Journal of Nuclear Medicine, published on November 9, 2023 as doi:10.2967/jnumed.122.265000

Artificial Intelligence for PET and SPECT Image Enhancement

Vibha Balaji[1], Tzu-An Song[1], Masoud Malekzadeh[1], Pedram Heidari[2], and Joyita Dutta[1]

[1]*Department of Biomedical Engineering, University of Massachusetts Amherst, Amherst, Massachusetts; and* [2]*Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts*

Nuclear medicine imaging modalities such as PET and SPECT are confounded by high noise levels and low spatial resolution, necessitating postreconstruction image enhancement to improve their quality and quantitative accuracy. Artificial intelligence (AI) models such as convolutional neural networks, U-Nets, and generative adversarial networks have shown promising outcomes in enhancing PET and SPECT images. This review article presents a comprehensive survey of state-of-the-art AI methods for PET and SPECT image enhancement and seeks to identify emerging trends in this field. We focus on recent breakthroughs in AI-based PET and SPECT image denoising and deblurring. Supervised deep-learning models have shown great potential in reducing radiotracer dose and scan times without sacrificing image quality and diagnostic accuracy. However, the clinical utility of these methods is often limited by their need for paired clean and corrupt datasets for training. This has motivated research into unsupervised alternatives that can overcome this limitation by relying on only corrupt inputs or unpaired datasets to train models. This review highlights recently published supervised and unsupervised efforts toward AI-based PET and SPECT image enhancement. We discuss cross-scanner and cross-protocol training efforts, which can greatly enhance the clinical translatability of AI-based image enhancement tools. We also aim to address the looming question of whether the improvements in image quality generated by AI models lead to actual clinical benefit. To this end, we discuss works that have focused on task-specific objective clinical evaluation of AI models for image enhancement or incorporated clinical metrics into their loss functions to guide the image generation process. Finally, we discuss emerging research directions, which include the exploration of novel training paradigms, curation of larger task-specific datasets, and objective clinical evaluation that will enable the realization of the full translation potential of these models in the future.

Key Words: artificial intelligence; denoising; superresolution; PET; SPECT

J Nucl Med 2024; 00:1–9
DOI: 10.2967/jnumed.122.265000

PET and SPECT are nuclear medicine–based molecular imaging modalities that generate 3-dimensional (3D) visualizations of the biodistribution of exogenous radiotracers. These modalities provide functional and physiological information and are vital for disease diagnostics, staging, treatment planning, and therapeutic evaluation for a wide range of disorders, including many cancer types, neurodegenerative disorders, cardiovascular disease, and musculoskeletal disorders (*1–6*). Recent advances in hardware and software have greatly enhanced the quantitative capabilities of PET and SPECT imaging, addressing issues related to both high noise and low spatial resolution, while also augmenting their traditionally semiquantitative clinical utility. The emergence of artificial intelligence (AI) has brought forth a multitude of image enhancement techniques for denoising, deblurring, and partial-volume correction of PET and SPECT images. AI-based enhancement methods can be implemented after reconstruction into existing PET/SPECT clinical workflows to achieve purely software-based improvement in image quality without expensive hardware upgrades. These models that learn image representations directly from data benefit from the increasing volume (i.e., more training examples) and variety (i.e., a diverse training population) of training datasets. AI-based image enhancement techniques accomplish a range of tasks, including boosting the signal-to-noise ratio, enhancing spatial resolution, shortening scan times, and reducing radiotracer dose. In this review, we discuss emerging denoising and deblurring techniques that can be potentially transformative for PET and SPECT imaging.

Most AI-based image enhancement techniques rely on a deep-learning model that receives a corrupt image as its input and generates a clean image as its output. For denoising, the corrupt input image is noisy, whereas for deblurring, it is low resolution. Deblurring efforts for PET and SPECT encompass partial volume correction approaches that seek to mitigate the partial volume effect. The latter arises from the blurring of tissue boundaries (the predominant factor for modalities such as PET and SPECT) and discretizing the image space (*7*). Unlike image reconstruction, AI-based image enhancement models do not require raw data and can be readily trained and validated by existing image repositories. These methods are thus rapidly gaining popularity in nuclear medicine, where large image-domain datasets are much more accessible than list-mode or sinogram datasets. AI models for image enhancement have consistently outperformed filtering, deconvolution, and other traditional analytic or model-based iterative approaches for denoising or partial volume correction. AI has led to new approaches for multimodality fusion (*8*) that can provide improved cross-modality anatomic guidance to PET and SPECT using information from high-resolution MRI or CT. The evolution of deep neural network architectures, training strategies, and data requirements over the past several years has contributed to the accuracy, usability, robustness, and versatility of these models.

Figure 1 presents a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart illustrating this review's systematic article selection process, and Figure 2 offers a breakdown of the selected articles. We exclude articles that involve

Received Apr. 14, 2023; revision accepted Oct. 10, 2023.
For correspondence or reprints, contact Joyita Dutta (jdutta@umass.edu).
Published online Nov. 9, 2023.

