observed in the low-dose imaging [17]. Numerous methods have been proposed to reduce the image noise, including non-local mean filters, block-matching 3D filters, and deep-learning based denoising methods [18-20]. The latter approach leads to streak artifacts, due to reduced angular sampling. Some groups are investigating the use of image prior information to alleviate the artifact [21]. Recently, deep learning-based methods have shown promises in enhancing

https://doi.org/10.62347/MLFB9278

reduced projection reconstruction through sinogram synthesis [22-24]. Tobias Rydén et al. proposed a convolutional U-Net-shaped synthetic intermediate projections (CUSIP) approach, which generated the projections of 120 angles from those of 30 angles with a minor bias for clinical dosimetry protocols in ^{177}Lu SPECT [25]. Si Li et al. also proposed a deep neural network combining LSTM and U-Net to synthesize sparse-view sinogram for SPECT reconstruction of 1/2-, 1/3- and 1/4-view projection data, which was then evaluated by a simulation study [26]. Different from synthesizing sinograms in the projection domain, Xiongchao Chen et al. proposed a deep-learning-based dual-domain sinogram synthesis (DuDoSS) method, which utilized the information of both projection and image domains to synthesize the projection of 60 angles from the projection of 15 angles, and demonstrated its superior performance than other end-to-end network based on a training dataset of 250 patient studies [27].

However, these methods still required at least 15 angles' projection for SPECT MPI, meaning that the gantry rotation remains unavoidable. As a result, their application in dynamic SPECT MPI is limited due to the slow gantry rotation. In this study, we proposed a deep learning-based reconstruction method for dual-view projection obtained from conventional dual-head SPECT/CT scanner without gantry rotation. The performance of the proposed method was evaluated using clinical dataset. Meanwhile, we investigated the impact of the data organization schemes for network input and the dimensionality of the neural network on reconstruction performance.

Materials and methods

Subjects and SPECT/CT scanning protocols

This study included 69 subjects, who underwent routine SPECT/CT MPI using ^{99m}Tc-tetrofosmin tracer on a GE NM/CT 640 SPECT/CT scanner. These subjects included 33 males and 36 females with an average body mass index (BMI) of 30.3±6.0 kg/m^2 and ages ranging from 31 to 86 years. The study was approved by the institutional review board of Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, and all subjects provided informed consent. A total of 116 SPECT/CT scans were obtained, with 47 subjects undergoing a two-day rest and stress protocol and 22 subjects undergoing a one-day stress protocol. The average injection dose across all the scans was approximately 15 mCi. Low-energy high-resolution parallel-hole collimators were utilized for the SPECT scanning. The SPECT projections were obtained with an energy window of 126.5 to 154.5 keV in L mode, with 60 angles over 180 degrees from the right anterior oblique (RAO) view to the left posterior oblique (LPO) view. The dimension of the obtained projection was 64×64 and the pixel size was 6.79×6.79 mm^2. Subsequently, CT scans were performed with 120 kVp and 30 mAs. The attenuation map (ATT MAP) for 140 keV gamma photons was generated from the well-registered CT images, and then used for the attenuation correction in SPECT image reconstruction.

SPECT image reconstruction and dataset preparation

The image reconstruction was performed using a self-developed in-house reconstruction software based on maximum likelihood expectation maximization (MLEM) algorithm with resolution recovery, attenuation correction and scatter correction [28, 29]. The iteration number was 60 for the MLEM reconstruction. The SPECT images had dimensions of 64×64×64 with a voxel size of 6.79×6.79×6.79 mm^3. Then the SPECT images obtained from the MLEM reconstruction using the full-view projection (60 angles), were used as label images.

To obtain the dataset for network training and testing, we generated 30 pairs of dual-view projections and label images from the full-view projection for each scan. While the full-view projection ranged from RAO to LPO, we used the projections of the anterior view and the left view as the dual-view projection for the reconstruction to obtain the SPECT images. Additionally, 29 other pairs of two-projection views with a 90-degree interval can also be used to form 29 dual-view projections. However, the corresponding label image must be rotated by the relative angle to the anterior view. Meanwhile, we obtained the ATT MAP using the same rotation operation for each dataset, which can be included as additional input for the neural network.

Network architecture

We adopted both 2D and 3D U-Net neural networks for the dual-view projection reconstruction [30]. The U-Net architecture consists of a contracting path, a bottleneck path and an expansive path, as shown in **Figure 1**. In the contracting path, we have an initial convolution layer with a rectified linear unit (ReLU), followed by four contracting layers. Each contracting layer is composed by two convolution layers with ReLU and a half-size max pooling layer. The bottleneck path includes two convolution layers with ReLU. The expansive path consists of four expanding layers, each comprising a twice-size up-convolution layer and two convolution layers with ReLU, followed by a final convolution layer with ReLU. Meanwhile, feature maps obtained from the up-convolution layer in the expanse path are concatenated with those obtained from the corresponding layers in the contracting path. The kernel size for the initial and the final convolution layers is 1×1 for 2D U-Net and 1×1×1 for 3D U-Net, while the kernel size for other convolution layers is 3×3 for 2D U-Net and 3×3×3 for 3D U-Net. The feature number in the first layer is 32.

Data organization scheme for the network input

Since U-Net requires that the input and output image sizes to be identical, we designed three different projec-

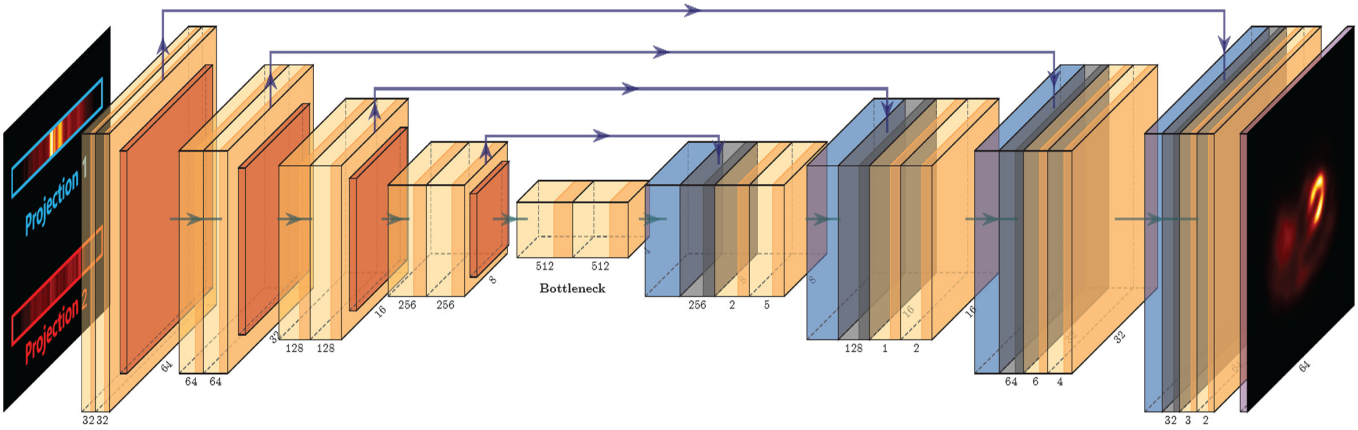

Figure 1. The U-Net architecture.

tion organization schemes to assemble the dual-view projection as the network input. **Figure 2A** shows the two projections to be assembled in our proposed dual-view projection method, each with a size of 64×64×1. **Figure 2B** shows the paved projection with a size of 64×64×64, which was obtained by repeating the two projections by 32 times and then concatenating them. **Figure 2C** shows the interleaved projection with a size of 64×64×64, which was generated by interleaving the two projections by 32 times. **Figure 2D** shows the stacked projection with a size of 64×64×64×2. To generate the stacked projection, each of the two projections was repeated by 64 times to obtain two 64×64 images. Then, these two images were rotated according to its projection angle and then were stacked together.

