Fast-MC-PET: A Novel Deep Learning-Aided Motion Correction and Reconstruction Framework for Accelerated PET

Bo Zhou[1(✉)], Yu-Jung Tsai[2], Jiazhen Zhang[1], Xueqi Guo[1], Huidong Xie[1], Xiongchao Chen[1], Tianshun Miao[2], Yihuan Lu[3], James S. Duncan[1,2], and Chi Liu[1,2]

[1] Department of Biomedical Engineering, Yale University, New Haven, US
bo.zhou@yale.edu
[2] Department of Radiology and Biomedical Imaging, Yale University, New Haven, US
[3] United Imaging Healthcare, Shanghai, China

Abstract. Patient motion during PET is inevitable. Its long acquisition time not only increases the motion and the associated artifacts but also the patient's discomfort, thus PET acceleration is desirable. However, accelerating PET acquisition will result in reconstructed images with low SNR, and the image quality will still be degraded by motion-induced artifacts. Most of the previous PET motion correction methods are motion type specific that require motion modeling, thus may fail when multiple types of motion present together. Also, those methods are customized for standard long acquisition and could not be directly applied to accelerated PET. To this end, modeling-free universal motion correction reconstruction for accelerated PET is still highly under-explored. In this work, we propose a novel deep learning-aided motion correction and reconstruction framework for accelerated PET, called Fast-MC-PET. Our framework consists of a universal motion correction (UMC) and a short-to-long acquisition reconstruction (SL-Reon) module. The UMC enables modeling-free motion correction by estimating quasi-continuous motion from ultra-short frame reconstructions and using this information for motion-compensated reconstruction. Then, the SL-Recon converts the accelerated UMC image with low counts to a high-quality image with high counts for our final reconstruction output. Our experimental results on human studies show that our Fast-MC-PET can enable 7-fold acceleration and use only 2 min acquisition to generate high-quality reconstruction images that outperform/match previous motion correction reconstruction methods using standard 15 min long acquisition data.

Keywords: Accelerated PET · Universal Motion Correction · Deep Reconstruction

A. Frangi et al. (Eds.): IPMI 2023, LNCS 13939, pp. 523–535, 2023.
https://doi.org/10.1007/978-3-031-34048-2_40

1 Introduction

Positron Emission Tomography (PET) is a commonly used functional imaging modality with wide applications in oncology, cardiology, neurology, and biomedical research. However, patient motion during the PET scan, including both involuntary motions (i.e. respiratory, cardiac, and bowel motions) and voluntary motions (i.e. body and head motions), can lead to significant motion artifacts, degrading the downstream clinical tasks. Moreover, the long acquisition time that easily exceeds 15 min, will lead to increased patient motion, patient discomfort, and low patient throughput.

In previous works of PET motion correction (MC), a variety of external device-aided and data-driven MC methods have been developed for correcting specific motion types. For example, in respiratory MC, Chan *et al.* [4] developed a non-rigid event-by-event continuous MC list-mode reconstruction method. Lu *et al.* [10] further improved their method by generating matched attenuation-corrected gate PET for respiratory motion estimation. In body MC, Andersson *et al.* [1] proposed to divide the PET list-mode data into predefined temporal frames for reconstructions, where the reconstructions of each frame are registered to a reference frame for body MC. Later, Lu *et al.* [11] further developed a reconstruction-free center-of-distribution-based body motion detection and correction method. In cardiac MC, cardiac cycle tracking/gating using electrocardiography (ECG) is still the gold-standard [12]. While providing efficient MC solutions to reduce motion artifacts for different motion types, these methods usually require prior knowledge of the motion type and need motion-type-specific modeling. Thus, these previous MC methods may lead to sub-optimal image quality or fail when multiple motion types are present simultaneously. There are also recent attempts in using ultra-fast list-mode reconstruction of short PET frames to estimate motion during the PET scan [18,21]. However, these methods may not adapt well to many motion types with non-rigid motion [18], and extending to non-rigid motion is computationally infeasible, i.e. requiring non-rigid registration of thousands of frames for a single scan using traditional registration algorithms [21]. In addition, it still requires the standard long acquisition to collect sufficient events to achieve a reasonable signal-to-noise ratio (SNR) in the final reconstruction. On the other hand, previous works have also investigated the feasibility of reducing the PET acquisition time. Lindemann *et al.* [9] and Lasnon *et al.* [8] found that one can reasonably maintain the PET image quality and lesion detectability with two-fold acquisition time reduction using traditional reconstructions. Weyts *et al.* [19] show that a deep learning-based denoising model can enable two-fold PET acquisition time reduction and provide image quality that matches with the full acquisition. However, these works only show the feasibility of a 2-fold time reduction and did not consider the residual motions during the accelerated acquisition.

