Automatic reorientation by deep learning to generate short-axis SPECT myocardial perfusion images

Fubao Zhu, PhD,a Guojie Wang, MS,a Chen Zhao, PhD,b Saurabh Malhotra, MD,c,d Min Zhao, MD, PhD,e Zhuo He, PhD,b Jianzhou Shi, MD, PhD,f Zhixin Jiang, MD, PhD,f and Weihua Zhou, PhDb,g

a School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan, China
b Department of Applied Computing, Michigan Technological University, Houghton, MI
c Division of Cardiology, Cook County Health and Hospitals System, Chicago, IL
d Division of Cardiology, Rush Medical College, Chicago, IL
e Department of Nuclear Medicine, Xiangya Hospital, Central South University, Changsha, China
f Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China
g Center for Biocomputing and Digital Health, Institute of Computing and Cybersystems, and Health Research Institute, Michigan Technological University, Houghton, MI

Received Sep 25, 2022; accepted Jan 30, 2023
doi:10.1007/s12350-023-03226-2

Background. Single photon emission computed tomography (SPECT) myocardial perfusion images (MPI) can be displayed both in traditional short-axis (SA) cardiac planes and polar maps for interpretation and quantification. It is essential to reorient the reconstructed transaxial SPECT MPI into standard SA slices. This study is aimed to develop a deep-learning-based approach for automatic reorientation of MPI.

Methods. A total of 254 patients were enrolled, including 226 stress SPECT MPIs and 247 rest SPECT MPIs. Fivefold cross-validation with 180 stress and 201 rest MPIs was used for training and internal validation; the remaining images were used for testing. The rigid transformation parameters (translation and rotation) from manual reorientation were annotated by an experienced nuclear cardiologist and used as the reference standard. A convolutional neural network (CNN) was designed to predict the transformation parameters. Then, the derived transform was applied to the grid generator and sampler in spatial transformer network (STN) to generate the reoriented image. A loss function containing mean absolute errors for translation and mean square errors for rotation was employed. A three-stage optimization strategy was adopted for model optimization: (1) optimize the translation parameters while fixing the rotation parameters; (2) optimize rotation parameters while fixing the translation parameters; (3) optimize both translation and rotation parameters together.

Results. In the test set, the Spearman determination coefficients of the translation distances and rotation angles between the model prediction and the reference standard were 0.993 in X.
axis, 0.992 in Y axis, 0.994 in Z axis, 0.987 along X axis, 0.990 along Y axis and 0.996 along Z axis, respectively. For the 46 stress MPIs in the test set, the Spearman determination coefficients were 0.858 in percentage of profusion defect (PPD) and 0.858 in summed stress score (SSS); for the 46 rest MPIs in the test set, the Spearman determination coefficients were 0.9 in PPD and 0.9 in summed rest score (SRS).

Conclusions. Our deep learning-based LV reorientation method is able to accurately generate the SA images. Technical validations and subsequent evaluations of measured clinical parameters show that it has great promise for clinical use. (J Nucl Cardiol 2023;30:1825–35.)

Key Words: SPECT MPI • reorientation • deep learning • convolutional neural networks

Abbreviations
SPECT Single photon emission computed tomography
MPI Myocardial perfusion imaging
CNN Convolutional neural networks
STN Spatial transformer network
LV Left ventricular
SA Short-axis
ECTb Emory Cardiac Toolbox

INTRODUCTION
Myocardial perfusion imaging (MPI) with single photon emission computed tomography (SPECT) is regarded as one of the most utilized non-invasive cardiac imaging modalities in the diagnosis of coronary artery disease. The display of SPECT MPI in the transaxial view is patient-specific because of the differences in left-ventricular (LV) orientation between patients, which complicates the visual interpretation of images. It is essential to reorient the reconstructed transaxial SPECT MPI into standard short-axis slices.

