

SPASHT: An image-enhancement method for sparse-view MPI SPECT

Zezhang Yang¹, Zitong Yu², Nuri Choi³, Janice Tania², Wenxuan Xue²,
Barry A. Siegel³, and Abhinav K. Jha^{1,2,3}

¹Department of Electrical and Systems Engineering, Washington University,
St. Louis, MO, USA

²Department of Biomedical Engineering, Washington University, St. Louis,
MO, USA

³Mallinckrodt Institute of Radiology, Washington University, St. Louis,
MO, USA

Abstract

Single-photon emission computed tomography for myocardial perfusion imaging (MPI SPECT) is a widely used diagnostic tool for coronary artery disease. However, the procedure requires considerable scanning time, leading to patient discomfort and the potential for motion-induced artifacts. Reducing the number of projection views while keeping the time per view unchanged provides a mechanism to shorten the scanning time. However, this approach leads to increased sampling artifacts, higher noise, and hence limited image quality. To address these issues, we propose sparse-view SPECT image enhancement (SPASHT), a method that incorporates our understanding of the human visual system within a deep-learning approach to process the

sparse-view SPECT images, inherently training the algorithm to improve performance on defect-detection tasks. We objectively evaluated SPASHT on the clinical task of detecting perfusion defects in a retrospective clinical study using data from patients who underwent MPI SPECT, where the defects were clinically realistic and synthetically inserted. The study was conducted for different numbers of fewer projection views, including 1/6, 1/3, and 1/2 of the typical projection views for MPI SPECT. Performance on the detection task was quantified using area under the receiver operating characteristic curve (AUC). Images obtained with SPASHT yielded significantly improved AUC compared to those obtained with the sparse-view protocol for all the considered numbers of fewer projection views. To further assess performance, a human observer study on the task of detecting perfusion defects was conducted. Results from the human observer study showed improved detection performance with images reconstructed using SPASHT compared to those from the sparse-view protocol. The results provide evidence of the efficacy of SPASHT in improving the quality of sparse-view MPI SPECT images and motivate further clinical validation.

Keywords: Myocardial perfusion imaging, Single-photon emission computed tomography, Deep learning, Sparse-view tomography, Image enhancement, Task-aware deep learning

1 Introduction

Myocardial perfusion imaging (MPI) with single-photon emission computed tomography (SPECT) plays a crucial role in diagnosing, measuring severity, and tracking the progression of coronary artery disease ([1](#)). In clinical practice, MPI SPECT is typically performed by rotating gamma cameras around the chest of the patient in a step-and-shoot protocol to acquire projection data. In these exams, patients go through both stress and rest tests. The scanning procedures are generally lengthy (15-20 minutes for each of these tests), presenting challenges to efficient and accurate clinical exams. Patient movement during a long scan can cause motion artifacts, leading to significant image-quality degradation and reduced diag-

nostic accuracy (2,3). Excessive movements may require repeat scans, thereby increasing radiation exposure. Moreover, sometimes a CT scan is acquired along with the SPECT scan for attenuation compensation. As SPECT and CT images are acquired sequentially, this lengthy acquisition increases the probability of patient movement leading to misalignment between SPECT and CT images, resulting in potentially inaccurate diagnosis (4,5). Additionally, patients may experience discomfort due to the prolonged scanning time, increasing the risk of noncompliance and resulting in suboptimal scans and potentially inaccurate image interpretation (6). Finally, in a high-volume clinical institution, shorter scanning time can improve patient throughput (7). Thus, there is an important need to shorten the imaging time for MPI SPECT.

Reducing the number of projections while keeping the time spent in each projection unchanged offers a solution for shortening scanning time. However, such a reduction of the number of projections introduces severe artifacts and the loss of high-frequency components, compromising diagnostic accuracy (8). Thus, developing methods to enhance the quality of these sparse-view images is an active area of research. Recently, several deep learning (DL) methods have shown promise in addressing the severe artifacts of the sparse-view images (9–12). Shiri et al. trained a residual neural network to recover visual image quality from images acquired with either half the number of projections or half the acquisition time per projection. (9). Amirrashedi et al. developed a neural network to recover missing projection data in small animal PET, showing promising artifact reduction (10). Li et al. proposed LU-Net, a hybrid network that combines Long Short-Term Memory network (LSTM) and U-Net to synthesize missing sinogram views for sparse-view SPECT, showing improved reconstruction image quality and artifact suppression (11). In addition, Xie et al. introduced a diffusion model for 3D cardiac SPECT under low-dose and few-view conditions, observing generalizability across different acquisition settings and highlighting the effectiveness of their model in improving the image quality from low-dose or few-view SPECT data (13). Meanwhile, there are also several studies that develop DL methods to address the

loss of high-frequency components issue (14, 15). Li et al. introduced a method leveraging high-frequency enhancement and dual-domain attention mechanisms to combat detail loss in sparse CT reconstructions. Their approach incorporates a high-frequency regularization term to differentiate noise from textures, significantly improving detail reconstruction accuracy under sparse-view conditions (14). 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 (15).

These methods typically are evaluated by fidelity-based metrics such as mean squared error (MSE), structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR). However, fidelity-based evaluations may not correspond to performance on clinical tasks (16–20). For example, in a recent study, a discrepancy was observed between evaluations of a DL-based denoising approach for MPI SPECT on fidelity-based figures of merit (FoMs) versus on the cardiac defect-detection task (17, 21). Further, in another study, Zhang et al observed that DL-based algorithms for processing low-resolution images to estimate high-resolution images, when evaluated using fidelity-based FoMs, yielded interpretations inconsistent with evaluation based on clinical tasks (22). To address this issue, recent studies have developed DL approaches that preserve features that are relevant to the detection task (23–25). Findings from these studies support the importance of considering features that assist in the clinical task in the loss function. However, to the best of our knowledge, no prior work has implemented task-aware neural networks for sparse-view MPI SPECT images and evaluated the neural network enhanced sparse-view images based on task-specific performance.

In this manuscript, we propose SPASHT, an image enhancement method for sparse-view MPI SPECT images. Inspired by a technique that incorporates a task-specific term during the training of the neural network for denoising low-dose MPI SPECT images (26), our approach integrates a task-specific observer term into the loss function alongside the mean

squared error (MSE) term to improve the quality of sparse-view MPI SPECT images. This task-specific observer term is designed based on the channelized Hotelling model observer framework (27) and 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 (26). The performance of the SPASHT method was evaluated on the clinical task of detecting myocardial perfusion defects in a retrospective clinical study both with anthropomorphic model observer (27) and trained human observers.

2 Methods

2.1 SPASHT method

The proposed method is designed to process sparse-view images to obtain the estimated full-view images, such that the processed images have improved performance in the task of detecting myocardial perfusion defects. The neural network is parameterized by the vector Θ and represented by the operator E_Θ . \hat{f}_{SV} is the sparse-view image. This operator generates an estimated full-view image, denoted by \hat{f}_{FV}^{est} . The enhancement of sparse-view images can thus be formulated as:

$$\hat{f}_{FV}^{est} = E_\Theta(\hat{f}_{SV}) \quad (1)$$

To take the performance of the myocardial perfusion defect detection task into consideration during the neural network training and improve the overall noise properties of the estimated full-view images, we employ a hybrid loss function consisting of two parts: a fidelity term and an observer term.

The fidelity term is the mean squared error (MSE) between the estimated full-view image and the full-view image. The second term in our loss function is designed to capture

features of the image in the myocardial perfusion defect detection task. In clinical scenarios, these tasks are performed by physicians, i.e. human observers. Psychophysical studies have shown that the human observer can be modeled as a model observer that processes images through frequency-selective channels (28). The channelized Hotelling observer (CHO) with rotationally symmetric frequency channels has been observed to emulate human observer performance in MPI SPECT defect detection (29, 30). Thus, we design the task-specific observer term which penalizes the MSE between the CHO channel vectors of the estimated full-view image and the corresponding full-view image.

