Title

Unsupervised deep learning framework for data-driven gating in positron emission tomography

Journal

Medical Physics, 50(10)

ISSN

0094-2405

Authors

Li, Tiantian

Xie, Zhaoheng

Qi, Wenyuan

et al.

Publication Date

2023-10-01

DOI

10.1002/mp.16642

Peer reviewed

HHS Public Access

Author manuscript

Med Phys. Author manuscript; available in PMC 2023 October 23.

Published in final edited form as:

Med Phys. 2023 October ; 50(10): 6047–6059. doi:10.1002/mp.16642.

Unsupervised Deep Learning Framework for Data-Driven Gating in Positron Emission Tomography

Tiantian Li[a], **Zhaoheng Xie**[a], **Wenyuan Qi**[b], **Evren Asma**[b], **Jinyi Qi**[a,*]

[a]Department of Biomedical Engineering, University of California - Davis, Davis, CA 95616, USA

[b]Canon Medical Research USA, Inc., Vernon Hills, IL 60061, USA

Abstract

Background: Physiological motion, such as respiratory motion, has become a limiting factor in the spatial resolution of positron emission tomography (PET) imaging as the resolution of PET detectors continue to improve. Motion-induced misregistration between PET and CT images can also cause attenuation correction artifacts. Respiratory gating can be used to freeze the motion and to reduce motion induced artifacts.

Purpose: In this study, we propose a robust data-driven approach using an unsupervised deep clustering network that employs an autoencoder (AE) to extract latent features for respiratory gating.

Methods: We first divide list-mode PET data into short-time frames. The short-time frame images are reconstructed without attenuation, scatter or randoms correction to avoid attenuation mismatch artifacts and to reduce image reconstruction time. The deep AE is then trained using reconstructed short-time frame images to extract latent features for respiratory gating. No additional data are required for the AE training. K-means clustering is subsequently used to perform respiratory gating based on the latent features extracted by the deep AE. The effectiveness of our proposed deep clustering method was evaluated using physical phantom and real patient datasets. The performance was compared against phase gating based on an external signal (External) and image based principal component analysis (PCA) with K-means clustering (Image PCA).

Results: The proposed method produced gated images with higher contrast and sharper myocardium boundaries than those obtained using the External gating method and Image PCA. Quantitatively, the gated images generated by the proposed deep clustering method showed larger center of mass displacement and higher lesion contrast than those obtained using the other two methods.

Conclusions: The effectiveness of our proposed method was validated using physical phantom and real patient data. The results showed our proposed framework could provide superior gating than the conventional External method and Image PCA.

*Corresponding author: J Qi, qi@ucdavis.edu.

CONFLICT OF INTEREST

UC Davis has a research agreement with Canon Medical Research USA, Inc and a provisional patent application has been filed for this work.

Keywords

respiratory gating; data-driven; deep clustering; unsupervised learning

1. Introduction

Positron emission tomography (PET) imaging is an important image modality that provides quantitative information of various physiological and biochemical processes. Over the past few decades, efforts have been made by scientists and engineers to push the sensitivity and resolution of PET imaging further [1–3]. As spatial resolutions of PET systems improve, motion blurring increasingly becomes a limiting factor for resolution in many PET applications. For instance, image blurring caused by patient cardiac and respiratory motions during a PET scan would lead to the distortion of organ shape, loss of signal-to-noise ratio, attenuation mismatch artifacts, and also cause inaccurate parameter estimates in dynamic studies.

To eliminate motion-induced artifacts, respiratory gating has been proposed to collect data acquired at the same respiratory position [4]. The position information is traditionally acquired by an external device, such as a pressure belt or a motion camera along with gating software to process it. However, this requires extra setup time and can fail due to user errors, e.g. mispositioning of the sensor, detachment of marker, or poor calibration and synchronization [5]. To address these problems, many data-driven gating (DDG) algorithms have been proposed to extract the respiratory signal directly from the raw PET data or pre-reconstructed images. Some previous clinical studies have demonstrated that DDG can provide comparable or even superior performance compared to external devices [6,7]

Klein *et al.* proposed to use the axial component of the center of mass (COM) for list mode events as an indicator of motion[8]. The performance of this method could be further improved by incorporating TOF information to provide more accurate event localization and by applying a volume-of-interest (VOI) to reduce contributions from the stationary background [8–12]. Since this method does not require reconstruction of images, it may be applied to motion detection at a very fine temporal scale. However, good performance for these COM based methods usually requires extra computation time or additional user inputs. Also, the drawback of above-mentioned COM based methods is that they rely on band pass filtering to get a smooth curve, which may not be able to preserve irregular or deep breathing cycles. Alternatively, the respiratory signal can also be extracted by applying principal component analysis (PCA) or spectral analysis to the sinogram or histogram of PET data sampled at short time intervals [13–15]. As an alternative to sinogram-space approaches, image-space approaches were also proposed for respiratory signal acquisition. Compared with the sinogram-space methods based on the line integral of the radio-activity along each LOR, image-based methods can separate moving and stationary regions to achieve a higher SNR for extracting the motion signal. Moreover, the current list-mode reconstruction enables ultra-fast short time frame reconstruction, which makes image-based methods more attractive [16]. As an alternative, Yang *et al* proposed to extract the respiratory motion signal from the summed singles rate without sorting the singles events into coincidences [17]. While

this method is promising since it is simple and fast, its performance still needs improvement when the breathing pattern is irregular due to the low signal-to-noise ratio (SNR) in singles data for detecting respiratory motion using a standard PET scanner.

After the motion signal extraction, respiratory gating can be conducted based on either phase or amplitude of the motion signal [18]. In phase-based gating, each respiratory cycle is divided into a set of approximately equal time bins. For irregular breathing patterns, phase-based gating could mis-assign data from different respiratory positions to the same gate. In amplitude-based gating, each respiratory gate is formed by combining data corresponding to the same amplitude of the motion signal. The performance of amplitude gating could be affected by a baseline shift in the motion signal. A common limitation of the above-mentioned methods is that respiratory gating is based only on a single respiratory signal, which cannot capture complicated motion patterns, such as hysteresis. Methods using multiple signals have been proposed and reviews of techniques on gating strategy and motion correction in PET can be found in [19,20].

Recently, deep clustering networks that use autoencoders (AE) to extract latent features and exploit intrinsic similarities without using supervisory labels for clustering have been found to be successful in many computer vision tasks [21,22]. A similar idea based on utilizing clustering algorithms for gating has been demonstrated in brain PET applications, [23,24] although the features used for clustering in the literature are still based on the COM or sinogram domain PCA. In this study, we propose a data-driven approach using an unsupervised deep clustering network that employs an AE to extract latent features from list-mode reconstructed short-time frames for respiratory gating. This method overcomes the limitation of PCA and spectral analysis mentioned above and the effectiveness of our proposed method was evaluated using physical phantom and real patient data. The results showed that the proposed approach (Deep Clustering) outperforms traditional phase gating based on an external signal (External) as well as image-domain PCA with K-means clustering (Image PCA).

2. Methods

2.1 Deep cluster framework

The proposed deep cluster framework includes two parts: a deep AE network to extract latent features and K-means clustering to perform respiratory gating based on the latent features. The AE comprises of an encoder $\boldsymbol{q}_{\omega}(\boldsymbol{z}|\boldsymbol{x})$ and a decoder $\boldsymbol{p}_{\theta}(\boldsymbol{z}|\boldsymbol{x})$ (Fig.1 (a)). It aims to find a low-dimensional representation for each input image by minimizing the mean squared errors (MSE) between its input and output over all training datasets. To perform data-driven gating, list-mode data are first binned into short time frames and reconstructed frame-by-frame. The short-time frame images are reconstructed without attenuation, scatter or randoms correction to avoid attenuation mismatch artifacts and to reduce image reconstruction time. The reconstructed images are used to train the AE and to extract latent features. The input of the network is $\boldsymbol{x} \in \mathbb{R}^{X \times Y \times Z}$, where X, Y, Z denote the three spatial dimensions in the image. Unsupervised feature learning is performed by a 3D convolutional AE, in which only 3D or elementwise operations, such as 3D

convolution, batch normalization (BN), and rectified linear units (ReLU), are used to exploit spatial structure information for spatial feature extraction (Fig. 1(b)). We further include a self-attention (SA) module in the AE framework [25,26]. It calculates response at a position as a weighted sum of the features at all positions in the feature map, which allows the network to learn to suppress irrelevant features in an input image while highlighting desired features [27]. It should be noted that this network is trained by using the same input images as the training labels and it does not require any additional data. As a result, spatial structure information can be encoded and learned in an unsupervised fashion. Once the AE is trained, the latent features $\boldsymbol{z}$ are extracted from the bottleneck layer of the network and are fed into a K-means clustering algorithm to group the short time frame images to a fixed number of respiratory gates. The K-means clustering results are then used to re-sort the list-mode data into respiratory gates for the final image reconstruction.