Traditional manual reorientation is time-consuming and less reproducible. A number of studies have shown that significant artifacts, and quantitative analysis may be misleading due to incorrect reorientation. Therefore, precise and automated methods are needed to obtain satisfactory SA images and reliable quantitative results. Automated methods for reorientation have been investigated and used in clinical routine. Nevertheless, they rely on the integrity of the myocardium. Therefore, current commercial software also provides manual alternatives to prevent automated methods from failing in reorientation as well. Recently, deep learning has shown great performance in medical image processing, including tissue and organ localization and reorientations. Zhang et al. developed a method using convolutional neural networks (CNN) for automatic reorientation of MPI. It achieved a high technical performance, but clinical parameters measured from their LV reorientation results were not evaluated and further improvement in accuracy is still needed for clinical use.

This study is aimed to develop a deep-learning-based approach for automatic reorientation of MPI and evaluate its values in MPI quantitative analysis for clinical use.

MATERIALS AND METHODS

Data acquisition
We retrospectively enrolled 254 patients (226 stress and 247 rest MPIs) from the First Affiliated Hospital of Nanjing Medical University. Two hundred and nineteen paired stress/rest MPIs were acquired using a two-day protocol; for the remaining, 28 rest-only and 7 stress-only MPIs were acquired. For either stress or rest MPI, Tc-99m sestamibi doses ranged from 25 to 30 mCi according to the body mass index (BMI). The MPI images were acquired on a dual-headed camera (CardioMD, Philips Medical Systems) with a standard protocol. The imaging parameters included a 20% energy window around 140 keV, 180° orbit, 32 steps with 25 seconds per step, 8-bin gating, and 64 planar projections per gate. All SPECT planar images were reconstructed by Emory Reconstruction Toolbox (ERTb v2.0; Syntermed, Atlanta, GA) with 3 iterations and 10 subsets of ordered subset expectation maximization, and then low-pass filtered with Butterworth at a cutoff frequency of 0.3 cycles/mm to order 10 subsets.

The image size was 64 × 64 in each slice, LV slice numbers range from 22 to −36, and the voxel size was 6.4 mm × 6.4 mm × 6.4 mm. All ungated reconstructed transaxial images were reoriented into SA images by an experienced nuclear cardiologist using Emory Reconstruction Toolbox (ERTb v2.0; Syntermed, Atlanta, GA) with 3 iterations and 10 subsets of ordered subset expectation maximization, and then low-pass filtered with Butterworth at a cutoff frequency of 0.3 cycles/mm to order 10 subsets.

See related editorial, pp. 1836–1837
Ninety percent of the stress and rest MPIs (180 stress and 201 rest SPECT MPIs) were used for the fivefold cross-validation, and the remaining images were used for testing and clinical evaluation. In addition, 177 stress/rest SPECT MPIs from the same patient were included in the cross-validation dataset. The minimum, maximum, mean, and standard deviation (STD) of the transformation parameters of both the training and the internal validation are shown in Table 1. Institutional review board approval by the Institutional Ethical Committee of the First Affiliated Hospital of Nanjing Medical University was obtained with no informed consent required for this HIPAA-compliant retrospective analysis.

### Data augmentation

To increase the sample size and avoid overfitting, several image processing techniques were applied. First, a Gaussian distribution was calculated for all the transformation parameters. Secondly, each set of data generates an additional 40 sets of transformation parameters from the truncated Gaussian distributions. Finally, SA images were inversely transformed with Gaussian-generated parameters to generate new images, regarded as the augmented transaxial images.

After data augmentation, a total of 15,621 MPIs were available for training and cross-validation. The remaining data (46 rest and 46 stress SPECT MPIs) were used in the test set.

### Deep-learning-based method

Our 3D network architecture for LV reorientation is shown in Fig. 2. All 3D MPIs were rescaled to $64 \times 64 \times 32$ to facilitate the network training and prediction. To enhance the contours in the SPECT MPI, gradient images of the transaxial images were calculated, and combined with the transaxial images to form a two-channel network. As a result, the network consists of two channels ($64 \times 64 \times 32 \times 2$) as the input for the learning process.