Denote the fidelity term as $L_{fid}(\Theta)$ and the task-specific observer term as $L_{ob}(\Theta)$. The loss function $L(\Theta)$ is given by:

$$L(\Theta) = L_{fid}(\Theta) + L_{ob}(\Theta) \quad (2)$$

Let J represent the total number of training samples, and denote the j -th sample of the sparse-view and full-view images by the N -dimensional vectors $\hat{\mathbf{f}}_{SV}^j$ and $\hat{\mathbf{f}}_{FV}^j$, respectively. Additionally, let $\hat{\mathbf{f}}_{FV}^{est,j}$ denote the full-view images estimated by the neural network when the sparse-view image $\hat{\mathbf{f}}_{SV}^j$ is used as input. Thus,

$$\hat{\mathbf{f}}_{FV}^{est,j} = E_\Theta(\hat{\mathbf{f}}_{SV}^j) \quad (3)$$

Given J patient images in the training set, the fidelity-loss term is defined as:

$$L_{fid}(\Theta) = \frac{1}{J} \sum_{j=1}^J \|\hat{\mathbf{f}}_{FV}^j - \hat{\mathbf{f}}_{FV}^{est,j}\|_2^2 \quad (4)$$

Denote the number of voxels in each image slice by N_{2D} and the channelization operator by a $C \times N_{2D}$ matrix \mathbf{U} , where C represents the number of channels. Then, we denote S as the operator that shifts each channel to the centroid of the defect. In addition, denote the image slice by a N_{2D} -dimensional vector $\hat{\mathbf{f}}_{2D}$. To compute the feature vector for the j th patient sample, we first perform a 2D shift on each channel, aligning the center of the

channel profile with the centroid of the defect. Since defect locations vary between patients, let S^j denote the shift operator for the j^{th} patient. Let $\hat{\mathbf{f}}_{2D,s}^j$ denote the s^{th} slice of the image $\hat{\mathbf{f}}^j$. Therefore, the j^{th} patient, s^{th} slice of the full-view image is $\hat{\mathbf{f}}_{FV,2D,s}^j$ and the j^{th} patient, s^{th} slice of the estimated full-view image is $\hat{\mathbf{f}}_{FV,2D,s}^{est,j}$, respectively. The task-specific observer term $L_{ob}(\Theta)$ is thus defined as:

$$L_{ob}(\Theta) = \frac{1}{J} \sum_{j=1}^J \sum_{s=s1}^{s2} \| (\mathcal{S}^j \mathbf{U})^T \Delta \mathbf{f} \|_2^2 \quad (5)$$

$$\Delta \mathbf{f} = \hat{\mathbf{f}}_{FV,2D,s}^j - \hat{\mathbf{f}}_{FV,2D,s}^{est,j} \quad (6)$$

We designed an encoder-decoder architecture to minimize the loss function defined in Eq.3. A schematic of this architecture is presented in Fig.1. The network's input is the sparse-view short-axis images $\hat{\mathbf{f}}_{SV}$, and its output is the enhanced (estimated full-view) short-axis images $\hat{\mathbf{f}}_{FV}^{est}$.

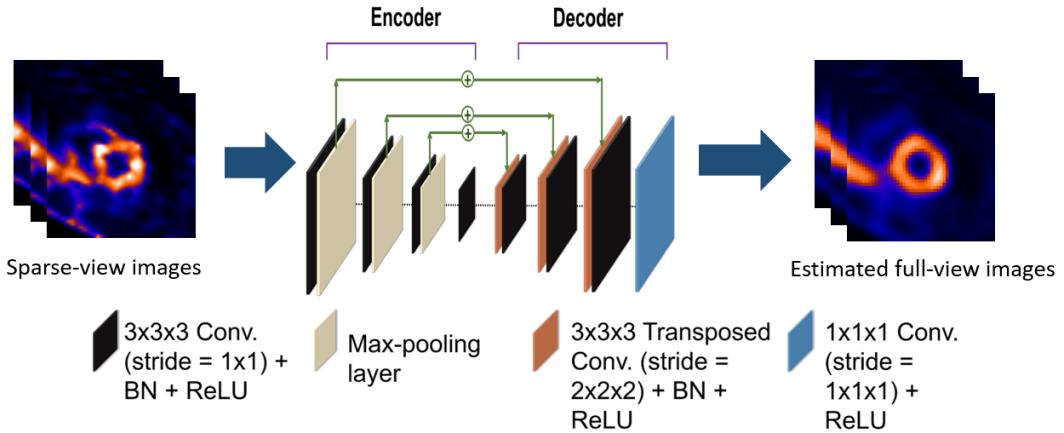


Figure 1: 3D neural network architecture of the SPASHT method (Conv: convolutional neural network layer, BN: batch normalization, ReLU: rectified linear unit activation function)

The encoder extracts local spatial features from the sparse-view image, generating a set of lower-dimensional features used to reconstruct the estimated full-view image. Skip connections were incorporated to add features learned in the encoder to those generated by

the decoder, enhancing feature retention across layers. Dropouts were also applied to prevent overfitting.

3 Evaluation

We evaluated the SPASHT method in an IRB-approved retrospective study using data obtained from patients who underwent an one-day MPI SPECT study. This evaluation followed best practices for assessing AI algorithms in nuclear medicine (RELAINCE guidelines (20)).

3.1 Data curation

We collected data from $N = 449$ patients, including SPECT projection and CT images along with patient sex and diagnosis. These images were originally acquired at full-view and normal-dose levels. To simulate sparse-view images, we evenly sampled projection views within the projection domain. These patients were divided into the training dataset of $N = 184$ patients and the test dataset of $N = 265$ patients. The datasets were acquired from two scanners: the GE Discovery 670 Pro NaI and the GE Discovery 670 CZT.

To create a dataset with known ground truth about the presence or absence of defects, we selected images that were clinically interpreted as normal. We inserted synthetic defects into the SPECT projections corresponding to these images to create our defect-present population. In the training dataset, we designed 12 types of defects with two different locations (inferior and anterior walls) in the left ventricle (LV) wall, extents (30 and 60 degrees around the LV wall), and severities (10%, 17.5%, and 25% less activity than the normal myocardium). In the test dataset, we designed 18 types of defects with two different locations (inferior and anterior walls) in the LV wall, extents (30, 60 and 90 degrees around the LV wall), and severities (10%, 17.5%, and 25% less activity than the normal myocardium). The defect insertion pipeline is based on a strategy proposed by Narayanan et al (31), and was described in detail by Rahman et al (26).

To create sparse-view projection data, we sampled 5, 10, and 15 projection views evenly from an original set of 30 projection views, yielding the sparse-view projections. Both full-view and sparse-view projection data were reconstructed using an ordered subsets expectation maximization (OSEM)-based approach with 8 iterations and 5 subsets for sparse-view projection data and 6 subsets for full-view projection data (32). The reconstructed images were then reoriented to the short-axis view. From these short-axis images, we extracted a $48 \times 48 \times 48$ volume centered on the left ventricle.

