

# A detection-task-specific deep-learning method to improve the quality of sparse-view myocardial perfusion SPECT images

Ze Zhang Yang<sup>a</sup>, Zitong Yu<sup>b</sup>, Nuri Choi<sup>c</sup>, and Abhinav K. Jha<sup>a,b,c</sup>

<sup>a</sup>Department of Electrical and Systems Engineering, Washington University, St. Louis, MO, USA

<sup>b</sup>Department of Biomedical Engineering, Washington University, St. Louis, MO, USA

<sup>c</sup>Mallinckrodt Institute of Radiology, Washington University, St. Louis, MO, USA

## ABSTRACT

Myocardial perfusion imaging (MPI) with single-photon emission computed tomography (SPECT) is a widely used and cost-effective diagnostic tool for coronary artery disease. However, the lengthy scanning time in this imaging procedure can cause patient discomfort, motion artifacts, and potentially inaccurate diagnoses due to misalignment between the SPECT scans and the CT-scans which are acquired for attenuation compensation. Reducing projection angles is a potential way to shorten scanning time, but this can adversely impact the quality of the reconstructed images. To address this issue, we propose a detection-task-specific deep-learning method for sparse-view MPI SPECT images. This method integrates an observer loss term that penalizes the loss of anthropomorphic channel features with the goal of improving performance in perfusion defect-detection task. We observed that, on the task of detecting myocardial perfusion defects, the proposed method yielded an area under the receiver operating characteristic (ROC) curve (AUC) significantly larger than the sparse-view protocol. Further, the proposed method was observed to be able to restore the structure of the left ventricle wall, demonstrating ability to overcome sparse-sampling artifacts. Our preliminary results motivate further evaluations of the method.

**Keywords:** Myocardial perfusion imaging, Single-photon emission computed tomography, Deep-learning, Sparse-view image

## 1. INTRODUCTION

Myocardial perfusion imaging (MPI) by single-photon emission computed tomography (SPECT) plays an important role in diagnosing, monitoring, and assessing coronary artery disease.<sup>1</sup> The procedure typically involves a step-and-shoot protocol with scanning time of up to 15 minutes. However, prolonged scanning presents several challenges. First, patients may experience discomfort and anxiety during the prolonged scanning procedure.<sup>2</sup> Second, the lengthy scanning procedure causes potential motion artifacts that degrade image quality.<sup>3,4</sup> Additionally, in this process, often an extra CT scan is acquired for attenuation compensation. This long scanning time can result in patient motion, which can cause misregistration between the SPECT and CT scans, compromising diagnostic accuracy.<sup>5-7</sup> Furthermore, the lengthy scanning time reduces patient throughput in high-volume clinical institutions. Hence, there is an important need for methods that can shorten the scanning time.<sup>8</sup>

One potential method to shorten scanning time is to reduce the number of projection angles while maintaining the same acquisition time per angle. However, the sparse-view acquisition yields reconstructed images with limited image quality for the task performance.<sup>9</sup> Therefore, there is an important need for methods to process sparse-view SPECT MPI images such that their quality for the task performance is improved. Recent advancements in deep-learning (DL) have shown promise in processing sparse-view images,<sup>10-12</sup> with most approaches being trained using image fidelity loss functions and evaluated through figures of merit such as root mean squared error (RMSE) and structural similarity index measure (SSIM). However, fidelity-based evaluations may not correspond to performance on clinical tasks.<sup>13-17</sup> For instance, previous literature has found discrepancies between fidelity-based assessments of a DL-based denoising method for MPI SPECT and performance on the task of cardiac defect detection.<sup>18</sup> Of most relevance to this study, Zhang et al. observed inconsistencies in

fidelity-based evaluations versus detection-task-based evaluations of DL algorithms that processed low-resolution images to predict high-resolution samples.<sup>19</sup> These findings highlight the need for investigating strategies that can incorporate the clinical task when developing DL-based imaging methods.

