EJNMMI Physics

ORIGINAL RESEARCH Open Access

SPECT-MPI iterative denoising during the reconstruction process using a two-phase learned convolutional neural network

Farnaz Yousefzadeh[1], Mehran Yazdi[2*], Seyed Mohammad Entezarmahdi[3], Reza Faghihi[4], Sadegh Ghasempoor[5], Negar Shahamiri, Zahra Abuee Mehrizi[3] and Mahdi Haghighatafshar[3]

*Correspondence:
yazdi@shirazu.ac.ir

[1] Department of Computer Science and Engineering and IT, School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran
[2] School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran
[3] Department of Nuclear Medicine, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran
[4] Department of Nuclear Engineering, Shiraz University, Shiraz, Iran
[5] Department of Nuclear Medicine, Alzahra Hospital, Shiraz University of Medical Sciences, Shiraz, Iran

Abstract

Purpose: The problem of image denoising in single-photon emission computed tomography (SPECT) myocardial perfusion imaging (MPI) is a fundamental challenge. Although various image processing techniques have been presented, they may degrade the contrast of denoised images. The proposed idea in this study is to use a deep neural network as the denoising procedure during the iterative reconstruction process rather than the post-reconstruction phase. This method could decrease the background coefficient of variation (COV_bkg) of the final reconstructed image, which represents the amount of random noise, while improving the contrast-to-noise ratio (CNR).

Methods: In this study, a generative adversarial network is used, where its generator is trained by a two-phase approach. In the first phase, the network is trained by a confined image region around the heart in transverse view. The second phase improves the network's generalization by tuning the network weights with the full image size as the input. The network was trained and tested by a dataset of 247 patients who underwent two immediate serially high- and low-noise SPECT-MPI.

Results: Quantitative results show that compared to post-reconstruction low pass filtering and post-reconstruction deep denoising methods, our proposed method can decline the COV_bkg of the images by up to 10.28% and 12.52% and enhance the CNR by up to 54.54% and 45.82%, respectively.

Conclusion: The iterative deep denoising method outperforms 2D low-pass Gaussian filtering with an 8.4-mm FWHM and post-reconstruction deep denoising approaches.

Keywords: SPECT-MPI, Image denoising, Iterative image reconstruction, Convolutional neural networks, Deep learning

Introduction

Single-photon emission computed tomography (SPECT) myocardial perfusion imaging (MPI) reveals information concerning the uptake of a radioactive compound within the heart in three dimensions. It is considered the most frequently performed procedure in nuclear cardiology [1]. While the American society of nuclear cardiology (ASNC) pointed out that decreasing radiation exposure is essential [2], dose reduction degrades image quality due to the inherent noise of these images [3].

Reducing the dose of radiopharmaceuticals or shortening the acquisition time can increase the inherent noise. However, iterative image reconstruction methods, which are statistical approaches, can decrease the noise to some degree and preserve the accuracy of diagnoses. It has been shown that the injected dose or acquisition time of SPECT-MPI can be reduced by a factor of two or higher via advanced iterative image reconstruction algorithms [4–6]. To this end, Ramon et al. investigated the accuracy of perfusion-defect detection in three reconstruction strategies and showed that the administrated dose could be reduced without sacrificing diagnostic performance [4].

Conventional approaches use filters in the spatial or frequency domain to suppress the noise [7]. One example of these methods is non-local mean filtering which Arabi et al. presented in order to perform the denoising process [8]. Moreover, in another study, the authors proposed a new hybrid dual-domain filtering approach, combining domain and frequency filtering methods to improve signal-to-noise ratio (SNR) and quantitative accuracy [7].

In addition to the aforementioned methods, deep learning (DL) algorithms have shown promising results in decreasing noise [9]. Various types of neural networks have been used in different studies to suppress noise in SPECT. For instance, Ramon et al. investigated two types of neural networks, i.e., convolutional neural network (CNN) and auto-encoder network (AE), to reduce the inherent noise of SPECT images and compared the results. They showed that the DL approach could allow for further dose reduction compared to conventional methods, especially when skip connections are used [10]. In another study, Song et al. used a CNN. They showed that marked improvement can be obtained both in the noise level of the reconstructed myocardium and in the spatial resolution of the left ventricular (LV) wall [11]. Aghakhan Olia et al. used a generative adversarial network (GAN) based on the U-Net structure and demonstrated that it can effectively recover the underlying information in 1/2-dose and 1/4-dose SPECT images [12]. Hashimoto et al. discussed multiple studies demonstrating that some researchers attempt to use deep networks to directly reconstruct images from sinograms, aiming to produce quick and accurate results. However, this approach may generate artifacts or false structures in the reconstructed images [13]. Consequently, other researchers prefer to focus on noise reduction using traditional reconstruction methods such as OSEM, which is the focus of this study.

All the above mentioned denoising methods used filters and neural networks as post-processing tools to suppress random noise in SPECT. However, Mustafovic et al. demonstrated that by using inter-iteration filtering, reconstructed images can achieve nearly object-independent and uniform resolution [14], which formed the basis of our study. In this paper, we propose inter-iteration filtering that uses a GAN as the denoising network after each iteration of the ordered subset expectation maximization (OSEM) algorithm, an iterative image

reconstruction process, to reduce the inherent random noise of SPECT images while keeping their contrast. The GAN is trained using image pairs obtained from low-counts (input) and high-counts (label) acquisitions of the same patients. Instead of training the network in one phase by feeding the image pairs, we innovatively considered two training phases to enhance the efficiency of the network. The proposed method is evaluated using real data.

Methods and materials

Problem formulation

In SPECT reconstruction, the measured data in the *j*th projection element, p_j, can be formulated as below:

$$p_j \approx \sum_i M_{i,j} q_i \tag{1}$$

where q_i is the intensity (or activity concentration) within the *i*th pixel in the image, and $M_{i,j}$ is the probability that emitted photon from the *ith* pixel will be detected by the *j*th projection element. M is known as detection probability matrix or system matrix which is very large. $M \in \mathbb{R}^{l \times k}$ where l equals the full set of pixel elements ($l = m \times n \mid m, n$ are the image dimensions) and k equals all the projection elements ($k = d \times \alpha \mid d, \alpha$ are the detector elements and projection angles respectively). The objective is to find the maximum likelihood estimation of the image $\widehat{Q} \in \mathbb{R}^{m \times n}$, which can be written as:

$$\widehat{Q} = \underset{Q \geq 0}{argmax} L(P|Q) \tag{2}$$

where L represents likelihood estimation function, $P \in \mathbb{R}^{d \times \theta}$ is the projection count distribution that was measured from an estimated intensity distribution in the image space, labelled Q.

Assuming the measured data to be Poisson distributed, by using iterative reconstruction to find the maximum likelihood estimation problem, function (2) is rewritten as below [15]:

$$\widehat{q}_i^{n+1} = \frac{\widehat{q}_i^n}{\sum_j M_{i,j}} \times \sum_j M_{i,j} \times \frac{p_j}{\left(\sum_t M_{t,j} \times \widehat{q}_t^n\right)} \tag{3}$$

where, n is the iteration number, the *ith* pixel element of $\widehat{Q}$ image after n iterations, $\widehat{q}_i^n$, is an estimation of q_i, and t is the pixel index similar to i (since the denominator must be computed before the rest of formula, t is indexed instead of i to avoid confusion).

After n iterations, each estimated pixel value $\widehat{q}_i^n$ is related to the real pixel value q_i, as follows:

$$\widehat{q}_i^n = q_i + \varepsilon_i^n \tag{4}$$

where ε_i represents the error magnitude after n iterations, which is influenced by both deterministic and random noise. It is anticipated that in the absence of random noise, by accounting for all sources of deterministic noise and patient-induced attenuation and scattering in the system matrix M, ε_i^n is significantly reduced as the number of iterations increases [16]. However, increasing the number of reconstruction iterations introduces higher frequency components into the reconstructed image [16]. In practical scenarios,

random noise contains high-frequency components. Consequently, increasing the number of iterations results in higher noise levels in the reconstructed image [16]. One of the simplest strategies to suppress both deterministic and random noise is to increase the number of iterations while applying a random noise-controlling filter, such as a low-pass filter, after the reconstruction [16]. The corresponding equation can be expressed as follows:

$$\Theta = f(\widehat{Q}) \tag{5}$$

where $f : \mathbb{R}^{n\times m} \rightarrow \mathbb{R}^{n\times m}$ ($n \times m$ representing the image dimension) is a filtering function to regulate the noise in the estimated image $\widehat{Q}$. $\Theta \in \mathbb{R}^{m\times n}$ is a noise-suppressed estimation of the real image. Conventionally, f is modeled by a low-pass filter to reduce noise. However, the primary drawback of using simple low-pass filters is that they blur the image, suppressing image edges and altering contrast values during filtering. Recent studies have demonstrated that deep neural networks provide superior noise filtering performance while retaining spatial resolution and quantitative accuracy compared to classical filtering methods [13].