2.2 Reference phase selection and 4D phase matched CT generation

A fast helical non-gated CT protocol with free breathing is usually used for attenuation correction (AC), which could cause attenuation correction artifacts if it is used for all respiratory gates. To eliminate such attenuation mismatch artifacts, we propose to generate phase matched CT images for attenuation correction. We first reconstruct the gated PET images without AC and SC using ordered subset expectation maximization (OS-EM) with 10 subsets and 2 iterations. The voxel size was 2.04× 2.04×2.04 mm^3 and the image array size is 272×272×96. We computed the normalized cross-correlation (NCC) between the reconstructed PET images and the CT image and selected the reference phase with the maximum NCC. Then we apply our previously developed unsupervised deep learning model for deformable motion estimation [28]. The estimated motion fields between the non-attenuation corrected reference and moving gate reconstructions were then used to deform the CT image to obtain phase-matched CT images for each moving gate.

The overall diagram of our proposed framework is shown in Fig.2.

2.3 Implementation details

The list-mode PET data was first divided into short frames (200-500 ms) and were reconstructed using a list-mode Maximum Likelihood Expectation Maximization (ML-EM) with 2 iterations. We used a simplified geometric projection matrix for all reconstructions in this study [29]. The ungated image was used for initialization and images were reconstructed without scatter and random correction to speed up the reconstruction. The voxel size was 4.08× 4.08×4.08 mm^3 and the image array size was 152×152×48.

The deep clustering network was implemented in Keras 2.2.4 with Tensorflow backend and trained on a NVIDIA GTX 1080Ti GPU. The learnable parameters of the network were updated using the adaptive moment estimation (ADAM) optimizer [30]. The learning rate was 0.0001 and the batch size was 1. For each training, the network was trained from scratch and all the trainable parameters were randomly initialized [31]. The network was trained using reconstructed short-time frames with 3 epochs. The extracted latent features were fed into a K-means clustering algorithm [32]. To force the clustering to focus more on respiratory motion, we weighted each latent feature $\boldsymbol{z}_i$ by the maximum magnitude of the frequency

component contained by z_i inside the human breathing frequency range (0.14-0.33 Hz in this study). For the K-means initialization, PCA was first applied to the latent features and phase-based gating was performed using the first principal component to compute the initial cluster centers in the latent feature space. As with most DDG methods, the extracted movement direction from PCA could be opposite of the true motion direction. For a better comparison with the results from the external phase gating method, sign correction was applied to the first principal component based on the assumption that humans generally spend more time at expiration in each respiratory cycle [12]. The peaks of the corrected signal represent end-inspiration and valleys represent end-expiration. After that, the cluster assignments can be obtained by iteratively optimizing the clustering objective by minimizing the sum of squared errors (SSD).

For phase-matched CT image generation, we used our previously proposed motion estimation neural network [28] to estimate the motion field between different gates. The network was fine-tuned for 200 epochs using the image pairs between the end-inspiration and end-expiration gated images from the External method. Eight moving-reference image pairs (4 patients each with 2 bed position) were fed into the network for the fine-tuning process. Negative NCC was chosen as the loss function with an L2 norm of the gradient of deformation field as regularization to encourage a smoothed motion field. The patient-specific fine-tuning improved the motion estimation accuracy and could be avoided if we had enough respiratory gated patient data for network training. For the purpose of comparing Image PCA and Deep Clustering, we used the same motion estimation network that was fine-tuned using the External gated images. In practice, the network can be fine-tuned using gated reconstructions from any gating method. Motion fields for all moving-reference image pairs were estimated for External, Image PCA and Deep Clustering methods, respectively (4 patients × 2 bed position × 6 moving-reference image pairs × 3 gating methods). These estimated motion fields were then used to deform the CT image to each moving gate to obtain 4D phase-matched CT images.

A uniform cylinder scan was used to obtain normalization factors and the matched phase CT images were used to compute the attenuation factors [33]. Scattered events were estimated using the single scatter simulation algorithm and randoms were estimated using the delayed random sinogram. The scatters and randoms were estimated based on ungated data and scaled based on the time duration of each gate. Finally, the gated PET data was reconstructed with full quantitative corrections using OS-EM with 10 subsets and 2 iterations.

2.4 Experimental data acquisition

2.4.1 Physical phantom data—A motion phantom, as shown in Fig. 3 (a), was used to evaluate the effectiveness of the proposed method. The phantom had a rotating disc inside a cylinder filled with uniform activity as shown in Fig. 4 (a). The disc holds four fillable spheres, as shown in Fig. 4 (b), and rotated at a constant speed of 4.66 sec/cycle. The activity concentration ratio between sphere A and the uniform background (pointed by the thick red arrow in Fig. 4 (a)) was 4:1, and the other slots (B-D) had the same activity concentration as the background. The phantom dataset was scanned on a Canon

Celesteion whole-body TOF PET/CT scanner (Canon Medical Corporation, Tochigi, Japan) with a total of 166.5 MBq ^{18}F-FDG for 10 mins (Fig. 3 (b)). A CT helical scan with 120 kVp was acquired for AC. Since the physical phantom was entirely filled with water equivalent material, a single CT was used for all gated reconstructions. An Anzai system (Anzai Medical Co, Ltd., Tokyo, Japan) was used to record the external gating signal. The list-mode data were divided into 200 ms frames and were grouped into eight gates using the three methods. Images reconstructed using the gated data were evaluated qualitatively based on visual assessment and quantitatively based on gate assignments.

2.4.2 Real patient data—Four patient datasets with 216.0-249.9 MBq ^{18}F-FDG administration were used for evaluation. List mode data was obtained from a Canon Celesteion whole-body TOF PET/CT scanner for two bed positions, each lasting 14 mins. By attaching the Anzai belt to the patient′s abdominal region, the movement of the abdominal surface due to respiration was recorded as the respiratory signal and subsequently used for gating. A free-breathing low-dose CT image was acquired at 120 kVp for each subject for AC. The list-mode data were divided into 500 ms frames and grouped into seven respiratory gates using the three different methods. Images reconstructed using the gated data were evaluated qualitatively based on visual assessment and quantitatively based on lesion contrast and maximum COM displacement.

2.5 Data analysis

2.5.1 Physical phantom—For the phantom data, since the disc rotated at a constant speed, the Anzai signal worked very well and was chosen as the ground truth. To avoid discretization error, the external gating was performed using a finer temporal resolution (25 ms). For quantitative evaluation, we computed the percentage gating accuracy and the absolute gate shift for each cluster. First, the gate center was obtained by averaging the true phase (in degrees) of all the frames assigned to each gate. Then the absolute phase shift of each frame was computed as the absolute phase difference between the frame phase and the gate center. Since we had eight gates per rotation, each gate covers an angular range of 45 degrees. Therefore, a frame was counted as correctly assigned if its absolute phase shift is less than 22.5 degrees. The average gating accuracy of gate i was then computed as

$$\text{Percentage Gating Accuracy} = \frac{\text{number of correctly assigned frames in gate } i}{\text{total number of frames in gate } i} \times 100\% .$$

The mean absolute gate shift of gate i was computed as

$$\text{Mean Absolute Gate Shift} = \frac{1}{\text{total number of frames in gate } i} \sum_{f \in \text{gate } i} \left| \phi_f - \bar{\phi}_i \right|$$

Where ϕ_f denotes the true phase in degrees of frame f and $\bar{\phi}_i$ denotes the gate center of gate i.

2.5.2 Real patient data—For real datasets, both visual assessment and quantitative region of interest (ROI) analysis were performed. Among four patient datasets, three of

them had high myocardial uptake, while the fourth patient had a pulmonary lesion that was affected by the respiratory motion. A box ROI was drawn on the ungated image covering the whole moving trajectory of either the myocardium or pulmonary lesion. Due to the lack of the ground truth, maximum COM displacement for different gating methods was used for comparison. Displacement along x- y- z- directions were derived as the largest difference in the COM measurements obtained from all gated images and the length of the moving trajectory over gates was calculated. Statistical significance of the observed differences between different gating methods were analyzed using paired t-tests. A p value of < 0.05 (*) was considered to be statistically significant and < 0.01 (**) was considered highly significant. In addition, the average wall thickness of the left ventricle was computed by averaging the full-width-at-half-maximum of the intensity profiles at 25 locations along the inferior ventricular walls. For the purpose of assessing respiratory motion blurring, the profiles were taken parallel to the z-axis. For the patient with a pulmonary lesion, we also plotted the average contrast of the lung lesion of all gated images as a function of the ROI size.