In this work, we aim to address these challenges by developing a PET reconstruction framework that can 1) reduce the acquisition time, i.e. 7-fold acceleration, and 2) correct the residual motion, regardless of the motion type, in the accelerated acquisition. Specifically, we propose a novel deep learning-aided

data-driven motion reduction and accelerated PET reconstruction framework, called Fast-MC-PET. In the Fast-MC-PET, we first design a universal motion correction method aided by deep learning to reconstruct a motion-reduced image from the short acquisition. While reducing the motion artifacts given the accelerated acquisition and our motion correction, the reconstructed image still suffers from high noise levels due to low event counts. Thus, in the second step of Fast-MC-PET, we also deploy a deep generative network to convert the low-counts images to high-counts images. Our experimental results on real human data demonstrate the Fast-MC-PET can generate high-quality images with reduced motion-induced errors while enabling 7-fold accelerated PET acquisition.

2 Methods

Our Fast-MC-PET consists of two key components, including a universal motion correction (UMC) module and a short-to-long acquisition reconstruction (SL-Recon) module. In UMC, we first partition the list-mode data into ultra-short list-mode data, i.e. every 500 ms, and estimate a quasi-continuous motion over the short acquisition. Given the motion and the original list-mode data, a motion-corrected short-acquisition image is then reconstructed by a motion-compensated OSEM list-mode reconstruction. Finally, a deep generative model is devised to transform the motion-corrected short-acquisition image into a high-count long-acquisition image, thus providing a motion-corrected high-count image using only accelerated short-acquisition. In the following sections, we will describe these steps in detail (Fig. 1).

2.1 Universal Motion Correction

With the short acquisition data, the UMC aims to generate a motion reduced low-count reconstruction. The UMC consists of three steps, including point cloud image (PCI) & paired gated image generation, quasi-continuous motion estimation, and motion-compensated OSEM list-mode reconstruction.

Point Cloud and Paired Gated Image Generation. To estimate a continuous motion, the list-mode data is first partitioned into a series of ultra-short list-mode data, i.e. every 500 ms. For every 500 ms list-mode data, we back-project the Line-of-Response (LOR) of each event within the time-of-flight (TOF) bin, and all the back-projected LORs form a PCI for this short time frame. The PCI reconstruction can be formulated as

$$P_{j,t} = \sum_i \frac{c_{i,j,t} L_{i,t}}{Q_j}, \tag{1}$$

where $c_{i,j,t}$ is the system matrix that represents the contribution of an annihilation originating from pixel j being detected on LOR i at time t, accounting for geometry, resolution, and solid angle effects. $L_{i,t}$ is the decay correction factor. Q_j is the sensitivity of voxel j that is pre-computed via $Q_j = \sum_i c_{i,j}$, and $P_{j,t}$ is the back-projected value of voxel j at time t with sensitivity correction.

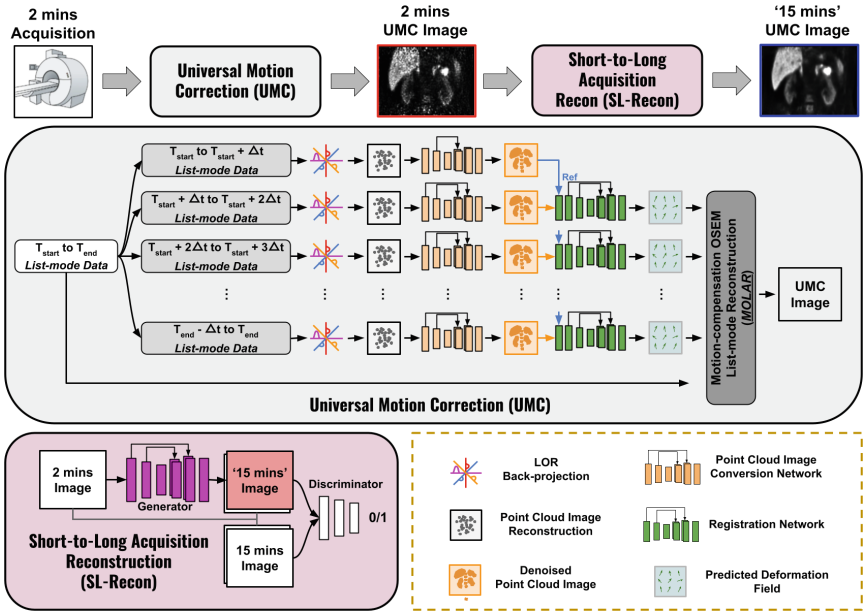

Fig. 1. The overall pipeline of Fast-MC-PET. The Universal Motion Correction (UMC) module (grey box) reconstructs motion-reduced image from the short acquisition data. The Short-to-Long Acquisition Reconstruction (SL-Recon) module (pink box) converts the UMC image from short acquisition to long acquisition. (Color figure online)