If the f function is used after completing the iterations [17], it is known as the post-reconstruction filtering/denoising method. In our work, we propose to apply the f function after each iteration. In this manner, the number of times when f is applied equals (or less than) the number of iterations. This idea can be formulated as follows:

$$\widehat{q}_i^{n+1} = \frac{\theta_i^n}{\sum_j M_{i,j}} \times \sum_j M_{i,j} \times \frac{p_j}{\left(\sum_t M_{t,j} \times \theta_t^n\right)} \tag{6}$$

where, θ^n is the *ith* pixel element of the Θ^n and $\Theta^n = f\left(\widehat{Q}^n\right)$.

Here, we have applied the proposed f function two times lesser than the number of iterations. (f is not applied in the first two iterations). We have also explored the effects of using specific f for each iteration. We have proposed a GAN to estimate the f as a denoising function in this work. To do that, the GAN has been trained with a pair of high- and low-noise images in three training approaches. Moreover, to reduce the number of iterations, the OSEM [18] reconstruction method is used. In this way after performing all of the sub-iterations and finishing an iteration, f is applied.

Network architecture

Our implemented network, which we call a two-phase learned convolutional neural network (TPL-CNN), is based on a GAN consisting of two parts; a generator and a discriminator. After exploring various architectures, the CNN network model was proposed for both the generator and the discriminator, as shown in Fig. 1.

In the generator, two convolutional and two deconvolutional layers were used. Ten 3×3 convolutional kernels with no padding and a stride of one were utilized throughout all layers. Each convolutional layer was followed by a batch normalization layer (BN) and a leaky rectifier linear unit (ReLU) layer, with a slope of 0.3, as the activation function. Except for the last convolution layer, ReLU was used for the activation function. We also considered including skip connections, which directly concatenate the values from an input layer to the values at the output layer having the same dimension. These skip

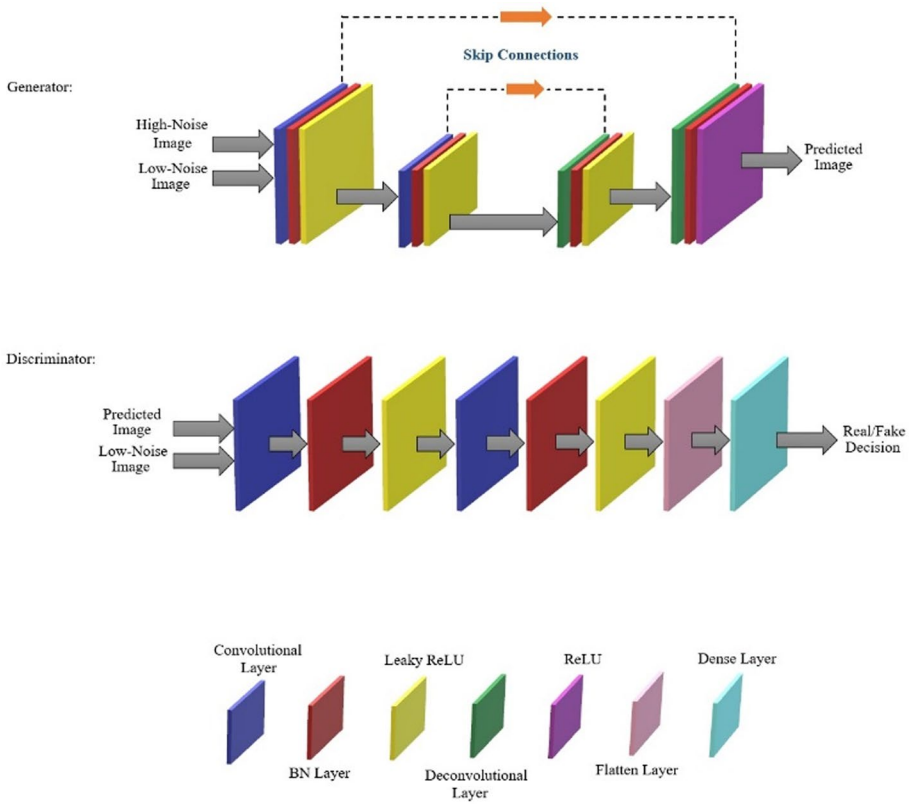

Fig. 1 Architecture of the generator and discriminator networks in the GAN model of the TPL-CNN

connections were introduced to address the potential risk of sacrificing essential image details when increasing the number of encodings. They facilitate a parallel pathway for conveying image features from shallower encoding layers to deeper decoding layers, effectively resolving this concern [19].

The discriminator comprised two convolutional layers with ten kernel sizes of 3×3. Each convolution layer was connected to a BN, and a leaky ReLU with a slope of 0.3 was used as the activation function. A flatten layer was used to convert the spatial dimensions of the previous layers into a single one-dimensional vector that can be used as input for the last layer, which is a dense layer.

The information about the architecture of the applied network has been represented in appendix with the caption of "TPL-CNN Characteristics."

Network implementation

In SPECT-MPI, the myocardium is of interest, which is much smaller than the entire reconstructed image, and the vast majority of the image pixels correspond to background regions (or other organs) outside the heart volume, which reduces the accuracy of a network. Hence, the network can be trained only with this ROI to improve its accuracy [10, 12]. On the other hand, the network must see the entire reconstructed images to improve its generalization. Therefore, the network is trained by a two-phase algorithm presented at the SNMMI 2022 conference [20].

In the first phase of this algorithm, the aforementioned network is trained with reconstructed images from high-noise and low-noise data that contain only the cardiac region as the input and output of the network. These smaller image pairs, each of size 32×24 pixels, are automatically prepared using QCard-NM package as proposed in [21, 22]. This package identifies a fixed-size rectangle encompassing the cardiac region. The ROI is then presented to the user. If the ROI does not properly surround the left ventricle, the user has the option to manually adjust its position.

Subsequently, the cropped images are rotated with angles of $+20°$ and $-20°$ and translated along the x and y axes in both positive and negative directions using Python libraries. This expands the image set to seven times its original amount. The image order is then randomly shuffled before being fed as input to the neural network.

When the random initial weights of the first network are updated, they serve as the initial weights of the second network in a process called transfer learning. The second network has the same architecture as the first network, except for the input and output layer sizes, which remain the same as the original reconstructed image sizes, i.e., 64×64 pixels. In transfer learning, the size of the filters (kernels) and weights within the convolutional layers generally do not need to be altered if the input size changes. This process is illustrated in Fig. 2.

In our implementation, we first applied one iteration of OSEM with four sub-iterations and used it as the input for the iterative denoising process. The iterative denoising process tries to reduce the noise of the input image in four iterations. Hereby, the input data is denoised in each iteration of the iterative denoising process by the TPL-CNN before being reconstructed for one more iteration with OSEM. The overall algorithm flowchart is presented in Fig. 3.