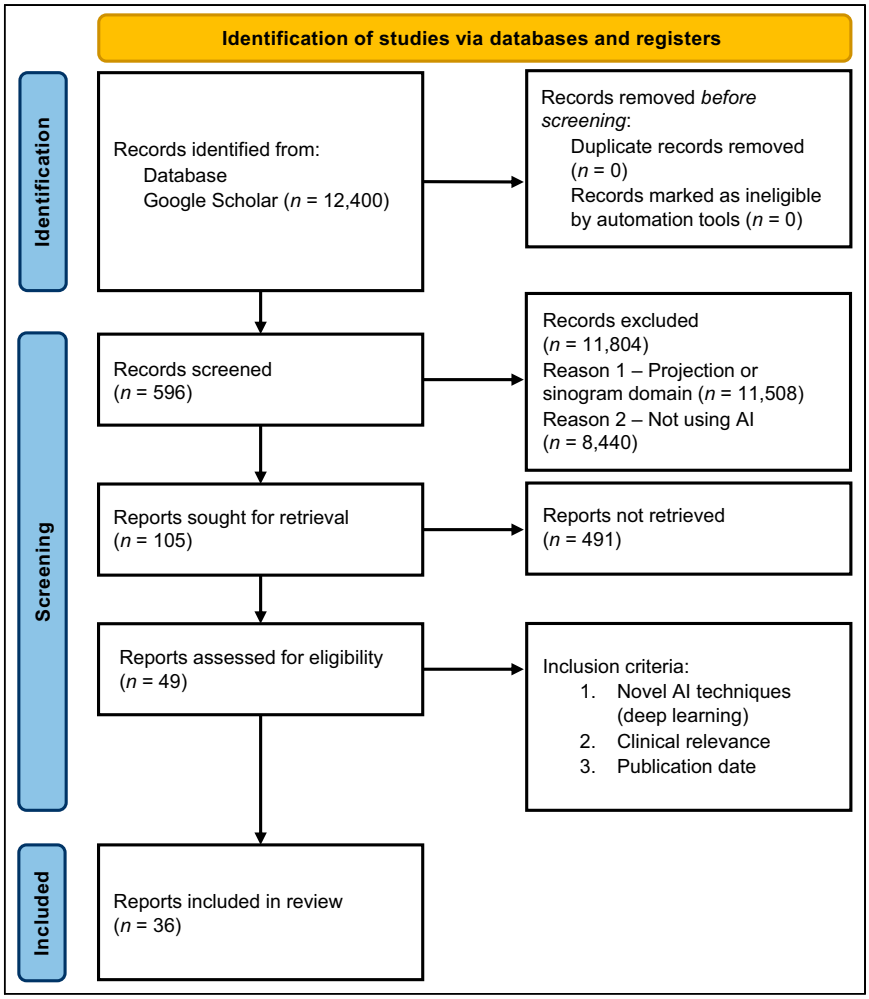

FIGURE 1. PRISMA flow diagram demonstrating selection strategy of research articles included in review.

projection- or sinogram-domain approaches, data correction techniques, and motion compensation methods. In the subsequent sections, we present a survey of recent works on PET and SPECT image enhancement and highlight emerging areas in this field. We provide an overview of predominant deep-learning architectures, loss functions, and training strategies relevant to PET and SPECT image enhancement. We then present and chronologically tabulate a selection of related articles for each modality, emphasizing publications from the last 2 y. A discussion of emerging directions concludes the review.

NOTEWORTHY

- A variety of recent advances in deep neural network architectures, loss functions, and training strategies have facilitated the application of AI models to PET and SPECT image enhancement.
- Unlike supervised learning models, which require paired corrupt and clean images for training, emerging unsupervised approaches obviate paired training data and are better suited for most clinical image enhancement applications.
- Task-based objective clinical evaluation of AI-based approaches for PET and SPECT image enhancement is required to ensure their future clinical and diagnostic use.

TECHNICAL CONSIDERATIONS FOR AI-BASED IMAGE ENHANCEMENT

Deep-learning models are characterized by multilayered network architectures that learn complex feature representations at various levels of abstraction directly from the data. Figure 3 illustrates a typical supervised learning setup for an image-denoising task. In this setup, the neural network's layer weights are iteratively adjusted during the training phase to minimize a loss function that compares the denoised image with a target low-noise or noiseless image. The denoised image is assessed using evaluation metrics in the subsequent validation phase.

Network Architectures

The current state of the art in PET and SPECT image enhancement features a variety of network architectures. Early implementations used convolutional neural networks (CNNs) that reduce computational complexity via parameter sharing. Many CNNs discussed here have an encoder–decoder structure, wherein an encoder estimates a latent representation through downsampling operations and a decoder upsamples it to match the input image's dimensions. Skip connections are often used to recover finer details.

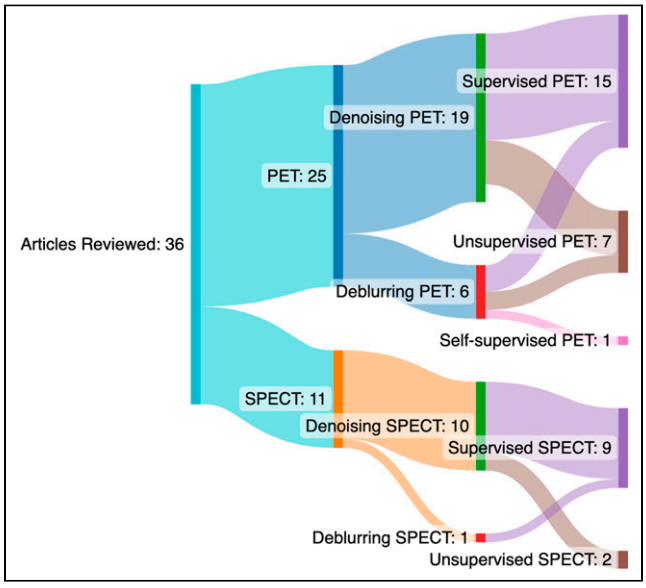

FIGURE 2. Categorywise split of selected articles reviewed here.

