Novel deep learning-based CT synthesis algorithm for MRI-guided PET attenuation correction in brain PET/MR imaging

Hossein Arabi, Guodong Zeng, Guoyan Zheng and Habib Zaidi†, *IEEE Fellow*

Abstract*– MRI-guided synthetic CT (sCT) generation is one of the main challenges hampering quantitative PET/MR imaging as well as MRI-only radiation planning. Deep learning-based approaches have recently gained momentum in a variety of medical imaging applications. In this work, a novel synthetic CT generation algorithm based on deep convolutional neural network is proposed for MRI-guided attenuation correction in PET/MRI. The proposed algorithm (AsCT) exploits adversarial semantic structure learning implemented as a CT segmentation approach to constrain the adversarial synthetic CT generation process. The proposed technique was trained using 50 pairs of CT and MR brain scans under a two-fold[1] cross validation scheme. The AsCT method was compared to an atlas-based method (Bone-Atl), previously developed for MRI-only radiation planning, as well as the commercial segmentation-based approach (2-class) implemented on the Philips TF PET/MRI system. The evaluation was performed using clinical brain studies of 40 patients who have undergone PET/CT and MRI scanning. The accuracy of the CT value estimation and cortical bone identification were assessed for the three different methods taking CT images as reference. Bias of tracer uptake (SUV) was measured on attenuation corrected PET images using the three techniques taking CT-based attenuation corrected PET as reference. Bone-Atl and AsCT exhibited similar cortical bone extraction (using an intensity threshold of 600 Hounsfield Unit (HU)) resulting in Dice coefficient (DSC) of 0.78±0.07 and 0.77±0.07, respectively. Bone-Atl method performed slightly better in terms of accuracy of CT value estimation where a mean absolute error of 123±40 (HU) was obtained for the whole head region while AsCT and 2-class methods led to 141±40 and 230±33 (HU), respectively. Quantitative analysis of brain PET images demonstrated competitive performance of AsCT and Bone-Atl methods where mean relative errors of 1.2±13.8% and 1.0±9.9% were achieved in bony structures, respectively, while the 2-class approach led to a mean SUV error of -14.7±8.9%. The proposed AsCT algorithm showed competitive performance with respect to the atlas-based method and outperformed the segmentation-based (2-class) method with clinically tolerable errors.

I. Introduction

The development of algorithms for synthetic CT (sCT) generation from MRI is crucial for accurate attenuation correction on PET/MRI systems [1] as well as MRI-only radiation planning . The aim of this work is to develop and evaluate a novel deep learning–based algorithm aiming at generating sCT images from MRI. Deep learning-based approaches have recently gained momentum in a variety of medical imaging applications [2, 3], including synthetic CT generation for the purpose of PET attenuation correction [4, 5] and MRI-only radiation planning [6-8]. The aim of this work is to propose a novel deep learning method to generate synthetic CT from single structural MR image for the attenuation correction in brain PET/MR imaging. The proposed algorithm (AsCT) is based on adversarial learning that has been shown to have superior performance in synthetic data generation [9]. This algorithm (in which the semantic structure training is explicitly incorporated) is implemented as a CT segmentation approach to constrain adversarial MRI to CT synthesis. The proposed algorithm was compared to our atlas-based sCT generation approach (Bone-Atl), previously developed for MRI-only radiation planning [10, 11]. The proposed algorithm was trained in the head region to convert MR brain images into sCTs. The comparison between the proposed learning based algorithm, the state-of-the-art atlas-based and the segmentation-based methods was evaluated to provide valuable insight into the accuracy of MRI-guided attenuation correction in PET/MR brain imaging.

II. Material and Methods

A. Proposed deep learning-based approach (AsCT)

The proposed AsCT method consists of two parts: (1) segment the target MRI into 4 tissue classes (AsCT-seg), namely, air, soft-tissue, bone and background. Then, the segmented image is used to regularize the main MRI-to-CT synthesis process (AsCT-syn), achieved by back-propagating gradients from the segmentation part. The AsCT-seg consists of two compartments, referred to as segmentation generator network (G-seg) and segmentation discriminator network (D-seg). (2) AsCT-syn generates sCT from the input MR images. The AsCT-syn, similarly, has two compartments called generator network (G-syn) and synthesis discriminator network (D-syn). The G-syn is the main block that generates sCT images from the input MRI and the outcome is fed into the D-syn and D-seg. These two blocks will then verify whether the generated sCT is real or false in terms of CT intensity and tissue segmentation, respectively. Prior to the discrimination analysis by the D-seg block, the generated sCT (output of the G-syn block) is fed into the G-seg for tissue segmentation.