Due to the ultra-low-counts level, the signal-to-noise ratio (SNR) of PCI is low and is unsuitable for motion estimation tasks, as demonstrated in Fig. 2's 1st row. Thus, we deploy a deep learning-based denoising network, i.e. UNet [16], that aims to convert PCIs to gated OSEM images with high SNR. To train the denoising network, we first reconstruct the amplitude-based respiratory gated OSEM images [15] using the body motion free list-mode data, extracted by the Centroid-Of-Distribution (COD)-based body motion detection method [11]. Then, within each gate, we randomly extract 10% PCIs to construct the training pairs of PCI and the corresponding gated image. $\mathcal{L}_2$ loss is used for the network training, and can be formulated as

$$\mathcal{L}_{dn} = ||\gamma_g - f_{dn}(P_g)||_2^2 \tag{2}$$

where γ_g is the gated OSEM image and P_g is the randomly extracted PCI that lies in the same gate. With a trained denoising model $f_{dn}(\cdot)$, the series of PCIs can then be converted to a series of high-quality denoised PCI (dPCI) via:

$$\gamma_t = f_{dn}(P_t) \tag{3}$$

where γ_t is the denoised images with $t = (0 \sim \Delta t, \Delta t \sim 2\Delta t, ..., T - \Delta t \sim T)$. Here, we set $\Delta t = 0.5s$ and $T = 120s$ here, thus generating 240 3D images. Examples of dPCIs are illustrated in Fig. 2's 2nd row.

Quasi-Continuous Motion Estimation. A quasi-continuous motion can be estimated using the series of dPCIs from the previous step. Within the first 5 s, the dPCI in the expiration phase, i.e. with the highest COD coordinates in the z-direction, is chosen as the reference frame γ_{ref} for all the other frames γ_t, resulted in 239 dPCI pairs requiring registration. Conventional registration methods [14, 20] are time-consuming, and it is prohibitively long to register hundreds of 3D pairs here. Thus, we propose to use a deep learning-based registration method for fast motion estimation [3] in our framework. Given the reference dPCI image γ_{ref} and the source dPCI image γ_t, we use a motion estimation network, i.e. UNet [16], to predict the motion deformation $M_t = f_m(\gamma_{ref}, \gamma_t)$. The network is trained by optimizing the following loss function:

$$\mathcal{L}_m = ||\gamma_{ref} - M_t \circ \gamma_t||_2^2 + \beta||\nabla M_t||_2^2 \tag{4}$$

where the first term measures the image similarity after applying the motion prediction M_t, and the second term is a deformation regularization that adopts a L2-norm of the gradient of the deformation. The regularization's weight is set as $\beta = 0.001$. During training, γ_{ref} and γ_t are randomly selected from the gated images. With a trained motion estimation network $f_m(\cdot)$, we can then estimate the quasi-continuous motion using $M_t = f_m(\gamma_{ref}, \gamma_t)$ with $t = (0 \sim \Delta t, \Delta t \sim 2\Delta t, ..., T - \Delta t \sim T)$.

Motion-compensated OSEM List-mode Reconstruction. To reconstruct a single image λ at the reference location γ_{ref} using all the coincidence events, we can deform the system matrix at each time t to the reference location, generating new deformed system matrixs $c_{i,j}^{t\to ref}$ using M_t from the previous step. Deforming the system matrix can be seen as "bending" the LORs into curves of response (CORs), where both forward and back-projections are traced along the CORs. In list-mode notation, for event k occurring on LOR $i(k)$ at time $t(k)$, we replace indexes i by k, and substitute $c_{k,j}$ in the previous TOF-MOLAR [6] by $c_{k,j,\tau_k}^{t\to ref}$. The OSEM updating equation can thus be formulated as:

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{Q_j} \sum_{k=1}^{K} \frac{c_{k,j,\tau_k}^{t\to ref} L_k A_k N_k}{T(\sum_{j'} c_{k,j',\tau_k}^{t\to ref} L_k A_k N_k \lambda_{j'}^n + R_{k,\tau_k} + S_{k,\tau_k})} \tag{5}$$

$$Q_j = \frac{1}{n_T} \sum_{t'=1}^{n_T} \sum_{i=1}^{I} \sum_{\tau=1}^{n_\tau} c_{i,j,t',\tau}^{t\to ref} L_{i,t'} A_{i,t'} N_i \tag{6}$$

where n is the number of iteration, k is the index of each detected event, $c_{k,j,\tau_k}^{t\to ref}$ is the deformed system matrix element with τ_k denoting the TOF bin for event k. L_k is the decay factor and A_k is the attenuation factor derived from CT. N_k is the sensitivity term, R_{k,τ_k} is the randoms rate estimate, and S_{k,τ_k} is the scatter rate estimate in counts per second in TOF bin τ_k. The random events are estimated from the product of the singles rates of the two detectors for each LOR, and then uniformly distributed across all TOF bins. Here, Q is the sensitivity image that is pre-computed by back-projecting randomly sampled events along the CORs to

account for the motion on voxel sensitivity. When calculating Q, each time frame of duration T is divided into n_T short time bins, i.e. t'. Moreover, n_τ denotes the total number of TOF bins ($n_\tau = 13$ for the Siemens mCT PET scanner used in this study). Here, we set the number of iteration to 2 and the number of subsets to 21 for our UMC reconstructions.

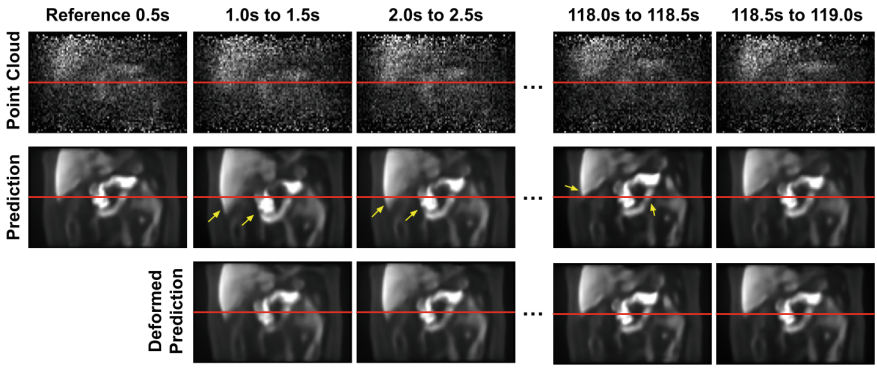

Fig. 2. Examples of the Point Cloud Images (PCIs), the denoised PCIs (dPCIs), and the deformed dPCIs using estimated motion fields.

2.2 Short-to-Long Acquisition Reconstruction

Even though the UMC reduces the motion effects in the reconstruction, the UMC image still suffers from low SNR due to the limited counts from the short acquisition, as compared to the long acquisition. Thus, we propose to use a short-to-long acquisition reconstruction (SL-Recon) to convert the UMC image from a short-acquisition to a long-acquisition one. Here, we use a conditional generative adversarial network for this reconstruction. Given a UMC image λ_s from the short acquisition, we can use a generative network, i.e. UNet [16], that directly predicts the UMC image λ_l from a long acquisition from it. The SL-Recon network is trained using both a pixel-wise L2 loss and an adversarial loss defined as:

$$\mathcal{L}_2 = ||G(\lambda_s) - \lambda_l||_2^2 \tag{7}$$

$$\mathcal{L}_{adv} = -log(D_{gan}(\lambda_l|\lambda_s)) - log(1 - D_{gan}(G(\lambda_s)|\lambda_s)) \tag{8}$$

where G is the SL-Recon generative network and D is the discriminator network. Here, we simply use OSEM reconstructions from long acquisitions (15 min), paired with OSEM reconstructions from short acquisitions (2 min in the center period), for the network's training.

2.3 Evaluation on Human Data

We included 26 pancreatic ^{18}F-FPDTBZ [13] PET/CT patient studies. All PET data were obtained in list mode using the 4-ring Siemens Biograph mCT scanners

equipped with the AZ-733V respiratory gating system (Anzai Medical, Tokyo, Japan). The Anzai respiratory trace was recorded 40 Hz for all subjects. The average dose administered to the patients is 9.13±1.37 mCi. 15 min of the list-mode acquisition were used for each patient study. We used 23 patients to generate the training data for the PCI denoising model, the motion estimation model, and the SL-Recon model. Extensive evaluations were performed on the remaining 3 patients with different motion types. For training the PCI denoising model and the motion estimation model, we generated 5 gated images for each patient using OSEM (21 subsets and 2 iterations). For training the SL-Recon model, the training pairs of long/short acquisition images were reconstructed using the same OSEM protocol without gating. All the images were reconstructed into $200 \times 200 \times 109$ 3D volumes with a voxel size of $2.032 \times 2.032 \times 2.027\ mm^3$.