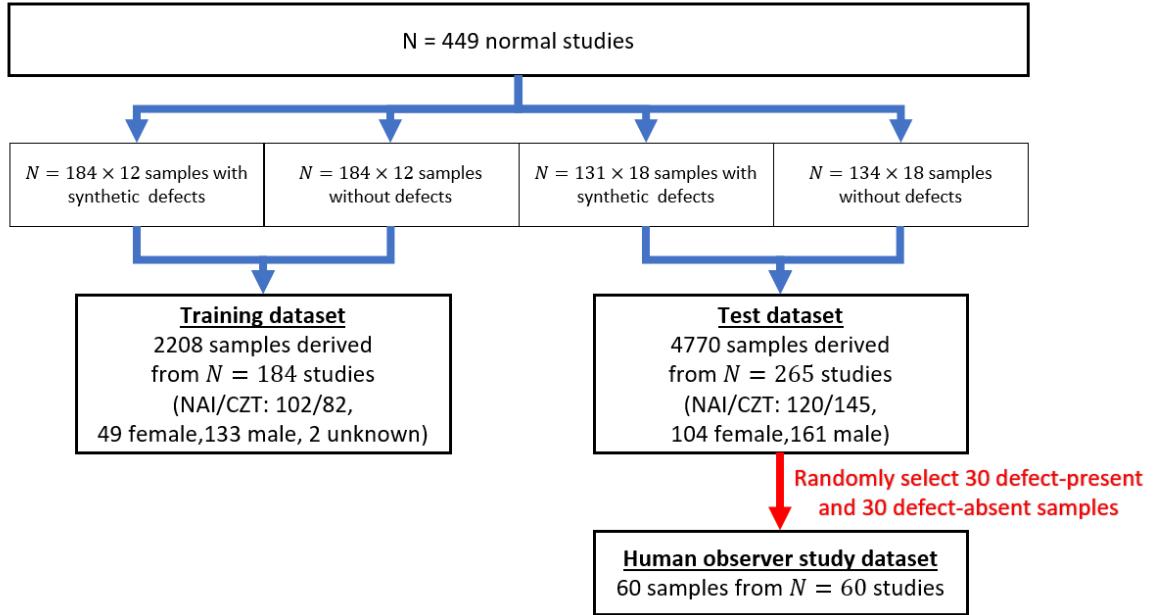


Figure 2: Flow chart of the data curation process

3.2 Training and testing dataset

A total of 184 patient studies are included in the training dataset. As mentioned earlier, we simulated 12 different types of defects. Thus, in the defect-present population, there were a total of 2208 defect-present samples (184×12). For the defect-absent population, each patient yielded a cluster of 12 samples, yielding 2208 defect-absent samples (184×12). These two populations, comprising a total of $N = 4416$ samples, were used to train the network.

The network was trained by minimizing the loss in Eq.3 using the ADAM optimization algorithm (33). The kernel weights of the network were initialized using the Glorot normal initializer (34). Biases were initialized to a constant of 0.03. To avoid overfitting, the dropout rate is set as 0.1 (35). We conducted a five-fold cross-validation to determine the optimal number of epochs. During the fold generation, we ensured that all samples from a single patient were contained within the same fold. After identifying the optimal number of epochs, we retrained the network using the entire training dataset. During the training, to extract channel vectors from defect-present images, each channel profile in U was shifted to align with the defect centroid. For defect-absent images, channel vectors were obtained by centering the profiles in U on the location where a synthetic defect would be inserted otherwise. Using these aligned channel profiles, we extracted channel vectors from both the estimated and full-view images to compute the task-specific loss term. A five-fold cross-validation was conducted to optimize the network. The training and validation were performed using Tensorflow 2.2.4 (36). The training took around 10 hours. The number of epochs was chosen where the minimum averaged five-fold cross-validation loss was reached. The training was performed on an NVIDIA Quadro P6000 GPU with 24 GB of memory.

A total of 265 patient studies are included in the test dataset. There were a total of 2208 defect-present samples (131×18) and a total of 2412 defect-absent samples (134×18). For the model observer study, all 4620 samples were used. For the human observer study, we randomly selected 30 defect-present samples from the test set, with each sample from one patient study with one synthetic defect type. The synthetic defects in the human observer study varied across eight types, with two extents (30° and 60° along the left ventricle wall), two severity levels (25% [moderate] and 50% [severe] perfusion reduction), and two anatomical locations (anterior and inferior walls of the left ventricle). We also select 30 defect-absent samples from the test dataset. A total of $N = 60$ samples were selected for the human observer study.

3.3 Performing the detection task

3.3.1 Anthropomorphic model observer study

In the clinical setting, the task of the myocardial perfusion defect detection task is performed by physicians. Thus, it would be ideal for the task-specific evaluation of SPASHT to also be performed by physicians; However, conducting human observer studies at the early stage of method development is often impractical because they are resource-intensive, requiring significant physician time and coordination. Such studies are also labor-intensive and time-consuming, from organization to data collection and analysis

To address these limitations during early-stage evaluation, we first performed an anthropomorphic model observer study using the CHO. As referenced above, it has been shown in literature that the CHO can serve as a surrogate for human observers in the defect detection task. Specifically, CHO with rotationally symmetric frequency channels has been validated as effective emulators of human observer performance for detecting location-known perfusion defects in MPI SPECT (37). Additionally, they provide the benefit of quantitative, consistent, and repeatable performance, unlike human observers, whose responses may vary due to fatigue, bias, or subjective interpretation. Therefore, to objectively evaluate the SPASHT for detecting perfusion defects, we used the CHO to conduct the anthropomorphic model observer study.

We evaluated the SPASHT’s performance on the clinical task of detecting perfusion defects, as well as using fidelity-based metrics. Performance was compared to sparse-view images without enhancement, referred to as the sparse-view protocol. To assess the effect of our task-specific enhancement approach, we also compared results to images denoised with a widely used deep learning (DL) denoising method (38), trained with a loss function that included only the fidelity term. We refer to this as the task-agnostic DL method.

The estimated full-view images were produced from the test set using the trained model. These images were then used to perform task-specific image quality evaluation on the de-

tection of perfusion defects in short-axis SPECT images. In our study, we selected the 2-D short-axis slice containing the defect centroid along with the two adjacent slices for each 3-D test sample. From each slice, we extracted a 32×32 region ($21.76 \text{ cm} \times 21.76 \text{ cm}$) centered on the defect centroid. Pixel values of each extracted region were mapped to the range [0,255] for the purpose of normalization. The channel operator was then applied to these images to extract feature vectors. The template of CHO was learned from the feature vectors of defect-present and defect-absent populations using a leave-one-out strategy (39, 40). Subsequently, test statistics were calculated and used for ROC analysis. A stratified analysis was also conducted based on sex, defect severity, and defect extent.

3.3.2 Human observer study

After anthropomorphic model observer study, we organized a human observer study to evaluated the SPASHT method using a human observer study on the task of detecting perfusion defects in MPI-SPECT images and further confirm that the performance results observed with the CHO are also the similar when interpreted by human readers. The study was based on the 10 sparse-view level setting.

Training of the human observers

Four experienced human observers participated in this study. The observers are hereafter referred to as "readers." Before the formal human observer study, all the readers had a training session to let the readers become familiar with the human observer study. In the training session, all the readers are taught how to conduct the myocardial perfusion defect detection task following guidance from physicians. The possible types of defects, including variations in the severity, extent, and location of the defect, were shown to all the readers. Some examples of MPI SPECT images were shown to the readers with the ground truth of the defect and the ground truth of the image description.

Readers were instructed on how to perform the task specific defect-detection task and the

usage of the web-based application. All readers were asked to perform the defect detection task with specific examples to confirm adequate training prior to the formal observer study.

Finally, before the formal human observer study, all readers were asked to perform the myocardial perfusion defect detection task with some examples to confirm that all readers are prepared for the human observer study.

Reading Procedure

The study was conducted using a fully crossed design, in which each reader reviewed the entire image set. All image interpretations were performed using a web-based application that simulated a clinically relevant MPI SPECT reading interface (Fig. 3). The application presented clinically standardized MPI SPECT image displays and allowed readers to adjust image contrast and switch between grayscale and GE color scales. Additionally, rotating views of SPECT photopeak projection images were provided. Readers evaluated one sample at one time and assigned a likelihood rating of perfusion defect presence using a seven-point ordinal scoring scale.

