Brain MR Imaging Segmentation Using Convolutional Auto Encoder Network for PET Attenuation Correction

Imene Mecheter[1(✉)], Abbes Amira[2], Maysam Abbod[1], and Habib Zaidi[3]

[1] Department of Electronics and Computer Engineering, Brunel University London, Uxbridge, UK
imene.mecheter@brunel.ac.uk

[2] Institute of Artificial Intelligence, De Montfort University, Leicester, UK

[3] Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland

Abstract. Magnetic resonance (MR) image segmentation is one of the most robust MR based attenuation correction methods which have been adopted in clinical routine for positron emission tomography (PET) quantification. However, the segmentation of the brain into different tissue classes is a challenging process due to the similarity between bone and air signal intensity values. The aim of this work is to study the feasibility of deep learning to improve the brain segmentation with the application of data augmentation. A deep convolutional auto encoder network is applied to segment the brain into three tissue classes: air, soft tissue, and bone. The dice similarity coefficients of air, soft tissue, and bone tissues are 0.96 ± 0.01, 0.86 ± 0.02, and 0.63 ± 0.06 respectively. Despite the small datasets used in this work, the results are promising and show the feasibility of deep learning with data augmentation to perform accurate segmentation.

Keywords: Magnetic resonance imaging · Segmentation · PET attenuation correction · Deep learning

1 Introduction

Positron emission tomography (PET) is a well-known imaging modality that provides direct imaging of molecular information. However, the process of PET acquisition leads to an inhomogeneous bias which affects the image resolution. Thus, this attenuation should be properly corrected before PET reconstruction. The attenuation maps are routinely obtained from computed tomography (CT) images since there is a direct transformation between CT intensity and attenuation coefficients [1].

On the other hand, magnetic resonance (MR) imaging is considered nowadays the leading imaging modality for structural brain analysis thanks to its excellent soft tissue contrast, high spatial resolution, anatomical and functional

K. Arai et al. (Eds.): IntelliSys 2020, AISC 1252, pp. 430–440, 2021.
https://doi.org/10.1007/978-3-030-55190-2_32

information, and lack of ionizing radiation. Additionally, MR images have been extensively used for diagnosis, treatment planning, and follow-up of a variety of neurological conditions such as brain tumor [2], Alzheimer disease [3], multiple sclerosis [4], brain stroke [5], and other neurodegenerative diseases.

Hybrid PET/MR imaging is an emerging modality which has been recently commercialized and adopted in clinical domain. This type of scanners provides quantitative anatomical and functional information [6]. However, the main challenge is that there is no direct correlation between MR signal intensity and attenuation coefficients which opens the doors for researchers to explore new research questions and propose new complicated methods to correct PET attenuation using MR imaging without the need to CT images.

Different methods have been applied to address the PET attenuation correction problem using MR imaging which can be categorized into: segmentation based, atlas based, and emission based methods [7]. Although atlas based method outperforms other methods, it is still not considered robust enough to be adopted in clinical domain [8]. Segmentation based is the most robust and simple method which has been applied in commercial scanners [7].

Deep learning networks have been recently applied on different computer vision applications after the successful application of convolutional neural network on ImageNet dataset [9]. Deep learning has been also employed in various medical applications using different imaging modalities [10]. Different deep network architectures have been proposed for medical images segmentation such as fully convolutional network [11], Segnet [12], and U-Net [13].

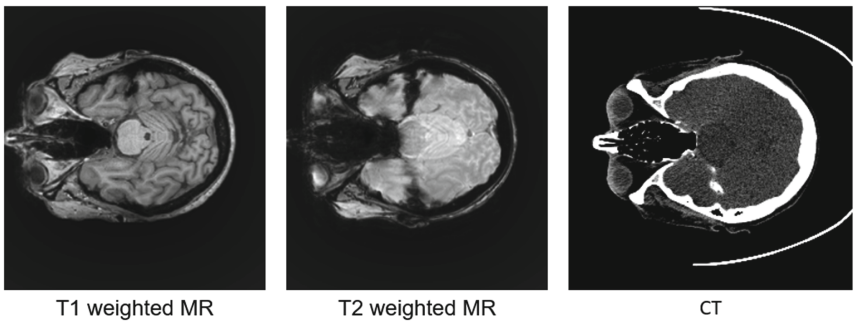

Fig. 1. Example of training datasets: T1-w and T2-w MR images with their corresponding CT image.

Deep networks are performing greatly with the assumption that the training and testing datasets follow the same data distribution. However, this is very challenging with medical datasets due to the variation of commercial scanners and the availability of different imaging modalities and protocols. This variation is called domain shift [14]. Figure 1 shows the visual comparison between T1 weighted (T1-w) MR protocol, T2 weighted (T2-w) MR protocol, and CT image where T1-w MR image looks darker than T2-w in most tissue classes.

