FBP-Net for direct reconstruction of dynamic PET images

Bo Wang\textsuperscript{1} and Huafeng Liu\textsuperscript{1,∗}

\textsuperscript{1} State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, 310027 Hangzhou, People's Republic of China

∗ Author to whom any correspondence should be addressed

E-mail: liuhf@zju.edu.cn

Keywords: reconstruction, FBP, denoiser, dynamic PET

Abstract

Dynamic positron emission tomography (PET) imaging can provide information about metabolic changes over time, used for kinetic analysis and auxiliary diagnosis. Existing deep learning-based reconstruction methods have too many trainable parameters and poor generalization, and require mass data to train the neural network. However, obtaining large amounts of medical data is expensive and time-consuming. To reduce the need for data and improve the generalization of network, we combined the filtered back-projection (FBP) algorithm with neural network, and proposed FBP-Net which could directly reconstruct PET images from sinograms instead of post-processing the rough reconstruction images obtained by traditional methods. The FBP-Net contained two parts: the FBP part and the denoiser part. The FBP part adaptively learned the frequency filter to realize the transformation from the detector domain to the image domain, and normalized the coarse reconstruction images obtained. The denoiser part merged the information of all time frames to improve the quality of dynamic PET reconstruction images, especially the early time frames. The proposed FBP-Net was performed on simulation and real dataset, and the results were compared with the state-of-art U-net and DeepPET. The results showed that FBP-Net did not tend to overfit the training set and had a stronger generalization.

1. Introduction

Positron emission tomography (PET) provides biological metabolism information \textit{in vivo}. In organisms, metabolic abnormalities occur earlier than changes in tissue structure (Quigley \textit{et al} 2011), so PET is conducive to detect and treat lesions earlier, and relieves patients’ pains. Dynamic PET imaging divides the overall scan time into several short time frames, and data are collected independently in each time frame, so it can provide both temporal and spatial information. This spatiotemporal information can be used for further quantitative analysis, such as estimation of kinetic parameters (Pan \textit{et al} 2017, Sari \textit{et al} 2018), dual tracer separation (Ruan and Liu 2017, Xu and Liu 2019, Tong \textit{et al} 2019), etc. Unfortunately, dynamic reconstruction is more challenging than static reconstruction. One the one hand, a single frame of dynamic scan data has shorter scan time and lower counts than static scan, resulting in higher level of noise in the reconstructed images than static reconstructed images, especially the early time frames. On the other hand, the dynamic scan obtains 4-dimensional data, including multiple slices and multiple time frames. The reconstruction workload of each time frame is equivalent to a static reconstruction, which leads to a long time for dynamic reconstruction. In this context, how to improve the quality and speed of dynamic PET reconstruction images is a subject worth studying.

Traditional reconstruction algorithms include filtered back-projection (FBP) algorithm and iterative reconstruction methods (Shepp and Vardi 1982, Green 1990, Wang \textit{et al} 2015, Sudarshan \textit{et al} 2018, Merlin \textit{et al} 2018, Wang \textit{et al} 2014). The FBP algorithm is fast, but suffers from high noise and ugly strip artifacts (Zeng 2012). Iterative algorithms use the system model and regularizations to obtain better reconstructed images than FBP algorithm (Green 1990, Wang \textit{et al} 2015, Sudarshan \textit{et al} 2018, Merlin \textit{et al} 2018, Wang \textit{et al} 2014). However, these iterative algorithms have several common problems. Firstly, the accuracy of the system
model affects the final reconstructed images. The geometry-based system model is simple but not accurate, while the accurate system model requires post sources to measure the point spread function (PSF) of the scanner, which is time-consuming and expensive (Qi et al 1998, Pan et al 2006). Secondly, the choice of regularization terms is an open question, and it is difficult to say which regularization is the best ( Häggström et al 2018). Finally, the calculation amount of each iteration of iterative algorithms is equivalent to one FBP operation, so iterative algorithms usually much slower than FBP algorithm (Bendriem and Townsend 2013).

Recently, deep learning has made great progress in problems such as detection and segmentation, and has been gradually applied to the field of medical imaging. Reconstructing PET images with deep learning has two advantages. One advantage is that deep neural network can learn features automatically from the data without manually selecting features for regularization in advance. Another advantage is that well-trained neural network only needs one forward pass to get the final results, so it is expected to be much faster than traditional iterative algorithms which need to perform projection and back-projection operations many times. Many scholars have used deep learning techniques for PET image reconstruction (Xu et al 2017, Gong et al 2018b, Kim et al 2018, Wang et al 2018, Yang et al 2018, Chen et al 2019, Gong et al 2018a, Häggström et al 2018, Kaplan and Zhu 2019, Liu et al 2019, Ouyang et al 2019, Yokota et al 2019). Most of these works treated neural networks as post-processing tools for enhancing low-quality reconstruction maps obtained by traditional methods (Chen et al 2019, Ouyang et al 2019, Wang et al 2018, Xu et al 2017, Yang et al 2018, Kim et al 2018, Kaplan and Zhu 2019). Although these methods are simple to implement, the final reconstruction results are easily affected by traditional reconstruction algorithms. Furthermore, the traditional iterative algorithms are not fast, and the post-processing of the neural networks is bounded to further increase the reconstruction time. Liu et al 2019 and Häggström et al 2018 both adopted neural networks to obtained images directly from sinograms. Their methods do not consider known relationship between sinogram and reconstruction images, such as inverse radon transform, system matrix and so on. The training of networks is completely dependent on data, and needs large amount of training data to avoid over-fitting. Affected by factors such as patient privacy and collection costs, it is unrealistic to obtain a large amount of real medical data. Yokota et al 2019 and Gong et al 2018a, Gong et al 2018b used U-nets as deep image prior or regularizations, but their methods are based on an iterative framework, which means that their methods are as slow as the iterative algorithms. In addition, there are several regularization coefficients that need to be determined manually in advance, and the choice of these regularization coefficients might have a crucial impact on the final results.

The methods based entirely on neural networks are fast, but their generalizations are not good, and they can only handle the situation where the training set and the test set are similar. For example, the network trained on the brain dataset can only be used to reconstruct the brain images instead of images of other organs. Besides, large amount of data are required to train the networks to avoid over-fitting, but data are scarce and precious in the field of medical imaging. As for methods based on iterative framework, they are time-consuming, because they need to perform forward projection and back-projection many times, and they also need to carefully choose regular coefficients to ensure the quality of the reconstructed images. These deep learning methods are either poorly generalized and require a lot of training samples, or the speed is as slow as iterative algorithms and require artificial selection of regularization parameters. The purpose of this work is to propose a method that has a strong generalization and low time cost, and can automatically obtain the final results without manual adjustments.