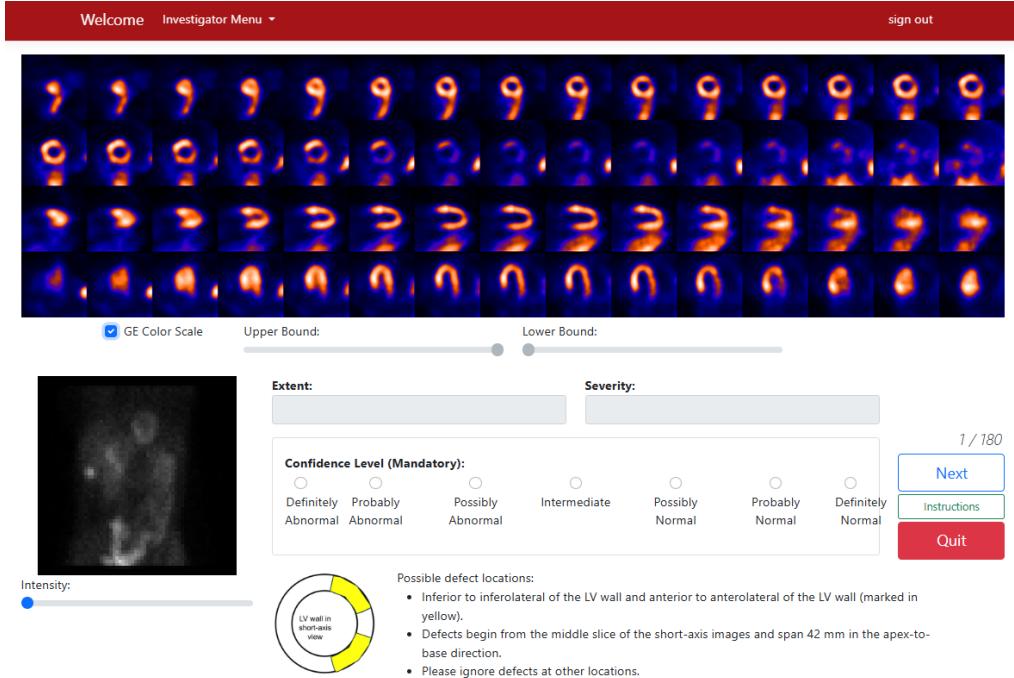


Figure 3: The web-based application used for image display and scoring. The top section displays rows of short-axis, horizontal long-axis, and vertical long-axis SPECT-MPI images. A rotating view of the SPECT photopeak projection image is shown at the bottom left. The seven-point ordinal scoring scale is explained in the center of the interface. An illustration of the two possible defect locations (anterior and inferior walls) is provided at the bottom.

Images of the same study obtained by the three methods (full-view, the SPASHT method, sparse-view) were randomized within the reading session. Readers were blinded to the reconstruction method used for each image to minimize potential bias.

3.3.3 Figure of merit

Performance evaluation was conducted using the receiver operating characteristic (ROC) and the area under the ROC curve (AUC). We performed the ROC analysis and calculated AUCs with 95% confidence intervals (CIs) for sparse-view protocol, full-view protocol, the task-agnostic DL method and the SPASHT method, using a nonparametric strategy that accounted for the correlated nature of the data (41, 42), where the correlation between data

samples was obtained from Hanley and McNeil's method (43).

We also performed stratified analyses for male and female populations, as well as across defect extents and severity levels. In addition, we also evaluated the generalizability of the SPASHT method across the two different scanners used to acquire the images. We trained SPASHT using a mixed dataset comprising both CZT-camera and NaI-camera data and tested it separately on CZT-camera and NaI-camera datasets. Performance was quantitatively compared on both the defect-detection task and the metrics of RMSE and SSIM. For all statistical tests in this study, a p-value < 0.05 was used to infer statistical significance with using a nonparametric strategy that accounted for the correlated nature of the data (41,42),

4 Results

4.1 Evaluation on the perfusion defect detection task

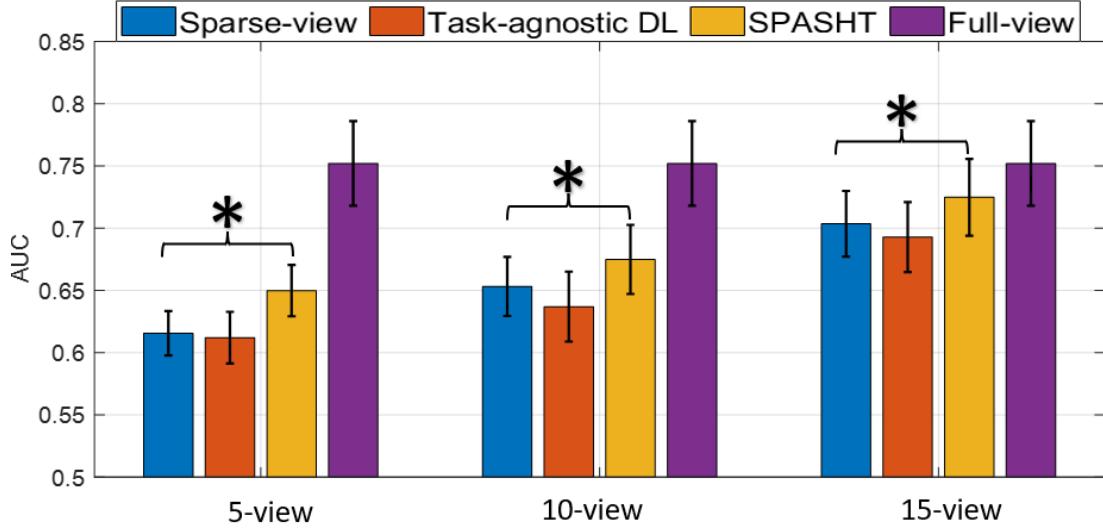


Figure 4: AUC values were obtained for the full-view and sparse-view protocol images, and the images were enhanced using the SPASHT method and task-agnostic DL method at various sparse-view levels with Channelized Hotelling Observer (CHO). Error bars denote 95% confidence intervals. Stars indicate p-values less than 0.05.

Fig.4, shows the AUC values obtained with the sparse-view protocol, task-agnostic DL,

and SPASHT methods at all the considered sparse-view angle levels, and with the full-view protocol. At all sparse-view levels, the SPASHT method significantly outperformed the sparse-view protocol (Delong's test p-value < 0.05).

Fig.5 shows the qualitative comparison of the different protocols at 10 sparse-view level setting. We observe in these cases that the task-agnostic DL protocol tends to wash out the defect in the predicted full-view images and that the sparse-view images have a high noise level. This observation is consistent with the findings reported in previous studies (17, 23). However, in the SPASHT method, some of the defect contrast in the estimated full-view image can be restored.

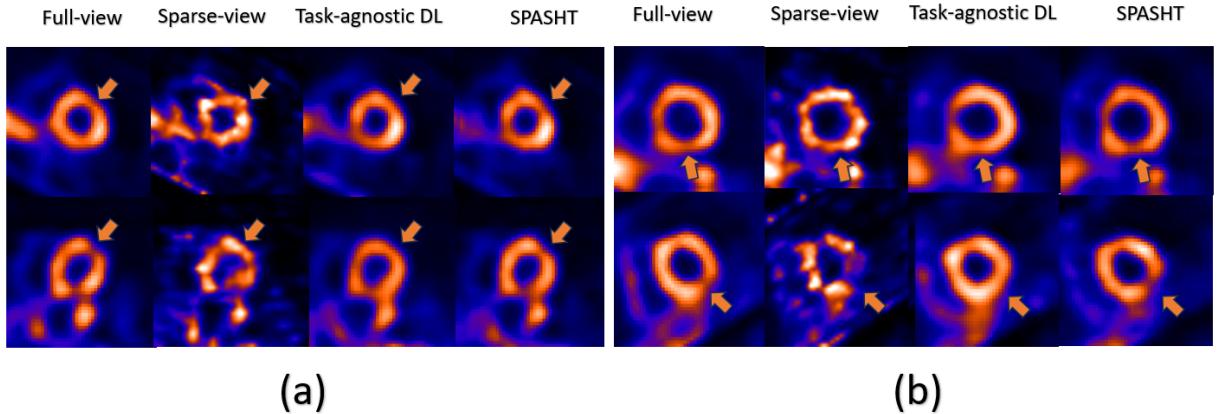


Figure 5: Four representative test cases qualitatively demonstrate the performance of the task-agnostic method compared to the SPASHT method. In each case, the sparse-view level was set to 10. In cases (a) and (b), defects were located in the anterior and inferior walls, respectively. For all four cases, the defects had an extent of 30 degrees and a severity of 25%.

Fig.6 demonstrates the AUC values for male and female populations, showing that the SPASHT method significantly improved detection-task performance at most sparse-view levels compared to the sparse-view protocol (with the exception of female patients at 5 sparse-view levels). Additionally, in 5 out of 6 configurations of 3 methods and 2 sex groups the SPASHT method showed significant performance gains over the task-agnostic method. In contrast, the task-agnostic DL method did not enhance (and in some cases reduced)

performance compared to the sparse-view protocol.