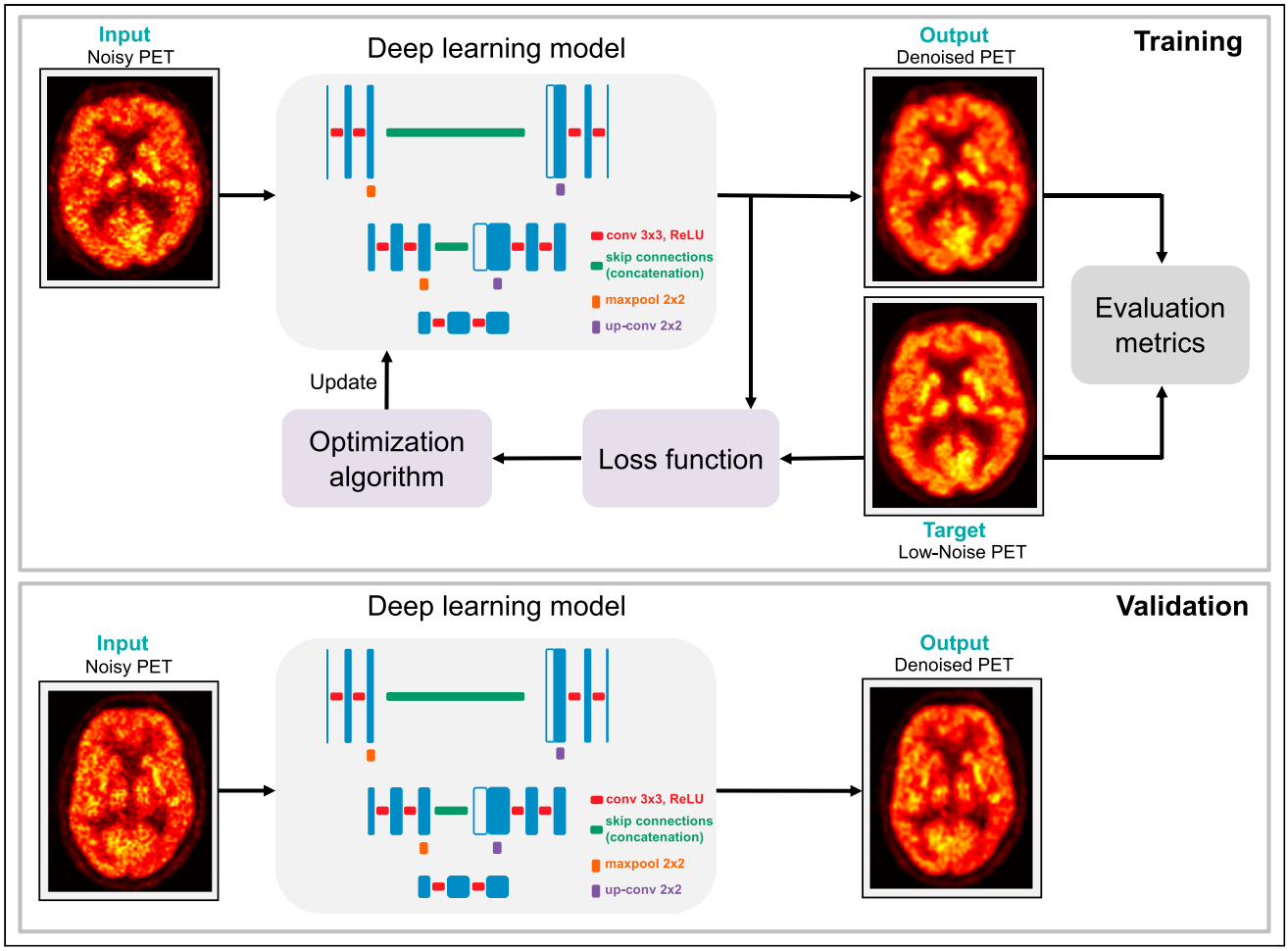

FIGURE 3. Typical supervised framework for PET image denoising using deep learning with training phase that minimizes loss function and validation phase that evaluates deep-learning model's performance. conv = convolution; ReLU = rectified linear unit; maxpool = maximum pooling; up-conv = upsampling convolution.

The U-Net (*9*), which evolved from fully convolutional architectures, has a U-shaped structure with a contracting path followed by a symmetric expanding path. It is widely used in image enhancement models, including those for PET and SPECT, with most being 3D because of the nature of the input images (*10*). Promising variants include conditional U-Nets, capturing mutual conditional dependence across modalities (*11*), and coupled U-Nets, containing modified single U-Nets that are interconnected for reduced learning redundancy (*12*).

The deep image prior (DIP) (*13*) is a widely used convolutional architecture for medical image enhancement that relies on a generator to learn clean image characteristics directly from noisy data without prior training. DIP architectures often use U-Net-like generators. In the case of PET image synthesis, anatomical images can be used for DIP initialization.

Generative adversarial networks (GANs) consist of a generator network that synthesizes an enhanced version of a subject's corrupt input image and a discriminator network that assesses how realistic the synthetic image is by comparing it with a clean image from the same subject or an unpaired clean image from a different subject (*14*). Both networks are jointly trained in competition with each other. Various GAN variants have been applied to PET and SPECT image enhancement, including conditional GANs (cGANs), which use additional prior information to guide image synthesis (*15*), and cycleGANs, which use 2 generator–discriminator pairs and can be trained with unpaired datasets (*16*).

Transformer architectures have shown potential in enhancing PET images by capturing long-range dependencies between different image regions (*17*). Additionally, diffusion models, which progressively contaminate the training data with increasing noise levels and then reverse the process to recover the data, are gaining popularity in medical imaging (*18*).

Loss Functions

Alongside network architecture, the loss function, which compares the target and predicted output, has a profound impact on model performance. Standard loss functions used in training image enhancement models include mean-squared error and mean absolute error. These functions compute the L2 and L1 norms, respectively, of voxelwise differences between the enhanced and target images. Mean-squared error is more sensitive to outliers in the training data. These loss functions lack sensitivity to visual perception, as they ignore voxel interactions and overall image structure. Perceptual loss functions address this limitation by using a pretrained network to assess high-level content and global structure in the enhanced and target images. GANs use adversarial loss

functions, which are a type of loss function used to determine whether images synthesized by a generator network have characteristics comparable to target images. However, joint generator–discriminator training can be unstable. Among GAN variants, Wasserstein GANs implement adversarial losses based on Wasserstein distances and have more training stability and less sensitivity to network architecture and parameter selection than regular GANs. CycleGANs use cycle-consistency loss functions to reduce the number of mappings between corrupt and clean image domains.

Training Strategies

Conventional supervised learning frameworks, such as the one illustrated in Figure 3, require paired clinical datasets for training, which are easy to simulate but challenging to obtain clinically as they require dual scans or access to raw data for synthesizing low-count images from standard-count ones. Furthermore, supervised learning models may not generalize well across datasets. Unsupervised approaches are thus gaining traction as they obviate the need for paired training data. Certain approaches rely solely on corrupt input data. For example, Noise2Noise (*19*) uses noisy inputs exclusively during training. Methods such as DIP benefit from the addition of anatomic information or from population-based unsupervised pretraining, which has a regularizing effect (*20*).