AsCT-Seg is trained together with AsCT-syn in an end-to-end fashion. In AsCT-syn, there is a loss function (L-syn) which combines an adversarial term from D-syn and a voxel-wise L2 reconstruction loss comparing the intensities in sCT and the ground truth CT. The second term in the loss function is designated to encourage accurate sCT intensity estimation.

H. Arabi is with the Division of Nuclear Medicine & Molecular Imaging, Geneva University Hospital, CH-1211, Geneva, Switzerland (e-mail: hossein.arabi@unige.ch).

G. Zeng and G. Zheng are with the Institute for Surgical Technology and Biomechanics, University of Bern, CH-3014, Bern, Switzerland.

†H. Zaidi is with the Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland, Geneva, Neuroscience Center, Geneva University, CH-1205 Geneva, Switzerland, and Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands (e-mail: habib.zaidi@hcuge.ch).

978-1-5386-8494-8/18/$31.00 ©2018 IEEE

Moreover, a volume-wise loss function from D-syn is designed to enforce a higher order consistency in the appearance space.

On the other hand, AsCT-seg consists of a loss function with an adversarial term from D-seg and a voxel-wise multi-class cross-entropy loss between the estimated segmentation and the reference segmentation.

D-syn and D-seg share the same convolutional neural network containing three convolutional and three fully connected layers where the output dimensionality of the fully connected layers is 512, 128 and 1, respectively. The convolutional and deconvolutional layers in G-syn and G-seg networks are followed by a batch normalization layer and a ReLU layer, apart from the last one before the final output layer. All the filters employed in the convolutional and deconvolutional layers in the G-syn and G-seg blocks are 3×3×3 in size.

For the training of the proposed method, a training database was constructed using 50 pairs of T1-weighted MRI and CT brain images. Prior to training, the T1-weighted MR images were non-rigidly aligned to the corresponding CT images. The quality of image alignment was visually verified to avoid any gross mis-registration errors.

B. Comparison to atlas-guided approach

The proposed learning method was evaluated against a state-of-the-art atlas-guided approach called Bone-Atl [10]. The Bone-Atl framework consists of two consecutive steps. In the first step, a bone segmentation from the target MRI is optimized based on morphological similarity between the target MRI and atlas MRIs using a local weighting scheme. Then, given the segmented bone from the target MRI, fusion of the atlas CTs, which are aligned to the target MR image, is achieved through similarity of bony structures in the atlas CTs to the segmented bone from target MRI as well as morphological similarity between target MRI and atlas MRIs. To obtain the bony structures from the atlas CT images, a simple intensity threshold of 140 HU was applied. The main objective in defining image morphology similarity between target MR and the atlas MR images is to assign higher weights to similar atlases, where the phase congruency map was chosen as it can detect structural characteristics of an image in a way that is invariant to image intensity and robust to noise.

Moreover, the commercial segmentation-based approach (2-class) implemented on the Philips TF PET/MRI system was evaluated in this study. To generate this attenuation map (2-class), the head contour was segmented from the target MR image followed by assigning Hounsfield Unit of water (HU=0) to all voxels within the head contour and HU=-1000 to the background air [12].

C. Clinical PET/CT-MRI studies

PET/CT and MRI brain scans of 40 patients were retrospectively employed for the quantitative evaluation of the proposed MRI-based sCT generation methods. The patients underwent a 3D T1-weighted magnetization prepared rapid gradient-echo, MP-RAGE and a T2-weighted turbo spin-echo. Thereafter, the patients underwent an ^{18}F-FDG PET/CT scan 20 minutes after injection of 210.2 ± 13.9 MBq ^{18}F-FDG. The CT scan was used for CT-based PET attenuation correction (CTAC), considered as reference to assess the proposed algorithm. For the MRI-based attenuation map generation, only T1-weighted images were used.

D. Performance evaluation metrics

The performance of the proposed AsCT algorithm was compared with Bone-Atl and the commercial segmentation based (2-class) methods. The resulting attenuation maps were analyzed based on the accuracy of the extracted cortical bone in the head region using the Dice (DSC) and relative volume difference (RVD) metrics. The accuracy of the CT value estimation was evaluated through comparison between voxel intensity of reference CTs and the generated sCTs. Cortical bone was segmented from sCT images (A) and reference CT images (R) using an intensity threshold of 600 HUs. The mean error (ME) and mean absolute error (MAE) were calculated between the reference CTs (R_{CT}) and synthetic-CTs (A_{CT}) within bony structures. Moreover, the SUV bias in PET images was also estimated for the three MRI guided attenuation correction techniques.

$$DSC(A,R) = \frac{2|A \cap R|}{|A|+|R|} \qquad RVD(A,R) = 100 \times \frac{|A|-|R|}{|R|}$$

$$ME_{CT} = \frac{1}{P}\sum_{i=1}^{P} A_{CT}(i) - R_{CT}(i) \qquad MAE_{CT} = \frac{1}{P}\sum_{i=1}^{P} |A_{CT}(i) - R_{CT}(i)|$$

III. Results and Discussion

Representative slices of the different sCT images are presented in Fig. 1 together with the target MR and reference CT images. The cortical bone extracted from reference CT and sCT images are also illustrated below each image.