Three reorientation blocks using CNNs, as shown in Fig. 2 Top, were employed to predict transformation parameters for image reorientation. The CNN architecture is illustrated in Fig. 2 Bottom. A spatial transformer network (STN) was used after each reorientation block to generate the reoriented images. STN has been used in our previous studies to combine results predicted by CNN and prior knowledge for improved image segmentation. It provides three sub-modules: a predicted network of transformation matrices, a grid generator for coordinate mapping, and a sampler for pixel collection. As shown in Fig. 2 Top, skip links are transformations of images through samplers.

Mean absolute error (MAE) and mean square error (MSE) was used in the loss function of translation parameters and rotation parameters, respectively, as shown in Eqs. 1 and 2.

$$\text{LMSE}(\hat{y}; y) = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2,$$  \hspace{1cm} (1)

### Table 1. Statistical information for the transformation parameters of 381 SPECT MPIs in both the training and the internal validation

|       | $x_m$ | $y_m$ | $z_m$ | Alpha | Beta | Theta |
|-------|-------|-------|-------|-------|------|-------|
| MIN   | − 8.171 | − 7.064 | − 8.101 | − 0.432 | − 0.255 | − 0.340 |
| MAX   | 2.169 | 5.945 | 6.522 | 0.378 | 0.582 | 0.351 |
| MEAN  | − 0.365 | − 0.776 | − 1.556 | − 0.015 | 0.106 | − 0.048 |
| STD   | 2.9077 | 1.981 | 1.060 | 0.133 | 0.237 | 0.104 |

The units are radian for rotations and pixel for translations.
\[
\text{LMAE}(\tilde{y}, y) = \frac{1}{n} \sum_{i=1}^{n} |\tilde{y}_i - y_i|, \quad (2)
\]

where \(\tilde{y}_i\) and \(y_i\) are the predicted and reference standard and \(n\) indicates the number of parameters, i.e., \(n = 3\) for translation and rotation parameters.

To train both translation parameters and rotation parameters, a composite loss function that combines MAE with MSE was used:

\[
L_{\text{par}}(\tilde{y}, y) = \mu \text{LMSE}(\tilde{y}, y) \text{LMAE}(\tilde{y}, y), \quad (3)
\]

where \(\mu\) is an empirical weight, which balances the losses from MSE and MAE.

Both stress and rest MPIs were input into the network. Our proposed method was implemented in Python using PyTorch, and the model was trained on a workstation with an NVIDIA V100 GPU. An Adam Optimizer was used to fine-tune the weights of the network. The model was trained for 200 epochs with a batch size of 8. We trained only the translation parameters in the first 80 epochs, only rotation parameters in the 81–160 epochs, and both translation and rotation parameters in the 161–200 epochs.

### Evaluation and statistical analysis

Both technical and clinical evaluations were conducted in the test set with 46 stress and 46 rest MPIs. To technically evaluate the transformation parameters, the correlation was reported. To clinically validate the accuracy of clinical parameters measured from the SA images generated, both correlation and MAE were reported. All clinical parameters were measured by a commercial software package (Emory Cardiac Toolbox 4.0; Atlanta, GA).

To further confirm the effectiveness of STN in the LV reorientation, both 3DVGG and 3DResNet were implemented as the benchmark, and further we incorporated the three-stage optimization strategy with 3DVGG and 3DResNet to validate their values in optimizing the transformation parameters. For both of them, the first 80 epochs were used for training translation parameters, the 81–160 epochs for training rotation parameters, and 160–200 epochs for overall training. For 3DVGG and 3DResNet without the three-stage optimization, a total of 160 epochs were used for training translation and rotation parameters together. The same loss function as in our method was used for the training.