In this work, we combined FBP algorithm with neural network, and proposed FBP-Net to reconstruct the dynamic PET images directly from sinograms. Compared with the existing method, it has the following differences: firstly, unlike post-processing method, FBP-Net was a direct reconstruction method that combined traditional method and neural network. Secondly, FBP-Net was expected to have a speed comparable to the FBP algorithm, much faster than the deep learning methods based on iterative framework or traditional iterative methods, because it was based on framework of analytical algorithm. Finally, FBP-Net had stronger generalization, and lower requirements for the number of samples than methods based entirely on neural networks, such as U-net and DeepPET. The proposed FBP-Net was validated on simulation and real datasets. Experimental results showed that the proposed method was highly generalizable, and had advantages over state-of-art U-net and DeepPET in the case of a small number of samples.

2. Methods

The proposed FBP-Net consisted of two parts: FBP part and denoiser part. The FBP part of the proposed FBP-Net kept the parameters of back-projection unchanged, and only adaptively learned the frequency filter. The denoiser part was a convolutional neural network (CNN) that merged information from different time frames to improve reconstruction quality.
There were several novelties in this work: firstly, the proposed FBP-Net combined traditional methods and neural networks, and could directly reconstruct dynamic PET images from sinograms. Secondly, it not only had a strong generalization but also had a speed comparable to FBP algorithm. Finally, it had a small number of training parameters, and performed well in the case of a small number of samples.

2.1. Theoretical basis

In PET imaging, the positrons emitted by the tracer annihilate with the negative electrons in the organism, and emit a pair of photons flying back to back with an energy of 511 KeV. Thus, the beams detected by the PET scanner have a parallel geometry. The relationship between tracer activity images and projection data (sinograms) is discreted radon transform,

\[ y = Gx. \]  

(1)

\( y \in \mathbb{R}^{AB} \) denotes the sinogram, where \( A \) and \( B \) are the number of detector bins and angles respectively.

\( x \in \mathbb{R}^M \) denotes tracer activity distribution image, where \( M \) represents the number of pixels in the image.

\( G \in \mathbb{R}^{AB \times M} \) denotes the projection matrix.

Given the measured sinogram \( y \) and projection matrix \( G \), the objective function of the reconstruction problem is expressed as,

\[ \min_x \| Gx - y \|^2_2. \]  

(2)

The theoretical solution of this optimization problem is,

\[ x = (G^T G)^{-1} G^T y, \]  

(3)

where \( G^T \) denotes back-projection matrix, and \( (G^T G)^{-1} \) is an ideal Ramp filter. This equation can be interpreted as back-projection first and then filtering, which is equivalent to filtering before back-projection, the FBP algorithm.

2.2. Data normalization

For dynamic PET imaging, the detected data are affected by many factors such as injection dose, detection efficiency, individual physiological differences and so on. In order to ensure the generalization of the method, it is necessary to perform normalization before training, validation or testing. In dynamic PET images, the relative concentration between different time frames is meaningful, and the change in radioactive concentration of a voxel or region of interest (ROI) can be drawn as a curve, the time activity curve (TAC). This curve can be used for kinetic analysis, estimation of macro- or micro-parameters, separation of dual tracers, etc. Based on this, we tried to preserve the shapes of TACs when normalizing. We regarded the reconstruction images of the same slice at all time frames as a group, and each group contained \( T \)-frame images of the same slice. The normalization was performed at the group level, and all \( T \)-frame images in a group were divided by the maximum value of pixels in this group to be normalized to 0 to 1. Since the FBP part was designed by the FBP algorithm, the sinograms were not normalized before they were inputted into the FBP-Net. However, there was a normalization layer at the end of the FBP part, and this normalization layer normalized the coarse reconstruction images to 0 to 1 to avoid enter big values into the denoiser part.

2.3. Network structure

The proposed FBP-Net was designed according to traditional FBP algorithm. It consisted of two parts, the FBP part and the denoiser part, as shown in figure 1. Supposed that the dynamic PET data contained \( T \) time frames, and each sinogram had \( A \) angles and \( B \) bins, and the size of reconstruction image was \( \sqrt{M} \times \sqrt{M} \).

The sinograms of \( T \) frames were regarded as a sample and were inputted to the FBP-Net, while the outputs of the FBP-Net were the corresponding \( T \)-frame reconstruction images.

2.3.1. Filtered back-projection (FBP) part

The FBP part realized the transformation from the detector domain to the image domain, that was, reconstructing rough images from sinograms. The FBP part consisted of four steps: filtering, back-projection, removing negative values and normalization.

The theoretical basis of FBP algorithm is the central slice theorem. Due to uneven density of slices in the Fourier domain, unfiltered back-projection images are blurred. So ramp filters with windows are commonly used to correct this blurring effect (Zeng 2010). In theory, ramp filtering can be achieved by multiplication in the frequency domain or convolution in the spatial domain. In practical applications, ramp filtering is often implemented in the frequency domain, since the ramp filter has an infinite impulse response (Würfl et al
2018). Traditional FBP algorithm adopts a pre-determined filter for sinogram data. If the filter could be adaptively adjusted according to the data, the quality of reconstructed images might be improved (Pelt and Batenburg 2014). Inspired by this, we made the filter in the FBP part learnable. We tried filtering both in frequency domain and in space domain, and found that frequency domain filters were less sensitive to initial values than spatial filters. Thus, we finally selected frequency domain filtering. Supposed that the number of angles and detector bins in the sinogram were $A$ and $B$ respectively, the length of the filter was $2 \lceil \log_2(2B - 1) \rceil$. Since the frequency filters in the FBP algorithm are generally non-negative, we used the rectified linear units (ReLU) for the filter, before it was multiplied by the Fourier transform of sinogram data. To avoid frequency alias, the dimension of the detector bins of the sinogram was supplemented with 0 to expand its size to $2 \lceil \log_2(2B - 1) \rceil$. The zero-padding sinograms first underwent fast Fourier transform (FFT), and then they were multiplied with the frequency domain filter to perform filtering. Finally, the filled part was removed to obtain the filtered sinograms.