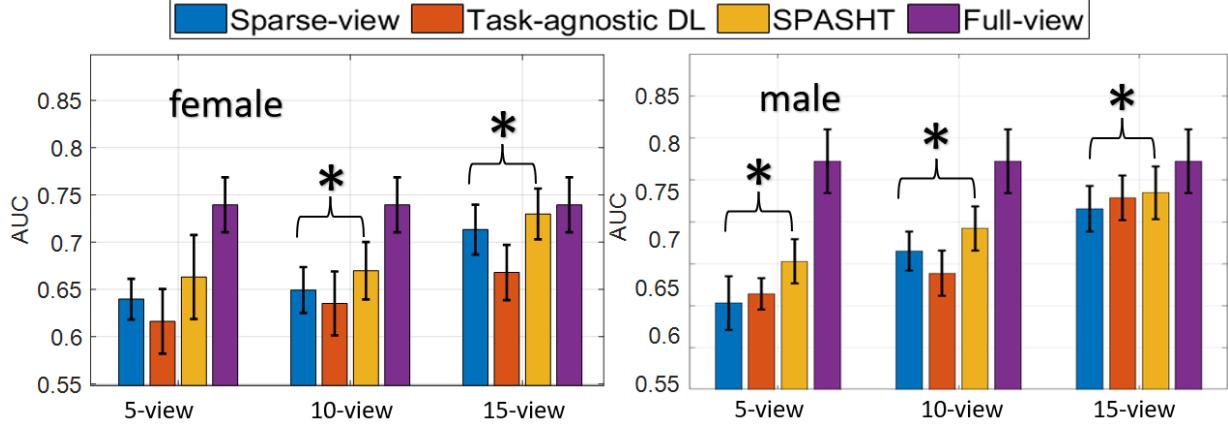


Figure 6: AUC values obtained for the different approaches and at various sparse-view levels with (a) male and (b) female patients using Channelized Hotelling Observer (CHO). Error bars denote 95% confidence intervals. Stars indicate p -values less than 0.05.

Fig.7 show the AUC values as a function of defect extent and severity at different sparse-view levels, respectively. We observe that at most sparse-view levels (except for 90-degree extent at 10 sparse-view level), the SPASHT method has improvement in AUC value which indicates improvement in the clinical task performance compared to the sparse-view protocol.

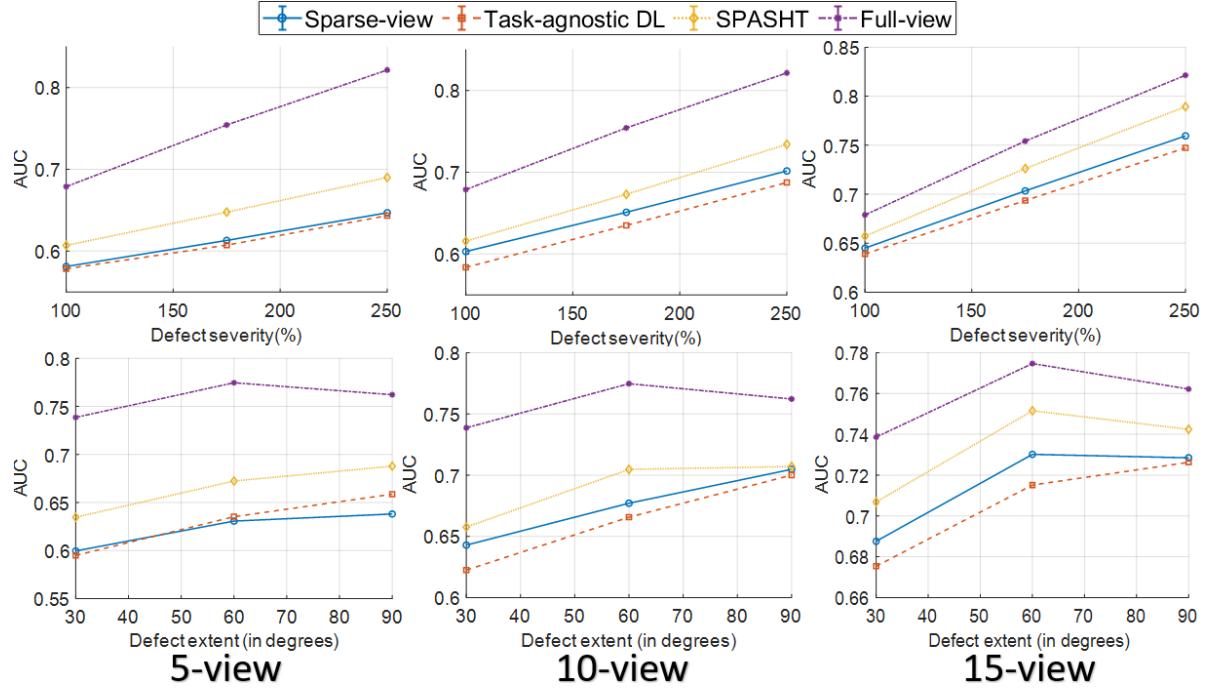


Figure 7: AUC values obtained using Channelized Hotelling Observer (CHO) for the various approaches as a function of different defect extents and defect severities.

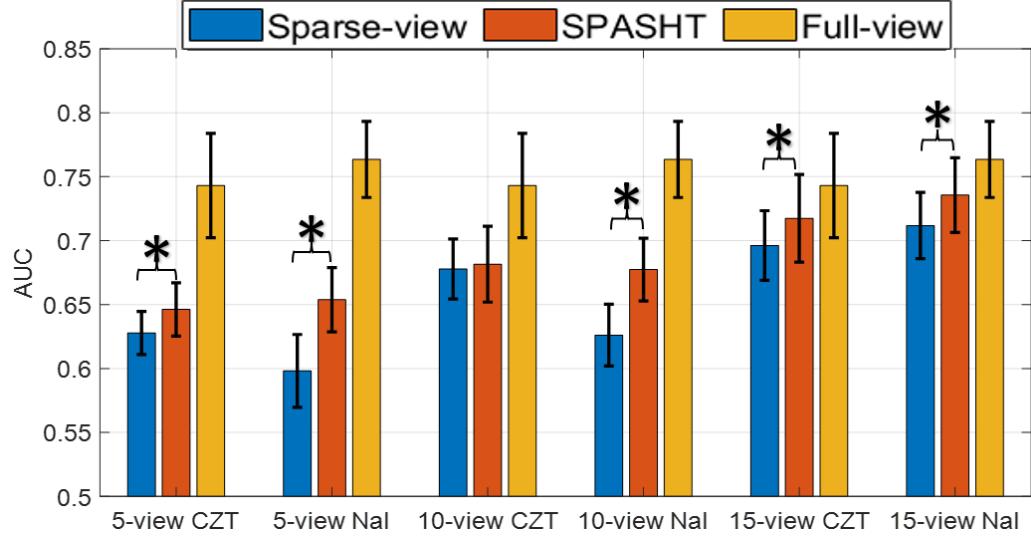


Figure 8: AUC obtained by sparse-view protocol, the SPASHT method, and full-view protocol across scanners and different sparse-view levels. Stars indicate p-values less than 0.05.

As referenced, above, our neural network was trained using a mixed dataset of CZT and

NaI detector data. To perform stratified analysis, we tested it separately on CZT-camera and NaI-camera datasets. Fig.8 shows results on the generalizability of the SPASHT method across two scanners. We observed that when trained on data mixed with the data from two cameras, the SPASHT method significantly outperformed the sparse-view protocol in most cases, with the exception of the 10 sparse-view level for CZT.

4.2 Evaluation based on fidelity-based figures of merit

The SSIM and RMSE metrics for the SPASHT method, the task-agnostic DL method, and the sparse-view protocol are shown in Fig.9. The SPASHT method yielded improved performance over the sparse-view protocol. Additionally, both the SPASHT method and the task-agnostic DL method generally produced very similar RMSE and SSIM values.