### RESULTS

The baseline characteristics of the study population were shown in Table 2. For all the patients, the age was 61 ± 11 years, and 164 (65%) patients were male.
Hypertension (52%), diabetes mellitus smoking (14%), and smoking (23%) were also shown in the baseline data. One hundred and ninety-eight patients (78%) had angina and 67 patients (26%) had Coronary artery computed tomography angiography (CCTA) stenosis > 70%.

Table 3 shows the correlations between the predicted reorientation parameters by different models and the reference standard, the number of network parameters, and the running/test time. Determination coefficients in our method were better than those by 3DVGG, 3DResNet, CNN + STN, and 3DVGG-ThreeStages. Compared to 3DResNet-ThreeStages, our method achieved almost the same accuracy but it required a much smaller number of network parameters and ran faster for training and testing.

Figure 3 shows the quantitative analysis of transformation parameters. There were excellent correlations in the reorientation parameters between our method and the reference standard. The determination coefficients were 0.993, 0.992, 0.994 for the translation parameters and 0.987, 0.99, 0.996 for the rotation parameters (all $P$ values < .01). The interclass correlation coefficients were 0.991, 0.991, 0.984 for the translation parameters, and 0.993, 0.997, 0.998 for the rotation parameters.

Figure 4 shows three examples by the visualization of three orthogonal orientations. The images obtained by our method were similar to those in the reference standard.

### Table 2. Baseline demographics of enrolled patients ($n = 254$)

|                        | Age, years (mean ± SD) | 61 ± 11 |
|------------------------|------------------------|---------|
| Males, n (%)           | 164 (65%)              |         |
| Hypertension, n (%)    | 131 (52%)              |         |
| Diabetes Mellitus, n (%)| 35 (14%)               |         |
| Smoking, n (%)         | 59 (23%)               |         |
| Angina, n (%)          | 198 (78%)              |         |
| CCTA Stenosis > 70%, n (%) | 67 (26%)         |         |

Data are presented as the mean ± standard deviation. CCTA Coronary artery computed tomography angiography.

### Table 3. The accuracy (measured as the determination coefficients of transformation parameters between prediction models and reference standard) and computational performance of different deep learning-based models for left-ventricular reorientation

|                        | 3DVGG  | 3DResNet | CNN + STN | 3DVGG-ThreeStages | 3DResNet-ThreeStages | Our method |
|------------------------|--------|----------|-----------|-------------------|----------------------|------------|
| $x_m$                  | 0.939  | 0.977    | 0.983     | 0.94              | 0.992                | **0.993**  |
| $y_m$                  | 0.979  | 0.985    | 0.987     | 0.98              | **0.995**            | 0.992      |
| $z_m$                  | 0.981  | 0.982    | 0.984     | 0.977             | **0.996**            | 0.994      |
| alpha                  | 0.867  | 0.967    | 0.975     | 0.966             | 0.983                | **0.987**  |
| beta                   | 0.845  | 0.983    | 0.984     | 0.985             | 0.987                | **0.99**   |
| theta                  | 0.855  | 0.966    | 0.98      | 0.967             | 0.995                | **0.996**  |
| Network complexity     | 44,278,022 | 33,185,030 | **360,262** | 88,555,654         | 66,366,982          | 719,897    |
| Training time          | 8.14   | 4.13     | **2.05**  | 9.23              | 5.07                 | 4.24       |
| Prediction time        | 0.9426 | 0.1168   | **0.1109** | 1.7902            | 0.2253              | **0.2143** |