In this work, 2D deep learning network is applied to perform brain MR image segmentation into three tissue classes using deep convolutional auto encoder architecture. The applied deep network follows the SegNet architecture. Two data augmentation techniques which are geometric augmentation and intensity augmentation are proposed to increase the size of training datasets and overcome the domain shift limitation. The combination of two different MR protocols: T1-w and T2-w is considered as a type of intensity augmentation. Despite the small available datasets, the obtained results are promising compared to other MR based attenuation correction methods [15–17].

The structure of the paper is as follows: Sect. 2 briefly presents the related work. The proposed methodology is illustrated in Sect. 3. Section 4 represents the obtained results and evaluation. The discussion is presented in Sect. 5 and finally the conclusion and future work is covered in Sect. 6.

2 Related Work

Deep learning has been proposed for several MR based attenuation correction studies [18–20]. However, few works have applied deep learning on brain MR images segmentation for PET attenuation correction.

Liu et al. [15] applied the deep convolutional encoder decoder network using T1-w MR images. This work required a co-registration between CT and MR images before the training process and the generation of ground truth. They used a dataset with 40 patients and achieved dice coefficients values of 0.97, 0.936, and 0.803 for air, soft tissue and, bone respectively.

Jang et al. [17] used UTE and out of phase MR images which were acquired using dual echo ramped hybrid encoding to segment the brain into three classes: air, soft tissue, and bone. UTE sequences used as an input to train a pretrained deep network with T1-w MR images. Transfer learning was applied to adopt the knowledge learnt from other MR to UTE sequences. The obtained segmented MR images were processed using conditional random field technique to refine the segmentation results. The dataset of the pre-trained model consists of 30 patients in addition to 14 new patients which used to train the new model. The achieved dice coefficient values are 0.76, 0.96, and 0.88 for air, soft tissue, and bone respectively.

Arabi et al. [16] proposed a deep learning generative adversarial network with two components: synthesis network and segmentation network to generate pseudo CT images. The synthesis part generated pseudo CT images from T1-w MR images and the segmentation network segmented the obtained pseudo CT images into four tissue classes which are bone, air, soft tissue, and background. This study used a dataset consists of 50 patients and recorded a cortical bone dice coefficient value of 0.77.

3 Proposed Methodology

3.1 Data Acquisition

Brain MR and PET/CT images were acquired as part of the clinical workup of patients. The dataset consists of 15 patients of T1-w MR images and 14 patients of T2-w MR images. The age range of the patients is 64.6 ± 11.7 years. The patients data show clinical indication of dementia (70%), epilepsy (25%) and brain tumors (5%). Firstly, the patients underwent an MRI scan on a 3T Siemens MAGNETOM Skyra scanner with a 64 channel head coil. The MR image scans used for this study are 3D T1- w magnetization prepared rapid gradient-echo, MP-RAGE (TE/TR/TI, 2.3 ms/1900 ms/ 970 ms, flip angle 8; NEX = 1, voxel size $0.8\times 0.8\times 0.8$ mm^3) and 3D T2-w turbo spin-echo, TSE (TE/TR, 100 ms/6200 ms, NEX = 2; voxel size $0.4\times 0.4\times 4$ mm^3). Afterwards, the patients underwent an 18F-FDG PET/CT scan on the Siemens Biograph mCT scanner for 20 min after injection of 210.2 ± 13.9 MBq 18F-FDG.

3.2 Data Preprocessing

Each MR volume of both T1-w and T2-w images is converted into 2D slices. 25 volumes have 149 slices and only 4 volumes have 112 slices. The dimension of each slice is 512×512. The background is removed by cropping each slice into an image with dimensions of 256×256 image. Then, each image is normalized using local contrast normalization technique. The same processing steps are applied to CT volumes. The total number of 2D images for the whole dataset is 2982 slices. Among them, only 2328 slices are used after discarding some of the first and last slices of each volume as they have small pixels that correspond to the brain tissue and the majority of the pixels correspond to the background.

3.3 Ground Truth Generation

CT images are used as ground truth for the supervised training process. MR images are firstly co-registered with CT images to determine a common coordinate system that enables the pixel-based comparison of images. Each MR slice is co-registered with its corresponding CT slice by applying the rigid Euler transformation followed by the non-rigid B-spline transformation using Elastix tool [21]. Afterwards, simple pixel intensity-based thresholding is applied to create the CT ground truth. The CT image is segmented into three tissue classes which are air, bone, and soft tissue. Hounsfield values which are: greater than 600 HU are classified as bone, lower than −500 HU are classified as air, otherswise are labeled as soft tissue.