Sparse-view level	Metric	Sparse-view	Task-agnostic DL	SPASHT
5	RMSE	39.00	22.90	23.52
	SSIM	0.19	0.40	0.39
10	RMSE	33.49	21.48	20.18
	SSIM	0.30	0.48	0.50
15	RMSE	30.36	18.43	18.40
	SSIM	0.37	0.57	0.56

Figure 9: Root Mean Square Error (RMSE) and Structural Similarity Index (SSIM) for different methods at different sparse-view levels.

4.3 Evaluation based on the human observer study

The reading phase of the human observer study lasted 2 days, during which reading scores were collected from all readers. Individual AUC values and their standard deviation obtained by readers are shown in Fig. 10

Reader	Sparse-view AUC (SD)	SPASHT AUC (SD)	Full-view AUC (SD)
R1	0.56 (0.05)	0.68 (0.07)	0.72 (0.06)
R2	0.55 (0.07)	0.66 (0.07)	0.77 (0.06)
R3	0.57 (0.06)	0.72 (0.07)	0.85 (0.05)
R4	0.51 (0.07)	0.65 (0.07)	0.78 (0.05)

Figure 10: Individual AUC estimate and standard deviation obtained by each method for each reader.

Fig.11 shows the empirical ROC curves obtained by full-view protocol, the SPASHT method, and sparse-view protocol. We observed that the ROC curve obtained by the SPASHT method outperformed that of sparse-view protocol.

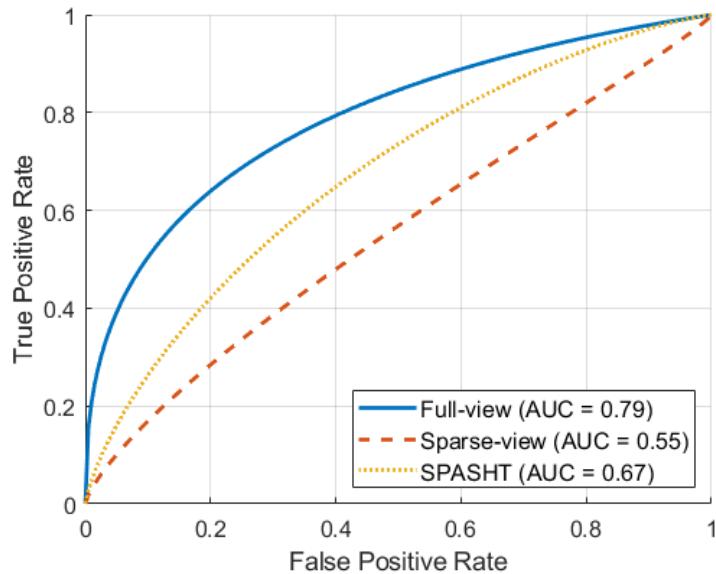


Figure 11: Empirical ROC curves obtained by full-view protocol, the SPASHT method, and sparse-view protocol.

5 Discussion

In this work, we proposed a method to enhance sparse-view MPI-SPECT images aiming to improve task-relevant defect detection task performance by incorporating a task-specific loss term. Results in Fig.4 demonstrate that this approach significantly improved defect-detection performance compared to using sparse-view images alone or enhancing with a standard task-agnostic DL method. Similar performance gains were observed across specific sub-groups based on sex and defect types, providing strong evidence that the task-specific loss term can enhance observer performance consistently across various defect characteristics. Fig.5 further illustrates the performance improvements achieved with the SPASHT method.

This study presents a task-specific method developed in the context of MPI SPECT. Meanwhile, the approach is also applicable to other medical imaging modalities. Potential applications include decreasing acquisition time in other medical imaging scenarios. While this method was designed for detection tasks, it may extend to other tasks reliant on extractable mathematical features from images.

Reducing the number of projection angles inherently leads to the loss of high-frequency features. Our results reveal that the SPASHT method attempts to address this challenge by utilizing the neural network’s capacity 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. This capability may enable SPASHT to mitigate the limitations traditionally associated with high-frequency feature loss in sparse-view images. In the work proposed by Snyder and Cox, the relationship between the number of projections and the maximum spatial frequency in reconstruction images is derived through an application of Fourier analysis and sampling theory (44). From Snyder and Cox’s conclusion, we can find that when the projection image has unlimited frequency bandwidth, the reconstruction image will have the maximum spatial frequency that is determined by the number of projections. The SPASHT method may be able to change the maximum spatial frequency. One area of future study can be inspecting the change of the maximum spatial

frequency of the enhanced images.

Another approach to reducing SPECT scanning time is to decrease the scanning time per projection angle while maintaining the total number of angles. This strategy preserves angular sampling density, thus retaining high frequency component. However, reducing the scanning time at each angle increases noise for each projection view. Another area for future research is a systematic comparison between reducing the number of projection angles and reducing the time spent per angle. Such a study would provide insights into the optimal scanning strategy, which may vary depending on the reconstruction algorithm, clinical application, and system hardware capabilities.

There are several limitations in this study. First, the SPASHT method was validated using anthropomorphic model observers and a limited number of non-physician human observers. However, the study results motivate evaluation of the method using large-sized physician-reader studies. Additionally, in our evaluation study, the defect location setting was limited to two locations. However, in clinical setting, the defect location is typically unknown and has more variations (45). Further human observer studies would be valuable for testing whether the SPASHT method improves performance in perfusion-defect detection when the defect location is unknown, offering insights into the method’s clinical translation potential.

Another limitation of this study is that the SPASHT method was trained on data in which defect-present cases were generated by synthetically inserting defects. While it would be ideal to validate the method using data with real perfusion defects, this approach was necessary due to the challenges associated with establishing reliable ground truth for both defect presence and centroid location in clinical images. Although clinical records can indicate the presence of defects, the accuracy of diagnosing and localizing these defects is hindered by inter and intra-reader variability. Training the network required prior knowledge of both defect presence and centroid location, which is not feasible with clinical images alone. Overcoming this limitation would require developing methods to annotate real defects

reliably in clinical data. Another approach to address this limitation is to train the network on synthetic-defect images, and fine-tune on a smaller number of clinical images where defect centroids are manually annotated, similar to a strategy proposed by Leung et al (46). Finally, the evaluation was conducted on data from a single center. The results encourage further testing across multiple centers to assess the method’s generalizability across different SPECT scanners and clinical settings.

6 Conclusion

We developed SPASHT, a task-specific DL method designed to enhance sparse-view MPI SPECT images by incorporating an anthropomorphic model observer term into the loss function. In retrospective evaluation, SPASHT yielded significant improvements in defect detection performance compared with sparse-view protocol, for different defect types, patient sexes, and scanner types. In addition, the proposed SPASHT method was observed to outperform the sparse-view protocol in the human observer study. As per the RELAINCE guidelines (20), we derive the following claim for our proposed SHASHT method:

A deep learning-based method to process sparse-view myocardial perfusion SPECT images provided improved performance on the task of detecting cardiac defects compared to sparse-view images. The method yielded statistical superiority to the sparse-view images on the task of detecting myocardial perfusion defects of AUC with significant level of 0.05 as evaluated in a retrospective clinical study with the anthropomorphic model observer and showed improved performance compared with the sparse-view images on the task of detecting myocardial perfusion defects of 0.12 higher in AUC value.

These studies provide evidence of the efficacy of the SPASHT method in improving the quality of sparse-view MPI SPECT images and motivate further validation of the method, including a large physician-involved human-observer study.