PET IMAGE ENHANCEMENT

Table 1 showcases many recent efforts that use AI for PET image enhancement. Most works on PET image enhancement focus on the image denoising task, the goal of which is to generate standard- or high-count PET images from noisy, low-count inputs. Early attempts using AI for PET denoising involved supervised CNNs. One study used an autocontext CNN with a sequence of convolutional modules to denoise ^{18}F-FDG PET images and examined the impact of additional anatomic T1-weighted MRI inputs on denoising performance (*21*). A dose reduction factor of 200× was reported using an encoder–decoder architecture that outperformed autocontext CNNs, nonlocal means filtering, block matching, and 3D filtering (*22*). A shift toward generative models helped overcome the limitations of traditional CNNs in capturing the underlying statistical distribution of PET images. One paper proposed a progressive refinement scheme based on concatenated 3D cGANs (*23*). Their network relied on a U-Net-like generator. Concatenated 3D cGANs were compared with single 3D cGANs, 2-dimensional cGANs, and U-Nets using ^{18}F-FDG PET brain scans from both healthy subjects and patients with mild cognitive impairment. One of the first applications of AI-based denoising to a non–^{18}F-FDG dataset was a cGAN-based ultralow-count PET imaging technique applied to ^{18}F-florbetaben scans for amyloid plaques in the brain (*24*). Importantly, the loss function in this work included a task-specific perceptual loss term that compares actual and predicted amyloid status determined by 2 expert radiologists. One paper proposed a locality-adaptive GAN model for PET image denoising in which the parametric weights are location-dependent and channel-dependent, providing a more economic way to fuse multimodal information than standard CNNs, where weights are shared across voxel locations and input channels (*25*). One work reported task-specific evaluations conducted by clinicians to determine overall image quality and lesion detectability for a denoising model based on a 3D U-Net architecture (*26*). Dilated convolutional kernels have been proposed in the context of PET image denoising to enable CNNs to capture a larger spatial context and detect features more robustly without the expensive downsampling and upsampling of internal representations (*27*). Several GAN refinements have improved GAN denoising performance in standard supervised learning scenarios. These include self-attention (*28*), cycleGAN implementations (*29*), and alternative loss functions such as the Wasserstein loss (*30*). As with other imaging modalities, there is currently great interest in diffusion models in the PET field. One paper proposed a diffusion model for PET denoising that leveraged an MRI-based prior and reported results based on ^{18}F-FDG and ^{18}F-MK-6240 radiotracers (*31*). A spatially adaptive technique and a transformer fusion network outperformed existing U-Net methods using a spatially adaptive block to extract features from both T1-weighted MRI and PET and a transformer network that established a pixelwise relationship between the 2 modalities (*32*). A Spach transformer was developed for PET denoising, which can capture long-range information efficiently, and outperformed other transformer networks and U-Nets (*33*). Notably, whereas the models were trained using ^{18}F-FDG and ^{18}F-ACBC (fluciclovine) data, the test dataset included 2 additional tracers, ^{18}F-DCFPyL and ^{68}Ga-DOTA-TATE, which were not used for model training.

In recent years, the research emphasis has largely shifted toward unsupervised models that can be trained using a single noisy image (no clean ground-truth images needed for training). The DIP has successfully performed unsupervised denoising using single noisy PET images (*34*). An extension of this idea showed improved results via population-level pretraining followed by individual fine-tuning (*35*). Noise2Void is another unsupervised approach applied for PET image denoising (*36*). It uses a single noisy input and is based on the idea of a blind spot network to estimate the intensity of a central pixel from its neighbors in a noisy image patch. Noise2Void has also been demonstrated to benefit from population-level pretraining and individual fine-tuning.

A key challenge with most supervised denoising approaches is their poor generalizability across different noise levels. A personalized denoising strategy has been proposed that uses different noise levels for training and incorporates a weighting factor that is based on the noise level in a task-dependent manner (*37*). A federated learning framework for PET image denoising was successfully tested with a simulated dataset with different noise settings corresponding to protocols from different institutions (*38*). Generalizability concerns also emphasize methods that can be adapted across scanners and tracers. One study customized a cGAN model for cross-scanner and cross-tracer optimization working with 3 scanner models (GE Healthcare Discovery MI, Siemens Biograph mCT, and Siemens Biograph Vision) and 3 radiotracers (^{18}F-FDG, ^{18}F-fluoroethyl-L-tyrosine (^{18}F-FET), and ^{18}F-florbetapir) (*39*). The results were independently assessed by 3 clinicians to ensure clinical utility.

Another key research theme for PET image enhancement centers around image deblurring and the related tasks superresolution and partial volume correction. A supervised approach for superresolving PET images by mapping from the lower-resolution Siemens HR+ scanner to the higher-resolution Siemens HRRT scanner used a very deep CNN with anatomic and spatial inputs (*40*). Later, a self-supervised solution to the same problem was proposed using a cycleGAN-like architecture and incorporating simulation guidance (*41*). This model was trained using unpaired low- and high-resolution images from the 2 scanners. A supervised cycleGAN framework was used to map PET image inputs not corrected for partial volume to outputs corrected for partial volume (*42*). The method was applied to ^{18}F-FDG, ^{18}F-flortaucipir, ^{18}F-flutemetamol, and ^{18}F-fluorodopa