To compare the performance of different data organization schemes, we trained three 2D U-Net models using various inputs, as follows: 2D-A: paved projection, 2D-B: interleaved projection, 2D-C: stacked projection. Additionally, we trained four more 2D U-Net models to evaluate the benefit of incorporating an ATT MAP as an additional input, with the following configurations: 2D-D: ATT MAP only, 2D-E: Paved projection with ATT MAP, 2D-F: interleaved Projection with ATT MAP, 2D-G: stacked Projection with ATT MAP.

To further enhance performance, we proposed using a 3D U-Net with stacked projection as input. As the dataset is limited for 3D U-Nets training, a patch-based strategy was employed for data augmentation during the training process of 3D U-Nets, which is compatible with the stacked projection data organization scheme. The patch-based strategy was implemented by randomly sampling a patch with the size of 48×48×48 from the original input. So, for each dataset we can generate 4,096 different patches, leading to a total of 11,059,200 patches for the training dataset. We trained four 3D U-Net models as follows: 3D-A: stacked Projection with ATT MAP, 3D-B: stacked Projection with ATT MAP with patch-based strategy, 3D-C: ATT MAP with patch-based strategy, 3D-D: reconstruction image from the dual-view projection with patch-based strategy.

Network training and testing

Among the 116 scans, 78% (90 scans) were used for training, 7% (9 scans) for validation, and 15% (15 scans) for testing. The subjects in the testing dataset were distinct from those in the training dataset. The input sizes were 64×64 for the 2D U-Net, 64×64×32 for the 3D U-Net, and 32×32×16 for the 3D U-Net with the patch-based strategy.

We employed an L1 loss function combined with the L2 regularization on the parameters of the U-Nets and used the Adam optimizer for training. The initial learning rate was set to 0.0005, with an exponential decay rate of 0.99. All networks were trained over 3,200 epochs, each consisting of 100 batches with a batch size of 16. The network with the best performance on the validation dataset was selected for testing. All the training and testing were conducted on a cluster equipped with six NVIDIA A6000 GPUs.

Performance evaluation and statistical analysis

The SPECT images obtained from the MLEM reconstruction using the full-view projection (60 angles), were denoted as reference images for further evaluation.

We evaluated each trained network using the normalized mean absolute error (NMAE), normalized mean square error (NMSE), structural similarity index measure (SSIM), and peak signal-to-noise ratio (PSNR) as follows:

$$NMAE = \frac{\| I - I^{REF} \|^1}{N \times (\max(I^{REF}) - \min(I^{REF}))} \quad (1)$$

$$NMSE = \frac{\| I - I^{REF} \|^2}{\| I^{REF} \|^2} \quad (2)$$

$$SSIM = \frac{(2\mu_x\mu_y + c1)(2\sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x^2 + \sigma_y^2 + c2)} \quad (3)$$

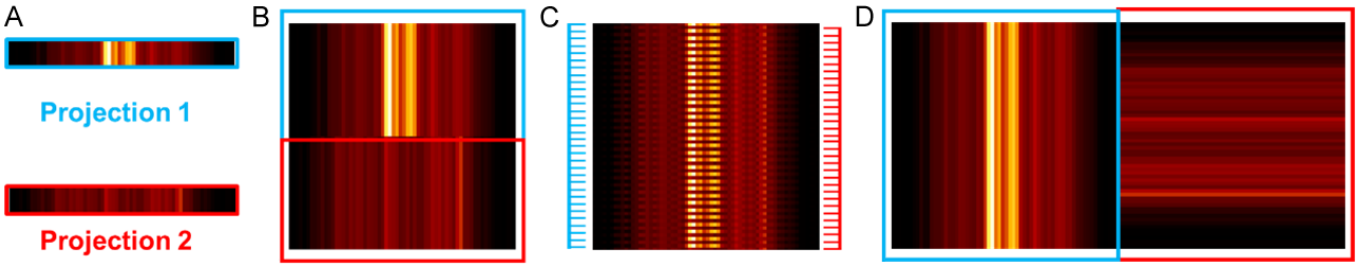

Figure 2. Three different data organization schemes for the network input. A. The two projections to be assembled; B. The paved projection; C. The interleaved projection; D. The stacked projection.

$$PSNR = 10 \times log_{10}\left(\frac{Max\left(I^{REF}\right)^2}{\| I - I^{REF} \|^2}\right) \quad (4)$$

Where I is the evaluated image; I^{REF} is the reference image; N is the total voxel number in the image; μ_x and μ_y are the mean values of the evaluated image and the reference image, respectively; σ_x and σ_y are the stand deviations of the evaluated images and the reference image, respectively; σ_{xy} is the cross-covariance for the two images. *c1* and *c2* are two small constants added to avoid division by zero or weak denominators.

We also generated bullseye polar maps from the images obtained using the 3D U-Nets with Carimas software [31], and compared these maps with those derived from the reference images. Additionally, we manually delineated regions of interest (ROIs) for the left ventricular (LV) cavity and myocardium. The regional percentage biases for the myocardium, LV cavity and three main coronary arteries, including Left anterior descending (LAD), right coronary artery (RCA) and left circumflex artery (LCX), were calculated based on AHA 17-segmental model [32] to further evaluate the images. The calculations were performed as follows:

$$Bias = \frac{\sum m \in \text{ROI}\, I_m^{Pred} - \sum m \in \text{ROI}\, I_m^{True}}{\sum m \in \text{ROI}\, I_m^{True}} \times 100\% \quad (5)$$

Where ROI is chosen for the myocardium, LV cavity, LAD, RAC and LCX. A paired t-test was used to determine whether the ROI biases are significantly different from zero.

For statistical analysis, a two-tailed paired t-tests were performed on the interested metrics between two testing groups. A p value lower than 0.05 means a significant difference of the metrics.

Results

2D U-Nets comparison

As shown in **Figure 3**, the predicted images using the seven networks appear reasonable in the transverse view when compared to the reconstruction images from full-view projections. In contrast, the conventional reconstruction algorithm based on the dual-view projection fails to produce satisfactory images due to the incomplete view sampling. Among the three U-Nets trained without ATT MAP as input, the image predicted by U-Net 2D-C, which was trained with stacked projection input, appears slightly better than those obtained with U-Nets 2D-A and 2D-B, trained with paved and interleaved projection input, respectively, particularly in the basal lateral regions. Since the images were predicted slice by slice, some discontinuities are observed in the axial direction. With the addition of ATT MAP, axial discontinuity was reduced for the four U-Nets, 2D-D, 2D-E, 2D-F, and 2D-G. Among these, the image obtained using U-Net 2D-G, trained by stacked projection with ATT MAP, appears superior to those obtained by the other three U-Nets.