3. Results

3.1 Physical phantom

We examined the frequency spectrum of the first 10 principal components from Image PCA and found that the second component has the strongest signal at the rotation frequency. Because the phantom motion is a circular rotation, if only one component was used for clustering, the result would be based on the amplitude only and cannot differentiate the hysteresis. Therefore, we used the first two principal components from Image PCA to perform the clustering for gating. For visualization purposes, we extracted the first principal component from the latent features of the deep clustering method and plotted it with the external motor signal and the 2^{nd} component from Image PCA in Supplemental Material Figure S1, overlaid with color-coded dots representing different gates. The signals were normalized to the range of $[-0.5, 0.5]$ by their min-max values. We can see that the single principal component of the latent features or images captures most of the motion signal. However, the clustering based on Image PCA cannot fully capture the phase information, even though we used two principal components. In comparison, the clustering based on the latent features can capture phase very well. Fig. 5 showed the reconstructed phantom images of the gated data for all three methods. Gated images with the Deep clustering gating visually provided comparable resolution and contrast to the External method, while some gates from Image PCA are very blurry. We acknowledge that conventional phase gating based on the second component of Image PCA would give a better gating result in this case, but phase gating has its own limitations, in particular in dealing with irregular motions.

For quantitative evaluation, we plotted the percentage gating accuracy and mean absolute phase shift for Image PCA and Deep Clustering methods in Fig. 6 (a–b). The percentage gating accuracy was 68% for Deep Clustering while it was only 24% for Image PCA. The mean absolute phase shift was 20.42 degrees (0.45 gate) and 61.82 degrees (1.37 gate) for Deep Clustering and Image PCA, respectively. Distribution of the count level variations over gates for three methods are also plotted in Fig. 6 (c), which showed the External and Deep

Clustering methods produced more even count distributions across gates, while Image PCA showed larger variations.

3.2 Real patient data

Fig. 7 shows sampled gated PET reconstructions at end-inspiration phase from the Deep Clustering method superimposed with (a) free-breathing helical CT image and (b) the generated phase matched CT image. We can clearly see that the phase-matched CT image aligns very well with the gated PET image in Fig. 7 (b). In Supplemental Material Figure S2, latent feature maps showed large activation on the voxels corresponding to the hot spot and liver boundary, which indicated the motion was highlighted by the latent vectors. Frequency analysis indicated that only the first principal component (PC) of Image PCA contains significant respiratory motion signal. Therefore, we used the first PC as the motion signal for gating in Image PCA. In Supplementary Figure S3, we plotted the motion signal from Image PCA and the first PC from latent features of Deep Clustering method with colored dots representing gate assignments. We also analyzed count distribution across seven gates. The mean count fraction is 0.14 for all gating methods and the standard deviations were 0.02, 0.05 and 0.03 for External, Image PCA and Deep Clustering methods, respectively. The Image PCA showed larger count variation across different gates. Higher noise in gates with fewer counts could reduce the motion estimation accuracy and hence affect the generation of the phase-matched attenuation map. Fig. 8 (a–c) shows sampled respiratory gated images for the External, Image PCA and Deep Clustering methods. The yellow line marks the top of the myocardium position in Gate 3 (the reference gate). Substantial myocardium displacement can be observed in other gates for all three methods. The average FWHMs of the left ventricle were plotted as a function of gate and the average FWHMs over all gates were also included in Fig.8 (d). The deep clustering method provides sharper boundaries in the myocardium region and larger displacement on gated PET images compared with the External method. Sample reconstructed images of respiratory gated pulmonary lesion images were shown in Fig. 9. The lesion in Gate0 was blurred by respiration in the External method. In contrast, Image PCA and Deep Clustering gating provided a clear lesion boundary and higher lesion contrast than the External method.

For quantitative analysis, the COM displacements in a box region covering the heart/lesion region were evaluated. Both maximum displacement and the length of the 3D moving trajectory across all gates were listed in Table 1. The standard deviation of each metric was computed over all 8 datasets (4 patients each with 2 bed positions). Clearly, gated images from Deep Clustering provided larger COM displacement than the images from the External method ($p < 0.05$ for COM displacement along three directions) and slightly better than those from Image PCA ($p < 0.01$ for COM displacement along the x-direction). Fig. 10 showed averaged contrast of 7 gated images for a pulmonary lesion as a function of lesion ROI size (number of voxels) for the External, Image PCA, Deep Clustering methods. Clearly Deep Clustering provided the highest lesion contrast for all ROI sizes, demonstrating the advantage of the proposed method.

To investigate the effectiveness of including the self-attention (SA) module, an ablation study was performed and the frequency spectrum of each element in latent features over

time before and after removing the SA module for a patient data was compared. Fig. 11 (b,c) show the spectrums of the latent feature with the maximum energy inside the respiratory frequency range extracted by the AE with and without the SA module, respectively. In comparison, the spectrum of the feature extracted with the SA module is less noisy and closer to the spectrum of the Anzai signal (Fig. 11 (a)) than the one without the SA module.

4 Discussion

In this study, we proposed a fully automatic data-driven gating approach using an unsupervised deep clustering network that exploits the autoencoder approach and clustering for respiratory gating. Convolutional neural networks can effectively transform data from a high-dimensional image domain into a lower-dimensional feature space while preserving contextual information, leading to improved clustering results. In comparison, the conventional PCA for dimension reduction is limited to linear decomposition which may lose useful features. Supplementary Figure S4 (a–c) shows the frequency spectrum of the first three principal components from Image PCA for the real data. We can see only the first component contains a prominent peak within the respiratory frequency range. The second principal component contains some second harmonics (around 0.4 Hz) of the fundamental frequency (around 0.2 Hz), but the SNR is very poor. No respiratory signal can be identified in the third principal component. Therefore, we did not expect that including the second or third component in Image PCA would result in improvements for real data. One potential way to improve Image PCA is to include the time derivative of the motion signal, which can be helpful to distinguish the phase between inspiration and expiration. Some studies show that the gradient can be used to characterize the complex breathing motion pattern [34,35]. There are also several more advanced nonlinear dimension reduction approaches beyond the linear PCA, such as kernel PCA (KPCA) [36]. KPCA computes the covariance matrix of the data after being transformed into a higher-dimensional space via a pre-selected kernel function. However, it is not trivial to find a good kernel, so KPCA does not guarantee good results for a specific dataset when using a fixed kernel.

It should be noted that the physical phantom was designed to have a circular motion, which is not representative of the respiratory motion in patients. For the image based PCA, the second component can capture the rotation signal. However, if only one component was used for clustering, the result would be based on the amplitude only and cannot differentiate the hysteresis. Using the image PCA derived signal with a phase-based gating or including its derivate as the second feature could likely be more suitable in this case.

Compared to other DDG methods, e.g., COM and sinogram based methods, we expect image-based methods to achieve a higher SNR by utilizing short-frame reconstructions to decouple the motion and non-motion components in the image. Figure S5 in the Supplemental Materials shows frequency spectra of the extracted motion signals from external Anzai system, our implementation of sinogram PCA, Image PCA and Deep clustering. We can see that the spectrum from sinogram PCA contains more noise at higher frequency compared to the Anzai signal and the image-based methods. However, our sinogram PCA method was not optimized, e.g., regarding downsampling or filtering

[37], and it could therefore be further investigated if extending the Deep AE method towards sinograms would be beneficial.

Image-based methods need many reconstructions. Indeed, long computational time of the sinogram based reconstruction would limit its application as a standard clinical protocol [16]. In this study, we use list-mode reconstruction and calculate the simple geometric projection matrix on-the-fly during forward and backward projection which means the computational time is proportional to the number of list-mode events. Hence, the short time frame images can be generated in a time period comparable to that of an ungated reconstruction. In this study, it took 0.11 sec for each 0.5s frame reconstruction (count level 56k).

In the training of our deep clustering network, the autoencoder was trained from scratch using the patient specific short-frame reconstructions in a self-supervised fashion. The training time is about 4 mins per epoch. The total workflow takes about 30 mins for gating each 14-min clinical dataset excluding the time spent on fine-tuning the image registration network. We can further improve the computational time by initializing the model weights with previous training results and fine-tune it using patient specific data to reduce training time. Alternatively, the network can be pre-trained with a sufficient amount of existing data and directly applied to new patient data without additional training. Besides, the cluster assignments were done with a basic K-means clustering and can also be clustered with other unsupervised algorithms such as Gaussian mixture model [38] or spectral clustering [39]. In addition, the use of cluster analysis for respiratory gating provides potential for the patient specific selection of an optimal number of gates for gating. For example, the SSD from the cluster centers of the respective clusters could be plotted as a function of number of clusters and the “elbow” (the point of inflection on the curve) can be detected as the optimum point. Moreover, the proposed method is not limited to respiratory gating. One can choose 1 or 2 sec frames for gross body motions or shorter frames, such as 0.1 sec frames, for cardiac and respiratory dual gating.