TABLE 1
Summary of AI Approaches for PET Image Enhancement

Paper	Data details	Architecture	Loss function
Xiang et al. (*21*)	16 brain ^{18}F-FDG PET + T1 MRI	Autocontext CNN	MSE
Xu et al. (*22*)	9 brain ^{18}F-FDG PET + T1 MRI	U-Net	MAE
Wang et al. (*23*)	16 brain ^{18}F-FDG PET	Concatenated 3D cGAN with a 3D U-Netlike generator	MAE; cGAN
Ouyang et al. (*24*)	39 brain ^{18}F-florbetaben PET	cGAN with U-Net-like generator	MAE; cGAN; perceptual; content
Wang et al. (*25*)	20 simulated + 16 clinical brain ^{18}F-FDG PET, T1 MRI, diffusion tensor imaging	Locality adaptive multimodality GAN	MAE; adversarial
Schaefferkoetter et al. (*26*)	31 lung ^{18}F-FDG PET	U-Net	MSE
Spuhler et al. (*27*)	35 brain ^{18}F-FDG PET	Dilated U-Net	MAE
Xue et al. (*28*)	10 whole-body ^{18}F-FDG PET	3D attention residual least-squares GAN	MSE; least-squares adversarial
Zhao et al. (*29*)	109 brain ^{18}F-FDG PET/CT	Supervised cycleGAN	Adversarial; cycle consistency; identity
Gong et al. (*30*)	9 cardiac torso ^{18}F-FDG PET	Parameter-transferred Wasserstein GAN	MSE; Adversarial
Gong et al. (*31*)	120 brain ^{18}F-FDG PET; 140 brain ^{18}F-MK-6240 PET + T1 MRI	Denoising diffusion probabilistic model	MSE
Zhang et al. (*32*)	70 brain ^{18}F-FDG PET + T1 MRI	Spatially adaptive and transformer fusion network	MAE
Cui et al. (*34*)	10 lung ^{68}Ga-PRGD2 PET/CT; 30 lung ^{18}F-FDG PET/T1 MRI	Modified 3D U-Net with DIP	MSE
Cui et al. (*35*)	10 lung ^{68}Ga-PRGD2 PET/CT; 30 lung ^{18}F-FDG PET/T1 MRI	Modified 3D U-Net with DIP	MSE
Song et al. (*36*)	20 simulated brain ^{18}F-FDG PET; 17 brain ^{18}F-FDG PET + T1 MRI	Noise2Void (U-Net with 3 resolution levels)	MSE
Liu et al. (*37*)	195 cardiac torso ^{18}F-FDG PET	3D U-Net	MSE
Zhou et al. (*38*)	Heterogeneous multiinstitutional PET	Deep attention residual U-Net	MSE
Xue et al. (*39*)	310 brain across ^{18}F-FDG, ^{18}F-FET, ^{18}F-florbetapir PET	cGAN	Conventional content; voxelwise
Jang et al. (*33*)	44 whole-body ^{18}F-FDG, 40 whole-body ^{18}F-ACBC, 10 whole-body ^{18}F-DCFPyL, 18 whole-body ^{68}Ga-DOTATATE	Spach transformer	Charbonnier
Song et al. (*40*)	20 simulated brain ^{18}F-FDG PET; 30 clinical brain ^{18}F-FDG PET + T1 MRI	Very deep superresolution CNN	MAE
Song et al. (*41*)	20 simulated brain ^{18}F-FDG PET; 30 clinical brain ^{18}F-FDG PET	Self-supervised superresolution (CycleGAN)	Adversarial; cycle consistency; total variation
Sanaat et al. (*42*)	50 brain ^{18}F-FDG, 50 brain ^{18}F-flortaucipir, 36 brain ^{18}F-flutemetamol, 76 brain ^{18}F-fluoro-DOPA + T1 MRI	CycleGAN	Adversarial
Sanaat et al. (*43*)	100 brain ^{18}F-FDG, 100 ^{18}F-flortaucipir, 100 brain ^{18}F-flutemetamol	CycleGAN	Adversarial
Azimi et al. (*44*)	160 brain ^{18}F-FDG PET/CT	Attention-based network (U-Net)	MSE
Mehranian et al. (*45*)	273 whole-body ^{18}F-FDG PET	3D residual U-Net	MSE

MSE = mean-squared error; MAE = mean absolute error; ^{18}F-ACBC = 1-amino-3-^{18}F-fluorocyclobutane-1-carboxylic acid.

datasets. In a joint denoising and partial volume correction framework, a cycleGAN variant was also trained in supervised mode to generate standard-count partial-volume–corrected PET images from low-count inputs for 3 tracers (^{18}F-FDG, ^{18}F-flortaucipir, and ^{18}F-flutemetamol) (*43*). A similar concept was also presented using a U-Net–based model for joint denoising and partial volume correction (*44*). Time-of-flight PET imaging has been shown to improve the image signal-to-noise ratio significantly. Although most denoising works focus on reducing scan time or tracer dose, one AI-based denoising approach computed time-of-flight–quality images from non–time-of-flight PET image inputs (*45*).

SPECT IMAGE ENHANCEMENT

Several recent efforts that use AI for SPECT image enhancement are highlighted in Table 2. Similar to PET reports, most papers on SPECT image enhancement focus on image denoising models, which generate standard- or high-count SPECT images from noisy, low-count inputs. One of the earliest applications of AI for SPECT myocardial perfusion imaging (MPI) using a ^{99m}Tc-sestamibi rest-and-stress protocol involved a 3D convolutional autoencoder to map low-count SPECT images (1/8 and 1/16 of standard) to standard-count images (*46*). An extension of this work reported comparisons of several convolutional autoencoder architectures and evaluated the denoising model for the clinical task of perfusion-defect detection at several successively reduced dose levels (1/2, 1/4, 1/8, and 1/16 of standard count) (*47*). The paper also showed that dose-specific models outperformed a one-size-fits-all model trained using inputs at different noise levels. Pix2Pix, a cGAN architecture, was applied to ^{99m}Tc-sestamibi stress scans with reduced counts (7/10 to 1/10 of standard) and led to improved denoising performance relative to convolutional autoencoders and conventional gaussian and Butterworth filters (*48*). A dual-gated (cardiac and respiratory) SPECT MPI study suggested that using a patient's own dataset for training a cGAN architecture was superior to conventional training based on cross-patient data (*49*). The cGAN led to the lowest noise level but also exhibited the poorest defect detection performance compared with CNN and U-Net. One recent study provided a theoretical framework for assessing signal detection accuracy for AI-based SPECT denoising and demonstrated the utility of virtual clinical trials in the evaluation of AI-based approaches (*50*). This study highlighted discrepancies between image-based and task-based evaluation outcomes and stressed the significance of task-based objective evaluation for denoising SPECT images.

Although most denoising studies focus on the reduction of the radiotracer dose, several studies specifically focus on the reduction of scan duration. One SPECT MPI study compared the denoising performance of a CNN with residual learning for half-time versus half-projection datasets (i.e., halving the scan duration vs. halving the number of projection views) and reported stronger performance for the former (*51*). Another study focused on scan-time reduction for pediatric patients with kidney disease imaged using ^{99m}Tc-dimercaptosuccinic acid and showed that a 3D residual U-Net for denoising led to good diagnostic performance for the detectability of defects in the renal cortex despite a reduction in the scan time (*52*). By using a U^{2}Net, a novel 2-layer nested U-shaped structure with a residual U-block that effectively captures contextual information on different scales, 1 study demonstrated good lesion detectability performance for ultra-high-speed (1/7 of standard scan time) SPECT bone imaging using ^{99m}Tc-methyl diphosphonate (*53*). Notably, the model incorporated a lesion attenuation loss function to enhance its accuracy at generating SUV measures for lesion regions.