Quantitatively, as shown in **Figure 4**, the 2D U-Net improved correlation compared to the dual-view reconstruction images. The addition of the ATT MAP further enhanced correlation while reducing variation. Although U-Net 2D-D did not yield quantitatively satisfactory images, exhibiting a notably high NMSE, the other six U-Nets significantly improved the quantitative metrics, as shown in **Table 1**, relative to the dual-view reconstruction images. Moreover, with the ATT MAP as an additional input, the quantitative metrics showed further improvement, evidenced by smaller NMAE and NMSE values, and larger SSIM and PSNR values. Notably, whether with and without the ATT MAP, the U-Nets trained by the stacked projection consistently outperformed the other U-Nets, indicating that the data organization scheme significantly impacts reconstruction performance.

3D U-Nets comparison

Based on the results obtained from the 2D U-Nets, we selected the stacked projection with ATT MAP for the dual-view projection 3D reconstruction.

As shown in **Figure 5**, the images predicted using 3D U-Nets exhibits improved performance regarding axial discontinuity compared to U-Net 2D-G. However, due to the limited size of the training dataset, discrepancies in the details of the myocardium are observed between the images predicted by the U-Net 3D-A and the reference

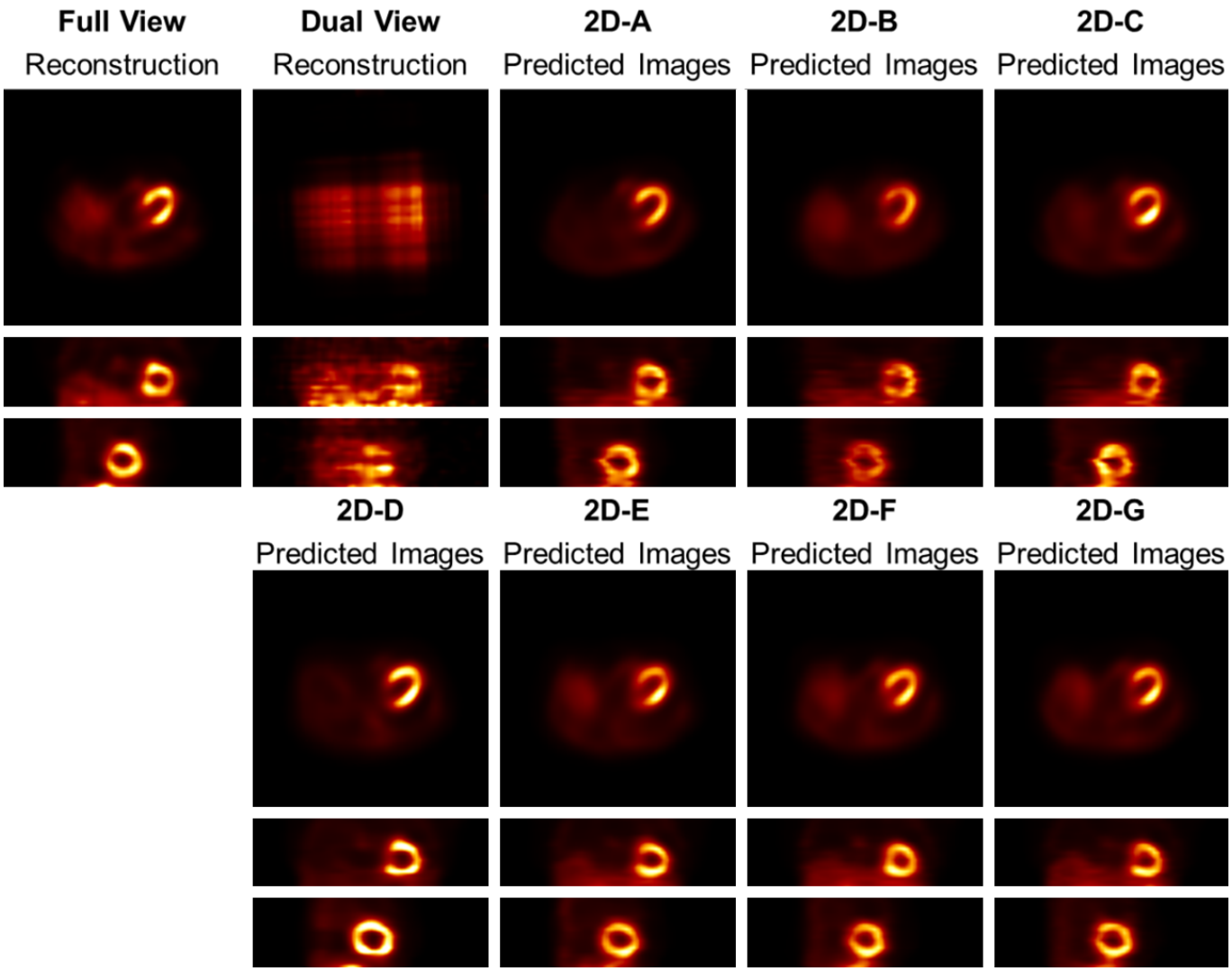

Figure 3. The sample slices in transverse view, coronal view and sagittal view (from top row to bottom row) of the full-view reconstruction images, dual-view reconstruction images, and the predicted images using various 2D U-Nets: 2D-A trained by paved projection, 2D-B trained by interleaved projection, 2D-C trained by stacked projection, 2D-D trained by ATT MAP, 2D-E trained by paved projection with ATT MAP, 2D-F trained by interleaved projection with ATT MAP, 2D-G trained by stacked projection with ATT MAP. All the slices were shown with the same color bar.

images. When employing the patch-based strategy, the images and the polar map predicted by U-Net 3D-B are the most similar to the reference, compared with those predicted by the U-Nets 3D-C and 3D-D. Additionally, U-Net 3D-C is unable to produce quantitatively accurate images due to the absence of SPECT information, which necessitated a different color bar for display.

As shown in **Figure 6**, the correlation for the 3D U-Nets, except for U-Net 3D-C, is much better than those for 2D U-Nets. No significant difference is observed between U-Nets 3D-A and 3D-B, both of which outperform U-Net 3D-D. Additionally, as indicated in **Table 2**, there is minimal difference in quantitative metrics between U-Nets 3D-A and 3D-B, except for a significant difference in SSIM ($P < 0.05$). However, U-Net 3D-B demonstrates significant differences compared to U-Nets 3D-C and 3D-D in terms of NMAE, SSIM, and PSNR (all $P < 0.05$).