The results of quantitative analysis in terms of cortical bone identification are summarized in Table 1. It can be observed that Bone-Atl exhibited slightly better bone extraction as well as more accurate CT value estimation within the bone structures.

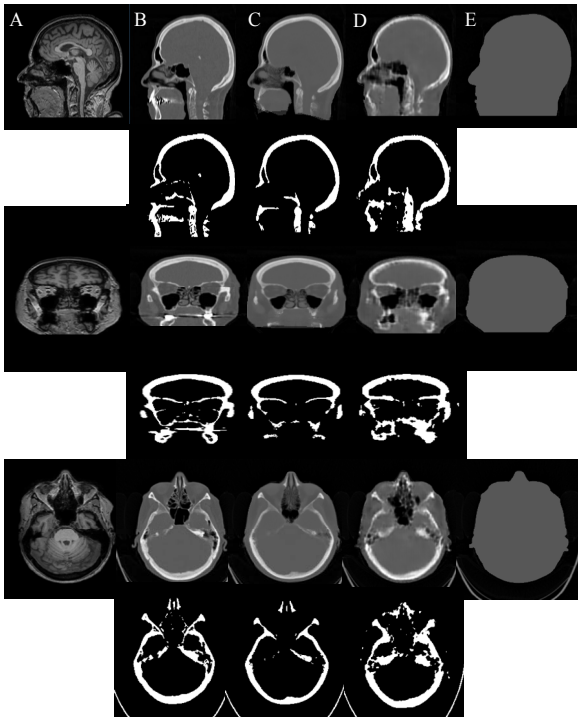

Fig 1. A) Target MRI. B) Reference CT. C) Bone-Atl. D) AsCT. E) 2-class. Cortical bone map are presented for reference CT, Bone-Atl and AsCT image.

Table 1. Accuracy of bone extraction (using intensity threshold of HU=600) and CT value estimation using Bone-Atl and AsCT methods.

Cortical bone	**DSC**	**RVD (%)**	**ME (HU)**	**MAE (HU)**
Bone_Atl	0.78±0.07	41.4±??	5±110	241±100
AsCT	0.77±0.07	46.3±16.3	-107±167	312±101
2-class	–	–	-1025±100	1025±100

For the whole head region, Bone-Atl resulted in ME±SD (MAE±SD) of -8±20 (123±40) while AsCT and the segmentation technique resulted in -14±51 (141±40) and -175±34 (230±33), respectively. The accuracy of CT value estimation for the air cavities and soft-tissue using different methods are presented in table 2.

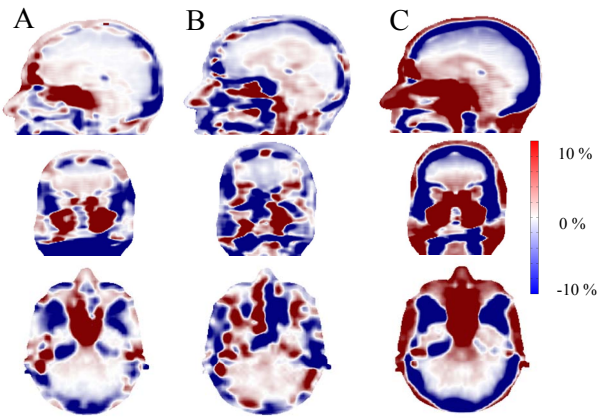

Fig 2. Error maps (relative bias) between PET-CTAC and PET images attenuation corrected using (A) Bone-Atl (B) AsCT, (C) and 2-class attenuation maps.

Table 2. Accuracy of CT value estimation in the air cavity and soft-tissue regions using the different MRI-based sCT generation methods.