TABLE 2
Summary of Deep-Learning Techniques for SPECT Image Enhancement

Paper	Data details	Architecture	Loss function
Ramon et al. (*46*)	930 cardiac torso ^{99m}Tc-sestamibi	3D convolutional autoencoder	MSE
Ramon et al. (*47*)	1,052 cardiac torso ^{99m}Tc-sestamibi	Convolutional autoencoder, CNN	MSE
Sun et al. (*48*)	100 simulated; 20 cardiac torso clinical ^{99m}Tc-sestamibi	Pix2Pix GAN	MAE; adversarial
Sohlberg et al. (*49*)	93 cardiac torso ^{99m}Tc-tetrososmin	CNN, residual network, U-Net, cGAN	MSE
Yu et al. (*50*)	4,800 simulated	CNN	MSE
Shiri et al. (*51*)	363 cardiac torso ^{99m}Tc-sestamibi	Deep residual neural network	MSE
Lin et al. (*52*)	112 cardiac torso ^{99m}Tc-DMSA	3D residual U-Net	MSE
Pan et al. (*53*)	20 cardiac torso ^{99m}Tc-MDP SPECT/CT	Lesion-attention weighted U^{2}Net	MAE; structural similarity index
Liu et al. (*54*)	895 cardiac torso ^{99m}Tc-sestamibi	Noise2Noise (U-Net)	MSE
Liu et al. (*55*)	1,050 cardiac torso ^{99m}Tc-sestamibi	3D-coupled UNet	MSE
Xie et al. (*56*)	28 cardiac ^{99m}Tc-RBC	Densely connected multidimensional dynamic U-Net	MAE; structural similarity index; Sobel operator; intramyocardial blood volume

MSE = mean-squared error; MAE = mean absolute error; DMSA = pentavalent dimercaptosuccinic acid; MDP = methyl diphosphonate; RBC = red blood cell.

Some SPECT MPI image denoising efforts leverage recent advances in unsupervised learning. One such effort uses Noise2-Noise, a deep-learning framework for denoising that is trained without clean images but requires 2 noisy realizations of a ground-truth image, one used as the input and the other as the training target (*54*). The study used a coupled U-Net architecture that incorporates multiple U-Nets to reuse feature maps within the network. To evaluate the detection performance for perfusion defects at multiple contrast levels, the authors used a bootstrap procedure to generate multiple noise realizations from list-mode clinical acquisitions. Furthermore, the study was extended to quantify perfusion defect detection accuracy using receiver operating characteristics on a large training and validation dataset for SPECT MPI, which included 1,050 human subjects (*55*). Notably, the results revealed significant discrepancies between image-based and task-based evaluation and underscored the importance of task-based objective evaluation in SPECT image denoising. They demonstrated that pretraining with subsequent fine-tuning can meaningfully enhance the detectability of perfusion defects.

Applications of AI for SPECT image enhancement tasks other than denoising are still emerging. One paper proposed a segmentation-free partial volume correction approach for SPECT MPI, which uses a densely connected multidimensional dynamic network that allows adaptive adjustment of convolutional kernels after training (*56*). Importantly, the approach incorporated intramyocardial blood volume into the loss function to add clinical relevance to the generated images.

DISCUSSION

We have presented here a summary of recent progress in AI-based PET and SPECT image enhancement. AI-based techniques have shown great promise in enhancing image quality by reducing levels of noise and blur and have shown clinical promise in many task-based evaluation studies. Importantly, many studies have suggested that AI-based denoising approaches can reduce radiotracer dose and scan times without sacrificing diagnostic accuracy. AI-based models have also been more successful than their predecessors at combining multimodal information (e.g., using CT or MRI for PET or SPECT image enhancement).

AI-based image enhancement is of great clinical significance. Denoising approaches can lead to reductions in radiotracer dose or scan duration. Whereas the former reduces patient radiation exposure and addresses challenges arising from radionuclide shortages, the latter enhances patient comfort, increases scanning throughput, and reduces motion artifacts that could compromise diagnostic accuracy. Several of the cited papers show that denoising could improve both image quantitation and lesion detectability in addition to improving scan logistics. Deblurring approaches can mitigate partial volume effects that can compromise the accuracy of quantitative image-based metrics such as SUV ratios computed from small regions of interest. This is of particular importance in the imaging of neurodegenerative diseases, where image-based quantitative metrics from small anatomic targets could have diagnostic or prognostic value. The growing clinical relevance of AI-based image enhancement is underscored by the availability of U.S. Food and Drug Administration–approved vendor-neutral commercial software such as SubtlePET (Subtle Medical) for AI-based denoising, as exemplified in a study using SubtlePET's CNN to enhance low-count scans to diagnostic quality (*57*).

Despite the field's initial focus on supervised learning techniques that require paired clean and corrupt images for model training, an array of promising unsupervised or weakly supervised alternatives has emerged in the PET and SPECT fields in recent years. Most of these approaches either use only corrupt images for training or use corrupt inputs with unpaired training targets. These methods are attractive because of their easy applicability to most clinical datasets when ground-truth images for training are not available. However, they tend to produce inferior image quality and are often slower than their supervised counterparts. Thus, there is active research interest in further developing unsupervised approaches.

Although AI-based methods have consistently outperformed traditional approaches in terms of image-based figures of merit, whether the improved image quality leads to a tangible clinical benefit remains a topic of continued research and investigation. Accordingly, there is an increased focus in the current literature on task-based objective clinical evaluation of these approaches. Interestingly, several of the noted approaches for both PET and SPECT have incorporated clinical metrics (such as amyloid positivity or lesion detectability) into their loss functions to encourage clinically meaningful solutions. Furthermore, the incorporation of multimodal fusion, which integrates information from different imaging modalities such as CT and MRI, holds promise for improving diagnostic accuracy.

Although a sizable fraction of existing research is focused on ^{18}F-FDG PET and SPECT MPI, applications to other tracers are rapidly expanding. Transfer-learning strategies are facilitating the application of data-hungry AI models to smaller datasets for newer radiotracers, which can enable model fine-tuning with limited data using cross-tracer pretraining (*39*,*58*,*59*). Unsupervised models have also leveraged transfer-learning paradigms using a combination of population-level pretraining and individual fine-tuning (*36*). Transfer learning has also aided cross-scanner image-mapping strategies that are enabling purely software-based generation of higher-resolution images mimicking the image characteristics of state-of-the-art scanner models (*41*).