ROI quantification for 3D U-Nets

Table 3 shows the ROI relative biases for the LV cavity, myocardium, LAD, RCA, and LCX for the predicted images from those 3D U-Nets and the dual-view reconstruction images. It is evident that the U-Net 3D-C, trained solely with ATT MAP, fails to produce quantitatively accurate images. In contrast, U-Nets 3D-A and 3D-B substantially reduce the ROI relative biases, outperforming U-Net 3D-D. Meanwhile, all ROI relative biases were not significantly different from zero ($P > 0.05$).

Figures 7 and **8** display the regression plots and Bland-Altman plots for the ROI intensity of LV cavity, myocardium, LAD, RCA and LCX across all 17 testing subjects. Both U-Nets 3D-A and 3D-B achieve higher correlation coefficients, with their regression lines from the two U-Nets being closer to the identity line compared to the U-Net

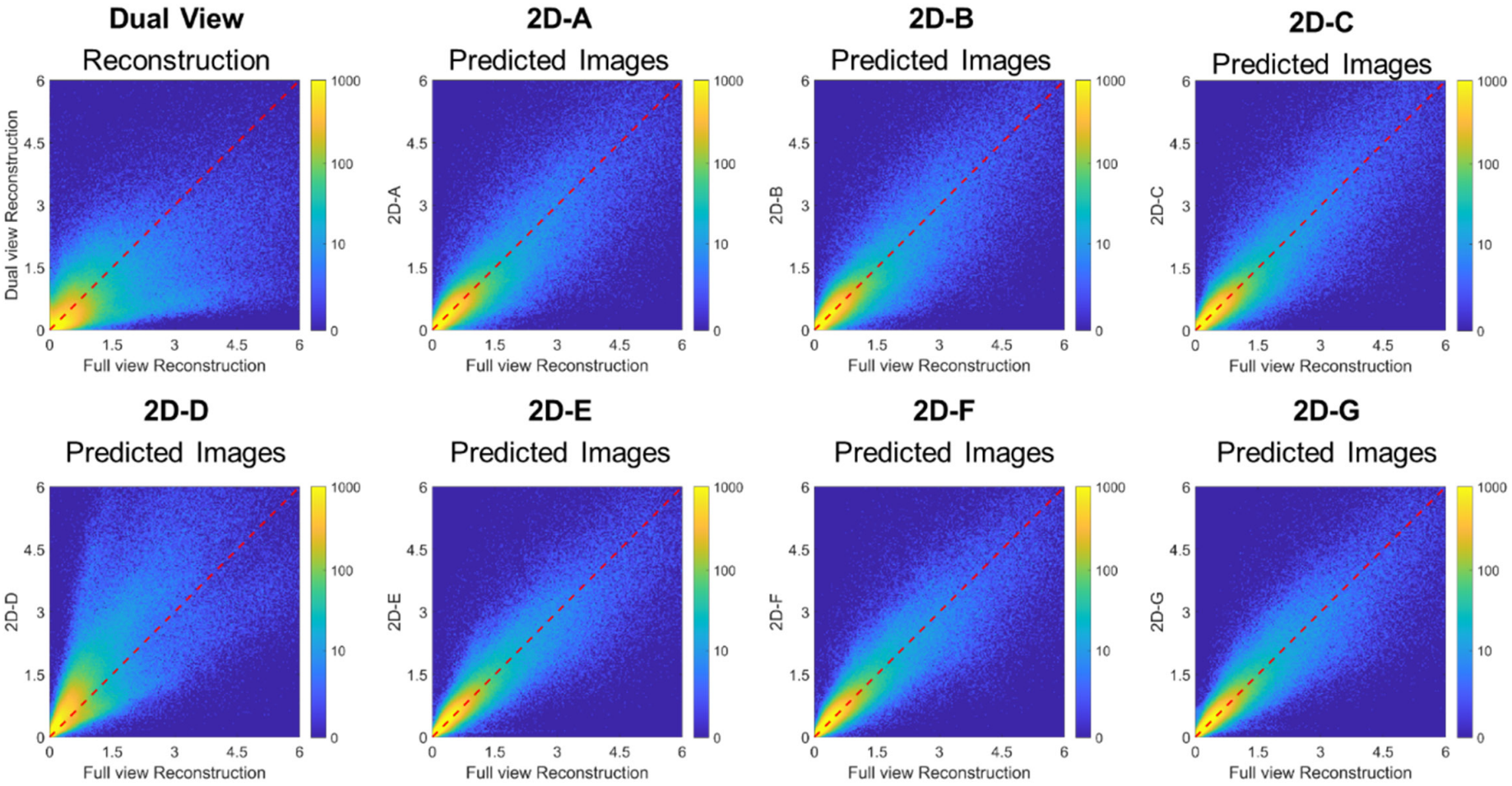

Figure 4. The correlation plots for dual view reconstruction image and the predicted images using those 2D U-Nets with the reference images for all the voxels in the 17 subjects. The dashed red lines are the identical line.

Table 1. The quantitative metrices for the 2D U-Nets

	Dual-View	2D-A	2D-B	2D-C	2D-D	2D-E	2D-F	2D-G
NMAE%	1.4±0.8	0.8±0.5	0.8±0.5	0.8±0.5	1.3±0.8	0.6±0.4	0.6±0.4	0.6±0.3
NMSE%	36.6±11.6	15.5±7.1	14.7±6.9	13.6±8.3	63.5±86.7	12.3±8.7	11.1±7.6	9.3±6.6
SSIM	0.71±0.07	0.86±0.07	0.87±0.04	0.91±0.04	0.85±0.10	0.89±0.06	0.92±0.04	0.93±0.04
PSNR	30.2±4.5	34.2±4.7	34.4±4.5	34.9±5.0	29.8±4.2	35.5.2±3.9	36.0±4.1	36.7±4.4