Recently, a detection-task-specific method for processing low-count MPI SPECT images was proposed by Rahman et al.<sup>20</sup> This was a deep-learning method that utilized our understanding of the human observer performance to preserve detection-task-specific features while processing the MPI SPECT images. In a retrospective clinical study, DEMIST yielded an improvement in performance on the task of detecting myocardial perfusion defects with an anthropomorphic model observer. Furthermore, DL approaches have been proposed for CT that typically utilize hybrid loss functions, combining image fidelity terms with task-specific components.<sup>21–23</sup> The importance of incorporating task-specific features in loss functions is supported by their demonstrated ability to enhance detection task performance, such as optimizing the signal-to-noise ratio,<sup>21</sup> using binary cross-entropy with DL-based observers<sup>22</sup> and using a trained network to extract features and define the feature-level loss as observer loss.<sup>23</sup>

Reducing the number of scanning angles leads to the loss of high-frequency features, impairing performance on the task of defect detection.<sup>24,25</sup> In this context, Li et al. developed a high-frequency enhancement method using dual-domain attention and a novel high-frequency regularization term, improving the ability of the method to recover high-frequency features under sparse-view conditions.<sup>24</sup> Similarly, in another study, Ayyoubzadeh et al. proposed a feature accentuation space enabling image restoration CNNs to enhance sharpness and high-frequency detail clarity. This method significantly improves the visual quality of restored images, particularly along edges and high-texture regions.<sup>26</sup> These studies provide promise that DL methods may help address high-frequency feature loss in imaging.

Building on prior work that considers task-specific features and the idea of the DEMIST approach in processing low-count MPI SPECT images,<sup>20</sup> we propose a detection-task-specific deep-learning method to enhance myocardial perfusion defect detection task performance for sparse-view MPI SPECT images. The proposed approach incorporates a task-specific term into the loss function alongside a fidelity term. The task-specific term is inspired by previous studies that have observed that the human observer performance can be modeled as an anthropomorphic model observer that processes information through a combination of frequency-selective channels.<sup>27</sup> The proposed method was evaluated on the clinical task of detecting myocardial perfusion defects using an anthropomorphic model observer.

## 2. METHODS

### 2.1 Proposed method

The architecture of the neural network for the proposed method is depicted in Fig.1. The proposed method takes sparse-view short-axis SPECT images as inputs and outputs the estimated full-view short-axis images.

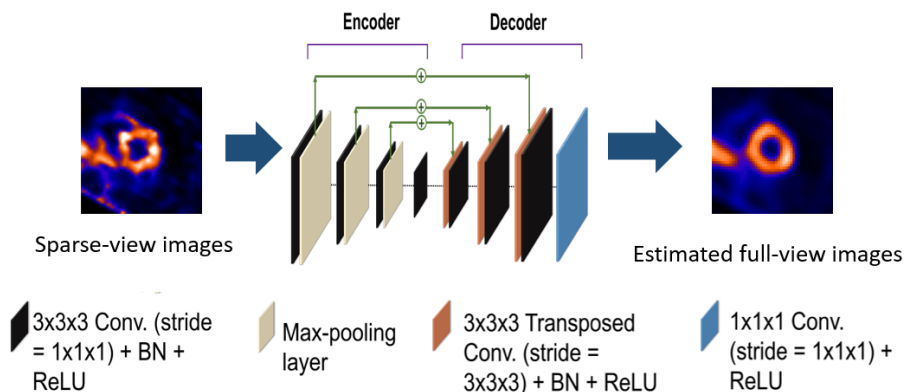


Figure 1: 3D neural network architecture of the proposed method

The goal of the proposed method is to process the sparse-view images such that the processed image yields improved performance on the task of detecting myocardial perfusion defects. For this purpose, we employ a loss function comprising two components: a fidelity term and an observer-based loss term. The fidelity term we incorporated is the mean squared error (MSE) between the estimated full-view images and the corresponding full-view images. Psychophysical studies have shown that the human observer can be modeled as a model observer that processes images through frequency-selective channels.<sup>28</sup> The channelized Hotelling observer (CHO) with rotationally symmetric frequency channels has been observed to emulate human observer performance in MPI SPECT defect detection.<sup>29,30</sup> Therefore, we incorporated an observer-based loss that penalizes the MSE between the channel vectors of the estimated full-view image and the corresponding full-view image.