For the back-projection layer, the simplest and most straightforward approach is to turn it into a fully connected layer, because each element in the sinogram contains information about many pixels in the reconstructed image. However, this approach requires at least $M \times A \times B$ learnable parameters. Too many parameters will not only increase the requirement of the amount of data, but also be prone to cause over-fitting. He et al used a sinusoidal back-projection (SBP) layer to simulate the back-projection (He and Ma 2019). Although this approach was very flexible, the number of parameters to be trained at this layer was $M \times B$. The number of parameters will still increase significantly as the image becomes larger. In Whiteley and Gregor (2019), sinograms were encoded to reduce the number of parameters in the radon inversion layer. However, this method has little significance in PET, because the numbers of bins and angles in PET data are much smaller than CT data. Reducing the dimensions of the data is bound to lose some information in PET data. Using a prior operations in neural networks has potential to significantly reduce the number of parameters (Maier et al 2018). Inspired by this, we fixed the parameters in the back-projection layer, making them unlearnable. In order to improve the accuracy of back-projection, we did not use the nearest neighbor interpolation, but linear interpolation, that was, the values of back-projection were calculated by the weighted sum of the two nearest bins.

After back-projection, the PET data changed from the detector domain to the image domain, but the reconstruction images at this time usually contained negative values. These negative values are often artifacts and need to be removed. The activation function ReLU of the neural network is as follows,

$$ F_{ReLU}(x) = \max(0, x). $$

This function can turn negative values into zeros, so we added it behind the back-projection layer to remove artifacts. After this, the rough reconstruction images were obtained.

In order to keep the outputs of the FBP part in a certain range, we introduced a normalization layer. Let $S$ denoted the $T$-frame non-negative reconstruction maps of the same slice obtained in the previous step, then
the output of the normalization layer, $F_{\text{norm}}(\cdot)$, could be expressed as,

$$F_{\text{norm}}(S) = \frac{S - \min(S)}{\max(S) - \min(S)},$$  \hspace{1cm} (5)

where $\min(S)$ and $\max(S)$ are the maximum and minimum values of $S$. With this formula, $S$ was normalized to 0 to 1 according to its maximum and minimum.

### 2.3.2. Denoiser

As mentioned earlier, directly learning back-projection process would lead to too many training parameters, so our FBP part only learned filters, and kept back-projection parameters fixed. But in this way, FBP part was unlikely to obtain high-quality reconstructed images. In order to improve the quality of the reconstructed images from FBP part, we introduced a denoiser part. The denoiser was inspired by denoising convolutional neural network (DnCNN) which combined convolutional neural network with residual learning to remove Gaussian noise in the natural images (Zhang et al. 2017). In our work, the denoiser was used for post-processing, that was, fusing the information from different time frames to improve the quality of dynamic reconstruction images.

When FBP part reconstructed images, the information in different time frames was independent, and the temporal and spatial correlations between different time frames were not given a consideration. If the motion is ignored, the anatomical structure of the scanned object is time-invariant. The main differences among reconstruction images in different time frames are bright and dark. For example, the blood pools of the heart are brighter in the early time frame, while the myocardium and the right ventricle are brighter in the later time frame. The photons of early time frames are few and sparsely distributed, resulting in reconstruction images with incomplete structures. Using the information of the late time frame might be able to help the early time frame recover some incomplete structural information. So the $T$-frame rough reconstruction images of the same slice obtained by the FBP part were considered as a sample and were inputted to the denoiser. The $T$ time frames of a sample were regarded as $T$ channels. After the first 2D convolutional layer, the information of these $T$ time frames was fused together. This fused information was propagated layer by layer through the feature maps, and it might be helpful for recovering early time frame using late time frames.

The denoiser contained eight convolutional layers with activation function Leaky-ReLU. The first seven convolutional layers had 64 filters of size $3 \times 3$, and batch normalization was used to speed up training. The eighth layer had $T$ filters of size $3 \times 3$, and the output of this layer was multiplied by -1, and then added to the output of the FBP part to remove noise and artifacts from the output of FBP part. The difference between the 1st and 8th convolutional layers might have negative values, which were meaningless in practical situations. In order to avoid negative values, and ensure the final reconstruction image within a certain range, an absolute value function and a normalization layer were added after the 8th convolution layer. This normalization layer was the same as the normalization layer in the FBP part.

### 2.3.3. Loss

The loss function used here was mean square error (MSE) between the estimated clear images and ground truths,

$$\text{loss} = \frac{1}{NMT} \sum_{i=1}^{N} \sum_{j=1}^{T} \sum_{m=1}^{M} (\hat{x}_{ijm} - x_{ijm})^2$$  \hspace{1cm} (6)

where $N$, $T$ and $M$ are the number of samples, the number of time frames and the number of pixels in a reconstructed image, respectively. $\hat{x}_{ijm}$ and $x_{ijm}$ are the $m$th pixel of the $j$th frame of $i$th sample by FBP-Net and the ground truth of the $m$th pixel of the $j$th frame of $i$th sample. During the training process, the loss function was minimized to adjust all learnable parameters in the FBP part and denoiser.

### 3. Data preparation

#### 3.1. Simulation datasets

The $128 \times 128$ 2D phantom used in the simulation was from Zubal phantom (Zubal et al. 1994), and a tumor was added. The phantom was divided into six regions of interest (ROIs) as shown in figure 2(a). Two simulation datasets were generated, and each dataset included 400 sets of data with different $k$ parameters. The first dataset used the phantom in figure 2(a), with only one shape. For the second dataset, we randomly rotated, translated, and scaled the phantom in figure 2(a), and obtained 400 new phantoms with different shapes. These phantoms would later be filled with time activity curves (TACs) to obtain true tracer activity images.
Figure 2. (a) The Zubal phantom used in the simulation. (b) Examples of time activity curves (TACs) of different regions of interest (ROIs).