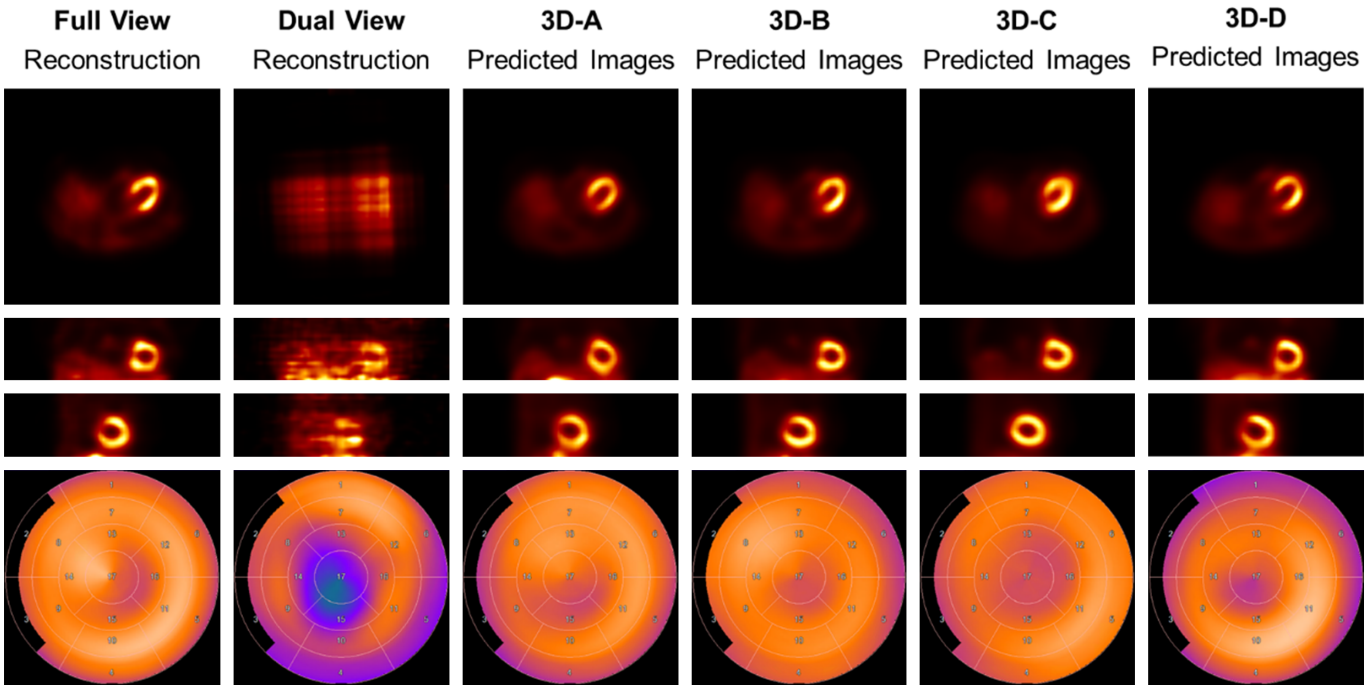

Figure 5. The sample slices in transverse view, coronal view and sagittal view and polar maps (from top row to bottom row) of the full-view reconstruction images, dual-view reconstruction images, and the predicted images using 3D U-Nets: 3D-A trained by stacked projection with ATT MAP, 3D-B trained by stacked projection with ATT MAP using patch-based strategy, 3D-C trained by ATT MAP using patch-based strategy, and 3D-D trained by the reconstruction images of the dual view projection using patch-based strategy. Those slices were shown with the same color bar, except the U-Net 3D-C.

3D-D. No significant difference is observed between U-Nets 3D-A and 3D-B.

Discussion

We have developed a deep learning-based image reconstruction algorithm for dual-view projection from conventional SPECT scanner to enable the non-rotational imaging, tailored for clinical MPI application. While the U-Net architecture has been widely and successfully used for image generation, we investigated the impact of the different input organization schemes for dual-view projection on reconstruction performance. The result demonstrated the stacked organization scheme, which treats the dual-view projection as two-channel data, outperforms the other two schemes. Due to limitations in the training time and the size of the training dataset, we initially employed 2D U-Nets. Consequently, we used 53,312 2D slices for training, ensuring the stability of the comparison. Practically, conventional SPECT imaging using parallel-hole collimators should be intrinsically shift-invariant in the axial dimension. However, when we aimed to obtain a 2D U-Net for image reconstruction, the results revealed discontinuity along the axial dimension. This discontinuity might be attributed to anatomical structure differences not captured in the dual-view projection. After including the attenuation map as an additional input channel, the discontinuity was significantly improved. Recently, with the development of instrumentation and quantitative reconstruction algorithms, quantitative SPECT imaging has gained significant progress in research and clinical practice [33]. The attenuation correction has been recommended for the quantitative SPECT image reconstruction, where the CT-based ATT MAP can be generated by an integrated SPECT/CT scanner. For a SPECT scanner without integrated CT, the ATT MAP could be acquired on a standalone CT scanner and then be aligned with the SPECT image for attenuation correction as well as for the application of our proposed networks. Deep-learning based ATT MAP generation from the SPECT data only has also been reported recently [29, 34], and future studies are needed to investigate the

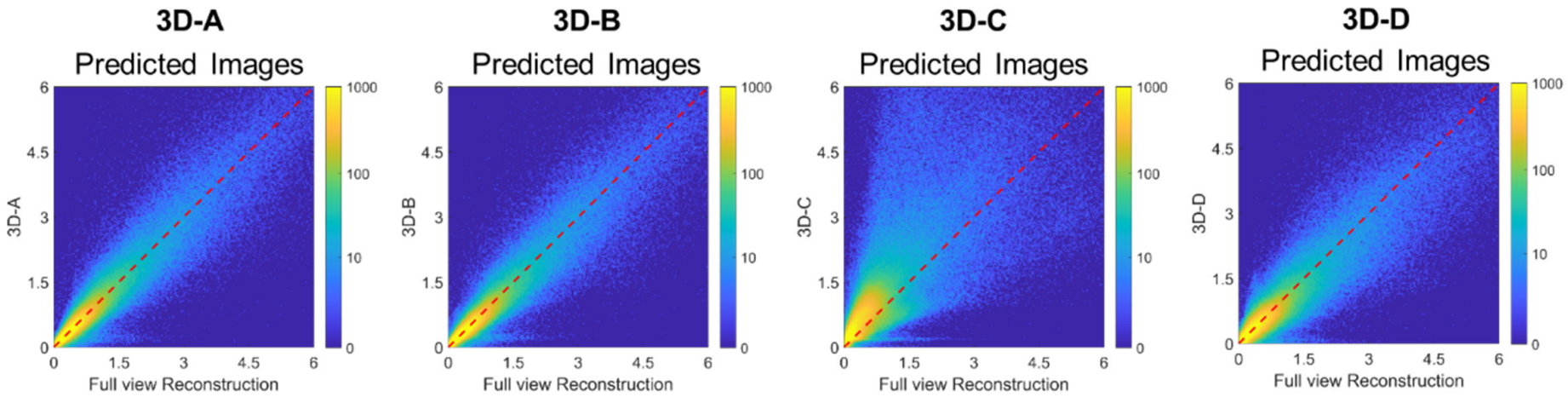

Figure 6. The correlation plots for the predicted images using those 3D U-Nets with the reference images for all the voxels in the 17 subjects. The dashed red lines are the identical line.

Table 2. The quantitative metrices for the 3D U-Nets

	Dual-View	3D-A	3D-B	3D-C	3D-D
NMAE%	1.4±0.8	0.6±0.3	0.6±0.3	1.9±1.9	0.8±0.4
NMSE%	36.6±11.6	9.5±7.1	11.0±15.3	82.5±90.1	10.9±6.4
SSIM	0.71±0.07	0.92±0.04	0.94±0.03	0.78±0.12	0.91±0.04
PSNR	30.2±4.5	36.7±3.9	36.9±4.2	28.5±6.2	35.8±4.8

Table 3. The ROI relative biases for the 3D U-Nets

	Dual-View	3D-A	3D-B	3D-C	3D-D
LV cavity%	72.5±65.7	5.5±15.7†	1.8±16.8†	100.5±87.6	3.6±25.9†
Myocardium%	-27.9±8.3	-2.8±5.5	-2.0±6.3†	50.4±49.9	-6.3±6.7
LAD%	-25.5±10.2	-6.4±6.5	-1.6±10.2†	47.0±48.9	-6.3±10.0
RCA%	-35.4±8.0	-2.8±6.9†	-2.7±8.5†	40.2±47.6	-7.7±8.7
LCX%	-29.5±10.1	-4.1±8.0	-3.3±6.6†	49.8±51.6	-11.1±7.4

†This indicates that the ROI relative bias was not significantly different with zero ($P > 0.05$).

feasibility of generating ATT MAP from dual-view SPECT projections.