7 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] Barry L Zaret and George A Beller. Clinical Nuclear Cardiology: State of the Art and Future Directions E-Book. 2010.
- [2] Vikram Agarwal and E Gordon DePuey. Myocardial perfusion SPECT horizontal motion artifact. *Journal of Nuclear Cardiology*, 21(6):1260–1265, 2014.
- [3] James Fitzgerald and Peter G Danias. Effect of motion on cardiac SPECT imaging: recognition and motion correction. *Journal of Nuclear Cardiology*, 8(6):701–706, 2001.
- [4] Guido Germano, Piotr J Slomka, and Daniel S Berman. Attenuation correction in cardiac SPECT: the boy who cried wolf? *Journal of nuclear cardiology*, 14:25–35, 2007.
- [5] Craig D Stone, James W McCormick, David R Gilland, Kim L Greer, et al. Effect of registration errors between transmission and emission scans on a SPECT system using sequential scanning. *The Journal of Nuclear Medicine*, 39(2):365, 1998.
- [6] Julie M Nightingale, Fred J Murphy, and Christine Blakeley. I thought it was just an x-ray: a qualitative investigation of patient experiences in cardiac SPECT-CT imaging. *Nuclear medicine communications*, 33(3):246–254, 2012.
- [7] Xiongchao Chen, Bo Zhou, Huidong Xie, Tianshun Miao, Hui Liu, Wolfgang Holler, MingDe Lin, Edward J Miller, Richard E Carson, Albert J Sinusas, et al. DuDoSS: Deep-learning-based dual-domain sinogram synthesis from sparsely sampled projections of cardiac SPECT. *Medical physics*, 50(1):89–103, 2023.

- [8] Xiongchao Chen, Bo Zhou, Huidong Xie, Xueqi Guo, Qiong Liu, Albert J Sinusas, and Chi Liu. 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*, pages 12–22. Springer, 2023.
- [9] Isaac Shiri, Kiarash AmirMozafari Sabet, Hossein Arabi, Mozhgan Pourkeshavarz, Behnoosh Teimourian, Mohammad Reza Ay, and Habib Zaidi. Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks. *Journal of Nuclear Cardiology*, 28(6):2761–2779, 2021.
- [10] Mahsa Amirrashedi, Saeed Sarkar, Hossein Ghadiri, Pardis Ghafarian, Habib Zaidi, and Mohammad Reza Ay. A deep neural network to recover missing data in small animal pet imaging: comparison between sinogram-and image-domain implementations. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 1365–1368. IEEE, 2021.
- [11] Si Li, Wenquan Ye, and Fenghuan Li. LU-Net: combining LSTM and U-Net for sinogram synthesis in sparse-view SPECT reconstruction. *Math Biosci Eng*, 19(4):4320–40, 2022.
- [12] Huidong Xie, Weijie Gan, Wei Ji, Xiongchao Chen, Alaa Alashi, Stephanie L Thorn, Bo Zhou, Qiong Liu, Menghua Xia, Xueqi Guo, et al. A Generalizable 3D Diffusion Framework for Low-Dose and Few-View Cardiac SPECT. *arXiv preprint arXiv:2412.16573*, 2024.
- [13] Huidong Xie, Weijie Gan, Wei Ji, Xiongchao Chen, Alaa Alashi, Stephanie L Thorn, Bo Zhou, Qiong Liu, Menghua Xia, Xueqi Guo, et al. A generalizable diffusion framework for 3D low-dose and few-view cardiac SPECT imaging. *Medical Image Analysis*, page 103729, 2025.

- [14] Guang Li, Zhenhao Deng, Yongshuai Ge, and Shouhua Luo. HEAL: high-frequency enhanced and attention-guided learning network for sparse-view CT reconstruction. *Bioengineering*, 11(7):646, 2024.
- [15] Seyed Mehdi Ayyoubzadeh and Xiaolin Wu. High frequency detail accentuation in CNN image restoration. *IEEE Transactions on Image Processing*, 30:8836–8846, 2021.
- [16] Zitong Yu, Md Ashequr Rahman, Thomas Schindler, Robert Gropler, Richard Laforest, Richard Wahl, and Abhinav Jha. AI-based methods for nuclear-medicine imaging: Need for objective task-specific evaluation, 2020.
- [17] Zitong Yu, Md Ashequr Rahman, Richard Laforest, Thomas H Schindler, Robert J Gropler, Richard L Wahl, Barry A Siegel, and Abhinav K Jha. Need for objective task-based evaluation of deep learning-based denoising methods: A study in the context of myocardial perfusion SPECT. *Medical physics*, 2023.
- [18] Kaiyan Li, Weimin Zhou, Hua Li, and Mark A Anastasio. Assessing the impact of deep neural network-based image denoising on binary signal detection tasks. *IEEE transactions on medical imaging*, 40(9):2295–2305, 2021.
- [19] Andreu Badal, Kenny H Cha, Sarah E Divel, Christian G Graff, Rongping Zeng, and Aldo Badano. Virtual clinical trial for task-based evaluation of a deep learning synthetic mammography algorithm. In *Medical Imaging 2019: Physics of Medical Imaging*, volume 10948, pages 164–173. SPIE, 2019.
- [20] Abhinav K Jha, Tyler J Bradshaw, Irène Buvat, Mathieu Hatt, KC Prabhat, Chi Liu, Nancy F Obuchowski, Babak Saboury, Piotr J Slomka, John J Sunderland, et al. Nuclear medicine and artificial intelligence: best practices for evaluation (the RELAINCE guidelines). *Journal of Nuclear Medicine*, 63(9):1288–1299, 2022.
- [21] P Hendrik Pretorius, Junchi Liu, Kesava S Kalluri, Yulei Jiang, Jeffery A Leppo, Seth T Dahlberg, Janusz Kikut, Matthew W Parker, Friederike K Keating, Robert Licho, et al.

- Observer studies of image quality of denoising reduced-count cardiac single photon emission computed tomography myocardial perfusion imaging by three-dimensional Gaussian post-reconstruction filtering and deep learning. *Journal of Nuclear Cardiology*, pages 1–11, 2023.
- [22] Xiaohui Zhang, Varun A Kelkar, Jason Granstedt, Hua Li, and Mark A Anastasio. Impact of deep learning-based image super-resolution on binary signal detection. *Journal of Medical Imaging*, 8(6):065501–065501, 2021.
- [23] Gregory Ongie, Emil Y Sidky, Ingrid S Reiser, and Xiaochuan Pan. Optimizing model observer performance in learning-based CT reconstruction. In *Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*, volume 12035, pages 55–59. SPIE, 2022.
- [24] Kaiyan Li, Hua Li, and Mark A Anastasio. A task-informed model training method for deep neural network-based image denoising. In *Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*, volume 12035, pages 249–255. SPIE, 2022.
- [25] Minah Han, Hyunjung Shim, and Jongduk Baek. Low-dose CT denoising via convolutional neural network with an observer loss function. *Medical physics*, 48(10):5727–5742, 2021.
- [26] Md Ashequr Rahman, Zitong Yu, Richard Laforest, Craig K Abbey, Barry A Siegel, and Abhinav K Jha. DEMIST: A deep-learning-based detection-task-specific denoising approach for myocardial perfusion spect. *IEEE transactions on radiation and plasma medical sciences*, 8(4):439–450, 2024.
- [27] Eric C Frey, Karen L Gilland, and Benjamin MW Tsui. Application of task-based measures of image quality to optimization and evaluation of three-dimensional

- reconstruction-based compensation methods in myocardial perfusion SPECT. *IEEE transactions on medical imaging*, 21(9):1040–1050, 2002.
- [28] Harrison H Barrett and Kyle J Myers. *Foundations of image science*. John Wiley & Sons, 2003.
- [29] Kyle J Myers and Harrison H Barrett. Addition of a channel mechanism to the ideal-observer model. *JOSA A*, 4(12):2447–2457, 1987.
- [30] Sharlini Sankaran, Eric C Frey, Karen L Gilland, and Benjamin MW Tsui. Optimum compensation method and filter cutoff frequency in myocardial SPECT: a human observer study. *Journal of Nuclear Medicine*, 43(3):432–438, 2002.
- [31] Manoj V Narayanan, Michael A King, Jeffrey Leppo, S Dahlbert, P Hendrik Pretorius, and Howard C Gifford. Optimization of regularization of attenuation and scatter-corrected/^{99m}Tc cardiac SPECT studies for defect detection using hybrid images. *IEEE Transactions on Nuclear Science*, 48(3):785–789, 2001.
- [32] Thibaut Merlin, Simon Stute, Didier Benoit, Julien Bert, Thomas Carlier, Claude Comtat, Marina Filipovic, Frédéric Lamare, and Dimitris Visvikis. CASToR: a generic data organization and processing code framework for multi-modal and multi-dimensional tomographic reconstruction. *Physics in Medicine & Biology*, 63(18):185005, 2018.
- [33] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [34] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feed-forward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.