Although most clinical applications of image enhancement techniques are currently aimed at diagnostics, given the growing significance of radiopharmaceutical therapy, clinical applications of AI-based image enhancement could span beyond diagnostics, as image-quality improvements due to AI could potentially lead to more accurate image-based dosimetry. Given the privacy and security concerns surrounding health care, there is also a growing interest in federated learning approaches for image enhancement, wherein code sharing can circumvent the many challenges associated with data sharing, thus enabling the creation of robust models trained and validated over multiple sites and data sources.

CONCLUSION

AI methods have shown great promise in improving the quality and utility of PET and SPECT images. From traditional CNNs to more advanced GANs and transformer networks, deep-learning architectures have been applied to a range of clinical applications. Although encouraging results based on both image-domain and task-based evaluations have been reported, several roadblocks linger for the clinical translation of AI tools. Accordingly, there is a pressing need for large disease-specific datasets, standardized evaluation metrics, and integration of image enhancement tools with

existing clinical workflows. The future of AI in PET and SPECT imaging holds great potential to improve diagnostic accuracy, enable novel clinical applications, and ultimately benefit patients.

DISCLOSURE

This research was supported by grants R01AG072669 and R03AG070750. No other potential conflict of interest relevant to this article was reported.

REFERENCES

1. Hicks RJ, Hofman MS. Is there still a role for SPECT–CT in oncology in the PET–CT era? *Nat Rev Clin Oncol.* 2012;9:712–720.
2. Hooker JM, Carson RE. Human positron emission tomography neuroimaging. *Annu Rev Biomed Eng.* 2019;21:551–581.
3. Best SRD, Haustrup N, Pavel DG. Brain SPECT as an imaging biomarker for evaluating effects of novel treatments in psychiatry: a case series. *Front Psychiatry.* 2022;12:713141.
4. Palestro CJ. Radionuclide imaging of musculoskeletal infection: a review. *J Nucl Med.* 2016;57:1406–1412.
5. Garcia EV, Slomka P, Moody JB, Germano G, Ficaro EP. Quantitative clinical nuclear cardiology, part 1: established applications. *J Nucl Cardiol.* 2020;27:189–201.
6. Slomka PJ, Moody JB, Miller RJH, Renaud JM, Ficaro EP, Garcia EV. Quantitative clinical nuclear cardiology, part 2: evolving/emerging applications. *J Nucl Cardiol.* 2021;28:115–127.
7. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–945.
8. Zhou T, Ruan S, Canu S. A review: deep learning for medical image segmentation using multi-modality fusion. *Array.* 2019;3–4:100004.
9. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. arXiv website. https://arxiv.org/abs/1505.04597. Published May 18, 2015. Accessed October 24, 2023.
10. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. arXiv website. https://arxiv.org/abs/1606.06650. Published June 21, 2016. Accessed October 24, 2023.
11. Zhang L, Zhang W, Japkowicz N. Conditional-UNet: a condition-aware deep model for coherent human activity recognition from wearables. arXiv webstie. https://arxiv.org/abs/2004.09376. Published April 15, 2020. Accessed October 24, 2023.
12. Tang Z, Peng X, Li K, Metaxas DN. Towards efficient U-Nets: a coupled and quantized approach. *IEEE Trans Pattern Anal Mach Intell.* 2020;42:2038–2050.
13. Ulyanov D, Vedaldi A, Lempitsky V. Deep image prior. *Int J Comput Vis.* 2020;128:1867–1888.
14. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger KQ, eds. *Advances in Neural Information Processing Systems.* Vol 27. Curran Associates, Inc.; 2014:2672–2680.
15. Mirza M, Osindero S. Conditional generative adversarial nets. arXiv website. https://arxiv.org/abs/1411.1784. Published November 6, 2014. Accessed October 24, 2023.
16. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Paper presented at: 2017 IEEE International Conference on Computer Vision (ICCV); October 22–29, 2017; Venice, Italy.
17. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. arXiv website. https://arxiv.org/abs/1706.03762. Published June 12, 2017. Accessed October 24, 2023.
18. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. arXiv website. https://arxiv.org/abs/2006.11239. Published June 19, 2020. Accessed October 24, 2023.
19. Lehtinen J, Munkberg J, Hasselgren J, et al. Noise2Noise: learning image restoration without clean data. arXiv website. https://arxiv.org/abs/1803.04189. Published March 12, 2018. Accessed October 24, 2023.
20. Erhan D, Manzagol P-A, Bengio Y, Bengio S, Vincent P. The difficulty of training deep architectures and the effect of unsupervised pre-training. Paper presented at: Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics; April 16-18, 2009; Clearwater Beach, FL.
21. Xiang L, Qiao Y, Nie D, An L, Wang Q, Shen D. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. *Neurocomputing.* 2017;267:406–416.
22. Xu J, Gong E, Pauly J, Zaharchuk G. 200x low-dose PET reconstruction using deep learning. arXiv website. https://arxiv.org/abs/1712.04119. Published December 12, 2017. Accessed October 24, 2023.
23. Wang Y, Yu B, Wang L, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *Neuroimage.* 2018;174:550–562.
24. Ouyang J, Chen KT, Gong E, Pauly J, Zaharchuk G. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. *Med Phys.* 2019;46:3555–3564.
25. Wang Y, Zhou L, Yu B, et al. 3D auto-context-based locality adaptive multi-modality GANs for PET synthesis. *IEEE Trans Med Imaging.* 2019;38:1328–1339.
26. Schaefferkoetter J, Yan J, Ortega C, et al. Convolutional neural networks for improving image quality with noisy PET data. *EJNMMI Res.* 2020;10:105.
27. Spuhler K, Serrano-Sosa M, Cattell R, DeLorenzo C, Huang C. Full-count PET recovery from low-count image using a dilated convolutional neural network. *Med Phys.* 2020;47:4928–4938.
28. Xue H, Teng Y, Tie C, et al. A 3D attention residual encoder–decoder least-square GAN for low-count PET denoising. *Nucl Instrum Methods Phys Res A.* 2020;983:164638.
29. Zhao K, Zhou L, Gao S, et al. Study of low-dose PET image recovery using supervised learning with cycleGAN. *PLoS One.* 2020;15:e0238455.
30. Gong Y, Shan H, Teng Y, et al. Parameter-transferred Wasserstein generative adversarial network (PT-WGAN) for low-dose PET image denoising. *IEEE Trans Radiat Plasma Med Sci.* 2021;5:213–223.
31. Gong K, Johnson KA, El Fakhri G, Li Q, Pan T. PET image denoising based on denoising diffusion probabilistic models. arXiv website. https://arxiv.org/abs/2209.06167. Published September 13, 2022. Accessed October 24, 2023.
32. Zhang L, Xiao Z, Zhou C, et al. Spatial adaptive and transformer fusion network (STFNet) for low-count PET blind denoising with MRI. *Med Phys.* 2022;49:343–356.
33. Jang S-I, Pan T, Li Y, et al. Spach transformer: spatial and channel-wise transformer based on local and global self-attentions for PET image denoising. arXiv website. https://arxiv.org/abs/2209.03300. Published September 7, 2022. Accessed October 24, 2023.
34. Cui J, Gong K, Guo N, et al. PET image denoising using unsupervised deep learning. *Eur J Nucl Med Mol Imaging.* 2019;46:2780–2789.
35. Cui J, Gong K, Guo N, et al. Populational and individual information based PET image denoising using conditional unsupervised learning. *Phys Med Biol.* 2021;66:155001.
36. Song T-A, Yang F, Dutta J. Noise2Void: unsupervised denoising of PET images. *Phys Med Biol.* 2021;66:214002.
37. Liu Q, Liu H, Mirian N, et al. A personalized deep learning denoising strategy for low-count PET images. *Phys Med Biol.* 2022;67:145014.
38. Zhou B, Miao T, Mirian N, et al. Federated transfer learning for low-dose PET denoising: a pilot study with simulated heterogeneous data. *IEEE Trans Radiat Plasma Med Sci.* 2023;7:284–295.
39. Xue S, Guo R, Bohn KP, et al. A cross-scanner and cross-tracer deep learning method for the recovery of standard-dose imaging quality from low-dose PET. *Eur J Nucl Med Mol Imaging.* 2022;49:1843–1856.
40. Song T-A, Chowdhury SR, Yang F, Dutta J. Super-resolution PET imaging using convolutional neural networks. *IEEE Trans Comput Imaging.* 2020;6:518–528.
41. Song T-A, Chowdhury SR, Yang F, Dutta J. PET image super-resolution using generative adversarial networks. *Neural Netw.* 2020;125:83–91.
42. Sanaat A, Shooli H, Böhringer AS, et al. A cycle-consistent adversarial network for brain PET partial volume correction without prior anatomical information. *Eur J Nucl Med Mol Imaging.* 2023;50:1881–1896.
43. Sanaat A, Boehringer A, Ghavabesh A, et al. Deep-PVC: a deep learning model for synthesizing full-dose partial volume corrected PET images from low-dose images. Paper presented at: 2021 Virtual IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC); October 16–23, 2021; Piscataway, NJ.
44. Azimi M-S, Kamali-Asl A, Ay M-R, Arabi H, Zaidi H. A novel attention-based convolutional neural network for joint denoising and partial volume correction of low-dose PET images. Paper presented at: 2021 Virtual IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC); October 16–23, 2021; Piscataway, NJ.
45. Mehranian A, Wollenweber SD, Walker MD, et al. Deep learning–based time-of-flight (ToF) image enhancement of non-ToF PET scans. *Eur J Nucl Med Mol Imaging.* 2022;49:3740–3749.
46. Ramon AJ, Yang Y, Pretorius PH, Johnson KL, King MA, Wernick MN. Initial investigation of low-dose SPECT-MPI via deep learning. Paper presented at: 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC); November 10–17, 2018; Sydney, NSW, Australia.