3.4 Deep Network Model

Segmentation Network Architecture. The deep segmentation network follows the convolutional auto encoder (CAE) architecture which connects an

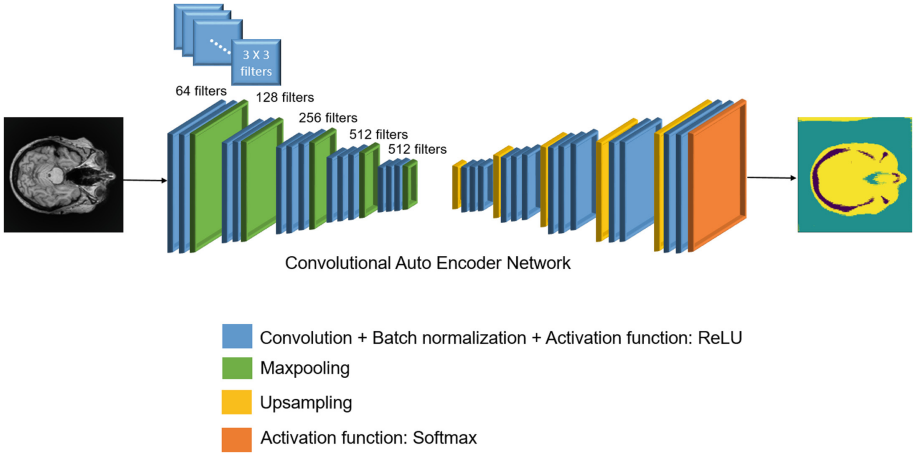

Fig. 2. An illustration of the SegNet architecture

encoder with a decoder. The encoder extracts high level features from raw images and the decoder deconvolves the extracted features and reproduces the original shape of the input data. This type of models preserves the same representation of the data but with more meaningful features. SegNet is one common deep CAE model that has been applied successfully to segment the brain MR images [15,17]. SegNet's encoder consists of 13 convolutional layers with its corresponding batch normalization layers, maxpooling layers, and rectified-linear unit (ReLU) activation function. Each convolutional layer consists of multiple 3×3 filters that extract the high level brain features throughout the network layers. The decoder is the mirror of the encoder with upsampling layers instead of down sampling. The final layer of the decoder is a multiclass softmax classifier that produces class probabilities for each pixel. An illustration of the proposed network is shown in Fig. 2.

Network Training. The deep network is trained from scratch using 12 T1-w MR patients and 11 T2-w MR patients. In the testing phase, 3 patients from T1-w dataset and 3 patients from T2-w dataset are selected for testing and prediction. A grid search strategy is followed to find the best hyperparameters. The network weights are initialized using He normal scheme [22] and updated using Adam optimizer with a fixed learning rate 0.001 where beta 1 and beta 2 are set to 0.9 and 0.999 respectively. At each step, the network reads a batch of 10 samples and calculates the multiclass cross-entropy loss for 50 epochs. In order to create randomization, the training data are shuffled after each epoch.

3.5 Computing Environment

The proposed framework is implemented using both MATLAB and Python programming languages. MATLAB is used for data preprocessing while Keras and TensorFlow libraries are used for deep network implementation. The network is trained using Tesla V100 GPU with 16 GB RAM which is part of a GPU cluster.

4 Experimental Results and Evaluation

The network is firstly trained and tested with T1-w MR images using 2 fold cross validation to ensure the model performance and select the optimal number of training epochs in order to avoid the overfitting. Afterwards, the best hyperparameters are used to train the network without the validation dataset.

The second experiment is performed by training and testing the network with T2-w MR images with the application of 2 fold cross validation.

The third experiment is conducted by combining both T1-w and T2-w MR images to train and test the network. The difference between the two MR protocols is the contrast and brightness of different tissue classes. This intensity values variation is considered as a type of data augmentation which in terms increases the size of training datasets while applying domain adaptation. The resultant trained model is able to segment both T1-w and T2-w MR images.

The segmentation results of both T1-w and T2-w images are illustrated in Figs. 3 and 4 with their corresponding ground truth. The evaluation metrics are recorded in Table 1. Data augmentation is applied to both MR protocols datasets to increase the size of the training datasets and generalize the network. The applied geometric augmentation techniques are rotation by 90°, translation by 20°, and reflection. Table 2 shows the mean dice similarity coefficient while applying data augmentation for training datasets. Data augmentation is not applied to testing datasets. Tables 3, 4, and 5 illustrate the segmentation evaluation metrics for the air, soft tissue, and bone classes after the application of data augmentation.