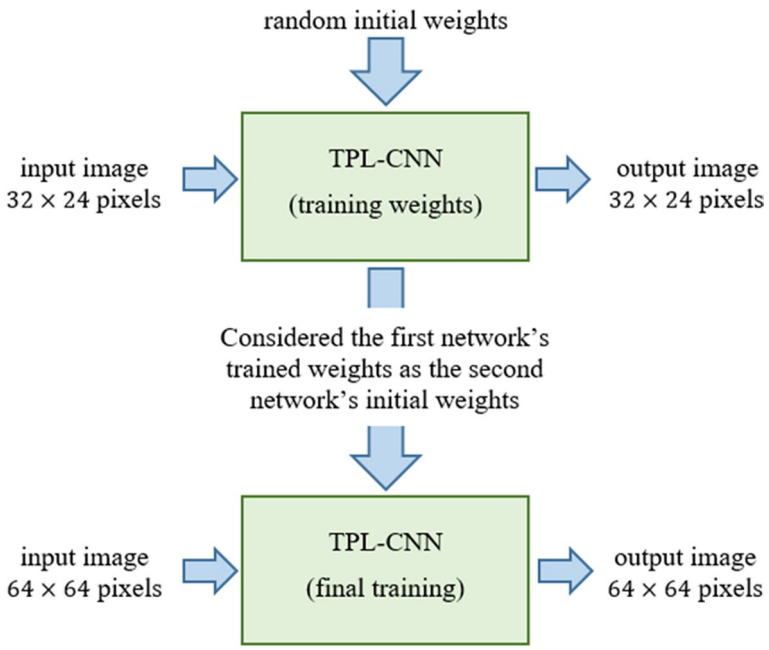

Fig. 2 Flowchart of the two-phase learning process of the TPL-CNN training

All image reconstruction methods in our work were done in fully two-dimensional mode. The neural network was implemented using the Keras deep learning framework 2.8.0 based on the TensorFlow libraries 2.8.2 in Python 3.9. All the experiments were conducted on NVIDIA GeForce GTX 1050 with a 4GB memory graphical processing unit. During training, the generator is updated via a weighted sum of both adversarial and L2-norm loss. The trainable parameters were updated to minimize the L2-norm loss calculated between the predicted images and the reference input images. The binary cross entropy loss function was used to train the discriminator model. Adaptive moment estimation (Adam) optimizer with a learning rate of $1 \times e^{-4}$ was used to minimize the loss functions.

Dataset

In total, 247 patients were referred for MPI, and two immediate serially high and low-noise SPECT data were acquired as the input and ground truth images. For each scan, 32 projections with matrix size of 64×64 were acquired and the data is reconstructed using OSEM algorithm for five iterations with four sub-iterations. The data were divided into three subsets: 1) one with 149 patients used for training, 2) one with 49 patients used as the validation set, and 3) the other 49 patients were used for performance evaluation (denoted as the test set). Patients were unmoved between acquisitions. Cases were randomly divided into training, validation, and evaluation sets. Acquired high and ground truth low-noise image-pairs were obtained from 10- and 30-s acquisition time per projection, respectively.

Experimental setup

Due to the device's inability to capture the list mode, we were compelled to conduct the high-noise scan in a serial manner. The experimental setup was meticulously designed, involving the standard scan (low-noise), which took 10 min, followed by the fast scan (high-noise), which took 3 min. Patients remained unmoved on the bed for 13 min after an hour of injection, during which the changes in activity distribution and decay were

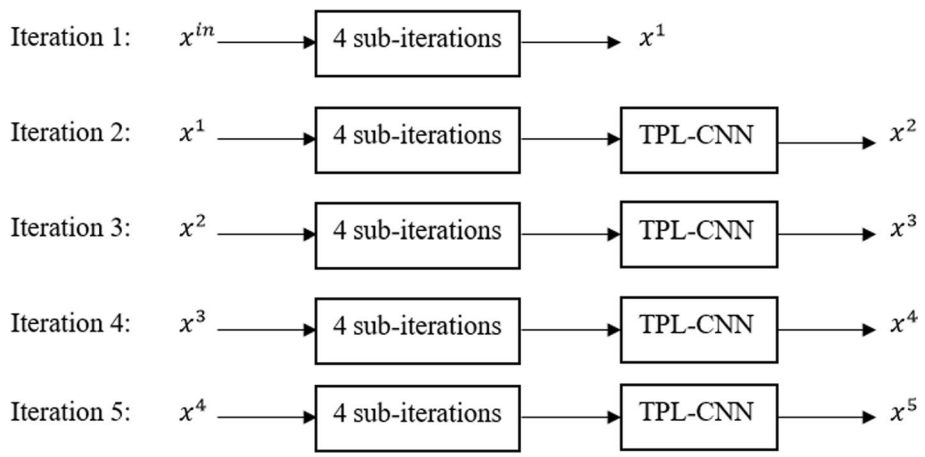

Fig. 3 Flowchart of the iterative denoising algorithm. OSEM is used as the reconstruction tool. Each iteration of the OSEM reconstruction method consists of four sub-iterations

deemed insignificant based on the assumptions of the myocardial scan. Tc-99 m sestamibi 2-days protocol was used to perform the rest/stress scan. The used dataset contains both the rest and stress phases. The administered doses were based on the weight of the patients (8 MBq/kg) [22]. Rest acquisition started 60 min after the tracer was injected. Pharmacological stress was induced by an infusion of dipyridamole, and 3 min after the dipyridamole slow infusion, the radiotracer was injected. Stress acquisition started 45–90 min after the injection.

The SPECT studies were acquired with a dual-head detector camera, with low-energy, high-resolution collimators, a 20% symmetrical window at 140 keV, and a 64×64 matrix size. For each scan, 32 projections with 10- and 30-s per projection were acquired.

CorCam Gamma Camera System (DDD-Diagnostic, Denmark) was used to acquire the images. No attenuation, scatter, and detector response corrections were applied to maintain the generality of the proposed approach.

Parameter selection

Parameter selection in neural networks involves determining various parameters to achieve optimal denoising effects. The learning rate is considered the most important hyperparameter in training neural networks that network converges and has been set to 0.001 in our experiments. We used Adam, an adaptive gradient descent optimizer, which generally works well with a fixed learning rate. Moreover, we used normalization, allowing good results even with exotic learning rates. Therefore, we set the batch size to 32 to prevent occupying a large amount of memory, which can slow down the training process. As the number of layers or filters increased, the computation time also increased, but with a slight improvement in the output. Therefore, setting fewer layers and filters was reasonable for finding the optimal balance between computational performance and cost, which was four and 32 for layers and filters, respectively. To select the epoch number, we used a validation set given to the network after each epoch, and the PSNR, RMSE, and SSIM metrics values were measured in each epoch. An acceptable epoch was selected when there was no significant improvement in the value of the mentioned metrics, and they were oscillating between the same values. The results are presented in Table S-1 in the Supplementary Information section.

TPL-CNN training approaches

We considered three different approaches to train the TPL-CNN. The neural network's architecture in all approaches is the same, though the input fed to the neural network to be trained is different.

Solo training

This method trains the TPL-CNN using reconstructed data images that have undergone two iterations of the OSEM, including both input and ground truth counterparts. The same neural network is used throughout the denoising process, as the algorithm calls it. This means that the function f in Eq. (5) remains constant throughout the whole process.

Mélange training

This method uses reconstructed image pairs acquired from different stages of the OSEM algorithm to train the TPL-CNN. The training data for this network is a mixture of image pairs acquired from two, three, four, and five iterations of the OSEM. However, throughout the entire process of denoising, this trained TPL_CNN network is consistently utilized, as stated by the algorithm. In a manner similar to the solo training approach, the function f in Eq. (5) remains constant throughout the entire procedure.

Distinctive training

This method employs several neural networks with identical architectures but distinct training data. Depending on the stage of the OSEM reconstruction, a neural network is explicitly trained. For instance, the neural network inputs consist of image pairs reconstructed using two iterations of OSEM for the second iteration of OSEM reconstruction, and a new network is trained using image pairs reconstructed using three iterations of OSEM for the third iteration of OSEM reconstruction. The remaining OSEM iteration stages also follow this same pattern. Accordingly, in contrast to two previous approaches, the function f in Eq. (5) is changed after each OSEM iteration.

Evaluation

We compared the proposed method with the conventional reconstruction method (OSEM), Gaussian post filtering method, iterative Gaussian filtering method, and CNN denoising method. All of the Gaussian filters used in this study are in two dimensions and were created using the MATLAB library named "imgaussfilt," which is mainly defined for two-dimensional objects like images.

Conventional reconstruction method

This method is the conventional application of reconstruction done with the OSEM algorithm on our acquired high- and low-noise data without any filtration.

Gaussian post filtering

This method uses a Gaussian filter directly to the reconstructed images. After images are completely reconstructed using OSEM, a Gaussian filter with Full width at half maximum (FWHM) of 8.4 mm, according to the device manufacturer's suggestion, is applied as the post-reconstruction filter to reduce their noise level.

Iterative gaussian filtering

In this method, a Gaussian filter is used as the smoothing filter though it is not used as the post-processing filter. This method is made up of repetitive application of reconstructing and filtering. In this way, after each step of the OSEM reconstruction algorithm, a Gaussian filter with FWHM of 8.4 mm is used to smooth the noise of the reconstructed image, and then the next step of OSEM is applied to this smoothed

image. This means that, like our proposed method, we use Eq. (5) to reconstruct the image. However, in this case, the function f is a Gaussian filter instead of TPL_CNN.