2.4 Implementation Details

We implemented our deep learning modules using Pytorch. We used the ADAM optimizer [7] with a learning rate of 10^{-4} for training the PCI denoising network, motion estimation network, and the SL-Recon network. We set the batch size to 3 for all networks' training. All of our models were trained on an NVIDIA Quadro RTX 8000 GPU. The PCI denoising network was trained for 200 epochs, and then fine-tuned for 10 epochs on the patient-specific gated images of the test patient during the test time. The motion estimation network was trained for 250 epochs, and the SL-Recon network was trained for 200 epochs. To prevent overfitting, we also implemented 'on-the-fly' data augmentation for the PCI denoising and SL-Recon networks. During training, we performed $64 \times 64 \times 64$ random cropping, and then randomly flip the cropped volumes along the x, y, and z-axis.

3 Results

The qualitative comparison of Fast-MC-PET reconstructions is shown in Fig. 3. As we can observe, the 2 min reconstruction with no motion correction (NMC) suffers from both motion blurring and high-noise levels due to low counts. The first patient has both body/torso motion and respiratory motion during the 2 min PET scan, thus introducing heavy blurring for major organ boundaries, i.e. liver and kidneys. The 2 min UMC image recovers the sharp organ boundaries by correcting those motions during the short acquisition. Based on the UMC image from 2 min acquisition, the final Fast-MC-PET image further reduces the noise thus providing a near motion-free and high-count image, matching the 15 min UMC image quality. The second patient with respiratory and bowel motion introduces significant image blurring for the pancreas (view 1) and intestines (view 2). The 2 min UMC image can recover the diminished details inside these organ regions. The final Fast-MC-PET image further reduces the noise, thus generating a high-quality image with motion correction and high counts. On the other hand, by reducing the acquisition time from 15 min to 2 min, we can see that the diminished organ structures, especially the intestine structure (view 2) in 15 min

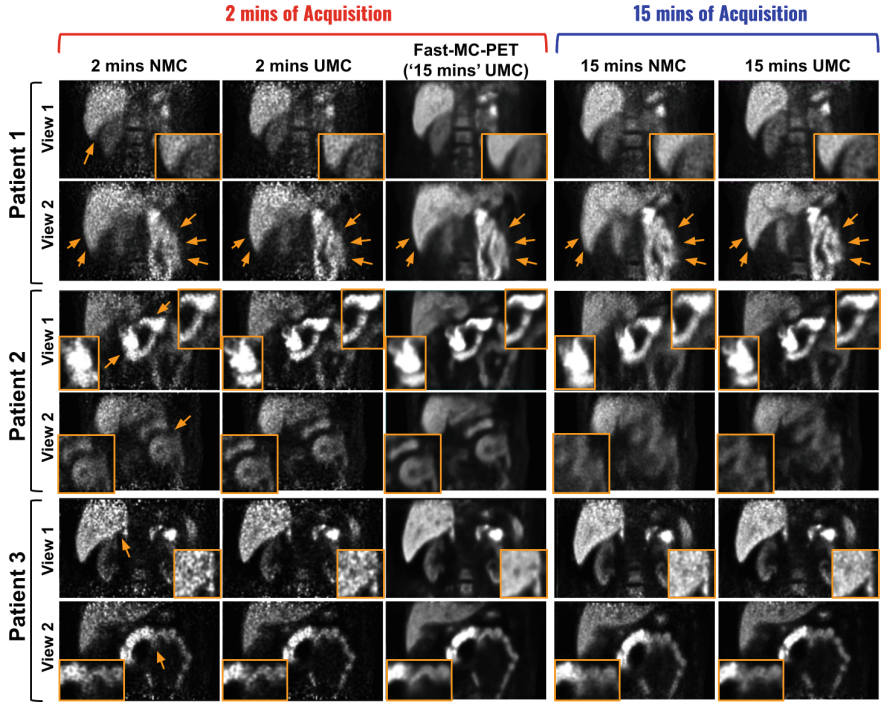

Fig. 3. Visualization of Fast-MC-PET reconstructions. The 2min UMC images (2nd column) contain less motion blurring, as compared to the no motion correction (NMC) images (1st column). The virtual 15 min UMC images (3rd column) predicted from 2 min UMC images (2nd column) provide image-quality that match the true 15 min images (last column).