Denote the size of the training dataset by  $J$ , and the  $j^{\text{th}}$  sample of the full-view image by  $\hat{\mathbf{f}}_{FV}^j$  and the  $j^{\text{th}}$  sample of the estimated full-view image by  $\hat{\mathbf{f}}_{FV}^{est,j}$ , respectively. Also, denote the channel operator by  $\mathbf{U}$ , a  $C \times N_{2D}$  matrix, where  $C$  represents the number of channels, and  $N_{2D}$  is the size of each image slice. In addition, denote the total number of slices picked from each image by  $s_2$  and the shift operation at  $j^{\text{th}}$  image as  $S^j$ . Finally, denote the  $j^{\text{th}}$  sample,  $s^{\text{th}}$  slice of the full-view image by  $\hat{\mathbf{f}}_{FV,2D,s}^j$  and the  $j^{\text{th}}$  sample,  $s^{\text{th}}$  slice of the estimated full-view image by  $\hat{\mathbf{f}}_{FV,2D,s}^{est,j}$ , respectively. Then, the loss function is given by:

$$L = \frac{1}{J} \sum_{j=1}^J \left\{ \|\hat{\mathbf{f}}_{FV}^j - \hat{\mathbf{f}}_{FV}^{est,j}\|_2^2 \right\} + \lambda \frac{1}{J} \sum_{j=1}^J \left\{ \sum_{s=1}^{s_2} \left\| (S^j \mathbf{U}) \left( \hat{\mathbf{f}}_{FV,2D,s}^j - \hat{\mathbf{f}}_{FV,2D,s}^{est,j} \right) \right\|_2^2 \right\}. \quad (1)$$

## 2.2 Objective evaluation of the proposed method

We evaluated the proposed method in an IRB-approved retrospective study. We followed best practices for developing<sup>31</sup> and evaluating<sup>17</sup> the proposed method, as recommended by the Society of Nuclear Medicine and Molecular Imaging (SNMMI) AI Task Force.

We collected data from  $N = 449$  patients, including SPECT projection and CT images. These images were originally acquired at the full-view and normal-dose level. To simulate the sparse-view images, we evenly sampled projection angles from full projection angles. These patients were divided into the training dataset of  $N = 184$  patients and the test dataset of  $N = 265$  patients. To ensure that the ground truth of the defects was known, we selected only projection scans of patients with no defects and inserted synthetic defects into these scans to create a population of defect-present cases. These defect types were defined based on their location in the left ventricle wall, extent, and severity. The defect insertion pipeline is described in detail in Rahman et al.<sup>20,32</sup> To generate sparse-view projections, we sampled 5 projection angles evenly from an original set of 30 projection angles. The full-view and sparse-view projection data were reconstructed using the ordered subsets expectation maximization (OSEM) algorithm with 8 iterations and 5 subsets for the sparse-view projection data and 6 subsets for the full-view projection data. The reconstructions were implemented in CASToR.<sup>33</sup> The resulting images were then reoriented to the short-axis view. From these reoriented images, we extracted a  $48 \times 48 \times 48$  volume centered on the left ventricle.

The neural network, as shown in Fig.1, was trained using the training samples. Further, five-fold cross-validation was conducted. The network was then tested with the test data. The test set comprised 134 healthy patients for the defect-absent population, while a separate set of 131 healthy patients with inserted defects was used to create the defect-present population.

Performance was evaluated on the task of detecting myocardial perfusion defects using an anthropomorphic model observer that, in previous studies, has been shown to emulate human-observer performance in MPI SPECT studies for defect detection tasks.<sup>30,34</sup> Performance on the detection task was quantified using the AUC values. The method was also compared with a method that only contained the MSE between the estimated and actual full-view images in the loss function and did not incorporate the observer-based loss term. We refer to this approach as the task-agnostic DL method (TADL). The statistical significance of the results was assessed using DeLong's test adjusted by Obuchowski's method.<sup>35,36</sup> A  $p$ -value less than 0.05 was used to infer statistical significance.