CNN denoising

This method uses a GAN with a similar architecture as the TPL-CNN as a post-processing noise reduction tool instead of conventional filters. To do this, first, a GAN is needed to be trained in a way to learn to reduce the noise level of the given images. Then, this trained GAN denoising filter is used after the OSEM reconstruction algorithm as the post filter, and the output images of the OSEM algorithm, which is completely done for five iterations, are fed to it to be filtered. High- and low-noise image pairs reconstructed with five iterations and four sub-iterations of the OSEM are fed as the input to train this denoising network. This network is also trained in two phases. First, images with the size of 32×24 that only contain cardiac regions are used to train the first network, and then its trained weights are used as the initial weights of the second network, which architecture is the same as the first one except for the input layer size that receives 64×64 image size. This network is the same as the network used in the Distinctive Training method for image pairs with five iterations of the OSEM. Therefore, the trainable parameters are updated using the exact loss function and optimizer with the same learning rate. Hence, the learned parameters are in their optimal values.

Qualitative and quantitative evaluations

Qualitative and quantitative evaluations of the proposed framework were performed on 49 subjects, referred to as the test dataset. Three SPECT images (acquired from solo, mélange, and distinctive training of the TPL-CNN) were compared to reconstructed reference images with the OSEM algorithm (five iterations with four sub-iterations) derived from high-noise projections. The performance of the TPL-CNN was evaluated through five metrics including root mean squared error (RMSE), the coefficient of variation in the background (COV_bkg), structural similarity index metrics (SSIM), contrast-to-noise ratio (CNR) and edge preservation index (EPI) all presented in Eqs. (7)–(11) respectively. Moreover, these metrics were also calculated for the high noise images to provide a baseline for the performance assessment of methods. After selecting the appropriate training method for TPL-CNN, its SNR, and contrast were assessed and compared with those of the other methods.

RMSE

RMSE measures the differences between the predicted and the observed values using Eq. (7).

$$RMSE = \sqrt{\frac{1}{w}\sum_{g=1}^{w}\left(I_u(\mathrm{g}) - I_v(\mathrm{g})\right)^2} \tag{7}$$

where I_v is the observed low-noise reconstructed image, I_u is the predicted image, w is the total number of pixels and g is the intended pixel.

Cov_bkg

Coefficient of variation (COV) measures the relative variability or dispersion of data around the mean in a sample or population. When computed over a uniform region of interest (ROI), it represents the amount of random noise. Therefore, the COV of the same ROIs (lung region) in the background of the estimated outputs and their low-noise counterparts was considered and calculated using Eq. (8):

$$COV\_bkg = \frac{\sigma_{background(ROI)}}{S_{background(ROI)}} \tag{8}$$

where S and σ represent the mean and standard deviation for ROI, respectively.

CNR

CNR is also used to assess image quality. However, CNR can consider a significant bias in the image by subtracting one term before calculating the ratio. As a result, an image may have a high SNR but a low CNR [23]. In this order, in one of the slices, the location of the left ventricle and its hole (cavity) are once selected manually and saved as the ROI and the background region, respectively. The slice number for each patient is also saved. Then, these saved locations of the specific slices are used to calculate CNR using Eq. (9), [23].

$$CNR = \frac{\left|S_{\text{myocardium (ROI)}} - S_{\text{cavity}}\right|}{\sqrt{\sigma_{\text{myocardium (ROI)}} + \sigma_{\text{cavity}}}} \tag{9}$$

where S and σ represent the mean and standard deviation for ROI, respectively.

SSIM

SSIM is a model that considers image degradation as perceived change in structural information, while also incorporating important perceptual phenomena, including both luminance masking and contrast masking terms. It is used to measure the similarity between two images and is defined as in Eq. (10).

$$\text{SSIM}(\text{I}_\text{u}, \text{I}_\text{v}) = \frac{\left(2\mu_{\text{I}_\text{u}}\mu_{\text{I}_\text{v}} + \text{c}_1\right)\left(2\delta_{\text{I}_\text{u}\text{I}_\text{v}} + \text{c}_2\right)}{\left(\mu_{\text{I}_\text{u}}^2 + \mu_{\text{I}_\text{v}}^2 + \text{c}_1\right)\left(\delta_{\text{I}_\text{u}}^2 + \delta_{\text{I}_\text{v}}^2 + \text{c}_2\right)} \tag{10}$$

where μ_{I_u} is mean of the predicted image, μ_{I_v} is mean of the low-noise image, $\delta_{\text{I}_\text{u}}^2$ is variance of the predicted image, $\delta_{\text{I}_\text{v}}^2$ is variance of the low-noise image, and $\delta_{\text{I}_\text{u}\text{I}_\text{v}}$ is covariance between the predicted and the low-noise images. c_1 and c_2 are regularization constants dependent on the image's luminance dynamic range, which depends on the data type of the input image and are equal to 0.01× luminance dynamic range and 0.03× luminance dynamic range, respectively. As in our implementation, this range equals [0, 1], c_1 and c_2 are equal to 0.01 and 0.03.

EPI

EPI can effectively evaluate the degree of edge preservation of denoised images. As our denoised estimations of the different methods are slice-based, which are two-dimensional images, this index has been measured in all two-dimensional slices of the test set that can be calculated using Eq. (11) [24].

$$\mathrm{EPI} = \frac{\sum \left(\Delta \mathrm{I}_u - \overline{\Delta \mathrm{I}_u}\right)\left(\Delta \mathrm{I}_n - \overline{\Delta \mathrm{I}_n}\right)}{\sqrt{\sum \left(\Delta \mathrm{I}_u - \overline{\Delta \mathrm{I}_u}\right)^2 \sum \left(\Delta \mathrm{I}_n - \overline{\Delta \mathrm{I}_n}\right)^2}} \tag{11}$$

where $\Delta \mathrm{I}_u$ and $\Delta \mathrm{I}_n$ represent the high pass filtered versions of the denoised and noisy images, obtained with a 3×3 pixel standard approximation of the Laplacian operator, respectively. $\overline{\Delta \mathrm{I}_u}$ and $\overline{\Delta \mathrm{I}_n}$ are the $\Delta \mathrm{I}_u$ and $\Delta \mathrm{I}_n$ mean values, respectively. The larger EPI value indicates a greater ability to preserve edges.

Statistical analysis

The mean and standard deviation of the above metrics were calculated and reported in the results. Additionally, the One-Way ANOVA test, a statistical method that compares the mean values of samples, was conducted for statistical analysis to compare the image-derived metrics using IBM SPSS Statistics software Version 26. The p values obtained from the statistical test were also presented in the results. A p value less than 0.05 was considered significant. Additionally, the post hoc Tukey's Honest Significant Difference (HSD) test was also used to assess the significance of differences between pairs of group means.

Results

Qualitative assessment

Figure 4 illustrates the estimated outputs of different approaches compared to their low- and high-noise counterparts. The visual inspection revealed that the TPL-CNN method could generally predict images with lower noise and higher resolution. Among the different training approaches of the TPL-CNN, it is observed that solo and distinctive training of the network can improve image quality better than the mélange training. However, TPL-CNN with solo training makes the ventricle wall thicker than the distinctive training of the TPL-CNN.

The conventional horizontal and vertical long-axis and short-axis sections of the myocardium for each training method have also been presented in the Fig. 1 of the Supplementary Information section.

Figure 5 illustrates a comparison between the chosen proposed method (TPL-CNN with distinct training) and other techniques, including conventional reconstruction method, Gaussian post filtering, iterative Gaussian filtering, and CNN denoising. The results show that iterative Gaussian filtering can decrease the noise level the most, while the contrast seems to be reduced. On the other hand, TPL-CNN (distinct training) significantly decreased the noise level, while causing the least amount of contrast degradation compared to the other method.