- [35] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.
- [36] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. Software available from tensorflow.org.
- [37] SD Wollenweber, BMW Tsui, DS Lalush, EC Frey, KJ LaCroix, and GT Gullberg. Comparison of Hotelling observer models and human observers in defect detection from myocardial SPECT imaging. *IEEE Transactions on Nuclear Science*, 46(6):2098–2103, 1999.
- [38] Albert Juan Ramon, Yongyi Yang, P Hendrik Pretorius, Karen L Johnson, Michael A King, and Miles N Wernick. Improving diagnostic accuracy in low-dose SPECT myocardial perfusion imaging with convolutional denoising networks. *IEEE transactions on medical imaging*, 39(9):2893–2903, 2020.
- [39] Xin He, E.C. Frey, J.M. Links, K.L. Gilland, W.P. Segars, and B.M.W. Tsui. A mathematical observer study for the evaluation and optimization of compensation methods for myocardial SPECT using a phantom population that realistically models patient variability. *IEEE Transactions on Nuclear Science*, 51(1):218–224, 2004.

- [40] E.C. Frey, K.L. Gilland, and B.M.W. Tsui. Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT. *IEEE Transactions on Medical Imaging*, 21(9):1040–1050, 2002.
- [41] Nancy A Obuchowski. Nonparametric analysis of clustered ROC curve data. *Biometrics*, pages 567–578, 1997.
- [42] Elizabeth R DeLong, David M DeLong, and Daniel L Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, pages 837–845, 1988.
- [43] James A Hanley and Barbara J McNeil. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, 148(3):839–843, 1983.
- [44] DL Snyder and JR Cox. An overview of reconstructive tomography and limitations imposed by a finite number of projections. *Reconstruction tomography in diagnostic radiology and nuclear medicine*, 28, 1977.
- [45] Janice Tania, Nu ri Choi, Wenxuan Xue, Zezhang Yang, Ashwin singh Parihar, Farhan Katchi, Barry Siegel, and Abhinav Jha. An LLM-based Approach to Derive Population Statistics of Myocardial Perfusion Defects from MPI SPECT Clinical Reports, 2025.
- [46] Kevin H Leung, Wael Marashdeh, Rick Wray, Saeed Ashrafinia, Martin G Pomper, Arman Rahmim, and Abhinav K Jha. A physics-guided modular deep-learning based automated framework for tumor segmentation in PET. *Physics in Medicine & Biology*, 65(24):245032, 2020.

Supplemental section

Evaluation based on the eigenanalysis-based treatment

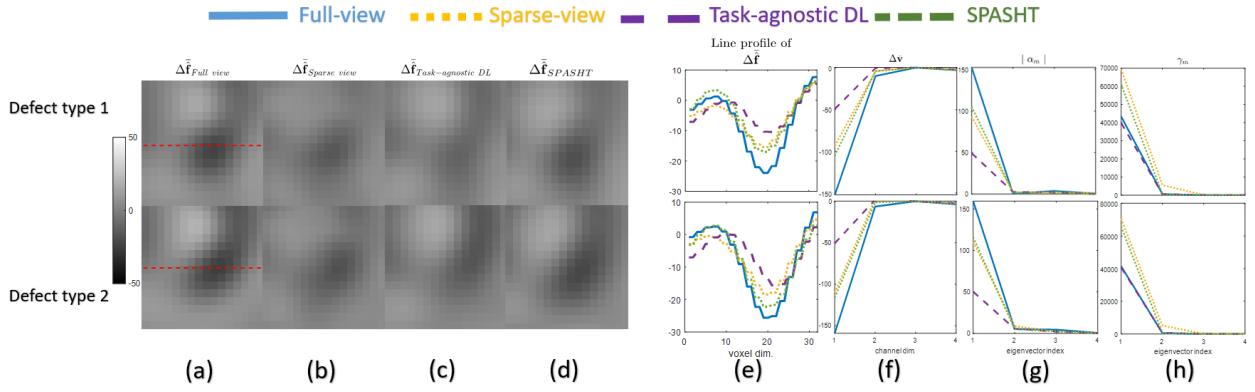


Figure 12: Mean difference reconstructed short-axis image between defect-absent and defect-present hypotheses for (a) full-view, (b) sparse-view, (c) task-agnostic DL, and (d) SPASHT method. The images in (a-d) are windowed to a region centered to the defect centroid. (e) Line profile of the mean difference reconstructed image of (a-d). (f) Mean difference channel vector $\Delta\bar{v}$ between defect-absent and defect-present hypotheses for various approaches. (g) Absolute value of coefficient α_m and (h) eigenvalue spectra of the noise covariance matrix. The sparse-view level was set to 5. (Defect type 1: 30 degrees extent, 25% severity, and anterior wall defect. Defect type 2: 60 degrees extent, 25% severity, and inferior wall defect.)

To quantitatively interpret the observed performance improvements, we conducted an analysis similar to that by Yu et al. (17). Specifically, we examined how sparse-view enhancement affects the first and second-order statistics of the channel vector within the test set, comparing both the SPASHT method and the task-agnostic DL method. This analysis was carried out separately for each defect type. Let $\Delta\bar{v}$ represent the mean difference channel vector between defect-present and defect-absent cases, and \mathbf{K}_v denote the channel-vector covariance matrix. The signal-to-noise ratio (SNR) for the CHO is then given by:

$$SNR^2 = \Delta\bar{v}^T \mathbf{K}_v^{-1} \Delta\bar{v} \quad (7)$$

If the test statistics for defect-present and defect-absent cases follow a normal distribution, the AUC and SNR of the observer are monotonically related (28), making observer SNR

analysis a valuable indicator of detection-task performance. Assume the reconstructed images are reoriented and windowed with the defect centroid centered. Let $\Delta\hat{\mathbf{f}}$ represent the mean difference reconstructed image between defect-present and defect-absent cases. Then, $\Delta\bar{\mathbf{v}} = U^T \Delta\hat{\mathbf{f}}$. According to Eq. 9, both the mean difference of the channel vector $\Delta\bar{\mathbf{v}}$ and the covariance matrix K_v influence observer performance. Eigenanalysis of the covariance matrix allows for a combined assessment of these factors on the observer SNR. Let \mathbf{u}_m and γ_m denote the m^{th} eigenvector and eigenvalue of K_v , respectively. We can express $\Delta\bar{\mathbf{v}}$ in terms of these eigenvectors as follows:

$$\Delta\bar{\mathbf{v}} = \sum_{m=1}^C \alpha_m \mathbf{u}_m \quad (8)$$

where the coefficient $\alpha_m = U_m^T \Delta\bar{\mathbf{v}}$. Thus, the SNR of the CHO is given by:

$$SNR^2 = \sum_{m=1}^C \frac{\alpha_m^2}{\gamma_m} \quad (9)$$

Thus, assessing the impact of enhancement on α_m and γ_m provides an interpretable approach to evaluate the effect of enhancement on observer performance, as shown in Fig.12, focusing on two defect types using the 5 sparse-view level. Initially, we plotted the mean difference between reconstructed images for defect-present and defect-absent cases. The neural network recovered some defect features, as indicated by the more intense defect signal at the center of the mean difference images compared to those of the sparse-view protocol. Both the task-agnostic DL and SPASHT method protocols demonstrated this capability. In Fig.12 (a)-(e), analyzing the mean difference channel vector between defect-present and defect-absent cases, we observed that the SPASHT method preserved and restored part of the mean difference originally present in the full-view images for these two defect types. In Fig.12 (f)-(h), the SPASHT method protocol showed an increase in the values of α_m compared to the sparse view angle protocol, leading to an overall improvement in defect detection performance. Simultaneously, γ_m decreased for both the SPASHT method and the

task-agnostic DL method relative to the sparse-view protocol, which is expected to positively influence observer performance.