47. 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;39:2893–2903.
48. Sun J, Du Y, Li C, Wu T-H, Yang B, Mok GSP. Pix2Pix generative adversarial network for low dose myocardial perfusion SPECT denoising. *Quant Imaging Med Surg.* 2022;12:3539–3555.
49. Sohlberg A, Kangasmaa T, Constable C, Tikkakoski A. Comparison of deep learning-based denoising methods in cardiac SPECT. *EJNMMI Phys.* 2023;10:9.
50. Yu Z, Rahman MA, Laforest R, et al. Need for objective task-based evaluation of deep learning-based denoising methods: a study in the context of myocardial perfusion SPECT. *Med Phys.* 2023;50:4122–4137.
51. Shiri I, Sabet KA, Arabi H, et al. Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks. *J Nucl Cardiol.* 2021;28:2761–2779.
52. Lin C, Chang Y-C, Chiu H-Y, Cheng C-H, Huang H-M. Reducing scan time of paediatric ^{99m}Tc-DMSA SPECT via deep learning. *Clin Radiol.* 2021;76:315.e13–315.e20.
53. Pan B, Qi N, Meng Q, et al. Ultra high speed SPECT bone imaging enabled by a deep learning enhancement method: a proof of concept. *EJNMMI Phys.* 2022;9:43.
54. 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;48:156–168.
55. Liu J, Yang Y, Wernick MN, Pretorius PH, Slomka PJ, King MA. Improving detection accuracy of perfusion defect in standard dose SPECT-myocardial perfusion imaging by deep-learning denoising. *J Nucl Cardiol.* 2022;29:2340–2349.
56. Xie H, Liu Z, Shi L, et al. Segmentation-free PVC for cardiac SPECT using a densely-connected multi-dimensional dynamic network. *IEEE Trans Med Imaging.* 2023;42:1325–1336.
57. Chaudhari AS, Mittra E, Davidzon GA, et al. Low-count whole-body PET with deep learning in a multicenter and externally validated study. *NPJ Digit Med.* 2021;4:127.
58. Chen X, Hendrik Pretorius P, Zhou B, et al. Cross-vender, cross-tracer, and cross-protocol deep transfer learning for attenuation map generation of cardiac SPECT. *J Nucl Cardiol.* 2022;29:3379–3391.
59. Liu H, Wu J, Lu W, Onofrey JA, Liu Y-H, Liu C. Noise reduction with cross-tracer and cross-protocol deep transfer learning for low-dose PET. *Phys Med Biol.* 2020;65:185006.