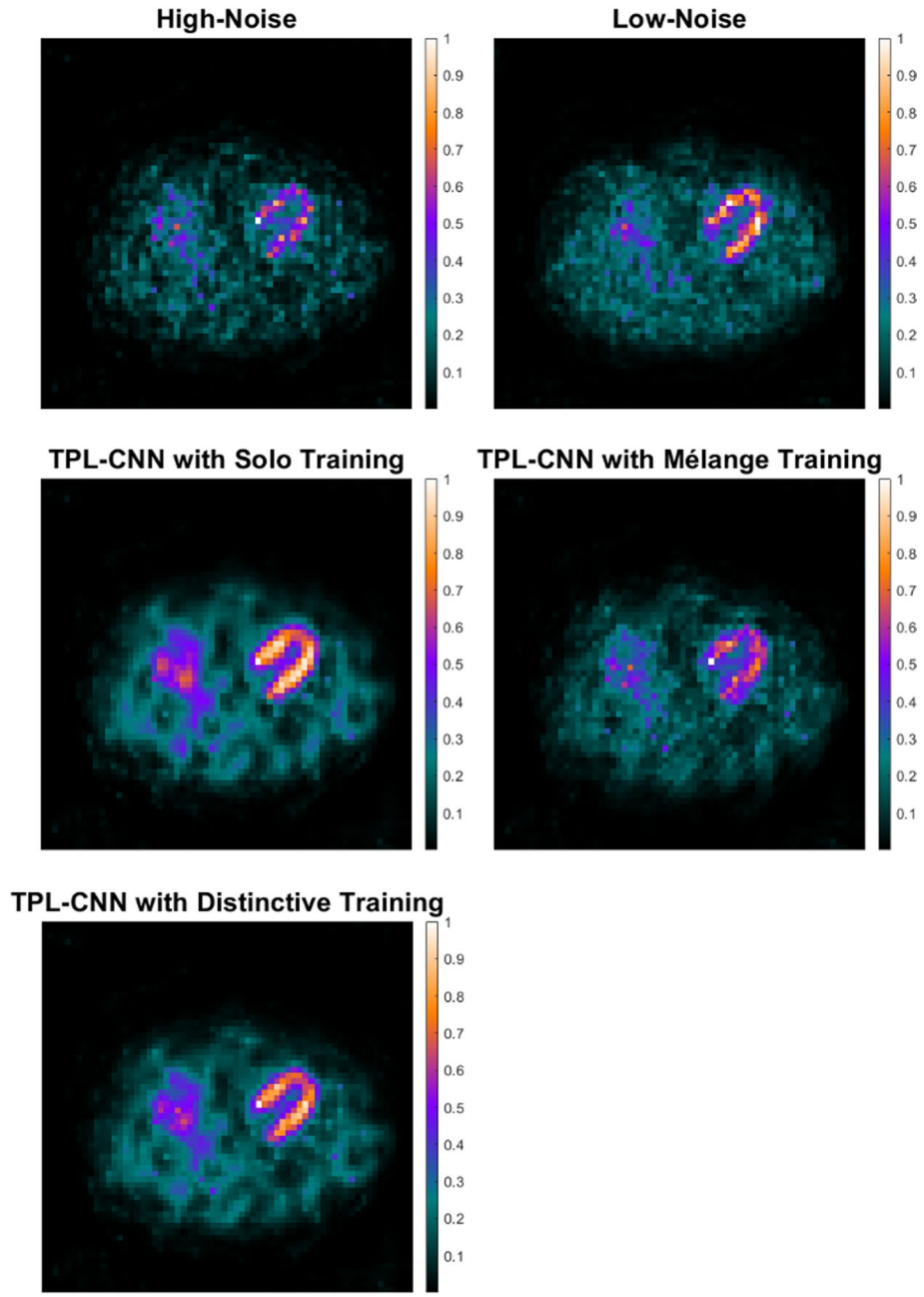

Fig. 4 Predicted outputs for a randomly selected patient from the test dataset acquired from solo, mélange, distinctive training of the TPL-CNN compared to the reference low-noise and high-noise images. All images are normalized to one and shown on the right side of each image

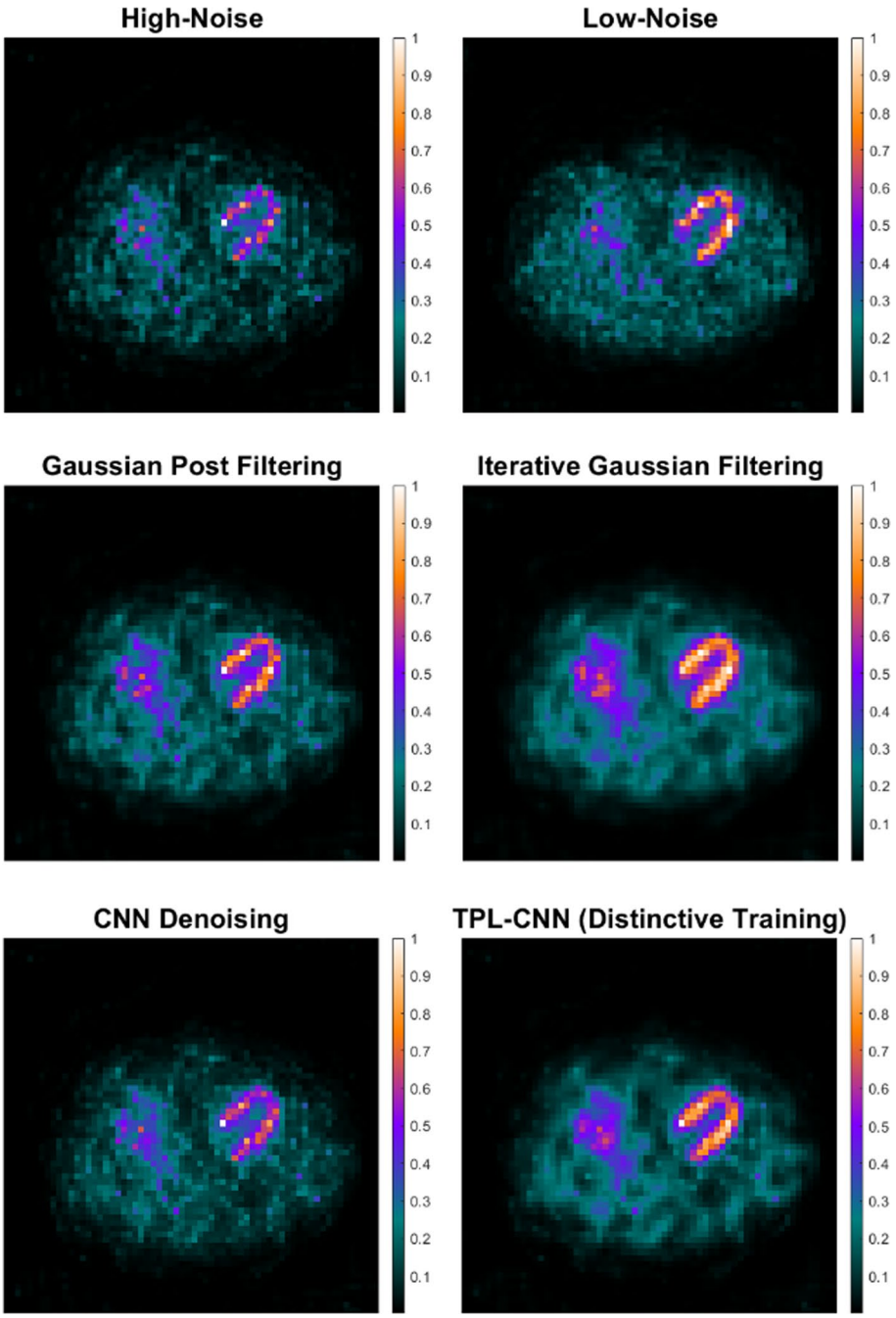

Fig. 5 Predicted outputs for a randomly selected patient from the test dataset acquired from Gaussian post filtering, iterative Gaussian filtering, CNN denoising, and distinctive training of the TPL-CNN compared to the reference low-noise and high-noise images. All images are normalized to one and shown on the right side of each image

The conventional horizontal and vertical long-axis and short-axis sections of the myocardium for each method have also been presented in the Fig. 2 of the Supplementary Information section.