NMC, can be preliminarily restored in 2 min NMC. Complex motion, e.g. bowel motion, in a 15 min long acquisition is extremely challenging to correct even with UMC. Thus, based on 2 min acquisition, the Fast-MC-PET here shows better reconstruction quality with better structural recovery. Similar observations can be found for the third patient with respiratory and bowel motion, where the 2 min-based Fast-MC-PET provides reconstruction quality matched the 15 min UMC reconstruction.

We compared our 2 min-based Fast-MC-PET reconstructions to previous correction methods that are long acquisition based, i.e. 15 min. The visual comparison is shown in Fig. 4. First, we compared with the classic respiratory motion correction method [2] that reduces the motion and noise by averaging the aligned amplitude-gated images, where non-rigid registration [14] is used for alignments. Then, we compared our method with the NR-INTEX [4] that compensates for the respiratory motion by estimating the continuous deformation field using internal-external motion correlation which is considered the current state-of-the-art method. Both previous methods require specific motion-type modeling, and

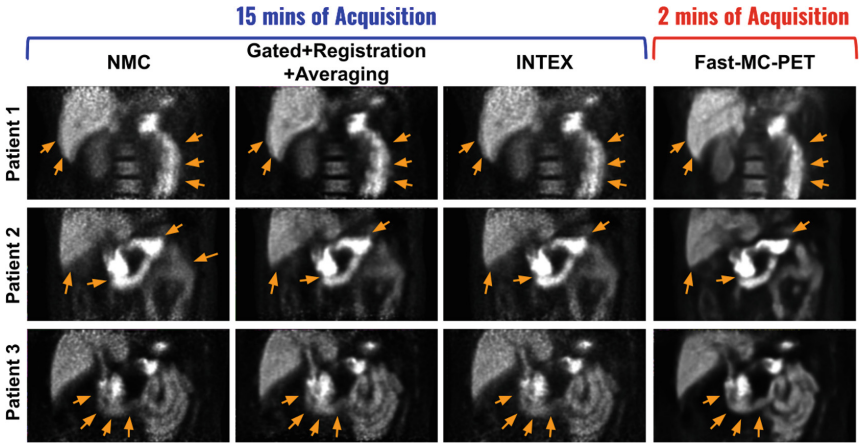

Fig. 4. Comparisons to previous motion correction methods. Our Fast-MC-PET with 2 min acquisition show improved structural details recovery (orange arrows), as compared to previous methods with 15 min acquisition.

thus fail when additional motion types are present, e.g. body motion (Patient 1) and bowel motion (Patient 3). The UMC module in the Fast-MC-PET is not specific to any motion type and thus can correct different types of motion together. Therefore, our Fast-MC-PET can provide consistently better results when multiple types of motion co-exist (Patients 1 and 3), and generate comparable reconstruction quality when respiratory motion is dominating (Patient 2).

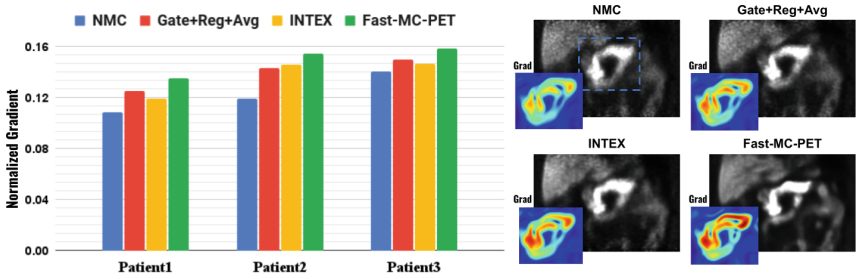

Fig. 5. Comparison of the gradient of reconstructions. Left: quantitative evaluation using the mean gradient value. Right: visual comparison of the reconstruction and the gradient.

For quantitative evaluation, we computed the mean normalized gradient of the reconstructions, where better reconstruction with sharper structure will have higher gradient values. The results are summarized in Fig. 5. The normalized gradient values of Fast-MC-PET are 0.159, 0.154, and 0.132 for Patients 1, 2, and 3, respectively, which are consistently higher than all previous methods. A

comparison example from Patient 2 is shown on the right. The gradient image of the Fast-MC-PET using only 2 min acquisition shows higher gradient values and more continuous structure patterns when compared to previous methods based on 15 min acquisition.