Table 1. Kinetic parameters of the compartmental model in the simulation. $V_b$ denotes blood volume ratio.

| ROI         | $K_1$   | $k_2$   | $k_3$   | $k_4$   | $V_b$ |
|-------------|---------|---------|---------|---------|-------|
| Gray matter | 1       | 0.100   | 0.140   | 0.170   | 0.013 | 0.103 |
| White matter| 2       | 0.050   | 0.110   | 0.100   | 0.006 | 0.266 |
| Caudate     | 3       | 0.120   | 0.170   | 0.190   | 0.016 | 0.101 |
| Putamen     | 4       | 0.130   | 0.160   | 0.170   | 0.010 | 0.092 |
| Thalamus    | 5       | 0.130   | 0.160   | 0.140   | 0.012 | 0.152 |
| Tumor       | 6       | 0.110   | 0.100   | 0.150   | 0.015 | 0.173 |

The simulated tracer was $[^{18}\text{F}]$-labeled fluorodeoxyglucose ($[^{18}\text{F}]$-FDG), and a three compartmental model was adopted to simulate the dynamic characteristics of $[^{18}\text{F}]$-FDG. Feng’s input function was used to drive this model (Feng et al. 1993),

$$C_p^{[^{18}\text{F}]\text{FDG}}(t) = (A_1 t - A_2 - A_3) e^{-\lambda_1 t} + A_2 e^{-\lambda_2 t} + A_3 e^{-\lambda_3 t}$$  \hspace{1cm} (7)

where $A_1, A_2, A_3, \lambda_1, \lambda_2, \lambda_3$ were 851.1225 $\mu$Ci/ml, 20.8113 $\mu$Ci/ml, 21.8798 $\mu$Ci/ml, 4.1339 min$^{-1}$, 0.0104 min$^{-1}$, 0.1191 min$^{-1}$. In order to generate more data, $k$ parameters in the compartmental model were selected randomly from Gaussian distribution. The means of $k$ parameters are shown in table 1, and standard deviations were set to 10% of the means (Cheng et al. 2015). For the first dataset with a single-shape phantom and the second dataset with phantoms of different shapes, 400 sets of $k$ parameters were randomly generated respectively. The sampling protocol was set to $3 \times 60$ s, $9 \times 180$ s, $6 \times 300$ s. After setting $k$ parameters, plasma input function and sampling protocol, the time activity curve (TACs) of ROIs were calculated by COMKAT toolbox (Muzic and Cornelius 2001), and filled into the aforementioned phantoms to obtain true radioactivity images. For either of the two datasets, 400 groups of true activity distribution maps were obtained. Examples of TACs for different ROIs were shown in figure 2(b).

The true tracer radioactivity images needed to be projected to sinograms to get training and test sinograms. This process was finished by Michigan Image Reconstruction Toolbox with a simple strip-integral system model (Fessler 1994). The number of projection angles and detector bins were 160 and 128 respectively. 20% Poisson random noise was added to the projected sinograms. Scattered photons were not considered. The total count of 18 time frames was around $7.2 \times 10^5$, and the count was only about $1.7 \times 10^4$ at the 1st frame, $5.2 \times 10^4$ at 18th frame. Finally, we got two dynamic PET datasets, one with a single shape and the other with many shapes. Each dataset had 400 groups dynamic PET images, and each group included 18 time frames. The size of images were all $128 \times 128$. For each dataset, 280 sets of data were randomly selected as the training set, 60 sets as the test set, and 60 sets as the validation set. The training set, validation set and test set had different $k$ parameters.

3.2. Rat datasets

Animal experiments were approved by the experimental animal welfare and ethics review committee of Zhejiang University, and were performed in compliance with the guidelines for animal experiments (World Medical Association and American Physiological Society 2002) and local legal requirements.

12 rats with gliomas were anesthetized and injected with 1 mCi FDG before 60-minute dynamic PET scans by Siemens Micro-PET/CT Inveon. The sampling protocol was set to $3 \times 60$ s, $9 \times 180$ s, $6 \times 300$ s. The
dynamic scans were started immediately after the injection. The data collected by the PET/CT scanner contained two types of sinograms: prompt sinograms and delay sinograms. The delay sinograms were subtracted from prompt sinograms to perform random correction. For real data, we did not consider random photons, because existing instruments could identify random photons and performed random correction. The randomly corrected 3D sinograms (michelograms, 128 bins, 160 views) of each rat can be reconstructed into 18 frames of 3D images, each frame contains 159 slices of size 128 × 128. In order to simulate low count, only the segments 0 of the michelograms were used as training and test 2D sinograms. Since the rats were large and only one bed was scanned, the obtained data only included the heads and abdomens of the rats. The slices with nothing were meaningless, so they were discarded when making the rat dataset. From now on, we referred to these data containing both head and abdomen as whole-body data to distinguish them from pure head data or abdominal data. Influenced by individual differences of rats, the minimum count of a single sinogram was between 590 and 1383, while the maximum count of a single sinogram ranged from $3.7 \times 10^4$ to $7.4 \times 10^4$. These low-count sinograms were taken as inputs to the FBP-Net, and the reconstructed PET images with full 3D counts and CT attenuation correction were used as labels.

In order to ensure that the data could accurately reflect the real situation, we did not carry out data enhancement. The training set and test set were divided based on individual level. Eight rats were selected randomly as training set, two rats as validation set, and two rats as test set. For all experiments, the training set, validation set, and test set remained unchanged, corresponding to the same rats respectively.

4. Experiments and results

4.1. Evaluation metrics

Peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) are common metrics for image quality evaluation (Würfl et al 2018, Häggström et al 2018). PSNR is based on the error of the corresponding pixels of the two images, and is related to mean square error between two images. The larger the value of PSNR, the better the image quality. SSIM compares the similarities between the two images in terms of brightness, contrast, and structure. The range of SSIM is 0 to 1, and larger the SSIM, the higher the similarity between the two images. Only when the two images are exactly the same, SSIM is equal to 1. In this work, we used PSNR and SSIM to evaluate the quality of the reconstructed images in the test sets,

\[ \text{PSNR} = 10 \log_{10} \left( \frac{\text{max}(X)^2}{||X - \hat{X}||^2_2} \right), \]

where $X \in \mathbb{R}^{\sqrt{33} \times \sqrt{33}}$ and $\hat{X} \in \mathbb{R}^{\sqrt{33} \times \sqrt{33}}$ denote the label and the reconstructed image respectively. max($X$) represents taking the maximum of $X$.

\[ \text{SSIM} = \frac{(2\mu_X\mu_{\hat{X}} + c_1)(2\sigma_{\hat{X}} + c_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + c_1)(\sigma_X^2 + \sigma_{\hat{X}}^2 + c_2)}, \]

where $X, \hat{X}, \mu$, and $\sigma$ denote the label, the reconstructed image, the mean of the image and the variance of the image. $c_1$ and $c_2$ are used to avoid the denominator being 0. Here, $c_1 = 0.01 \times \text{max}(X)$ and $c_2 = 0.03 \times \text{max}(X)$.