Table 1 The results of the evaluation metrics of the predicted test images in different training methods of the TPL-CNN; evaluation metrics are presented in mean ± std. (post-hoc test *p*-value)

Evaluation metrics	COV_bkg	CNR	SSIM	EPI	RMSE
High noise image	0.64 ± 0.03 (0.000)*	2.92 ± 2.16 (0.000)*	0.19 ± 0.04 (0.000)*	0.001 ± 0.08 (0.000)*	0.11 ± 0.01 (0.000)*
Solo training	0.45 ± 0.04 (0.001)*	23.10 ± 7.95 (0.435)	0.56 ± 0.07 (0.000)*	0.43 ± 0.13 (0.000)*	**0.06 ± 0.01 (0.000)***
Mélange training	0.48 ± 0.03 (0.998)	20.84 ± 6.84 (1.000)	0.53 ± 0.07 (0.000)*	0.37 ± 0.12 (0.000)*	**0.06 ± 0.01 (0.000)***
Distinctive training	**0.43 ± 0.04 (0.000)***	**30.48 ± 8.34 (0.000)***	**0.57 ± 0.08 (0.000)***	**0.46 ± 0.13 (0.000)***	**0.06 ± 0.01 (0.000)***
Low noise image	0.48 ± 0.04	20.619 ± 8.715	1 ± 0	1 ± 0	0 ± 0

COV_bkg coefficient of variation in the background, *CNR* contrast-to-noise ratio, *RMSE* root mean squared error, *SSIM* structural similarity index metrics, *EPI* edge preservation index, *p*-value: the significance of the mean difference between each method and the OSEM reconstruction of the low-noise data is calculated by the post-hoc test at the 0.05 level. The bolded values indicate the best performance achieved

Table 2 The results of the evaluation metrics of the estimated test images in different comparison methods

Method	COV_bkg Mean ± std (*p*-value)	CNR Mean ± std (*p*-value)
OSEM reconstruction (High-noise)	0.639 ± 0.032 (0.000)*	2.922 ± 2.157 (0.000)*
Gaussian post filtering	0.480 ± 0.036 (1.000)	13.854 ± 7.111 (0.000)*
Iterative Gaussian filtering	**0.391 ± 0.041 (0.000)***	12.140 ± 6.086 (0.000)*
Neural network post filtering	0.493 ± 0.033 (0.424)	16.514 ± 9.042 (0.063)
Distinctive training	0.431 ± 0.038 (0.000)*	**30.478 ± 8.339 (0.000)***
OSEM reconstruction (Low-noise)	0.477 ± 0.040	20.619 ± 8.715

COV_bkg coefficient of variation in the background, *CNR* contrast-to-noise ratio, *p*-value: the significance of the mean difference between each method and the OSEM reconstruction of the low-noise data is calculated by the post-hoc test at the 0.05 level. The bolded values indicate the best performance achieved

Quantitative analysis

Table 1 shows that the predicted images obtained from the TPL-CNN with the distinctive training approach exhibit the highest improvement in CNR (30.48 ± 8.34), SSIM (0.57 ± 0.08), and EPI (0.46 ± 0.13), while reaching the lowest COV_bkg (0.43 ± 0.04). Conversely, the TPL-CNN with the mélange training approach resulted in the lowest increase in CNR (20.84 ± 6.84), SSIM (0.53 ± 0.07), and EPI (0.37 ± 0.12) metrics, while reaching the highest COV_bkg (0.48 ± 0.033). The results exhibit that the TPL-CNN can decrease RMSE up to (0.06 ± 0.01), almost the same for all training approaches. In calculating SSIM, EPI, and RMSE, the ground truth was the low-noise data. Therefore, these three metrics show the best results in low-noise data since they compare the ground truth with itself.

The results of the ANOVA test show that there are statistically significant differences in the evaluation metrics across different approaches. The post hoc analysis conducted with the Tukey test suggests a significant difference between the TPL-CNN

with the distinctive and low-noise data in COV_bkg and CNR. When comparing TPL-CNN with distinctive and the other two training approaches, the results demonstrate a significant difference in CNR values. Therefore, TPL-CNN, with the distinctive training approach, is our chosen network for the rest of the study. The remaining metrics are detailed in the Table S-2 of the Supplementary Information section.

Table 2 summarizes the results of the statistical analysis of image quality metrics (mean ± SD), including COV_bkg, and CNR between the different SPECT images acquired from different comparison methods, and TPL-CNN with distinctive training approach. The proposed method shows the highest CNR value (30.48 ± 8.34) with a low COV_bkg (0.43 ± 0.04) compared to the other methods. Since the main purpose of this study is to reduce the noise while maintaining the contrast, these two metrics were solely discussed here. The remaining metrics are detailed in the Table S-3 of the Supplementary Information section.

The results of the ANOVA test show that there are statistically significant differences in the evaluation metrics among all six evaluated denoising approaches. According to the Tukey post-hoc test, although the iterative Gaussian filtering has the lowest COV_bkg, it has the worst CNR value. Conversely, while the COV_bkg of the proposed method with the chosen network (TPL-CNN with the distinctive training) was only slightly higher than that of iterative Gaussian filtering, making it the second most effective denoising method, it outperformed the others in terms of CNR values.

Discussion

This study aimed to reduce the amount of inherent noise of SPECT-MPI that was a result of low counting of gamma ray during data collection process of these images. Although several spatial filtering methods were applied in this order, it has been showed that neural networks could result in better denoised outputs. The crucial aspect of this study is that we proposed a method in which filtering is performed during the reconstruction process, while previous studies were mostly focusing on denoising before or after reconstructing the image. To this end, GAN networks were trained and used in each step of reconstruction process.

While conventional deep denoising methods are typically applied post-reconstruction, pre-reconstruction deep denoising has also been explored. Studies have shown that GAN networks can outperform other deep denoising models when applied pre-reconstruction [25]. However, pre-reconstruction denoising can alter the Poisson-nature of the acquisition data, potentially affecting the performance of maximum likelihood type reconstruction commonly used in SPECT [26]. Additionally, iterative deep denoising during reconstruction has been reported in positron emission tomography (PET) scans. One approach is to use a pre-trained network as a prior for maximum a posteriori type reconstruction [27]. Another method is to incorporate the network as an additional argument in the maximum likelihood maximization step of iterative reconstruction [28]. To the best of our knowledge, this is the first study to apply an iterative deep denoising network in SPECT and use it as an analytical filter within iterations.

Our findings indicate that the source of the training dataset significantly impacts the evaluation metrics. The results suggest that using a distinct denoising network for each iteration can be more effective than a single "one-size-fits-all" network. Results also show

that when considering TPL-CNN with solo training, the left ventricle wall of predicted images is estimated to be thicker than the low-noise images of their counterparts. However, this issue is addressed in TPL-CNN with distinctive training. Moreover, results show that when considering TPL-CNN with mélange, although the amount of noise has decreased, the shape of the predicted left ventricle has changed. This can be interpreted as the mélange network, which encompasses large differences in signal and noise levels, fails to adapt its filtering to the noise level in the input images. Similarly, since the solo network has only been trained on blurred low-noise images originating from the second OSEM iteration, it produces blurred low-noise images as output.

Previously reported deep learning denoising methods often suffer from a resolution-noise trade-off, which manifests as blurring of the image [26, 29]. However, it has been shown that by extracting the myocardium centered in the image dataset and training a network with this patch, the blurring effect can be reduced [10, 30]. Meanwhile manual extraction of the myocardium is time-consuming and limited when denoising of the entire image is needed. We have used two strategies to compensate for these limitations. First, we used an in-center developed automatic three-dimensional cardiac segmentation package to patch the myocardium (QCard-NM) [21, 22]. Second, we used a two-step training strategy to overcome the second limitation by transferring the learning of the network trained with the myocardium patch to another network that must denoise the full-size image. This improved the performance of the overall system and resulted in less blurring artifact [20].

In our research, we acquired high-noise images using a separate fast SPECT acquisition. This fast scan, which took approximately one-third of the standard acquisition time, followed immediately after the standard SPECT scan. However, other papers have reported various methods for generating high-noise images in the context of denoising SPECT data. Compared to one common method, bootstrapped sampling from recorded events during the standard SPECT acquisition, our proposed method exposes potential changes in radiotracer kinetics, particularly in extracardiac regions. Additionally, patient motion between adjacent scans, which is not a concern in bootstrapped sampling, may impact our approach. Another alternative involves manually adding Poisson noise to the standard scan. However, this method introduces unrealistic noise patterns that are easily discernible by artificial intelligence algorithms, making them straightforward to remove from the high-noise image. In contrast, our proposed method offers a more realistic noise profile, enhancing the fidelity of high-noise image acquisition. Hence, the denoising strategy based on our approach may prove more effective in real-world scenarios. Another approach for obtaining high-noise images involves low-dose injected activity imaging. This approach is applicable in preclinical laboratories (such as animal studies) and can be simulated using computer tools. As injected activity increases, recorded scatter counts rise, leading to higher noise levels during fast acquisitions compared to low-dose imaging. As a result, our accelerated scan may exhibit greater noise propagation, including additive noise from patient motion, when compared to an equivalent real or simulated high-noise scan [31].