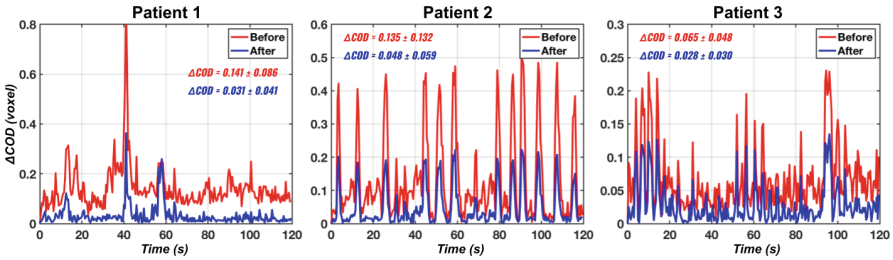

Fig. 6. The difference of COD trace between the reference frame and the current frame (ΔCOD) over the 2 min acquisition. The ΔCOD before (red) and after (blue) UMC correction are plotted for all three patients. The mean ΔCODs are reported in the plots. (Color figure online)

Ablative evaluation of motion correction is shown in Fig. 6. The difference of COD between the reference frame and the current frame (ΔCOD) over the 2 min acquisition is visualized. For Patient 1 with body motion and irregular breathing pattern, the ΔCOD curve before correction contains irregular steep changes leading to a mean ΔCOD of 0.141 ± 0.086. With the UMC in our Fast-MC-PET, the curve after correction is much more stable with a reduced mean ΔCOD of 0.031 ± 0.041 with significance ($p < 0.001$). For Patients 2 and 3 with more stable and regular motion patterns, the UMC can also reduce the mean ΔCOD from 0.135 ± 0.132 to 0.048 ± 0.059 and from 0.065 ± 0.048 to 0.028 ± 0.030, respectively. Both with significance ($p < 0.001$). A patient example of PCIs over the 2 min acquisition before and after applying the UMC correction is shown in Fig. 2.

4 Discussion

In this work, we propose a novel deep learning-aided data-driven motion correction and reconstruction framework for accelerated PET (Fast-MC-PET). The proposed method can accelerate the PET acquisition by nearly 7-fold and use only 2 min acquisition while providing high-quality reconstruction with motion correction. In this framework, we first devise a UMC module that estimates continuous motion based on PCIs and use this information to reconstruct motion-compensated images. Instead of using 15 min long acquisition that 1) inherits more motion due to long scanning time and 2) requires registrations of 1800 PCI pairs in UMC, we use 2 min accelerated acquisition with less motion and only requires registrations of 240 PCI pairs. The averaged registration inference

time for one pair is 0.41s, thus needing about 98.5s for all registration in UMC which is more manageable. The UMC reconstruction from accelerated acquisition can then be inputted into the SL-Recon module to directly generate the 15 min long acquisition motion-corrected reconstruction. With this simple yet efficient pipeline, we can generate high-quality motion corrected accelerated PET reconstruction that potentially outperforms previous methods with the standard long acquisition.

There are a few limitations and opportunities that are the subject of our ongoing work. First, our pilot study only tested on ^{18}F-FPDTBZ patients who were all scanned using Siemens mCT. The trained model may not directly generalize well to a different PET tracer/scanner. However, if the training data of different tracers/scanners is available, the Fast-MC-PET can be fine-tuned and potentially adapted to these distributions. Multi-institutional federated learning [22] may also be used to improve the adaptation. In the future, we will further evaluate the performance using patients scanned with different PET tracers/scanners. Second, we used a temporal resolution of 500 ms for PCI in UMC with a focus on abdominal region motion correction in this work. A higher temporal resolution, e.g. 100 ms, may be needed for cardiac motion correction in the chest region, which is an important direction in our future investigation. Third, the UMC correction performance is still not perfect, as shown in Fig. 6 blue curves, where the ΔCOD values are non-zero. The current implementation uses a simple 3-level UNet for motion prediction. Deploying a more advanced registration network, e.g. transformer-based network [5] and temporal registration networks [23], may potentially further reduce the registration error and improve the final reconstruction quality. Lastly, the PCI denoising step requires supervised training from paired gated images, which is time-consuming to prepare. In the future, we will also investigate self-supervised denoising methods, e.g. Noise2Void [17], for PCI denoising in our Fast-MC-PET.

5 Conclusion

This paper presents a deep learning-aided motion correction and reconstruction framework for accelerated PET, called Fast-MC-PET. The Fast-MC-PET consisting of UMC and SL-Recon, uses only 2 min accelerated PET acquisition data for high-quality reconstruction. The UMC reconstructs motion-corrected short acquisition image, regardless of the motion type in the abdominal region. The SL-Recon then converts the 2 min UMC image into virtual 15 min UMC image. The experimental results demonstrate that our proposed method can accelerate acquisition by nearly 7-fold and generate high-quality motion-corrected reconstruction for patients with different motions.