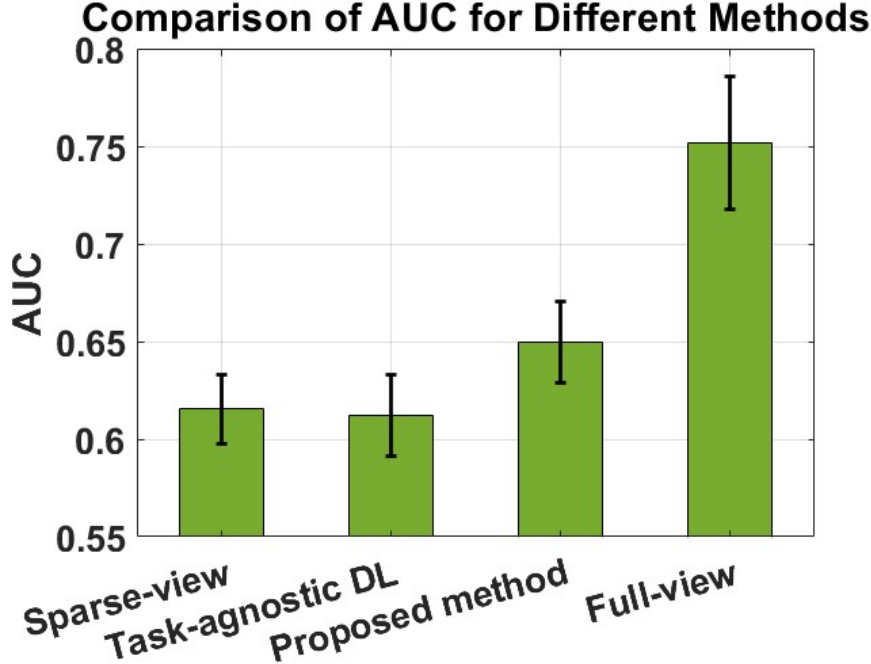


Figure 2: AUC values for the full-view and sparse-view protocol images, as well as the images processed using the proposed method and the task-agnostic DL method.

### 3. RESULTS

In this section, we present results with the proposed method and compare it with other protocols at the 5 sparse-view level setting. Fig. 2 shows the AUC values obtained with the sparse-view protocol, the task-agnostic DL method, and the proposed task-specific method at the 5 sparse-view level, along with the full-view protocol. The proposed method significantly outperformed the sparse-view protocol (DeLong’s test  $p$ -value  $< 0.05$ ).

Fig. 3 shows the qualitative comparison of the different protocols at the 5 sparse-view level setting. The proposed method demonstrates improved performance in reducing sparse-sampling artifacts. Notably, it restores the structure of the left ventricle wall while maintaining the visibility of the defect. The qualitative comparison in these two examples highlights the method’s ability to enhance defect visibility compared to sparse-view images, supporting its potential to improve defect detectability.

### 4. DISCUSSION

In this study, we developed a deep-learning-based method aimed at processing sparse-view MPI SPECT images with the goal of improving performance on the task of detecting myocardial perfusion defects. The method integrates an observer-based loss term alongside a fidelity-based term. The observer-based loss term specifically penalizes errors in features extracted from anthropomorphic channels. By minimizing these errors, the proposed method aims to improve the performance of sparse-view images on the task of detecting myocardial perfusion defects.

From the AUC results presented in Fig 2, the proposed task-specific method yields an improvement in detection task performance over sparse-view images. In particular, incorporating the observer-based loss term improves myocardial perfusion defect detection task performance from sparse-view images. From the preliminary results presented in Fig 3, we see that the proposed method shows promise in improving the performance of the sparse-view images in defect detection task. From the output images of the proposed method, the structure of the overall left ventricle is restored and the defect features are still visible. Therefore, the task of defect detection is expected to be more accurate in the estimated full-view images compared to the sparse-view images.