4.2. Experimental settings

4.2.1. Comparison method

The comparison methods were DeepPET (Häggström et al 2018) and U-net (Ronneberger et al 2015). There are many variations of U-net, and it is unrealistic to compare one by one. Here, we adopt the U-net in Ronneberger et al (2015). As for normalization, we regarded the sinograms of all time frames of the same slice as a group, and each group was divided by its maximum to be normalized to 0 to 1. Correspondingly, the images of the same slice at different time frames were also divided by the maximum of them. The input and output of DeepPET and U-net are single-frame data, so a sinogram was regarded as a sample, and all samples were shuffled before training. The loss functions of DeepPET and U-net were both MSE, and the optimizer was Adam. The initial learning rate was 0.0001, and the learning rate decayed to half of the original value after every 200 epochs. The batchsize was set to 128. The simulation experiment iterated 300 epochs, while the rat experiment iterated 200 epochs.

4.2.2. FBP-Net

Since FBP part was designed according to FBP algorithm, so the initial value of the learnable frequency filter could be determined by theory. We used a ramp function with a Landweber window to initialize the filter.
4.2.3. Environment
All methods were implemented with deep learning framework, Pytorch 1.2.0. The experiments were carried out on a Ubuntu 18.04 LTS server with a TITAN RTX 24G.

4.3. Results
4.3.1. Experiment 1 based on single-shape simulation data
The experiment 1 was carried out on the simulation dataset with a single shape. The curves of the loss of the training set and the validation set with epoch are shown in figure 3(a). For any of the three methods, the loss of the training set was close to the loss of the validation set. This meant that over-fitting was unlikely to occur when the similarity among training set, validation set and the test set was high. The means of PSNR and SSIM of test set varied with time frames are presented in figures 3(b) and (c). From the metrics, DeepPET and U-net obtained more accurate reconstruction images than FBP-Net. This might be because they have much more learnable parameters than FBP-Net. The FBP-Net only had about 0.25 million learnable parameters, while U-net had about 26 million learnable parameters, and DeepPET had about 62 million parameters. More training parameters meant stronger fitting ability. Quantitative metrics of the entire test set were shown in table 2. Although the PSNR and SSIM of FBP-Net seemed worse than DeepPET and U-net, they were more stable with smaller standard deviations. Representative reconstruction maps in figure 4 illustrated that DeepPET and U-net had better edge details than FBP-Net. This experiment showed that when the training set, the verification set and the test set were very similar, DeepPET and U-net could obtain higher quality reconstruction images than FBP-Net.

Table 2. The PSNR and SSIM of reconstruction images for the test set of simulation dataset with a single shape.

|       | PSNR (mean ± std) | SSIM (mean ± std) |
|-------|------------------|-------------------|
| FBP-Net | 25.63 ± 0.90     | 0.9678 ± 0.0057   |
| DeepPET | 33.09 ± 3.61     | 0.9845 ± 0.0079   |
| U-Net  | 33.98 ± 4.23     | 0.9848 ± 0.0096   |

This special window function was derived from the iterative Landweber algorithm, making the reconstruction images of FBP better than MLEM (Zeng 2012). The Landweber window function was expressed as,

\[
\text{window}(\omega) = 1 - (1 - \alpha \frac{(0.5 + 0.5 \cos(\omega))^\gamma}{|\omega|})^k,
\]

where \(\omega\) is the frequency variable, \(\alpha\), \(k\) and \(\gamma\) were corresponding to the step size, the number of iterations and the number of times that the low-pass filter is used. Here, \(\alpha = \pi/128\), \(k = 195\) and \(\gamma = 1\).

Since the input and output of the FBP-Net were data of all time frames, limited by the video memory, the batchsize was set to 16. The optimizer for FBP-Net was also Adam, and the setting of learning rates and the number of iterations were the same as DeepPET and U-net. The FBP part and the denoiser part were trained together without pre-training.
Figure 4. Comparison of reconstruction images of several time frames by different methods for simulation dataset with a single shape. The reconstruction images by the three methods are all close to the ground truths, but the images by DeepPET and U-net have shaper edges than the proposed FBP-Net.

Figure 5. Comparison of different methods for simulation dataset with many shapes. (a) Loss curves of training set and validation set. (b) The mean of $\text{PSNR}$ of the test set varied with frames. (c) The mean of $\text{SSIM}$ of the test set varied with frames.

4.3.2. Experiment 2 based on simulation data of many shapes
Experiment 2 was carried out on the simulation dataset with many shapes. Figure 5(a) presents the loss curves of training set and validation set. With the increase of epoch, the gaps of losses between the training set and the validation set of DeepPET and U-net were getting bigger and bigger, and over-fitting was getting more and more serious. Different from DeepPET and U-net, the loss gap between training set and validation set of FBP-Net had been small. This showed that DeepPET and U-net were more prone to overfit when there was a certain difference between the training set and the validation set, while FBP-Net was free of over-fitting problem. From the perspective of $\text{PSNR}$ and $\text{SSIM}$, FBP-Net was more accuracy than DeepPET and U-net over all time frames (figures 5(a) and (b)). Furthermore, the SDs of $\text{SSIM}$ and $\text{PSNR}$ of FBP-Net were obviously smaller than U-net and DeepPET. This showed that the performance of FBP-Net was more stable. Figure 6 presents several reconstruction maps. The images by DeepPET were blurred, and the images by U-net contained many noise blocks, while the images by FBP-Net were closest to the ground truths. Experiment 2 showed that the FBP-Net was more accuracy and stable when there was a certain difference between training set and test set.

4.3.3. Experiment 3 based on the whole-body data of rats
Experiment 3 was used to validate the feasibility of proposed FBP-Net on real dataset. Figures 7(a) and (c) show the boxplots of training set, test set and validation set by three methods. The metrics of training set, validation set and test set were very close for FBP-Net, while the metrics of training set were far better than the test and validation set for U-net and DeepPET. This might be because U-net and DeepPET overfitted the
Comparison of reconstruction images of several time frames by different methods for simulation dataset with many shapes. The reconstructed images by DeepPET are blurred, and the images by U-net contain many noise blocks, while the images by FBP-Net were closest to ground truths.

Table 3. The PSNR and SSIM of reconstruction images for the test set of simulation dataset with many shapes.

|                | PSNR    | SSIM     |
|----------------|---------|----------|
|                | (mean ± std) | (mean ± std) |
| FBP-Net        | 23.98 ± 2.30 | 0.9410 ± 0.0110 |
| DeepPET        | 20.83 ± 3.33 | 0.8683 ± 0.0667 |
| U-Net          | 20.11 ± 5.58 | 0.8159 ± 0.0943 |

training set, resulting in poor generalization on the test set and validation set. As shown in figures 7(b) and (d), the PSNR and SSIM of test set by FBP-Net were higher than those by DeepPET and U-net over 18 frames. The first frame had the lowest SSIM. This might be because the tracer concentration was very low at the beginning, and the count was particularly small.