3DVGG and 3DResNet are 3D implementations of VGG and ResNet architectures. CNN + STN is the implementation of our method without the three-stage optimization. 3DVGG-ThreeStages and 3DResNet-ThreeStages are the modifications of 3DVGG and 3DResNet incorporating three-stage optimization strategy. $x_m, y_m, z_m$ are translation parameters; alpha, beta, theta are degrees of rotation parameters. Network complexity is the number of parameters in the network; training time measures the time consumption for each epoch (in minutes); prediction time is the time consumption for each MPI (in seconds). The bold texts indicate the best results.
Figure 3. Linear regression analysis to evaluate the translation parameters (left column) and rotation parameters (right column) between our method and the reference standard.
Figure 6 shows the polar maps of three stress MPIs (top, normal; middle and bottom, mild perfusion defect) and three rest MPIs (top, normal; middle, mild perfusion defect; bottom, severe perfusion defect). The sPPD/rPPD and SSS/SRS of the 17-segment perfusion defect show strong consistency between the reference standard and our method.

**DISCUSSION**

In this study, we developed a deep learning network for LV reorientation on SPECT MPI to generate SA images. The correlations of all transformation parameters between our prediction and the reference standard are strong. Furthermore, our clinical evaluation for the quantification of MPI shows strong correlations for clinical parameters between the network generated images and reference standards.

The traditional methods require the extraction of LV myocardial contours or localization of LV apex and base before LV reorientation. This may lead to failure when the LV contours are not clear. However, the reorientation can be accomplished without extracting contours in our method. In addition, several reasons which may cause the failure of LV reorientation were mentioned: (1) the reorientations angles exceeding 45°; (2) severe overlap between LV and other organs. In a study of 200 patients by Germano et al. 1.5% failed while 6.4% failed in the study of 124 patients by Mullick et al. These problems were alleviated in our method, as illustrated in Fig. 7. The determination coefficients between our reorientation parameters and those by ECTb were 0.981, 0.984 and 0.985 for translation, and the determination coefficients were 0.954, 0.941 and 0.932 for rotation.

There are few studies using deep learning for LV reorientation so far. In the recent study by Zhang et al., it was reported that the determination coefficients were 0.928, 0.958, 0.994 for translation, and the determination coefficients were 0.973, 0.96, 0.970 for rotation. However, there was a difference in performance between their results and our implementations. When evaluated with our implementation and test dataset, the determination coefficients were 0.987, 0.99, 0.994 for translation, and the determination coefficients were 0.981, 0.984, 0.987 for rotation. The differences of the prediction accuracy confirm two major innovations in our method: (1) we used the combination of MAE and MSE in the loss function to train the translation and rotation parameters; (2) we designed an effective three-stage optimization for these transformation parameters. As shown in Table 3, the three-stage optimization...
significantly improved the performance for all the three networks (3DVGG, 3DResNet, and CNN + STN).

Technical validations and subsequent evaluations of measured clinical parameters show that our method for LV reorientation has great promise for clinical use. The standardized LV short-axis images by our proposed method will be used with our LV segmentation methods \(^{25-27}\) for MPI functional quantification \(^{6,28,29}\) and image fusion between cardiac functions from MPI and coronary anatomy from invasive coronary angiograms/ venograms to support clinical decision making for coronary revascularization and cardiac resynchronization therapy. \(^{30-33}\)

**Figure 5.** Linear regression analysis of PPD and summed scores between our method and the reference standard for stress/rest MPIs (a) SRS between our method and the reference standard, b: rest PPD between our method and the reference standard, c: SSS between our method and the reference standard, d: stress PPD between our method and the reference standard). There are 92 MPIs in each of these plots, and the plots show fewer samples because of overlapped data points.
Figure 6. Polar maps of SA images generated from the reference standard and our method. The percentage at the left bottom of each polar map is the PPD and the number at the right bottom is the summed score of the perfusion defect. The number in each region of 17 segments is the score of the segment perfusion defect.

Figure 7. Results of two MPI cases reoriented by our deep-learning-based method. Left, an MPI with a rotation angle of greater than 45°; Right, an MPI with significant overlap between the left ventricle and organs.
LIMITATIONS

This study enrolled a relatively small number of patients from a single medical center with the inherent limitation of such study design.