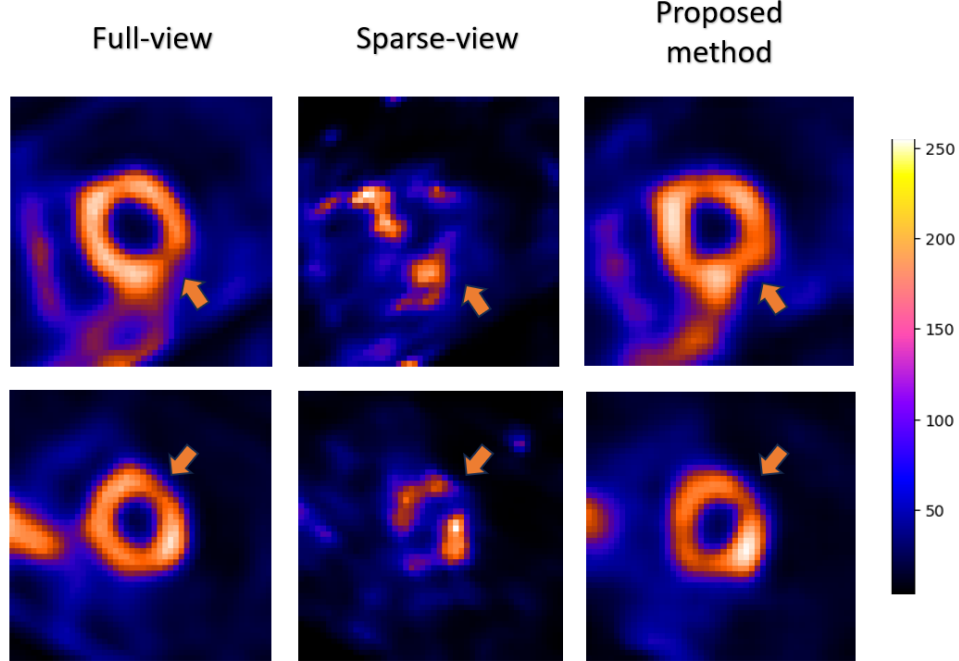


Figure 3: Two representative test cases qualitatively demonstrate the performance of the proposed method. In each case, the sparse-view level was set to 5. In the uppercase and lowercase cases, defects were located in the inferior and anterior walls, respectively. For all cases, the defects had an extent of 30 degrees and a severity of 25%.

This finding visually supports the potential of the proposed task-specific method to improve the performance of processed sparse-view images on the task of myocardial perfusion defect detection.

The sparse-view protocol loses high-frequency features. The proposed method addresses this issue by using the CNN neural networks to learn contextual priors and infer missing details by estimating full-view images from sparse-full images and comparing the estimated full-view images with the actual full-view images during training. The results in Fig. 3 show that the proposed method has the ability to address the sparse-sampling artifacts. These observations motivate further quantitative investigation.

There are several limitations that should be considered. First, the method was trained on data with inserted synthetic defects due to challenges in getting a reliable ground truth for the presence and location of the myocardial perfusion defect in clinical images. Therefore, one area of future research is to use clinical defect-present data with well-curated ground truth. Second, the defect locations were restricted to two regions. Expanding this to include additional regions, such as the septal and lateral walls, would further support the robustness of the method’s performance. Finally, the study used data from a single center. In the future, we will conduct multi-center validation as it is essential to evaluate the generalizability of the method across diverse clinical settings.<sup>37</sup>

## 5. CONCLUSION AND FUTURE WORK

We proposed a task-specific deep-learning method designed to enhance sparse-view myocardial perfusion SPECT images. The loss function in the method penalizes the loss of anthropomorphic channel features with the goal of improving performance in the perfusion defect-detection task. Our preliminary findings suggest that the proposed method can improve both the quality of the sparse-view images and the accuracy of defect detection, motivating further validation of the method.

## 6. ACKNOWLEDGMENTS

This work was supported in part by National Institute of Biomedical Imaging and Bioengineering of National Institute of Health (NIH) under grant number R01-EB031051 and R01-EB031962, and National Science Foundation (NSF) CAREER Award 2239707.