Figure 8 presents several representative reconstruction images. The hot spot of the suspected tumor in the 14th frame was not visible on the reconstruction images by DeepPET and U-net, but it was clearly visible on the image by FBP-Net. Looking closely, it was not difficult to find that the reconstruction maps by DeepPET and U-net were fatter than the labels. This might be caused by the individual differences of rats. Although these rats were all 9–10 weeks old and around 300 grams, their heads were not the same size, and their heads even existed a little oblique angles when scanning. So if the training set was overfitted, it might led to a poor generalization of the model on the test set.

4.3.4. Experiment 4 and 5 for model generalization

We divided the whole-body data of the rats into the thorax data and the brain data. In experiment 4, we trained the FBP-Net with the brain data (brain model), and test the well-trained FBP-Net with the thorax data. Conversely, in experiment 5, we trained the FBP-Net with the thorax data (thorax model), and test the well-trained FBP-Net with the brain data. As a comparison, the brain reconstruction images obtained by thorax model (untrained) were compared with those obtained by brain model (trained). Similarly, the thorax reconstruction images obtained by brain model (untrained) were compared with those obtained by thorax model (trained). Table 4 presents the metrics of the test set. Although the brain model did not use thorax data for training, it could still accurately reconstruct the thorax image, and vice versa. The brain and thorax images were presented in figures 9 and 10, respectively. The brain images obtained by the thorax model and the thorax images obtained by the brain model were all very close to the labels. Experiment 4 and experiment 5 showed that the proposed FBP-Net had a strong generalization, and could be competent for unseen shapes.
Figure 7. Comparison of metrics of different methods when training the networks with whole-body data of the rats. Comparison of boxplots of (a) $\text{PSNR}$ and (c) $\text{SSIM}$ for training set, test set and validation set. The mean of (b) $\text{PSNR}$ and (d) $\text{SSIM}$ of test set over 18 frames.

Figure 8. Comparison of reconstruction images by different methods for the test set of the whole body data of the rats. The hot spot of suspected tumor in the 14th is not visible on the reconstructed images by DeepPET and U-net, but it is visible on the image by FBP-Net. Besides, the reconstructed brains of DeepPET and U-net are a little fatter than the ground truths.
4.3.5. Necessity of FBP part and denoiser part

In order to justify the necessities of the FBP part and denoiser part in the FBP-Net, we compared the reconstruction images obtained from three structures, only FBP part, only denoiser part and FBP-Net. For

| Time Frame | Sinogram | Trained | Untrained | Label |
|------------|----------|---------|-----------|-------|
| 2nd        |          |         |           |       |
| 8th        |          |         |           |       |
| 14th       |          |         |           |       |

Figure 9. The 1st column, sinograms of different time frames. The 2nd column, the brain images were reconstructed by FBP-Net trained with brain data. The 3rd column, the brain images were reconstructed by FBP-Net trained with thorax data. The 4th column, the reference reconstruction images by Inveon Research Workplace.

| Time Frame | Sinogram | Trained | Untrained | Label |
|------------|----------|---------|-----------|-------|
| 2nd        |          |         |           |       |
| 8th        |          |         |           |       |
| 14th       |          |         |           |       |

Figure 10. The 1st column, sinograms of different time frames. The 2nd column, the thorax images were reconstructed by FBP-Net trained with thorax data. The 3rd column, the brain images were reconstructed by FBP-Net trained with brain data. The 4th column, the reference reconstruction images by Inveon Research Workplace.
B Wang and H Liu

Table 4. The PSNR and SSIM of the test set when training on specific organs.

|                  | PSNR (mean ± std) | SSIM (mean ± std) |
|------------------|-------------------|-------------------|
| Trained on brain |                   |                   |
| Tested on brain  | 32.57 ± 3.69      | 0.8950 ± 0.0384   |
| Tested on thorax | 31.13 ± 2.95      | 0.8787 ± 0.0460   |
| Trained on thorax|                   |                   |
| Tested on thorax | 28.35 ± 2.73      | 0.8909 ± 0.0559   |
| Tested on brain  | 27.38 ± 2.84      | 0.8839 ± 0.0564   |

Figure 11. Comparison of reconstruction images of 8th frame obtained from only FBP part, only denoiser part and FBP-Net. The 1st row, the reconstructed images of simulation dataset with a shape. The 2nd row, the reconstructed images of simulation dataset with various shapes.

any of these three structures, sinograms were taken as the inputs, and reconstructed images were considered as outputs. The training sets, test sets and the settings of training parameters (such as learning rate, epochs, etc) were all the same as experiment 1 and experiment 2. The representative reconstruction images are shown in figure 11. Whether the dataset had a single shape or various shapes, the performance of the only FBP part was similar, and blurred reconstruction images with basic shapes were obtained. In contrast, the only denoiser part acquired fuzzy reconstructed images when the dataset had a single shape, while failed to obtain images when the dataset had various shapes. This might be because the structure of denoiser part was simple without any encoder-decoder structure, and the number of training parameters was smaller than 0.25 million, much smaller than U-net or DeepPET, so the denoiser part was unable to fit the complex function relationship like inverse radon transform. Moreover, the backbone of the denoiser part was DnCNN which was originally designed for denoising tasks instead of complex reconstruction tasks. In denoising tasks, the similarity between the input noise image and the output clean image was very high, but in reconstruction tasks, the sinogram and the final reconstruction image were not similar at all. Therefore, it was reasonable for the only denoiser part to perform poorly for reconstruction tasks. To sum up, only the FBP part or the denoiser part was not enough to obtain clear reconstructed images, but if they were connected in series to form the FBP-Net, the final reconstruction images were clear and had more details. So both of the FBP part and the denoiser part were necessary for the FBP-Net.

5. Discussion

Both DeepPET and U-net are encoder-decoder convolutional neural network, but there are a few residual connection lines in U-net. During the experiments, the losses of U-net decreased significantly faster than DeepPET. This might be due to the residual connection, which sped up the convergence. DeepPET performed well in Häggström et al (2018), but performed poor on our rat data. We summarized a few possible reasons:

- **The amount of data.** The data used in this paper had complex shapes and a small number. The similarity between the training and test set was low. The small amount of data was not friendly to the network with many parameters, which easily led to over-fitting. It could be seen from the comparison of the results of the training set and the test set that DeepPET had sufficient learning ability, but overfitted the training set, resulting in poor performance on the test set. If the dataset was large enough, like in Häggström et al (2018),
this situation would be likely to be alleviated. However, it was costly to obtain a large amount of medical image data. If only a small dataset was available, a method such as FBP-Net that incorporates traditional methods would be more feasible.