Furthermore, we investigated the reconstruction performance of 3D U-Nets using the stacked projection and the attenuation map as inputs. Due to the limitation of the training dataset, we employed a patch-based strategy. Although the differences in quantitative metrics are minimal in terms of the NMAE, NMSE, SSIM, and PSNR after applying patch-based strategy, the reconstructed images appear improved, particularly regarding the myocardium, and the ROI relative biases also showed enhancement. Compared with the conventional U-Net 3D-D, which uses dual-view reconstruction image to generate full-view reconstruction image, our proposed U-Nets with and without patch-based strategy achieved superior performance in terms of quantitative results. This improvement might be may be attributed to the artifacts present in the dual-view reconstruction images. While the U-Net utilizing only the attenuation map as input can generate SPECT images, it does not yield quantitative results due to the absence of SPECT information. To compare with the previously reported deep learning approaches such as DuDoSS and CUSIP in reducing projections [25, 27], these prior approaches need at least 15 views of projection for SPECT image reconstruction, indicating that the gantry rotation remains unavoidable. By contrast, the proposed dual-view reconstruction algorithm can be used for non-rotational imaging, offering significant benefits for dynamic SPECT myocardium perfusion imaging.

Furthermore, we performed a preliminary clinical validation based on the reader study on the 19 testing datasets. The diagnostic gold standard was provided by the real-world CAD assessment of the full-view SPECT images. Then two experienced physicians were invited to diagnose the myocardial defect based on the predicted images using the four 3D U-Nets. The diagnostic accuracy was 70.5%, 73.5%, 26.5%, 67.6% for U-Nets 3D-A, 3D-B, 3D-C, and 3D-D, respectively. These results indicated that the U-Net 3D-B provided the highest diagnostic accuracy, while the U-Net 3D-C was not able to provide reliable diagnostic results, which are consistent with the quantitative metrics results.

For the training process, a total of 30 pairs of two-projection views with a 90-degree interval was used. While for the testing process, the anterior and left views were used for generating dual projections. The reason for the selection of anterior and left views for dual projections was that the image grid of the two projections aligns with the image grid of the reconstructed image. When using other pairs of views with a 90-degree interval for dual projections, the reconstructed image needs to be rotated and interpolated to align with the projections, which would decrease the quantification accuracy for the prediction of the reconstructed images.

In this study, all the training and testing processes were conducted on a cluster equipped with six NVIDIA RTX A6000 GPUs. For future clinical application of the proposed approach, only the testing process needs to be performed by using the well-trained 3D U-Net, and therefore could be handled by just one less-expensive GPU card such as RTX 3090, which may enable the deployment of the proposed 3D U-Net in clinical practice.

One limitation of the study is the limited dataset for 3D U-Nets training. To augment data in 3D U-Nets training, a patch-based strategy was implemented, which is a widely-used data augmentation scheme for deep-learning based 3D medical image synthesis networks. For future studies, we will incorporate more clinical datasets to further improve the generalizability and robustness of the proposed 3D U-Nets.

Conclusion

We proposed a deep learning-based image reconstruction algorithm for dual-view projection from a convention-

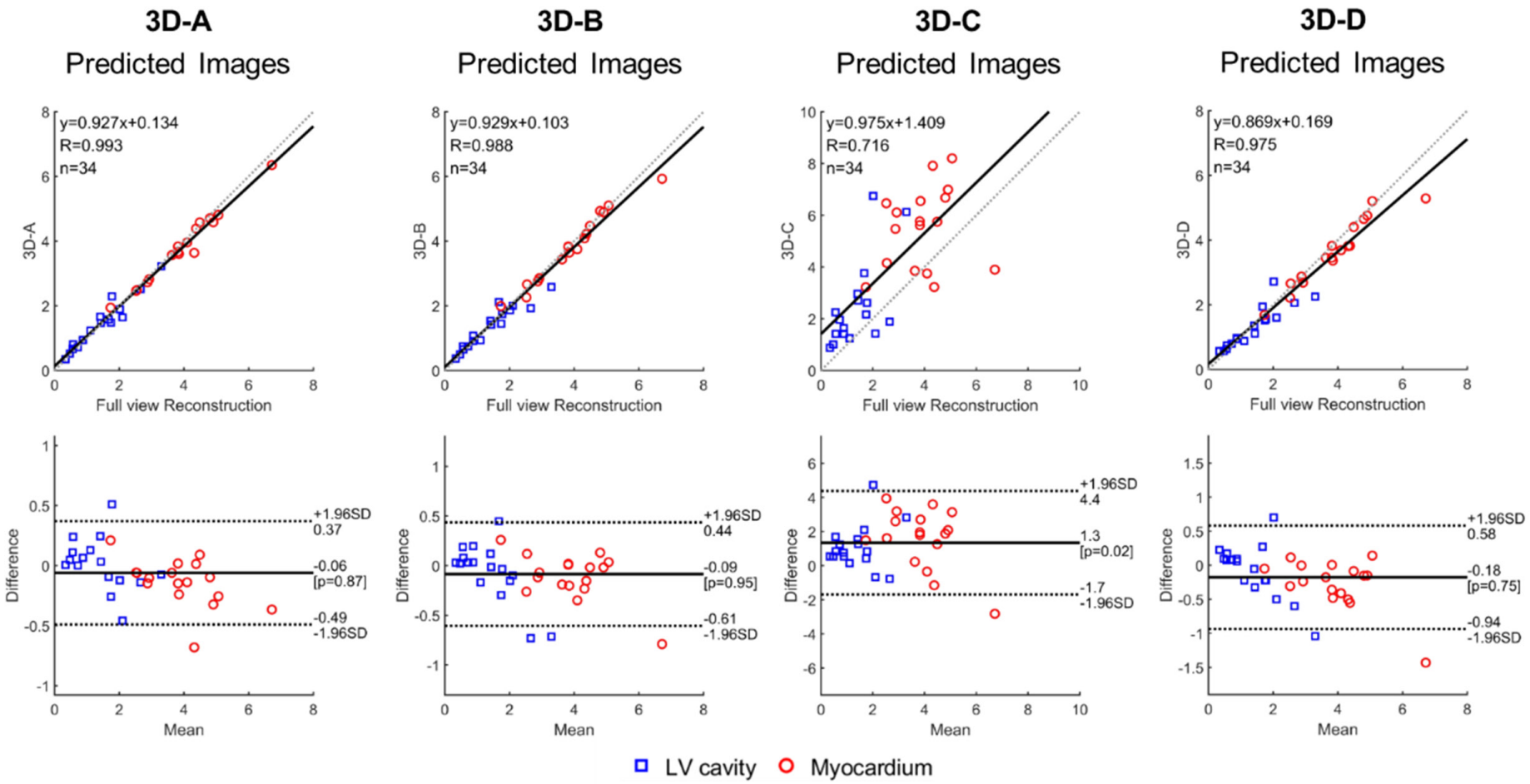

Figure 7. The regression plot (top row) and Bland-Altman plot (bottom row) between the predicted images using the U-Nets 3D-A, 3D-B, 3D-C and 3D-D, and the reference images for the ROIs of LV cavity and myocardium in the 17 subjects.