To address the attenuation mismatch, 4D phase-matched CT images can be generated based on gated PET-derived motion fields obtained from the non-rigid registration [40,41]. We utilize our previously developed unsupervised motion estimation network to perform non-rigid registration between the reference and moving gates. This approach provides greater flexibility as one can change the reference gate and quickly obtain the motion vectors instead of running a full new set of non-rigid registration algorithms. In future work, we will further incorporate the CT at reference phase into the motion estimation network training assuming the anatomical information would potentially improve the image registration. The image quality of different gating methods will be evaluated blindly by clinicians in the future.

5 Conclusions

In this study, we proposed a fully automatic data-driven gating approach using an unsupervised deep clustering network that exploits the autoencoder approach for respiratory gating. The effectiveness of our proposed method was validated using physical phantom and

real patient data. The result showed our proposed approach could provide superior gating than the conventional External and Image PCA methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

This work was supported in part by the US National Institute of Biomedical Imaging and Bioengineering under grant no R01EB000194 and P41EB032840 and research support from Canon Medical Research USA.

References

1. Carson R, Berg E, Badawi R, et al. Design of the NeuroEXPLORER, a next-generation ultra-high performance human brain PET imager. J Nucl Med. 2021;62(supplement 1):1120 LP–1120. http://jnm.snmjournals.org/content/62/supplement_1/1120.abstract
2. Zhang X, Zhou J, Cherry SR, Badawi RD, Qi J. Quantitative image reconstruction for total-body PET imaging using the 2-meter long EXPLORER scanner. Phys Med Biol. 2017;62(6):2465–2485. doi:10.1088/1361-6560/aa5e46 [PubMed: 28240215]
3. Prenosil GA, Sari H, Fürstner M, et al. Performance Characteristics of the Biograph Vision Quadra PET/CT system with long axial field of view using the NEMA NU 2–2018 Standard. J Nucl Med. Published online 2021:jnumed.121.261972. doi:10.2967/jnumed.121.261972
4. Daou D Respiratory motion handling is mandatory to accomplish the high-resolution PET destiny. Eur J Nucl Med Mol Imaging. 2008;35(11):1961–1970. doi:10.1007/s00259-008-0931-x [PubMed: 18787822]
5. Böning G, Todica A, Vai A, et al. Erroneous cardiac ECG-gated PET list-mode trigger events can be retrospectively identified and replaced by an offline reprocessing approach: First results in rodents. Phys Med Biol. 2013;58(22):7937–7959. doi:10.1088/0031-9155/58/22/7937 [PubMed: 24165267]
6. Büther F, Jones J, Seifert R, Stegger L, Schleyer P, Schäfers M. Clinical evaluation of a data-driven respiratory gating algorithm for whole-body PET with continuous bed motion. J Nucl Med. 2020;61(10):1520–1527. doi:10.2967/jnumed.119.235770 [PubMed: 32060218]
7. Walker MD, Morgan AJ, Bradley KM, McGowan DR. Data-driven respiratory gating outperforms device-based gating for clinical 18F-FDG PET/CT. J Nucl Med. 2020;61(11):1678–1683. doi:10.2967/jnumed.120.242248 [PubMed: 32245898]
8. Klein GJ, Reutter BW, Botvinick EH, Budinger TF, Huesman RH. Fine-scale motion detection using intrinsic list mode PET information. In: Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis. IEEE; 2001:71–78. doi:10.1109/mmbia.2001.991701
9. Xu J, Tsui BMW. Improved Intrinsic Motion Detection Using Time-of-Flight PET. IEEE Trans Med Imaging. 2015;34(10):2131–2145. doi:10.1109/TMI.2015.2423976 [PubMed: 25897950]
10. Feng T, Wang J, Dong Y, Zhao J, Li H. A Novel Data-Driven Cardiac Gating Signal Extraction Method for PET. IEEE Trans Med Imaging. 2019;38(2):629–637. doi:10.1109/TMI.2018.2868615 [PubMed: 30188816]
11. Bundschuh RA, Martínez-Moeller A, Essler M, et al. Postacquisition detection of tumor motion in the lung and upper abdomen using list-mode PET data: A feasibility study. J Nucl Med. 2007;48(5):758–763. doi:10.2967/jnumed.106.035279 [PubMed: 17475964]
12. Feng T, Wang J, Sun Y, Zhu W, Dong Y, Li H. Self-gating: An adaptive center-of-mass approach for respiratory gating in PET. IEEE Trans Med Imaging. 2018;37(5):1140–1148. doi:10.1109/TMI.2017.2783739 [PubMed: 29727277]
13. Thielemans K, Rathore S, Engbrant F, Razifar P. Device-less gating for PET/CT using PCA. In: IEEE Nuclear Science Symposium Conference Record. IEEE; 2011:3904–3910. doi:10.1109/NSSMIC.2011.6153742

14. Büther F, Dawood M, Stegger L, et al. List mode-driven cardiac and respiratory gating in PET. J Nucl Med. 2009;50(5):674–681. doi:10.2967/jnumed.108.059204 [PubMed: 19372491]
15. Schleyer PJ, O'Doherty MJ, Barrington SF, Marsden PK. Retrospective data-driven respiratory gating for PET/CT. Phys Med Biol. 2009;54(7):1935–1950. doi:10.1088/0031-9155/54/7/005 [PubMed: 19265207]
16. Spangler-Bickell MG, Deller TW, Bettinardi V, Jansen F. Ultra-fast list-mode reconstruction of short PET frames and example applications. J Nucl Med. 2021;62(2):287–292. doi:10.2967/jnumed.120.245597 [PubMed: 32646873]
17. Yang J, Khalighi M, Hope TA, Ordovas K, Seo Y. Technical Note: Fast respiratory motion estimation using sorted singles without unlist processing: A feasibility study. Med Phys. 2017;44(5):1632–1637. doi:10.1002/mp.12115 [PubMed: 28099995]
18. Qi W, Ponce S, Xia T, et al. Investigation of phase and amplitude respiratory gating for whole-body FDG-PET with different respiratory waveform patterns. J Nucl Med. 2016;57. https://www.embase.com/search/results?subaction=viewrecord&id=L72335361&from=export%0Ahttp://jnm.snmjournals.org/content/57/supplement_2/1892.abstract?sid=5ed5eed8-682b-4842-a6cc-6132b2448e77
19. Lamare F, Bousse A, Thielemans K, et al. PET respiratory motion correction: quo vadis? Phys Med Biol. 2022;67(3):03TR02.
20. Kyme AZ, Fulton RR. Motion estimation and correction in SPECT, PET and CT. Phys Med Biol. 2021;66(18):18TR02.
21. Xie J, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: 33rd International Conference on Machine Learning, ICML 2016. Vol 1. ; 2016:740–749.
22. Huang P, Huang Y, Wang W, Wang L. Deep embedding network for clustering. In: Proceedings - International Conference on Pattern Recognition. IEEE; 2014:1532–1537. doi:10.1109/ICPR.2014.272
23. Hong I, Burbar Z, Schleyer P. A Method to Estimate Motion Frames from PET Listmode by Merging Adjacent Clusters. In: 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2019:1–2.
24. Gillman AG. Motion Correction of PET for Improved Image Quality. PhD thesis, Univ Queensland, Brisbane. Published online 2019.
25. Xie Z, Baikejiang R, Li T, et al. Generative adversarial network based regularized image reconstruction for PET. Phys Med Biol. 2020;65(12):125016. doi:10.1088/1361-6560/ab8f72 [PubMed: 32357352]
26. Zhang H, Goodfellow I, Metaxas D, Odena A. Self-attention generative adversarial networks. In: International Conference on Machine Learning. PMLR; 2019:7354–7363.
27. Xie Z, Li T, Zhang X, Qi W, Asma E, Qi J. Anatomically aided PET image reconstruction using deep neural networks. Med Phys. 2021;48(9):5244–5258. [PubMed: 34129690]
28. Li T, Zhang M, Qi W, Asma E, Qi J. Motion correction of respiratory-gated PET images using deep learning based image registration framework. Phys Med Biol. 2020;65(15):155003. doi:10.1088/1361-6560/ab8688 [PubMed: 32244230]
29. Zhou J, Qi J. Efficient fully 3D list-mode TOF PET image reconstruction using a factorized system matrix with an image domain resolution model. Phys Med Biol. 2014;59(3):541–559. doi:10.1088/0031-9155/59/3/541 [PubMed: 24434568]
30. Kingma DP, Ba JL. Adam: A method for stochastic optimization. 3rd Int Conf Learn Represent. Published online 2015.
31. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision. Vol 2015 Inter. ; 2015:1026–1034. doi:10.1109/ICCV.2015.123
32. Pedregosa F, Weiss R, Brucher M, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825–2830. http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html
33. Kinahan PE, Townsend DW, Bailey DL, Sashin D, Jadali F, Mintun MA. Efficiency normalization techniques for 3D PET data. In: 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference Record. Vol 2. IEEE; 1995:1021–1025.