NEW KNOWLEDGE GAINED

We have developed a deep-learning-based method for LV reorientation on SPECT MPI. It can generate accurate SA images in a short time. There are high correlations for MPI quantification parameters measured from SA images generated by manual reorientation and by automatic reorientation using our model.

CONCLUSION

Our deep learning-based method is able to accurately generate the SA images. Technical and evaluations of measured clinical parameters show that it has great promise for clinical use.

Acknowledgements

This research was supported in part by a Michigan Technological Research Excellence Fund Research Seed grant (PI: Weihua Zhou), a seed grant from Michigan Technological University Health Research Institute (PI: Weihua Zhou), Henan Science and Technology Development Plan 2022 (Project Number: 222102210219), and the National Natural Science Foundation of China under Grant 62106233.

Disclosures

All authors declare that there are no conflicts of interest.

References

1. Hage FG, AlJaroudi WA. Review of cardiovascular imaging in the Journal of Nuclear Cardiology in, 2017 Part 2 of 2: Myocardial perfusion imaging. J Nucl Cardiol 2018;25:1390-9.
2. Kuhle WG, Porena G, Huang SC, Phelps ME, Schelbert HR. Issues in the quantitation of reoriented cardiac PET images. J Nucl Med 1992;33:1235-42.
3. deKemp RA, Nahmias C. Automated determination of the left ventricular long axis in cardiac positron tomography. Physiol Meas 1996;17:95-108.
4. Cauvin JC, Boire JY, Maublant JC, Bonny JM, Zanca M, Veyre A. Automatic detection of the left ventricular myocardium long axis and center in thallium-201 single photon emission computed tomography. Eur J Nucl Med 1992;19:1032-7.
5. van Hestenberg RP, Kemertink GJ, Hasman A. On the generation of short-axis and radial long-axis slices in thallium-201 myocardial perfusion single-photon emission tomography. Eur J Nucl Med 1996;23:924-31.
6. Zhou WH, Garcia EV. Nuclear image-guided approaches for cardiac resynchronization therapy (CRT). Curr Cardiol Rep 2016;18:89.
27. Zhu F, Zhao J, Zhao C, Tang S, Nan J, Li Y, et al. A new method incorporating deep learning with shape priors for left ventricular segmentation in myocardial perfusion SPECT images. ArXiv. 2022;abs/2206.03603.

28. He Z, Fernandes FdA, do Nascimento EA, Garcia EV, Mesquita CT, Zhou W. Incremental value of left ventricular shape parameters measured by gated SPECT MPI in predicting the super-response to CRT. J Nucl Cardiol. 2021.

29. He Z, Zhang X, Zhao C, Qian Z, Wang Y, Hou X, et al. A method using deep learning to discover new predictors of CRT response from mechanical dyssynchrony on gated SPECT MPI. 2021.

30. Zhao C, Vij A, Malhotra S, Tang J, Tang H, Pienta D, et al. Automatic extraction and stenosis evaluation of coronary arteries in invasive coronary angiograms. Comput Biol Med 2021;136:58.

31. Tang H, Bober RR, Zhao C, Zhang C, Zhou W. 3D fusion between fluoroscopy angiograms and SPECT myocardial perfusion images to guide percutaneous coronary intervention. J Nucl Cardiol. 2021.

32. Xu Z, Tang H, Malhotra S, Dong M, Zhao C, Ye Z, et al. Three-dimensional fusion of myocardial perfusion SPECT and invasive coronary angiography guides coronary revascularization. J Nucl Cardiol. 2022.

33. Zhou W, Hou X, Piccinelli M, Tang X, Tang L, Cao K, et al. 3D fusion of LV venous anatomy on fluoroscopy venograms with epicardial surface on SPECT myocardial perfusion images for guiding CRT LV lead placement. JACC-Cardiovascular Imaging 2014;7:1239-48.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.