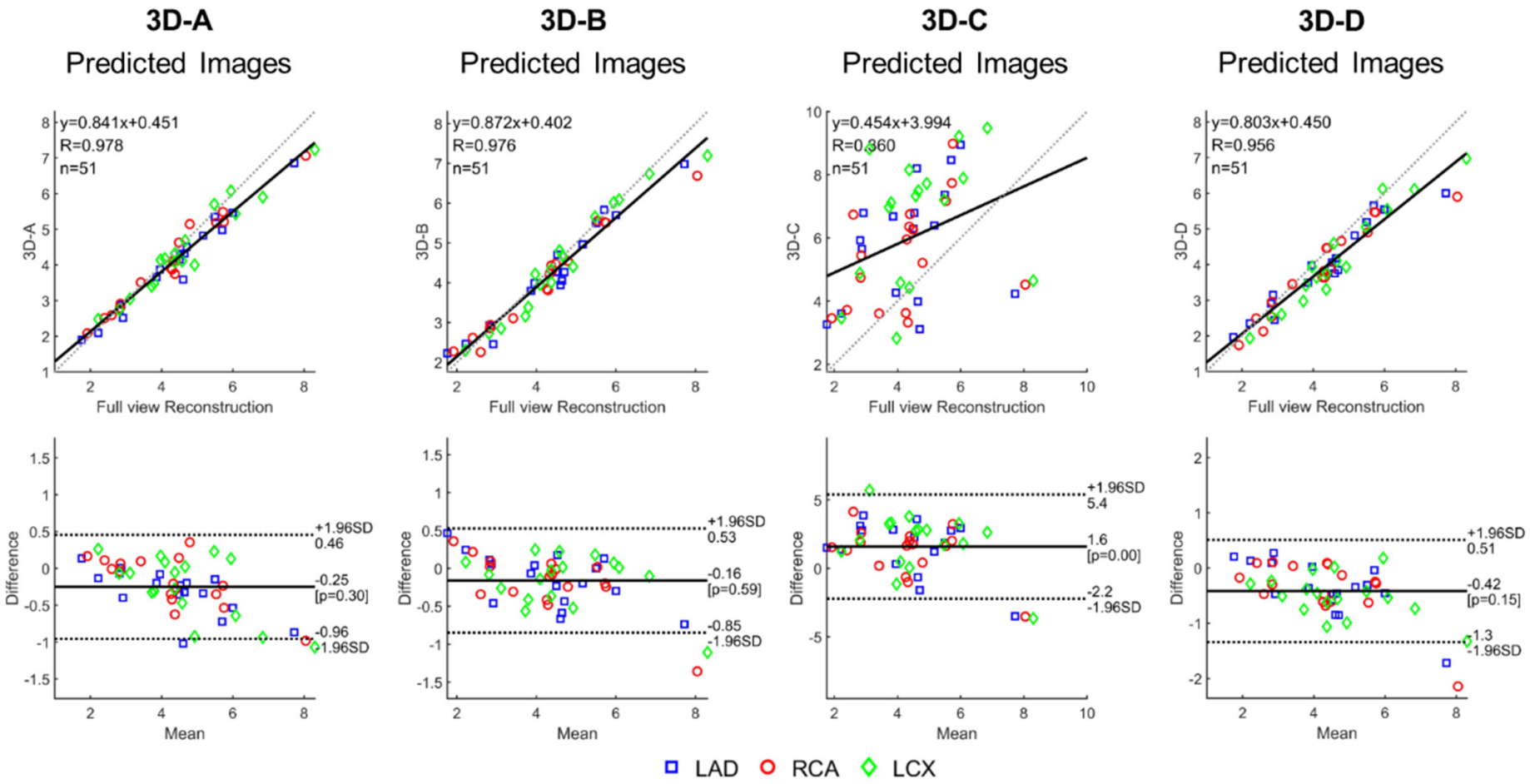

Figure 8. The regression plot (top row) and Bland-Altman plot (bottom row) between the predicted images using the U-Nets 3D-A, 3D-B, 3D-C and 3D-D, and the reference images for LAD, RCA and LCX in the 17 subjects.

al SPECT scanner and investigated the impact of different input organization schemes. The results indicated that utilizing the stacked projection with an attenuation map for the dual-view projection reconstruction is effective for non-rotational imaging, potentially benefiting dynamic myocardium perfusion imaging.

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China (No. 2022YFC2402002), National Natural Science Foundation of China (No. 12205166, No. 12305379), Tsinghua Precision Medicine Foundation (2022PY002) and CIRP Open Fund of Radiation Protection Laboratories (Shanxi Provincial Key Laboratory for Translational Nuclear Medicine and Precision Protection, TNMPP-2024-No.1).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Jing Wu, Center for Advanced Quantum Studies, School of Physics and Astronomy, Beijing Normal University, Beijing, China. E-mail: jingwu@bnu.edu.cn

References

[1] Holder L, Lewis S, Abrames E and Wolin EA. Review of SPECT myocardial perfusion imaging. J Am Osteopath Coll Radiol 2016; 5: 5-13.

[2] Task Force Members; Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, Bugiardini R, Crea F, Cuisset T, Di Mario C, Ferreira JR, Gersh BJ, Gitt AK, Hulot JS, Marx N, Opie LH, Pfisterer M, Prescott E, Ruschitzka F, Sabaté M, Senior R, Taggart DP, van der Wall EE, Vrints CJ; ESC Committee for Practice Guidelines; Zamorano JL, Achenbach S, Baumgartner H, Bax JJ, Bueno H, Dean V, Deaton C, Erol C, Fagard R, Ferrari R, Hasdai D, Hoes AW, Kirchhof P, Knuuti J, Kolh P, Lancellotti P, Linhart A, Nihoyannopoulos P, Piepoli MF, Ponikowski P, Sirnes PA, Tamargo JL, Tendera M, Torbicki A, Wijns W, Windecker S; Document Reviewers; Knuuti J, Valgimigli M, Bueno H, Claeys MJ, Donner-Banzhoff N, Erol C, Frank H, Funck-Brentano C, Gaemperli O, Gonzalez-Juanatey JR, Hamilos M, Hasdai D, Husted S, James SK, Kervinen K, Kolh P, Kristensen SD, Lancellotti P, Maggioni AP, Piepoli MF, Pries AR, Romeo F, Rydén L, Simoons ML, Sirnes PA, Steg PG, Timmis A, Wijns W, Windecker S, Yildirir A and Zamorano JL. 2013 ESC guidelines on the management of stable coronary artery disease: the task force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J 2013; 34: 2949-3003.