34. Manber R, Thielemans K, Hutton BF, et al. Joint PET-MR respiratory motion models for clinical PET motion correction. Phys Med Biol. 2016;61(17):6515. [PubMed: 27524409]

35. McClelland JR, Hawkes DJ, Schaeffter T, King AP. Respiratory motion models: a review. Med Image Anal. 2013;17(1):19–42. [PubMed: 23123330]

36. Schölkopf B, Smola A, Müller K-R. Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput. 1998;10(5):1299–1319.

37. Thielemans K, Schleyer P, Marsden PK, Manjeshwar RM, Wollenweber SD, Ganin A. Comparison of different methods for data-driven respiratory gating of PET data. In: 2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC). IEEE; 2013:1–4.

38. Reynolds DA. Gaussian mixture models. Encycl biometrics. 2009;741(659–663).

39. Ng A, Jordan M, Weiss Y. On spectral clustering: Analysis and an algorithm. Adv Neural Inf Process Syst. 2001;14.

40. Fayad HJ, Lamare F, Le Rest CC, Bettinardi V, Visvikis D. Generation of 4-dimensional CT images based on 4-dimensional PET-derived motion fields. J Nucl Med. 2013;54(4):631–638. doi:10.2967/jnumed.112.110809 [PubMed: 23471313]

41. Lu Y, Fontaine K, Mulnix T, et al. Respiratory motion compensation for PET/CT with motion information derived from matched attenuation-corrected gated PET data. J Nucl Med. 2018;59(9):1480–1486. doi:10.2967/jnumed.117.203000 [PubMed: 29439015]

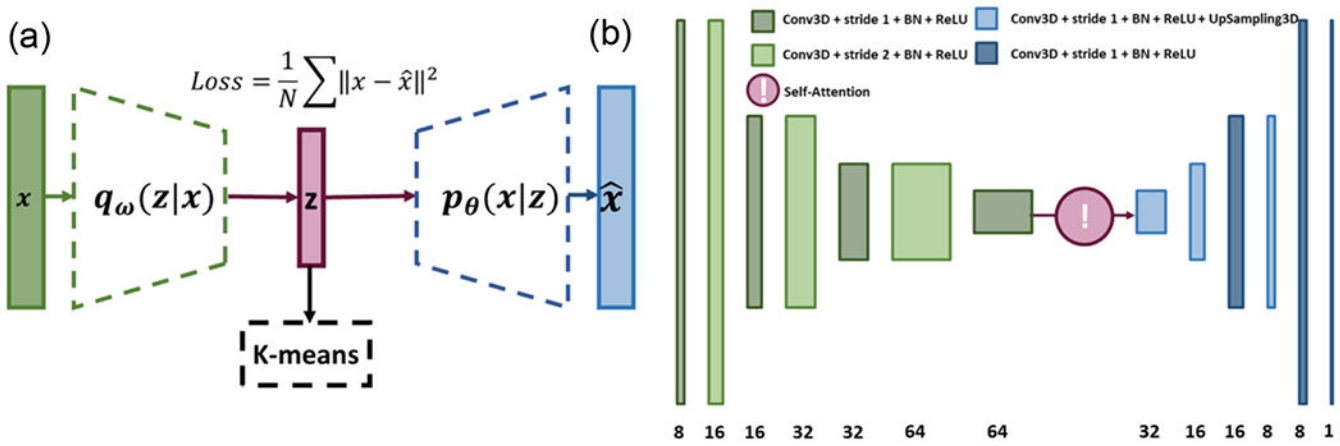

FIGURE 1.

(a) Autoencoder-based deep clustering network framework. (b) The schematic diagram of the AE with a self-attention gate. The number of features in each layer are indicated at the bottom of the graph.

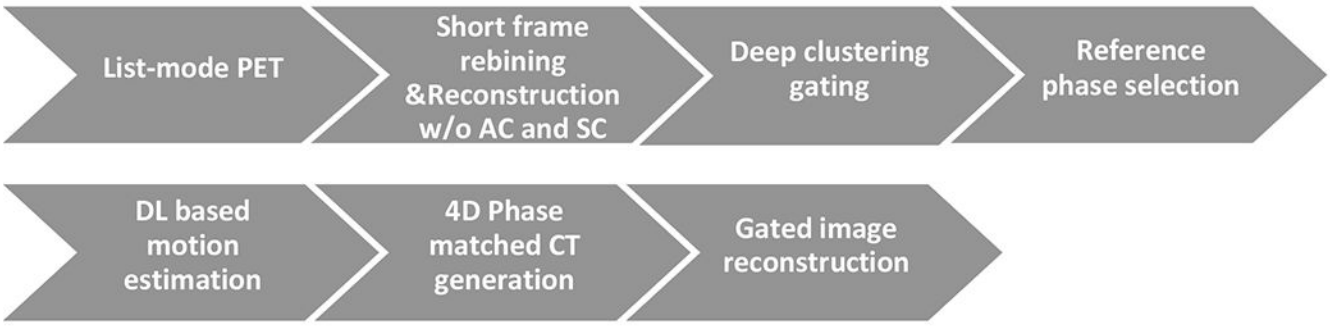

FIGURE 2.
The overall diagram of our proposed unsupervised deep learning framework for data-driven gating.

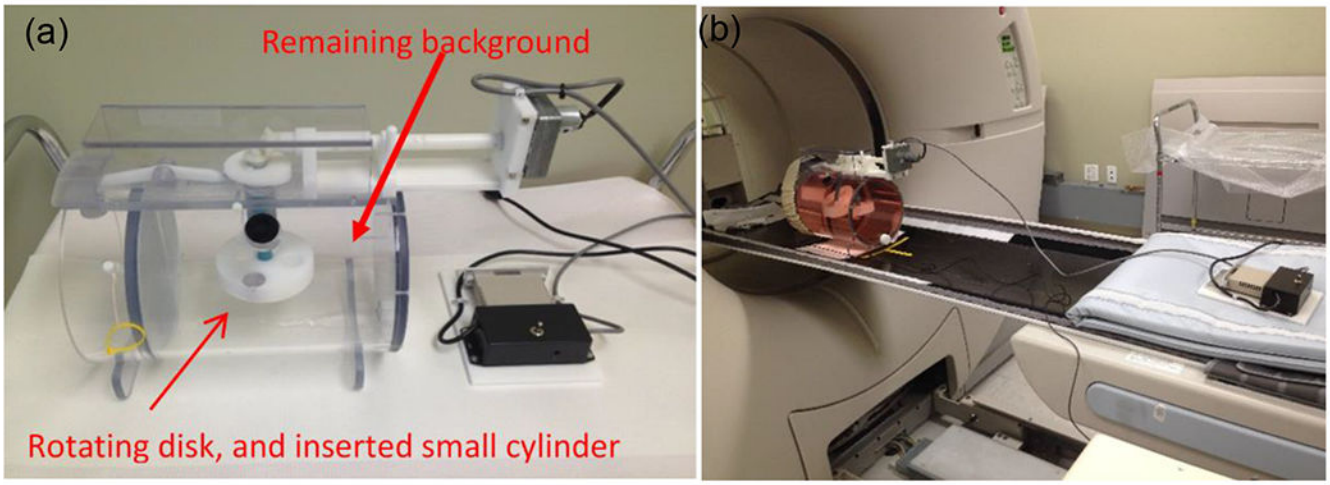

FIGURE 3.
(a) Picture of the respiratory motion phantom and (b) the acquisition setup with the Anzai belt attached to the phantom.