References

1. Andersson, J.L.: How to obtain high-accuracy image registration: application to movement correction of dynamic positron emission tomography data. Eur. J. Nucl. Med. **25**(6), 575–586 (1998)

2. Bai, W., Brady, M.: Regularized b-spline deformable registration for respiratory motion correction in pet images. Phys. Med. Biol. **54**(9), 2719 (2009)
3. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: VoxelMorph: a learning framework for deformable medical image registration. IEEE Trans. Med. Imaging **38**(8), 1788–1800 (2019)
4. Chan, C., et al.: Non-rigid event-by-event continuous respiratory motion compensated list-mode reconstruction for pet. IEEE Trans. Med. Imaging **37**(2), 504–515 (2017)
5. Chen, J., Frey, E.C., He, Y., Segars, W.P., Li, Y., Du, Y.: TransMorph: transformer for unsupervised medical image registration. Med. Image Anal. **82**, 102615 (2022)
6. Jin, X., et al.: List-mode reconstruction for the biograph MCT with physics modeling and event-by-event motion correction. Phys. Med. Biol. **58**(16), 5567 (2013)
7. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
8. Lasnon, C., et al.: How fast can we scan patients with modern (digital) PET/CT systems? Eur. J. Radiol. **129**, 109144 (2020)
9. Lindemann, M.E., Stebner, V., Tschischka, A., Kirchner, J., Umutlu, L., Quick, H.H.: Towards fast whole-body PET/MR: investigation of pet image quality versus reduced pet acquisition times. PLoS ONE **13**(10), e0206573 (2018)
10. Lu, Y., et al.: Respiratory motion compensation for PET/CT with motion information derived from matched attenuation-corrected gated pet data. J. Nucl. Med. **59**(9), 1480–1486 (2018)
11. Lu, Y., et al.: Data-driven voluntary body motion detection and non-rigid event-by-event correction for static and dynamic pet. Phys. Med. Biol. **64**(6), 065002 (2019)
12. Lu, Y., Liu, C.: Patient motion correction for dynamic cardiac pet: current status and challenges. J. Nucl. Cardiol. **27**(6), 1999–2002 (2020)
13. Normandin, M.D., et al.: In vivo imaging of endogenous pancreatic β-cell mass in healthy and type 1 diabetic subjects using 18f-fluoropropyl-dihydrotetrabenazine and pet. J. Nucl. Med. **53**(6), 908–916 (2012)
14. Papademetris, X., et al.: Bioimage suite: an integrated medical image analysis suite: an update. Insight J. **2006**, 209 (2006)
15. Ren, S., et al.: Data-driven event-by-event respiratory motion correction using TOF PET list-mode centroid of distribution. Phys. Med. Biol. **62**(12), 4741 (2017)
16. Ronneberger, Olaf, Fischer, Philipp, Brox, Thomas: U-Net: convolutional networks for biomedical image segmentation. In: Navab, Nassir, Hornegger, Joachim, Wells, William M.., Frangi, Alejandro F.. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
17. Song, T.A., Yang, F., Dutta, J.: Noise2void: unsupervised denoising of pet images. Phys. Med. Biol. **66**(21), 214002 (2021)
18. Spangler-Bickell, M.G., Deller, T.W., Bettinardi, V., Jansen, F.: Ultra-fast list-mode reconstruction of short pet frames and example applications. J. Nucl. Med. **62**(2), 287–292 (2021)
19. Weyts, K., et al.: Artificial intelligence-based pet denoising could allow a two-fold reduction in [18f] FDG PET acquisition time in digital PET/CT. Eur. J. Nucl. Med. Mol. Imaging **49**, 1–11 (2022). https://doi.org/10.1007/s00259-022-05800-1
20. Xu, Z., et al.: Evaluation of six registration methods for the human abdomen on clinically acquired CT. IEEE Trans. Biomed. Eng. **63**(8), 1563–1572 (2016)
21. Zhang, J., Fontaine, K., Carson, R., Onofrey, J., Lu, Y.: Deep learning-aided data-driven quasi-continous non-rigid motion correction in PET. In: 2021 28th IEEE Nuclear Science Symposium and Medical Imaging Conference (2021)

22. Zhou, B., et al.: Federated transfer learning for low-dose pet denoising: a pilot study with simulated heterogeneous data. IEEE Trans. Rad. Plasma Med. Sci. (2022)
23. Zhou, B., Tsai, Y.J., Chen, X., Duncan, J.S., Liu, C.: MDPET: a unified motion correction and denoising adversarial network for low-dose gated PET. IEEE Trans. Med. Imaging **40**(11), 3154–3164 (2021)