It has been found that among different deep network architectures (CNN, residual neural network (ResNet), UNet, GAN), GAN outperformed the others in SPECT or PET denoising tasks [12, 26]. However, contradictory findings have also been presented in

PET studies [32, 33]. Despite this, we chose to use GAN networks in this study because the high-noise image was acquired from a separate scan instead of bootstrap sampling of the standard scan. Since in our study there may be some deformation of signal content in the high-noise image compared to the low-noise image, the GAN network would outperform in this study. In a GAN network, the discriminator evaluates the content of the generated image by comparing it to a real low-noise image. In contrast, other convolutional networks typically evaluate paired images on a pixel-by-pixel basis.

Using skip connections in the generators of GAN networks makes deep denoising networks less sensitive to the number of layers and filters. This issue is associated with the concept of overfitting in the presence of shallow layers [10]. Our optimization analysis found that a network with a higher number of layers and filters leads to less accurate results, which has also been reported in previous studies [10]. Compared to other imaging modalities such as computerized tomography (CT) or Magnetic resonance imaging (MRI), SPECT-MPI images have lower image resolution and reduced anatomical detail, resulting in necessity of less complex deep denoising networks.

Limitations and further studies

In the TPL-CNN method proposed in this study, the architecture of the neural network remained constant from the lower to the last iterations of OSEM. However, since the reconstructed signal and noise vary among different iterations, it may improve overall system performance if the complexity of the network changes in accordance with the number of iterations passed. The effect of using deeper networks in higher iterations of OSEM could be investigated in future studies.

It should be noted that our results are based on a dual-head DDD dedicated cardiac-SPECT (CorCam). While we acknowledge the existence of other hardware methods for achieving low noise images with low-dose (or fast) scanning (such as high-sensitivity systems, cardio-centering converging collimation, CZT-based detectors, and full coverage detector rings [34]), we believe that the denoising approaches described in this paper could also be applied to these systems. This would likely result in even further reductions in dose without compromising diagnostic accuracy.

We reconstructed the images using our in-house simple two-dimensional OSEM code,[1] with respect to [18], without any specific correction methods, rather than applying commercially available reconstruction software. Despite this, our simple code shares a common theoretical foundation with the algorithms employed by vendors. The inclusion of attenuation, scatter, and resolution correction for OSEM has been shown to improve image quality and detection performance [4]. While this combination of corrections was not developed by us to maintain the generality, it may enhance the effectiveness of denoising methods. Moreover, since the effective number of iterations (iterations*subsets) can affect the convergence of image signal as well as the noise properties of the image, it appears optimal to adjust the reconstruction parameters for low-dose (or fast) scanning to achieve an perfect balance between noise level and signal convergence when using different denoising strategies [8]. Nevertheless, the ultimate

[1] https://github.com/FYousefzadeh/in-house-simple-OSEM2D-code.git.

conclusion of the proposed idea and the qualitative ranking of different denoising strategies evaluated in this study is independent of the specific details of the reconstruction implementation and correction techniques.

The total time taken for the reconstruction process using the simple two-dimensional OSEM code written in MATLAB with the computer's GPU (introduced in the Network Implementation section) is approximately about 2.488 s for an image with dimensions of 64×64 pixels. When the TPL_CNN is added as a filter in the middle of the reconstruction process, this time increases to about 4.441 s. However, this increase can be considered negligible due to the improved results achieved. It is worth noting that these estimated times can be reduced by using a more powerful GPU.

Currently only non-gated serial mode data sets were employed in this study. Moreover, the dataset is not variant due to different clinical defects. Although three-dimensional image reconstruction yields better image characteristics, this study utilizes the two-dimensional reconstruction mode to maintain generalization since the device's system matrix was unavailable. Further evaluations regarding gated data sets, listed mode ones, diverse datasets of all types of defects, as well as other reconstruction frameworks such as three-dimensional image reconstruction and PET are our future work. Additionally, the two-dimensional Gaussian filters used in comparison methods were not optimized in this study. Exploring this aspect has been left to the future work.

Conclusion

In this work, an iterative noise reduction algorithm that uses a CNN form, which was trained via GAN (TPL-CNN), during the reconstruction process was proposed to improve the resolution of non-gated SPECT-MPI images. In previous studies, the efficiency of a two-phase learned approach was demonstrated. In this study, we have shown that for the SPECT-MPI denoising approach, applying the CNNs iteratively during the reconstruction process is more beneficial than applying it post-reconstruction. Quantification results show that the proposed method performs better than post-processing methods, which use a spatial filter like a Gaussian filter or a neural network as a filter to decrease the noise of reconstructed images. Furthermore, we found that using a distinctive network (same architecture but different training dataset) after each reconstruction iteration is more effective than using a single 'one-size-fits-all' network (same architecture and training dataset) for the SPECT-MPI denoising approach. In order to enhance the results, other neural network structures and ways to optimize the training parameters using other types of data sets can be studied in the future.

Appendix

See Table 3.

Table 3 TPL-CNN Characteristics

Block no.	Layer no.	Layer name	Layer characteristics
Generator layers arrange and their specifications			
	1	Input	Phase1: 32 × 24 images Phase2: 64 × 64 images
1	2	Convolution	10 3 × 3 convolutions with stride 1 and padding 'valid'
	3	Batch normalization	momentum = 0.9
	4	Leaky ReLU	alpha = 0.3
2	5	Convolution	10 3 × 3 convolutions with stride 1 and padding 'valid'
	6	Batch normalization	momentum = 0.9
	7	Leaky ReLU	alpha = 0.3
3	8	Deconvolution	10 3 × 3 convolutions with stride 1 and padding 'valid'
	9	Skip connection	Concatenate output of the 4th layer with output of the 8th 00layer
	10	Batch normalization	momentum = 0.9
	11	Leaky ReLU	alpha = 0.3
4	12	Deconvolution	10 3 × 3 deconvolutions with stride 1 and padding 'valid'
	13	Skip connection	Concatenate the input layer with output of the 12th layer
	14	Batch normalization	momentum = 0.9
	15	ReLU	ReLU
Discriminator layers arrange and their specifications			
	1	Input	Phase1: 32 × 24 images Phase2: 64 × 64 images
1	2	Convolution	10 3 × 3 convolutions with stride 1 and padding 'valid'
	3	Batch normalization	momentum = 0.9
	4	Leaky ReLU	alpha = 0.3
2	5	Convolution	10 3 × 3 convolutions with stride 1 and padding 'valid'
	6	Batch normalization	momentum = 0.9
	7	Leaky ReLU	alpha = 0.3
	8	Flatten	Phase1: 1 × 5600 array Phase2: 1 × 36,000 array
	9	Dense	1 neuron
Training option			
		Max epochs	80
		Solver for generator network	Adam optimizer
		Learning rate	0.0001
		Loss function	Binary cross entropy

Abbreviations

Adam	Adaptive moment
AE	Auto-encoder network
ASNC	American society of nuclear cardiology
BN	Batch normalization
CNN	Convolutional neural network
CNR	Contrast-to-noise ratio
COV	Coefficient of variation
COV_bkg	Coefficient of variation in the background
CT	Computerized tomography
DL	Deep learning
EPI	Edge preservation index
GAN	Generative adversarial network
HSD	Honest significant difference
LV	Left ventricular
MPI	Myocardial perfusion imaging

MRI	Magnetic resonance imaging
OSEM	Ordered subset expectation maximization
PET	Positron emission tomography
QPS	Quantitative perfusion SPECT
ReLU	Rectifier linear unit
ResNet	Residual neural network
RMSE	Root mean squared error
ROI	Regions of interest
SNR	Signal-to-noise ratio
SPECT	Single-photon emission computed tomography
SSIM	Structural similarity index metrics
TPL-CNN	Two-phase learned convolutional neural network

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40658-024-00687-3.

Additional file 1.

Acknowledgements
This work is part of an MSc dissertation in biomedical engineering by Farnaz Yousefzadeh at Shiraz University. We would like to express our gratitude to AlZahra Hospital personnel who made this study possible. We also acknowledge all the participants in this study.