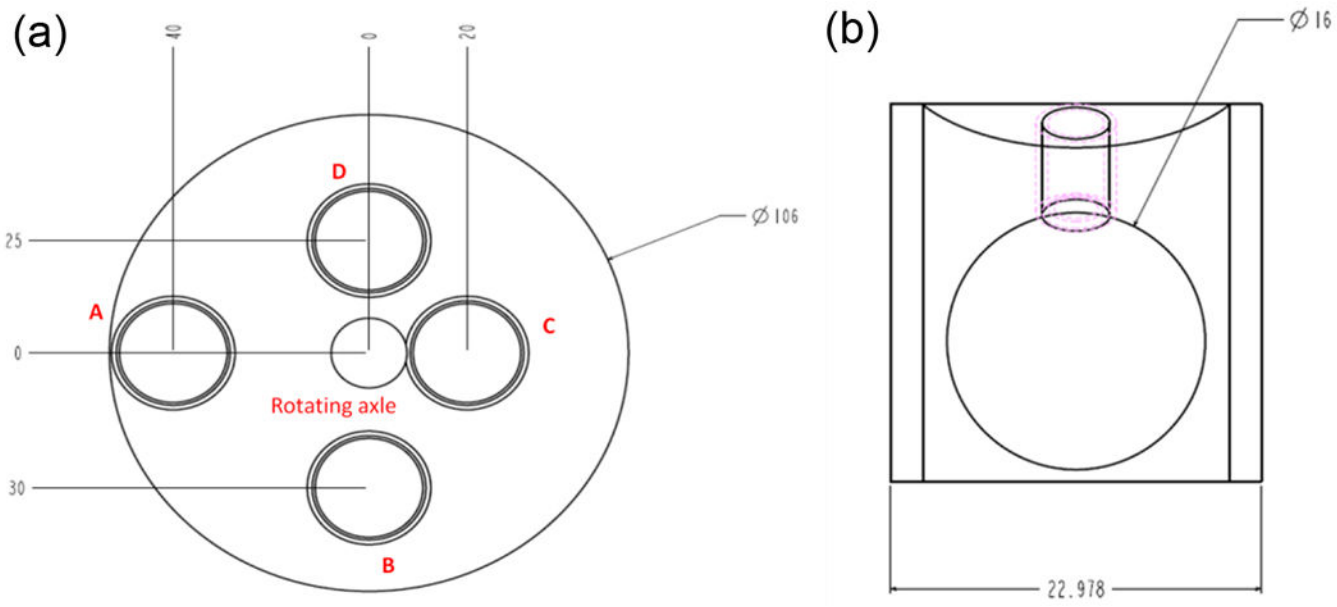

FIGURE 4.

(a) Mechanical drawing of the rotating disk. There are 4 holes in the rotating disc, and radial distance of each hole to the rotating center is 40 mm for “A”, 30 mm for “B”, 20 mm for “C” and 25 mm for “D”. (b) Mechanical drawing of each hole. A hot source (inner diameter: 16mm) was filled using a smaller cylinder in the hole “A”, and the other 3 holes were left empty (the diameter of each hole is 23mm). The empty holes are filled with the background activity.

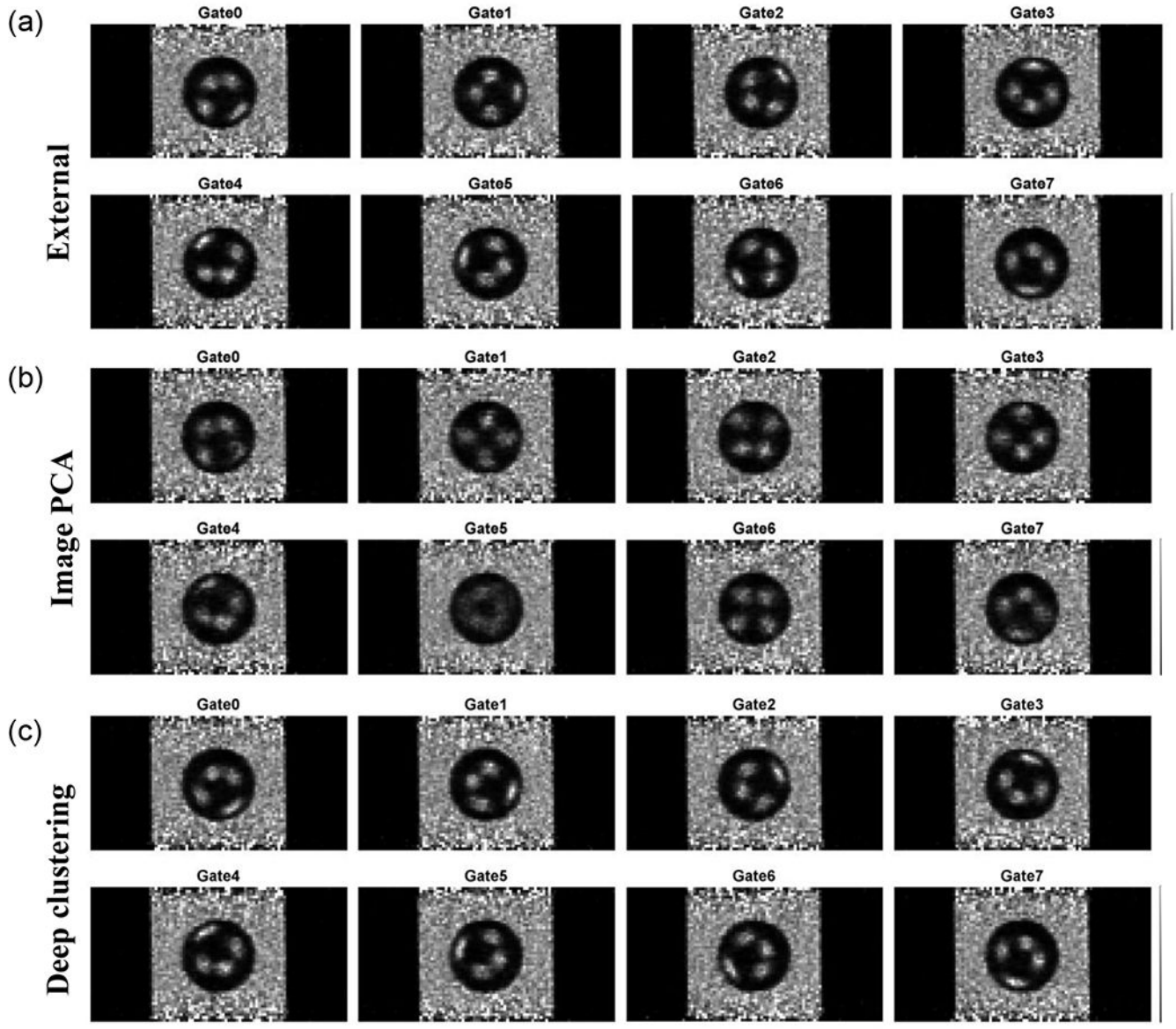

FIGURE 5.
Reconstruction of gated data using (a) External, (b) Image PCA and (c) Deep Clustering methods.