## REFERENCES

- [1] Zaret, B. L. and Beller, G. A., “Clinical Nuclear Cardiology: State of the Art and Future Directions E-Book,” (2010).
- [2] Nightingale, J. M., Murphy, F. J., and Blakeley, C., “I thought it was just an x-ray: a qualitative investigation of patient experiences in cardiac SPECT-CT imaging,” *Nucl. Med. Commun* **33**(3), 246–254 (2012).
- [3] Agarwal, V. and DePuey, E. G., “Myocardial perfusion SPECT horizontal motion artifact,” *J. Nucl. Cardiol* **21**(6), 1260–1265 (2014).
- [4] Botvinick, E. H., Zhu, Y., O’Connell, W. J., and Dae, M. W., “A quantitative assessment of patient motion and its effect on myocardial perfusion SPECT images,” *J. Nucl. Med* **34**(2), 303–310 (1993).
- [5] Goetze, S., Brown, T. L., Lavelly, W. C., Zhang, Z., and Bengel, F. M., “Attenuation correction in myocardial perfusion SPECT/CT: effects of misregistration and value of reregistration,” *J. Nucl. Med* **48**(7), 1090–1095 (2007).
- [6] Goetze, S. and Wahl, R. L., “Prevalence of misregistration between SPECT and CT for attenuation-corrected myocardial perfusion SPECT,” *J. Nucl. Cardiol* **14**, 200–206 (2007).
- [7] Saleki, L., Ghafarian, P., Bitarafan-Rajabi, A., Yaghoobi, N., Fallahi, B., and Ay, M. R., “The influence of misregistration between CT and SPECT images on the accuracy of CT-based attenuation correction of cardiac SPECT/CT imaging: Phantom and clinical studies,” *Iranian J. Nucl. Med* **27**(2), 63–72 (2019).
- [8] Heller, G. V., “Practical issues regarding the incorporation of PET into a busy SPECT practice,” *J. Nucl. Cardiol* **19**, S12–S18 (2012).
- [9] Chen, X., Zhou, B., Xie, H., Guo, X., Liu, Q., Sinusas, A. J., and Liu, C., “Cross-Domain Iterative Network for Simultaneous Denoising, Limited-Angle Reconstruction, and Attenuation Correction of Cardiac SPECT,” in *[International Workshop on Machine Learning in Medical Imaging]*, 12–22, Springer (2023).
- [10] Chen, X., Zhou, B., Xie, H., Miao, T., Liu, H., Holler, W., Lin, M., Miller, E. J., Carson, R. E., Sinusas, A. J., et al., “DuDoSS: Deep-learning-based dual-domain sinogram synthesis from sparsely sampled projections of cardiac SPECT,” *Med. Phys* **50**(1), 89–103 (2023).
- [11] 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* **62**(4), 528–535 (2021).
- [12] 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* **19**(4), 4320–40 (2022).
- [13] Yu, Z., Rahman, M. A., Schindler, T., Gropler, R., Laforest, R., Wahl, R., and Jha, A., “AI-based methods for nuclear-medicine imaging: Need for objective task-specific evaluation,” *J. Nucl. Med* **61**(Supplement 1), 575–575 (2020).
- [14] Yu, Z., Rahman, M. A., Laforest, R., Schindler, T. H., Gropler, R. J., Wahl, R. L., Siegel, B. A., and Jha, A. K., “Need for objective task-based evaluation of deep learning-based denoising methods: A study in the context of myocardial perfusion SPECT,” *Med. Phys* (2023).
- [15] Li, K., Zhou, W., Li, H., and Anastasio, M. A., “Assessing the impact of deep neural network-based image denoising on binary signal detection tasks,” *IEEE Trans. Med. Imag* **40**(9), 2295–2305 (2021).
- [16] Badal, A., Cha, K. H., Divil, S. E., Graff, C. G., Zeng, R., and Badano, A., “Virtual clinical trial for task-based evaluation of a deep learning synthetic mammography algorithm,” in *[Medical Imaging 2019: Physics of Medical Imaging]*, **10948**, 164–173, SPIE (2019).
- [17] Jha, A. K., Bradshaw, T. J., Buvat, I., Hatt, M., Prabhat, K., Liu, C., Obuchowski, N. F., Saboury, B., Slomka, P. J., Sunderland, J. J., et al., “Nuclear medicine and artificial intelligence: best practices for evaluation (the RELAINCE guidelines),” *J. Nucl. Med* **63**(9), 1288–1299 (2022).