Author contributions
FY had the main role in designing the idea and participated in implementing the research, performing the analysis, interpreting the results, and writing the manuscript. MY participated in implementing the research, interpreting the results, writing the manuscript, and supervising the research. SME had the main role in implementing the research and designing the idea, collecting the data, performing the analysis, interpreting the results, and writing the manuscript. RF participated in interpreting the results and supervised the research. NS participated in developing the algorithm and performing computational simulations. SG, ZAM, and MH participated in collecting and interpreting data and contributed to the clinical insights. All authors read and approved the final manuscript.

Funding
No funding was received for this article.

Availability of data and materials
The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
This study followed the ethical standards of the Shiraz University of Medical Sciences research committee and the 1964 Helsinki Declaration, with the ethics committee code number of IR.SUMS.MED.REC.1401.147.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 4 March 2024 Accepted: 26 September 2024
Published online: 08 October 2024

References

1. Malek H. Nuclear cardiology. In: Practical cardiology. Elsevier; 2018. p. 167–72. https://doi.org/10.1016/B978-0-323-51149-0.00009-2.
2. Jerome SD, Tilkemeier PL, Farrell MB, Shaw LJ. Nationwide laboratory adherence to myocardial perfusion imaging radiation dose reduction practices: a report from the intersocietal accreditation commission data repository. Cardiovasc Imaging. 2015;8:1170–6.
3. Wells RG. Dose reduction is good but it is image quality that matters. J Nucl Cardiol. 2020;27:238–40.
4. Juan Ramon A, Yang Y, Pretorius PH, Slomka PJ, Johnson KL, King MA, Wernick MN. Investigation of dose reduction in cardiac perfusion SPECT via optimization and choice of the image reconstruction strategy. J Nucl Cardiol. 2018;25:2117–28.
5. Ramon AJ, Yang Y, Wernick MN, Hendrik Pretorius P, Johnson KL, Slomka PJ, King MA. Evaluation of the effect of reducing administered activity on assessment of function in cardiac gated SPECT. J Nucl Cardiol. 2020;27(2):562–72. https://doi.org/10.1007/s12350-018-01505-x.

6. Zeraatkar N, Sajedi S, Farahani MH, Arabi H, Sarkar S, Ghafarian P, Rahmim A, Ay MR. Resolution-recovery-embedded image reconstruction for a high-resolution animal SPECT system. Phys Med. 2014;30:774–81.
7. Arabi H, Zaidi H. Improvement of image quality in PET using post-reconstruction hybrid spatial-frequency domain filtering. Phys Med Biol. 2018;63:215010.
8. Arabi H, Zaidi H. Non-local mean denoising using multiple PET reconstructions. Ann Nucl Med. 2021;35:176–86.
9. Cheng Z, Wen J, Huang G, Yan J. Applications of artificial intelligence in nuclear medicine image generation. Quant Imaging Med Surg. 2021;11:2792–822.
10. Ramon AJ, Yang Y, Pretorius PH, Johnson KL, King MA, Wernick MN. Improving diagnostic accuracy in low-dose SPECT myocardial perfusion imaging with convolutional denoising networks. IEEE Trans Med Imaging. 2020;39:2893–903.
11. Song C, Yang Y, Wernick MN, Pretorius PH, King MA, Low-dose cardiac-gated spect studies using a residual convolutional neural network. in 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), pp. 653–656, 2019.
12. Olia NA, Kamali-Asl A, Tabrizi SH, Geramifar P, Sheikhzadeh P, Farzanefar S, Arabi H, Zaidi H. Deep learning–based denoising of low-dose SPECT myocardial perfusion images: quantitative assessment and clinical performance. Eur J Nucl Med Mol Imaging. 2022;49(5):1508–22. https://doi.org/10.1007/s00259-021-05614-7.
13. Hashimoto F, Onishi Y, Ote K, Tashima H, Reader AJ, Yamaya T. Deep learning-based PET image denoising and reconstruction: a review. Radiol Phys Technol. 2024;17(1):24–46. https://doi.org/10.1007/s12194-024-00780-3.
14. Mustafovic S, Thielemans K. Object dependency of resolution in reconstruction algorithms with interiteration filtering applied to PET data. IEEE Trans Med Imaging. 2004;23(4):433–46.
15. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. IEEE Trans Med Imaging. 1982;1:113–22.
16. Gengsheng LZ. Medical image reconstruction: a conceptual tutorial. Beijing: Higher Education Press; 2010.
17. Shiri I, Sabet KA, Arabi H, Pourkeshavarz M, Teimourian B, Ay MR, Zaidi H. Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks. J Nucl Cardiol. 2021;28(6):2761–79. https://doi.org/10.1007/s12350-020-02119-y.
18. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imaging. 1994;13:601–9.
19. Mao X, Shen C, Yang Y-B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Adv Neural Inf Process Syst. 2016;29:2802–10.
20. Yousefzadeh F, Yazdi M, Entezarmahdi M, Faghihi R, Dehghan F, Shiri I. Improving the efficiency of deep learning-based denoising in SPECT myocardial perfusion image by two-phase learning algorithm. J Nucl Med. 2022;63:3227–3227.
21. Shahamiri N, Yazdi M, Entezarmahdi SM, Faghihi R, Dehghan F, Shiri I. An empirical update of left ventricular 3D segmentation algorithm in myocardial perfusion SPECT imaging. J Nucl Med. 2021;62:1691–1691.
22. Entezarmahdi SM, Faghihi R, Yazdi M, Shahamiri N, Geramifar P, Haghighatafshar M. QCard-NM: developing a semi-automatic segmentation method for quantitative analysis of the right ventricle in non-gated myocardial perfusion SPECT imaging. EJNMMI Phys. 2023;10:1–24.
23. Rahimian A, Etehadtavakol M, Moslehi M, Ng EYK. Myocardial perfusion single-photon emission computed tomography (SPECT) image denoising: a comparative study. Diagnostics. 2023;13(4):611. https://doi.org/10.3390/diagnostics13040611.
24. Sattar F, Floreby L, Salomonsson G, Lovstrom B. Image enhancement based on a nonlinear multiscale method. IEEE Trans Image Process. 1997;6(6):888–95.
25. Sun J, Jiang H, Du Y, Li C-Y, Wu T-H, Liu Y-H, et al. Deep learning-based denoising in projection-domain and reconstruction-domain for low-dose myocardial perfusion SPECT. J Nucl Cardiol. 2023;30:970–85.
26. Sohlberg A, Kangasmaa T, Constable C, Tikkakoski A. Comparison of deep learning-based denoising methods in cardiac SPECT. EJNMMI Phys. 2023;10:9.
27. Kim K, Wu D, Gong K, Dutta J, Kim JH, Son YD, et al. Penalized PET reconstruction using deep learning prior and local linear fitting. IEEE Trans Med Imaging. 2018;37:1478–87.
28. Gong K, Guan J, Kim K, Zhang X, Yang J, Seo Y, et al. Iterative PET image reconstruction using convolutional neural network representation. IEEE Trans Med Imaging. 2018;38:675–85.
29. Lyra M, Ploussi A. Filtering in SPECT image reconstruction. J Biomed Imaging. 2011;2011:1–14.
30. Liu J, Yang Y, Wernick MN, Pretorius PH, King MA. Deep learning with noise-to-noise training for denoising in SPECT myocardial perfusion imaging. Med Phys. 2021;48:156–68.
31. Sanaat A, Shiri I, Arabi H, Mainta I, Nkoulou R, Zaidi H. Deep learning-assisted ultra-fast/low-dose whole-body PET/CT imaging. Eur J Nucl Med Mol Imaging. 2021;48:2405–15.
32. Lu W, Onofrey JA, Lu Y, Shi L, Ma T, Liu Y, et al. An investigation of quantitative accuracy for deep learning based denoising in oncological PET. Phys Med Biol. 2019;64:165019.
33. Ladefoged CN, Hasbak P, Hornnes C, Højgaard L, Andersen FL. Low-dose PET image noise reduction using deep learning: application to cardiac viability FDG imaging in patients with ischemic heart disease. Phys Med Biol. 2021;66:054003.
34. Slomka PJ, Miller RJH, Lien-Hsin H, Berman DS. Novel techniques: solid-state detectors, dose reduction (SPECT/CT). In: Nekolla SG, Rischpler C, editors. Hybrid cardiac imaging. Cham: Springer International Publishing; 2022. p. 103–29. https://doi.org/10.1007/978-3-030-83167-7_7.

Publisher's Note