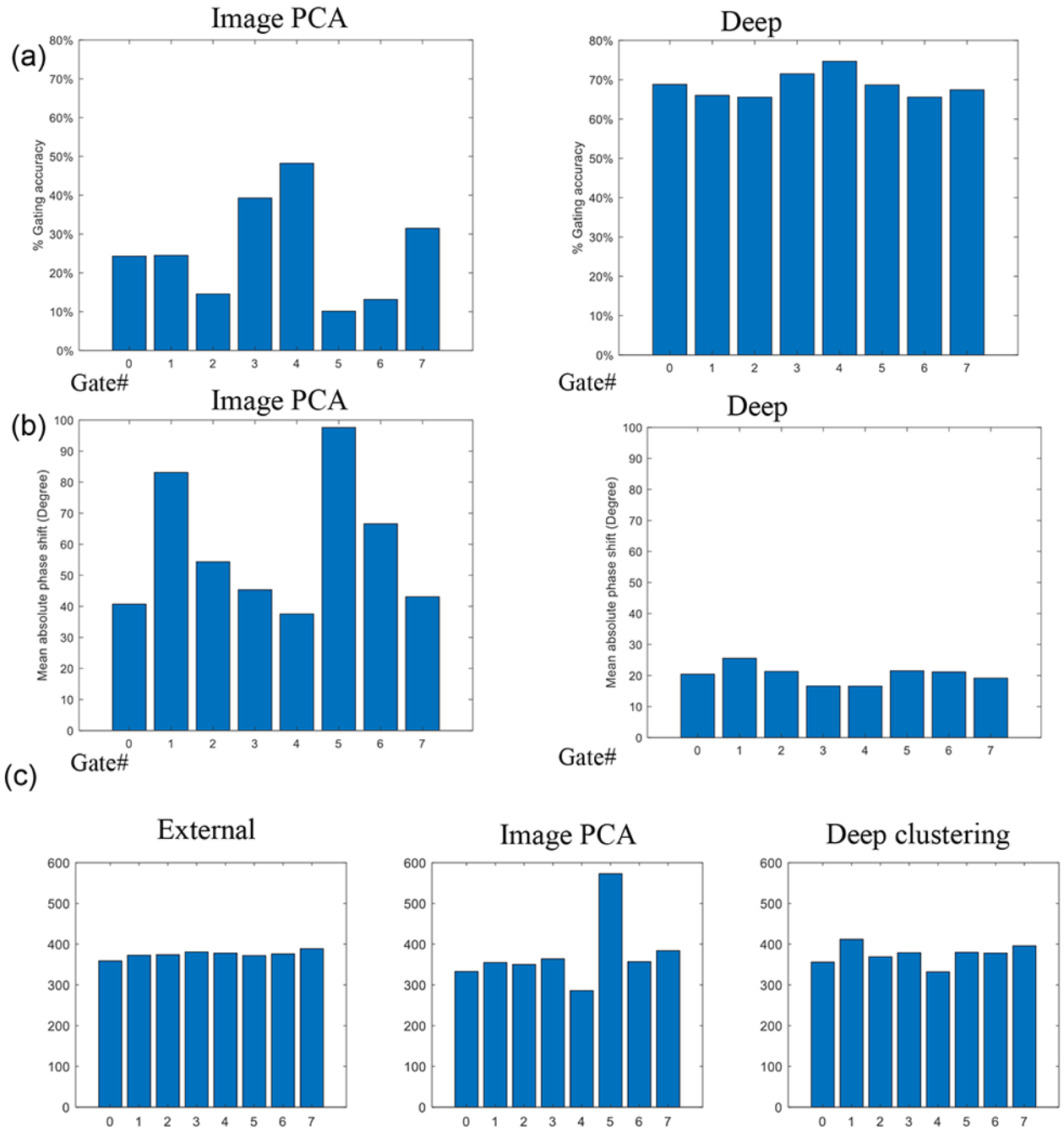

FIGURE 6.

Comparison between Image PCA and Deep Clustering methods. (a) Percentage gating accuracy. (b) Mean absolute phase shift relative to the gate centers. (c) Distribution of the count level over gates for the External, Image PCA and Deep Clustering methods.

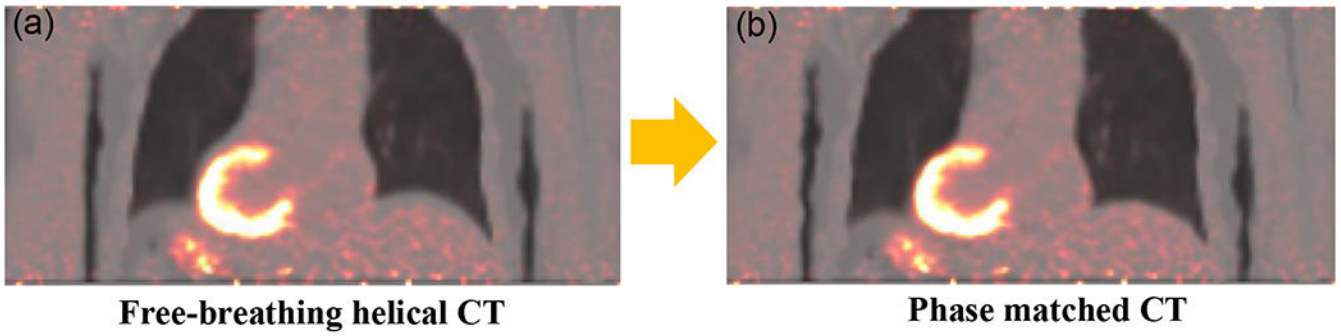

FIGURE 7.

Gated PET reconstructions at end-inspiration phase from Deep Clustering method superimposed with (a) free-breathing helical CT and (b) phase matched CT.

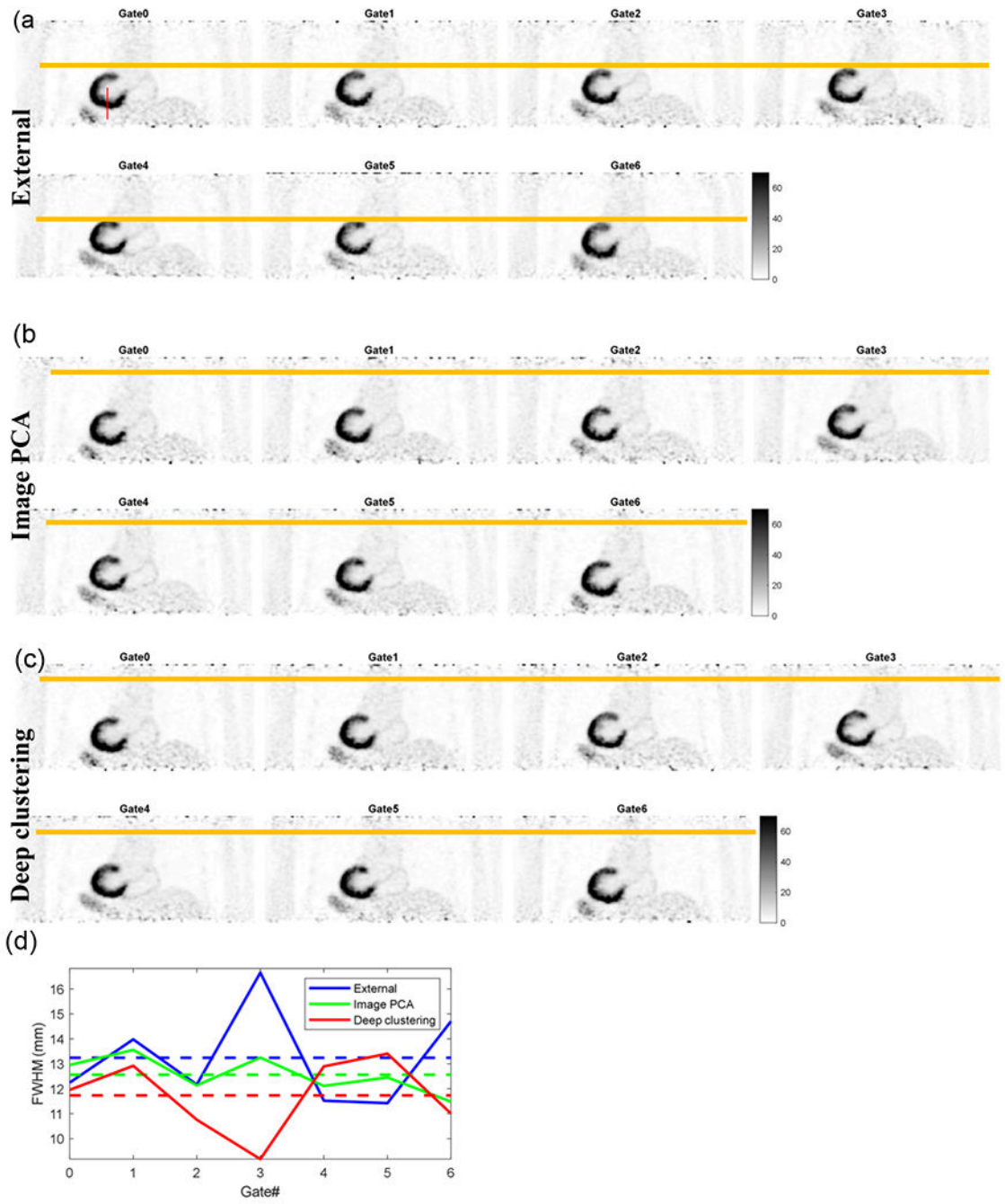

FIGURE 8.

Sampled reconstruction of gated images using (a) External, (b) Image PCA, (c) Deep Clustering methods. The yellow line marks the top of the myocardium position in Gate 3 (the reference gate). The average FWHMs over 25 different locations of the intensity profile (indicated by the red line) through the (d) inferior myocardial wall were plotted as a function of gate number (solid lines). The dashed lines denote the mean FWHMs over all gates of the three methods.