[3] Jiang N, Liu H, Xue M, Li C, Gao L, Liu F, Wu J and Liu Y. Performance evaluation of a novel multi-pinhole SPECT system. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 2024; 1059: 168976.

[4] Wu J and Liu C. Recent advances in cardiac SPECT instrumentation and imaging methods. Phys Med Biol 2019; 64: 06TR01.

[5] Slomka PJ, Berman DS and Germano G. New cardiac cameras: single-photon emission CT and PET. Semin Nucl Med 2014; 44: 232-251.

[6] Groch MW and Erwin WD. SPECT in the year 2000: basic principles. J Nucl Med Technol 2000; 28: 233-244.

[7] Dorbala S, Ananthasubramaniam K, Armstrong IS, Chareonthaitawee P, DePuey EG, Einstein AJ, Gropler RJ, Holly TA, Mahmarian JJ, Park MA, Polk DM, Russell R 3rd, Slomka PJ, Thompson RC and Wells RG. Single photon emission computed tomography (SPECT) myocardial perfusion imaging guidelines: instrumentation, acquisition, processing, and interpretation. J Nucl Cardiol 2018; 25: 1784-1846.

[8] Kyme AZ and Fulton RR. Motion estimation and correction in SPECT, PET and CT. Phys Med Biol 2021; 66.

[9] Fitzgerald J and Danias PG. Effect of motion on cardiac SPECT imaging: recognition and motion correction. J Nucl Cardiol 2001; 8: 701-706.

[10] Wells RG, Marvin B, Poirier M, Renaud J, deKemp RA and Ruddy TD. Optimization of SPECT measurement of myocardial blood flow with corrections for attenuation, motion, and blood binding compared with PET. J Nucl Med 2017; 58: 2013-2019.

[11] Garcia EV. Are SPECT measurements of myocardial blood flow and flow reserve ready for clinical use? Eur J Nucl Med Mol Imaging 2014; 41: 2291-2293.

[12] Liu H, Thorn S, Wu J, Fazzone-Chettiar R, Sandoval V, Miller EJ, Sinusas AJ and Liu YH. Quantification of myocardial blood flow (MBF) and reserve (MFR) incorporated with a novel segmentation approach: assessments of quantitative precision and the lower limit of normal MBF and MFR in patients. J Nucl Cardiol 2021; 28: 1236-1248.

[13] Pelletier-Galarneau M and Ruddy TD. The potential of regional myocardial blood flow measurement with SPECT. J Nucl Cardiol 2021; 28: 260-262.

[14] Bateman TM. Advantages and disadvantages of PET and SPECT in a busy clinical practice. J Nucl Cardiol 2012; 19 Suppl 1: S3-11.

[15] Gaemperli O and Kaufmann PA. Lower dose and shorter acquisition: pushing the boundaries of myocardial perfusion SPECT. J Nucl Cardiol 2011; 18: 830-832.

[16] Henzlova MJ and Duvall WL. The future of SPECT MPI: time and dose reduction. J Nucl Cardiol 2011; 18: 580-587.

[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] Lu W, Onofrey JA, Lu Y, Shi L, Ma T, Liu Y and Liu C. An investigation of quantitative accuracy for deep learning based denoising in oncological PET. Phys Med Biol 2019; 64: 165019.

[19] Chan C, Fulton R, Feng DD and Meikle S. Median non-local means filtering for low SNR image denoising: application to PET with anatomical knowledge. IEEE Nuclear Science Symposuim & Medical Imaging Conference. 2010. pp.3613-3618.

[20] Dabov K, Foi A, Katkovnik V and Egiazarian K. Image denoising with block-matching and 3D filtering. Image processing: algorithms and systems, neural networks, and machine learning. 2006. pp. 354-365.

[21] Wolf PA, Jørgensen JS, Schmidt TG and Sidky EY. Few-view single photon emission computed tomography (SPECT) reconstruction based on a blurred piecewise constant object model. Phys Med Biol 2013; 58: 5629-52.

[22] Xie H, Thorn S, Liu YH, Lee S, Liu Z, Wang G, Sinusas AJ and Liu C. Deep-learning-based few-angle cardiac SPECT reconstruction using transformer. IEEE Trans Radiat Plasma Med Sci 2023; 7: 33-40.
[23] Shao W, Pomper MG and Du Y. A learned reconstruction network for SPECT imaging. IEEE Trans Radiat Plasma Med Sci 2021; 5: 26-34.
[24] Ding Q, Zan Y, Huang Q and Zhang X. Dynamic SPECT reconstruction from few projections: a sparsity enforced matrix factorization approach. Inverse Problems 2015; 31: 025004.
[25] Rydén T, Van Essen M, Marin I, Svensson J and Bernhardt P. Deep-learning generation of synthetic intermediate projections improves 177Lu SPECT images reconstructed with sparsely acquired projections. J Nucl Med 2021; 62: 528-535.
[26] Li S, Ye W and Li F. LU-Net: combining LSTM and U-Net for sinogram synthesis in sparse-view SPECT reconstruction. Math Biosci Eng 2022; 19: 4320-4340.
[27] Chen X, Zhou B, Xie H, Miao T, Liu H, Holler W, Lin M, Miller EJ, Carson RE, Sinusas AJ and Liu C. DuDoSS: deep-learning-based dual-domain sinogram synthesis from sparsely sampled projections of cardiac SPECT. Med Phys 2023; 50: 89-103.
[28] Liu H, Ma T, Chen S, Liu Y, Wang S and Jin Y. Development of GPU based image reconstruction method for clinical SPECT. 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC). 2012. pp. 3415-3418.
[29] Liu H, Wu J, Shi L, Liu Y, Miller E, Sinusas A, Liu YH and Liu C. Post-reconstruction attenuation correction for SPECT myocardium perfusion imaging facilitated by deep learning-based attenuation map generation. J Nucl Cardiol 2022; 29: 2881-2892.
[30] Ronneberger O, Fischer P and Brox T. U-net: convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. 2015. pp. 234-241.
[31] Nesterov SV, Han C, Mäki M, Kajander S, Naum AG, Helenius H, Lisinen I, Ukkonen H, Pietilä M, Joutsiniemi E and Knuuti J. Myocardial perfusion quantitation with 15O-labelled water PET: high reproducibility of the new cardiac analysis software (Carimas). Eur J Nucl Med Mol Imaging 2009; 36: 1594-1602.
[32] Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, Pennell DJ, Rumberger JA, Ryan T and Verani MS; American Heart Association Writing Group on Myocardial Segmentation and Registration for Cardiac Imaging. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. Circulation 2002; 105: 539-542.
[33] Dickson JC, Armstrong IS, Gabiña PM, Denis-Bacelar AM, Krizsan AK, Gear JM, Van den Wyngaert T, de Geus-Oei LF and Herrmann K. EANM practice guideline for quantitative SPECT-CT. Eur J Nucl Med Mol Imaging 2023; 50: 980-995.
[34] Shi L, Onofrey JA, Liu H, Liu YH and Liu C. Deep learning-based attenuation map generation for myocardial perfusion SPECT. Eur J Nucl Med Mol Imaging 2020; 47: 2383-2395.