- [18] Yu, Z., Rahman, M. A., Laforest, R., Schindler, T. H., Gropler, R. J., Wahl, R. L., Siegel, B. A., and Jha, A. K., "Need for objective task-based evaluation of deep learning-based denoising methods: A study in the context of myocardial perfusion SPECT," *Med. Phys* (2023).
- [19] Zhang, X., Kelkar, V. A., Granstedt, J., Li, H., and Anastasio, M. A., "Impact of deep learning-based image super-resolution on binary signal detection," *J. Med. Imaging* **8**(6), 065501–065501 (2021).
- [20] Rahman, M. A., Yu, Z., Laforest, R., Abbey, C. K., Siegel, B. A., and Jha, A. K., "DEMIST: A deep-learning-based task-specific denoising approach for myocardial perfusion SPECT," *arXiv preprint arXiv:2306.04249* (2023).
- [21] Ongie, G., Sidky, E. Y., Reiser, I. S., and Pan, X., "Optimizing model observer performance in learning-based CT reconstruction," in [*Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*], **12035**, 55–59, SPIE (2022).
- [22] Li, K., Li, H., and Anastasio, M. A., "A task-informed model training method for deep neural network-based image denoising," in [*Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*], **12035**, 249–255, SPIE (2022).
- [23] Han, M., Shim, H., and Baek, J., "Low-dose CT denoising via convolutional neural network with an observer loss function," *Med. Phys* **48**(10), 5727–5742 (2021).
- [24] Li, G., Deng, Z., Ge, Y., and Luo, S., "HEAL: high-frequency enhanced and attention-guided learning network for sparse-view CT reconstruction," *Bioengineering* **11**(7), 646 (2024).
- [25] Qu, Z., Yan, X., Pan, J., and Chen, P., "Sparse view CT image reconstruction based on total variation and wavelet frame regularization," *IEEE Access* **8**, 57400–57413 (2020).
- [26] Ayyoubzadeh, S. M. and Wu, X., "High Frequency Detail Accentuation in CNN Image Restoration," *IEEE Trans. Image Process* **30**, 8836–8846 (2021).
- [27] Frey, E. C., Gilland, K. L., and Tsui, B. M., "Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT," *IEEE Trans. Med. Imaging* **21**(9), 1040–1050 (2002).
- [28] Barrett, H. H. and Myers, K. J., [*Foundations of image science*], John Wiley & Sons (2003).
- [29] Myers, K. J. and Barrett, H. H., "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Am. A* **4**(12), 2447–2457 (1987).
- [30] Sankaran, S., Frey, E. C., Gilland, K. L., and Tsui, B. M., "Optimum compensation method and filter cutoff frequency in myocardial SPECT: a human observer study," *J. Nucl. Med* **43**(3), 432–438 (2002).
- [31] Bradshaw, T. J., Boellaard, R., Dutta, J., Jha, A. K., Jacobs, P., Li, Q., Liu, C., Sitek, A., Saboury, B., Scott, P. J., Slomka, P. J., Sunderland, J. J., Wahl, R. L., Yousefirizi, F., Zuehlsdorff, S., Rahmim, A., and Buvat, I., "Nuclear Medicine and Artificial Intelligence: Best Practices for Algorithm Development," *J. Nucl. Med* (2021).
- [32] Narayanan, M. V., King, M. A., Leppo, J., Dahlbert, S., Pretorius, P. H., and Gifford, H. C., "Optimization of regularization of attenuation and scatter-corrected/sup 99m/Tc cardiac SPECT studies for defect detection using hybrid images," *IEEE Trans. Nucl. Sci* **48**(3), 785–789 (2001).
- [33] Merlin, T., Stute, S., Benoit, D., Bert, J., Carlier, T., Comtat, C., Filipovic, M., Lamare, F., and Visvikis, D., "CASToR: a generic data organization and processing code framework for multi-modal and multi-dimensional tomographic reconstruction," *Phys. Med. Biol* **63**(18), 185005 (2018).
- [34] Wollenweber, S., Tsui, B., Lalush, D., Frey, E., LaCroix, K., and Gullberg, G., "Comparison of Hotelling observer models and human observers in defect detection from myocardial SPECT imaging," *IEEE Trans. Nucl. Sci* **46**(6), 2098–2103 (1999).
- [35] Obuchowski, N. A., "Nonparametric analysis of clustered ROC curve data," *Biometrics* , 567–578 (1997).
- [36] DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L., "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics* , 837–845 (1988).
- [37] Yu, Z., Choi, N. R., Yang, Z., Obuchowski, N. A., Siegel, B. A., and Jha, A. K., "ISIT-GEN: An In Silico Imaging Trial to Assess the Inter-Scanner Generalizability of CTLESS for Myocardial Perfusion SPECT on Defect-Detection Task," in [*Proceedings of SPIE–The International Society for Optical Engineering*], **Accepted** (2025).