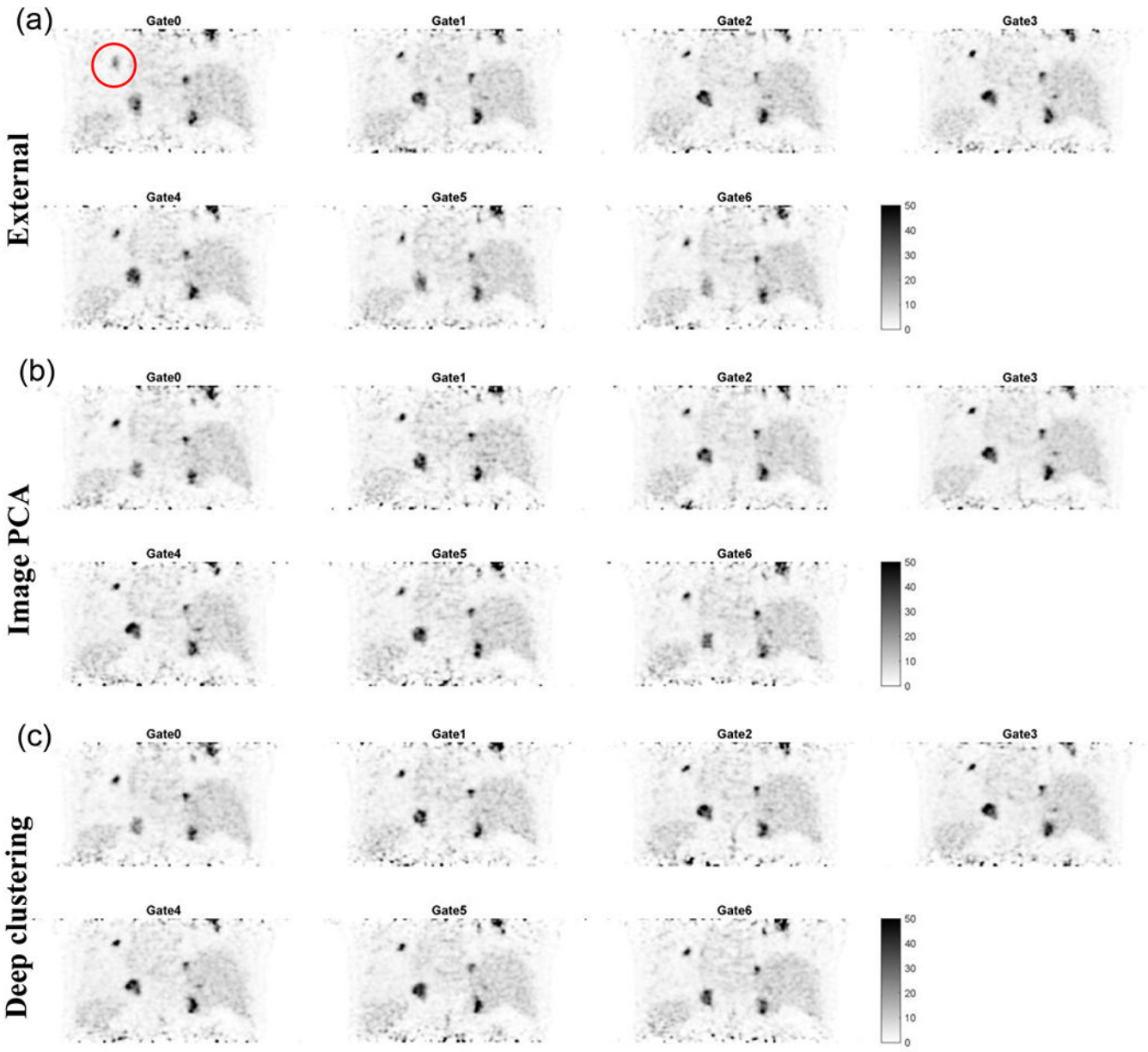

FIGURE 9.

Sampled reconstruction of gated images using External, Image PCA, Deep Clustering methods. The red circle marks the lesion used for quantification in Figure 10.

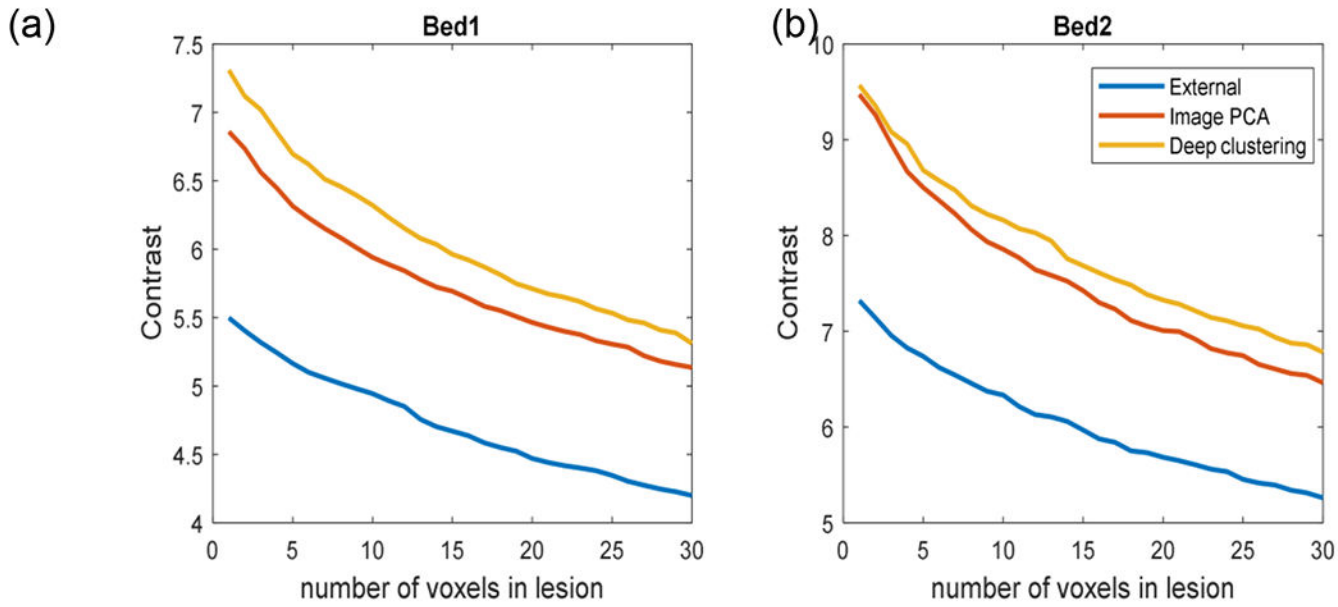

FIGURE 10.

Averaged contrast for a pulmonary lesion (marked by the red circle in Fig. 9) as a function of the lesion ROI size (number of voxels) for the External, Image PCA, Deep Clustering methods.

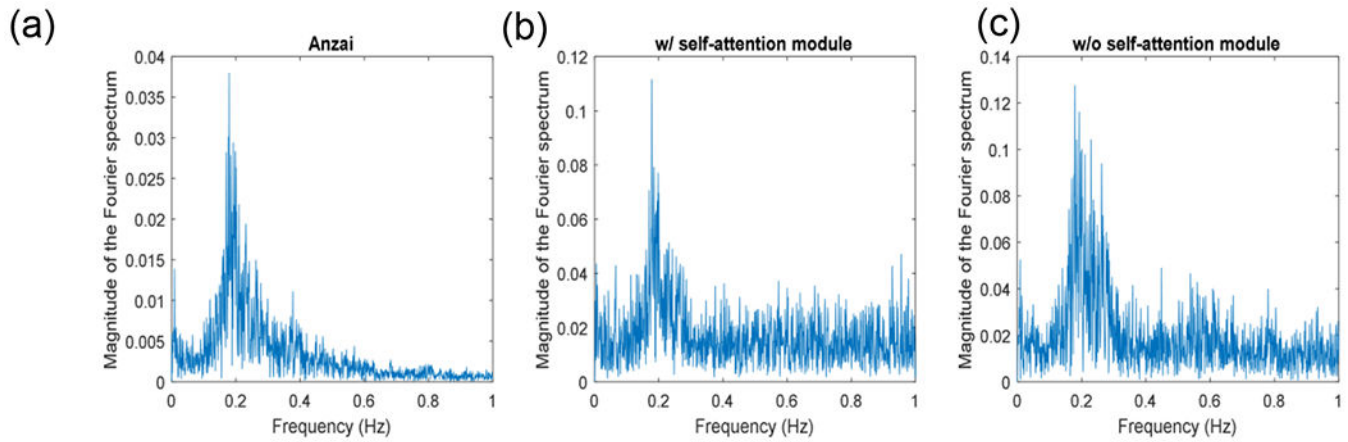

FIGURE 11.

Comparison of the frequency spectrums of (a) the Anzai signal and (b,c) the latent features with the maximum energy inside the respiratory frequency range extracted by the AE (b) with and (c) without the self-attention module.

TABLE 1.

Center of mass displacement in millimeters (mm) along x, y, z direction and the trajectory length for External, Image PCA, Deep clustering method.

	x	y	z	Trajectory length
External	0.71±0.18	1.08±0.61 *	3.49±1.58 **	6.88±2.87 **
Image PCA	0.80±0.40 *	1.89±1.09 *	5.13±1.59 **	10.38±3.47 **
Deep clustering	1.05±0.35 **	1.96±1.15	5.21±1.65	10.79±3.69

The p-values are indicated by '*' ($p < 0.05$) and '**' ($p < 0.01$), respectively.