Air cavity	**ME(HU)**	**MAE(HU)**	*Soft-tissue*	**ME(HU)**	**MAE (HU)**
Bone_Atl	317±294	459±240	**Bone_Atl**	1±5	8±4
AsCT	295±282	407±228	**AsCT**	2±6	10±5
2-class	805±47	805±47	**2-class**	-2±1	5±2

The Bone-Atl and AsCT methods performed similarly in terms of CT value estimation. Likewise, the quantitative PET analysis exhibited competitive performance of the Bone-Atl and AsCT methods. Fig. 2 shows representative slices of the PET bias map obtained from voxel-wise relative SUV error calculation between the reference PET-CTAC and the PETs attenuation corrected using Bone-Atl, AsCT and 2-class attenuation maps.

Table 3 summarizes the relative error between SUVs estimated on PET attenuation corrected using the different MRI-based sCTs and the reference CT images for the bone, soft-tissue and air regions.

Table 3. Relative SUV error (mean and absolute mean (%)) within different tissue classes for Bone-Atl, AsCT and 2-class methods.

	Soft-tissue mean ± SD (Abs. mean ± SD)	**Bone** mean ± SD (Abs. mean ± SD)	**Air cavity** mean ± SD (Abs. mean ± SD)
Bone-Atl	1.5±9.3 (3.5±8.7)	1.1±9.9 (6.0±8.3)	6.3±13.7 (7.1±12.6)
AsCT	3.2±13.6 (5.0±13.1)	1.2±13.8 (6.7±12.1)	3.2±13.6 (5.5±13.1)
2-class	-1.6±10.2 (5.4±8.8)	-14.7±8.9 (15.5±7.3)	40.8±10.6 (42.6±8.8)

The quantitative PET analysis in table 3 demonstrated slightly lower PET bias achieved by Bone-Atl method in comparison to AsCT. However, both these methods outperformed the commercial segmentation-based technique, particularly in bony structures.

IV. Conclusions

The proposed AsCT method based on the deep convolutional neural network, exhibited competitive performance to the state-of-the-art Bone-Atl method with clinically tolerable errors in MRI-guided PET attenuation correction.

Acknowledgments

This work was supported by the Swiss National Science Foundation under grant 320030_176052 and the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016.

References

[1] A. Mehranian, H. Arabi, and H. Zaidi, "Vision 20/20: Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities," *Medical physics,* vol. 43, pp. 1130-1155, 2016.

[2] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, *et al.*, "A survey on deep learning in medical image analysis.," *Med Image Anal,* vol. 42, pp. 60-88, 2017.

[3] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique.," *IEEE Transactions on Medical Imaging,* vol. 35, pp. 1153-1159, 2016.

[4] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, and A. B. McMillan, "Deep learning MR imaging–based attenuation correction for PET/MR imaging," *Radiology,* vol. 286, pp. 676-684, 2017.

[5] T. J. Bradshaw, G. Zhao, H. Jang, F. Liu, and A. B. McMillan, "Feasibility of Deep Learning–Based PET/MR Attenuation Correction in the Pelvis Using Only Diagnostic MR Images," *Tomography,* vol. 4, p. 138, 2018.

[6] H. Arabi, J. A. Dowling, N. Burgos, X. Han, P. B. Greer, N. Koutsouvelis, *et al.*, "Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI-guided radiation planning in the pelvic region," *Medical physics,* vol. 45, pp. 5218 - 5233, 2018.

[7] A. M. Dinkla, J. M. Wolterink, M. Maspero, M. H. Savenije, J. J. Verhoeff, E. Seravalli, *et al.*, "MR-only brain radiotherapy: Dosimetric evaluation of synthetic CTs generated by a dilated convolutional neural network," *International Journal of Radiation Oncology• Biology• Physics,* 2018.

[8] X. Han, "MR-based synthetic CT generation using a deep convolutional neural network method.," *Med Phys,* vol. 44, pp. 1408-1419, 2017.

[9] H. Emami, M. Dong, S. P. Nejad-Davarani, and C. Glide-Hurst, "Generating Synthetic CT s from Magnetic Resonance Images using Generative Adversarial Networks," *Medical physics,* 2018.

[10] H. Arabi, N. Koutsouvelis, M. Rouzaud, R. Miralbell, and H. Zaidi, "Atlas-guided generation of pseudo-CT images for MRI-only and hybrid PET–MRI-guided radiotherapy treatment planning," *Physics in medicine and biology,* vol. 61, p. 6531, 2016.

[11] A. Mehranian, H. Arabi, and H. Zaidi, "Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI," *NeuroImage,* vol. 130, pp. 123-133, 2016.

[12] H. Zaidi, N. Ojha, M. Morich, J. Griesmer, Z. Hu, P. Maniawski, *et al.*, "Design and performance evaluation of a whole-body Ingenuity TF PET-MRI system.," *Phys Med Biol,* vol. 56, pp. 3091-3106, 2011.