- **Data preprocessing.** In Häggström et al. (2018), the noise was removed from sinograms, and attenuation correction was also performed. However, there was no attenuation correction in this work. The attenuation correction was usually carried out using CT imaging. It is well known that CT has radioactivity, which is harmful to the health of patients and operators. Therefore, we hope to develop a PET reconstruction method that do not require attenuation correction.
- **Count rate.** The simulation sinogram in Häggström et al. (2018) had about 10⁶ photons, while the counts in our work were much smaller. For example, the sinograms of early frames even had several hundred photos. Low counts also had an impact on reconstruction methods.

### 5.1. FBP-based methods for computer tomography reconstruction

There are some structures similar to the proposed FBP-Net, but they are used to solve limited angle and sparse angle reconstruction problems on CT reconstruction (Würfl et al. 2018, He and Ma 2019, Li et al. 2019, Whiteley and Gregor 2019). Würfl et al. mapped parallel, fan and cone beam reconstruction problems on CT to neural networks respectively, but the network they proposed for parallel beams did not contain any learnable parameters (Würfl et al. 2018). So this parallel beam network was expected to obtain poor reconstruction images similar to traditional FBP algorithm when used for PET reconstruction. The other three studies used neural networks to learn back-projection parameters (He and Ma 2019, Li et al. 2019, Whiteley and Gregor 2019). The number of parameters of neural networks proposed by Li et al. and He et al. were on the order of \( o(M \times B) \) and \( o(M \times (A + 1)) \) respectively, where \( M, A \) and \( B \) denote the number of pixels, angles and detector bins. The number of parameters of methods proposed by Whiteley et al. was related to the size of masks and buffer area, and their methods was expected to have more parameters than the methods of He et al. or Li et al., because their method used patches instead of pixels when performing back-projection. For the data in this work, the method proposed by Whiteley et al., He et al. and Li et al. would have about 22 million, 3.5 million and 2.2 million training parameters respectively, while the proposed FBP-Net only had about 0.25 million parameters. If the size of image increased, the number of parameters of the proposed FBP-Net would be almost unchanged, because the back-projection weights were fixed. On the contrary, the numbers of parameters of Whiteley’s, He’s and Li’s neural network would increase significantly as the size of image increased. Although the flexibility of these three neural networks is high, too many parameters mean that a considerable number of samples are required to train the network to avoid over-fitting. In addition, large-scale neural networks also increase the difficulty of training and put forward high requirements for computing hardware. Different from these existing FBP-based networks that learned back-projection parameters, the FBP part of the proposed FBP-Net kept the parameters of back-projection unchanged, and only adaptively learnt the frequency filter, so the number of parameters in the FBP part was significantly less than the existing FBP-based deep learning methods. Besides the proposed FBP-Net also reduced the requirements for the number of training samples and computing hardware.

Due to limited real CT data, (He and Ma 2019) and (Li et al. 2019) use quantities of simulation data to pre-train the proposed networks, but these simulation data only retained shape information, lacking information on instrument structures, detector characteristics, etc. Compared with CT, it is more difficult to simulate a large amount of representative dynamic PET data, because PET is functional imaging, and a series of physiologically meaningful parameters are required to be set, such as plasma input functions, \( k \) parameters in the compartmental model, blood fraction and so on. To make matters worse, these parameters have individual differences, and there is no uniform standard which values these parameters should be set. For the above reasons, it is not feasible to pre-train the network with mass simulation data for dynamic PET data. So we proposed FBP-Net that only needed a small number of samples to train.

### 5.2. Limitations

This study had several limitations. Firstly, the reconstruction speed of the proposed FBP-Net might not be as fast as DeepPET. In Häggström et al. (2018), the authors showed that DeepPET was faster than the traditional FBP algorithm under the same conditions. The proposed FBP-Net were designed based on FBP algorithm, and had a speed comparable to the traditional FBP algorithm, so the FBP-Net was expected to be slower than DeepPET. Secondly, the simulation data were obtained by forward projection with Poisson noise, it would be more reliable to get data through Monte Carlo simulation. However, it is well known that Monte Carlo simulation is time consuming, and it may take several days or even weeks to get a group of data. So we adopted forward projection method to obtain simulation data. Finally, the sinograms used in FBP-Net were 2D sinograms, so 3D projection data needed to be rebinned to 2D sinograms before inputted into FBP-Net. Besides, the proposed FBP-Net did not consider time-of-flight (TOF) information that was expected to
improve the quality of reconstruction images. Incorporating TOF information into neural network design would be a direction worth exploring in the future.

6. Conclusion
In this work, we combined traditional FBP algorithm and neural network, and proposed FBP-Net with strong generalization. The FBP part adaptively learned filters from data, and reconstructed coarse images from sinograms. Then the denoiser part merged information from different time frames and enhanced the quality of the reconstructed coarse images. The simulation and real data experiments showed that the proposed FBP-Net could copy with the situation where the training set and the test set were quite different, and had advantages over U-net and DeepPET.

Acknowledgment
This work is supported in part by the National Natural Science Foundation of China (Grant Nos. 61525106, 61427807, U1809204), by the National Key Technology Research and Development Program of China (Grant Nos. 2017YFE0104000, 2016YFC1300302).

References

Bendriem B and Townsend D W 2013 The Theory and Practice of 3D PET vol 32 (Berlin: Springer) (https://doi.org/10.1007/978-94-017-3473-2)
Chen K T et al 2019 Ultra-low-dose 18f-florbetaben amyloid pet imaging using deep learning with multi-contrast MRI inputs Radiology 290 649–56
Cheng X, Li Z, Liu Z, Navab N, Huang S C, Keller U, Ziegler S I and Shi K 2015 Direct parametric image reconstruction in reduced parameter space for rapid multi-tracer PET imaging IEEE Trans. Med. Imaging 34 1498–1512
Feng D, Huang S C and Wang X 1993 Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography Int. J. Bio. Med. Comput. 32 95–110
Fessler J A 1994 Penalized weighted least-squares image reconstruction for positron emission tomography IEEE Trans. Med. Imaging 13 290–300
Gong K, Catana C, Qi J and Li Q 2018a PET image reconstruction using deep image prior IEEE Trans. Med. Imaging 38 1655–65
Gong K, Guan J, Kim K, Zhang X, Yang J, Seo Y, El Fakhri G, Qi J and Li Q 2018b Iterative pet image reconstruction using convolutional neural network representation IEEE Trans. Med. Imaging 38 675–85
Green P J 1990 Bayesian reconstructions from emission tomography data using a modified em algorithm IEEE Trans. Med. Imaging 9 84–93
He J and Ma J 2019 Radon inversion via deep learning Proc. SPIE 1094810
Hägström I, Schmidtlein C R, Campaella G and Fuchs T J 2018 DeepPET: a deep encoder–decoder network for directly solving the PET image reconstruction inverse problem Med. Image Anal. 54 253–62
Kaplan S and Zhu Y M 2019 Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study J. Digit. Imaging 32 773–8
Kim K, Wu D, Gong K, Dutta J, Kim J H, Son Y D, Kim H K, El Fakhri G and Li Q 2018 Penalized pet reconstruction using deep learning prior and local linear fitting IEEE Trans. Med. Imaging 37 1478–87
Li Y, Li K, Zhang C, Montoya J and Chen G H 2019 Learning to reconstruct computed medical images directly from sinogram data under a variety of data acquisition conditions IEEE Trans. Med. Imaging 38 2469–81
Liu Z, Chen H and Liu H 2019 Deep learning based framework for direct reconstruction of PET images Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Cham: Springer) pp 48–56
Maier A, Schebesch F, Syben C, Wurfl T, Steidl S, Choi J H and Fahrig R 2018 Precision learning: towards use of known operators in neural networks 2018 24th Int. Conf. on Pattern Recognition (ICPR) pp 183–8
Merlin T, Visvikis D, Fernandez P and Lamare F 2018 Dynamic PET image reconstruction integrating temporal regularization associated with respiratory motion correction for applications in oncology Phys. Med. Biol. 63 045012
Muzic R F and Cornelius S 2001 Compartment model kinetic analysis tool J. Nucl. Med. 42 636–45
Ouyang J, Chen K T, Gong E, Pauly J and Zaharchuk G 2019 Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss Med. Phys. 46 3555–64
Pan L, Cheng C, Haberkorn U and Dimitrakopoulou-Strauss A 2017 Machine learning-based kinetic modeling: a robust and reproducible solution for quantitative analysis of dynamic PET data Phys. Med. Biol. 62 3566
Panin V Y, Kehren F, Michel C and Casey M 2006 Fully 3-D PET reconstruction with system matrix derived from point source measurements IEEE Trans. Med. Imaging 25 907–21
Pelt D M and Rattenberg K J 2014 Improving filtered backprojection reconstruction by data-dependent filtering IEEE Trans. Image Process. 23 4750–62
Qi J, Leahy R M, Cherry S R, Chatziioannou A and Farquhar T H 1998 High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner Phys. Med. Biol. 43 1001
Quigley H, Colloby S J and O’Brien J T 2011 PET imaging of brain amyloid in dementia: a review Int. J. Geriatri. Psychiatry 26 991–9
Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional networks for biomedical image segmentation Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Cham: Springer) pp 234–41
Ruan D and Liu H 2017 Separation of a mixture of simultaneous dual-tracer PET signals: a data-driven approach IEEE Trans. Nucl. Sci. 64 2588–97
Sari H et al 2018 Non-invasive kinetic modelling of PET tracers with radiometabolites using a constrained simultaneous estimation method: evaluation with 11C-SB201745 EJNMMI Research 8 1–12
Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography IEEE Trans. Imaging 1 113–22
Sudarshan V P, Chen Z and Awate S P 2018 Joint PET + MRI patch-based dictionary for Bayesian random field pet reconstruction Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Cham: Springer) pp 338–46
Tong J, Chen Y and Liu H 2019 Single-scan dual-tracer separation network based on pre-trained GRU Int. Workshop on Multiscale Multimodal Medical Imaging (Cham: Springer) pp 43–50
Wang C, Hu Z, Shi P and Liu H 2014 Low dose pet reconstruction with total variation regularization 2014 36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Piscataway, NJ: IEEE) pp 1917–20
Wang Y et al 2015 Predicting standard-dose pet image from low-dose PET and multimodal MR images using mapping-based sparse representation Int. Workshop on Machine Learning in Medical Imaging (Cham: Springer) pp 127–35
Wang Y et al 2018 3D conditional generative adversarial networks for high-quality PET image estimation at low dose Neuroimage 174 550–62
Whiteley W and Gregor J 2019 Efficient neural network image reconstruction from raw data using a radon inversion layer 2019 IEEE Nuclear Symp. and Medical Conf. (NSS/MIC) IEEE pp 1–2
World Medical Association and American Physiological Society 2002 Guiding principles for research involving animals and human beings Am. J. Physiol. Regul. Integr. Comp. Phys. 283 R281–3
Würfl T, Hoffmann M, Christlein V, Breininger K, Huang Y, Unberath M and Maier A K 2018 Deep learning computed tomography: learning projection-domain weights from image domain in limited angle problems IEEE Trans. Med. Imaging 37 1454–63
Xu J, Gong E, Pauly J and Zaharchuk G 2017 200x low-dose PET reconstruction using deep learning (arXiv: 1712.04119)
Xu J and Liu H 2019 Three-dimensional convolutional neural networks for simultaneous dual-tracer PET imaging Phys. Med. Biol. 64 185016
Yang B, Ying L and Tang J 2018 Artificial neural network enhanced Bayesian PET image reconstruction IEEE Trans. Med. Imaging 37 1297–309
Yokota T, Kawai K, Sakata M, Kimura Y and Hontani H 2019 Dynamic PET image reconstruction using nonnegative matrix factorization incorporated with deep image prior Proc. IEEE Int. Conf. on Computer Vision pp 3126–35
Zeng G L 2010 Medical Image Reconstruction: A Conceptual Tutorial (Berlin: Springer) (https://doi.org/10.1007/978-3-642-05368-9)
Zeng G L 2012 Filtered backprojection algorithm can outperform iterative maximum likelihood expectation-maximization algorithm Int. J. Imaging. Syst. Technol. 22 114–20
Zhang K, Zuo W, Chen Y, Meng D and Zhang L 2017 Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising IEEE Trans. Image Process. 26 3142–55
Zubal I G, Harrell C R, Smith E O, Rattner Z, Gindi G and Hoffer P B 1994 Computerized three-dimensional segmented human anatomy Med. Phys. 21 299–302