Table 1. Dice similarity coefficient of the conducted experiments with different MR imaging protocols

MR imaging protocol	Bone	Soft tissue	Air
T1-w	0.49 ± 0.05	0.82 ± 0.02	0.95 ± 0.01
T2-w	0.54 ± 0.01	0.83 ± 0.02	0.95 ± 0.01
T1-w + T2-w	0.63 ± 0.05	0.86 ± 0.01	0.96 ± 0.01

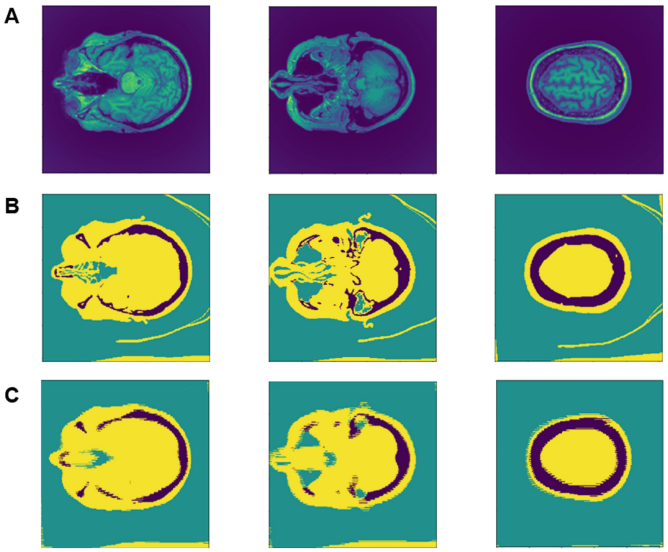

Fig. 3. Example of three MR T1-w slices (A) with their corresponding CT slices as ground truth (B) and the automatic segmented slices using deep learning(C). The colors in row B and C indicate that the green refers to the air tissue, the yellow represents the soft tissue, and the purple shows the bone tissue.

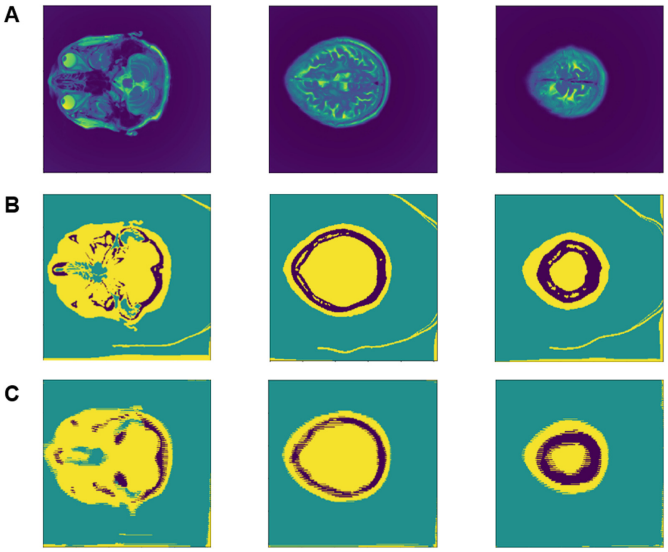

Fig. 4. Example of three MR T2-w slices (A) with their corresponding CT slices as ground truth (B) and the automatic segmented slices using deep learning(C). The colors in row B and C indicate that the green refers to the air tissue, the yellow represents the soft tissue, and the purple shows the bone tissue.

Table 2. Dice similarity coefficient of the conducted experiments with geometric data augmentation

MR imaging protocol	Bone	Soft tissue	Air
T1-w	0.56 ± 0.04	0.83 ± 0.02	0.95 ± 0.01
T2-w	0.58 ± 0.02	0.85 ± 0.02	0.95 ± 0.00
T1-w + T2-w	0.63 ± 0.06	0.86 ± 0.02	0.96 ± 0.01

Table 3. Evaluation metrics of air tissue class in the testing dataset

Air	Precision	Recall	Dice/F1-score	Jaccard
Patient T1-w 01	0.95	0.97	0.96	0.93
Patient T1-w 02	0.95	0.98	0.97	0.93
Patient T1-w 03	0.92	0.98	0.95	0.91
Patient T2-w 04	0.96	0.98	0.97	0.94
Patient T2-w 05	0.95	0.98	0.97	0.93
Patient T2-w 06	0.94	0.98	0.96	0.92

Table 4. Evaluation metrics of soft tissue class in the testing dataset

Soft tissue	Precision	Recall	Dice/F1-score	Jaccard
Patient T1-w 01	0.88	0.86	0.87	0.77
Patient T1-w 02	0.87	0.82	0.85	0.73
Patient T1-w 03	0.89	0.82	0.85	0.74
Patient T2-w 04	0.91	0.87	0.89	0.80
Patient T2-w 05	0.88	0.84	0.86	0.76
Patient T2-w 06	0.90	0.83	0.87	0.77

Table 5. Evaluation metrics of bone tissue class in the testing dataset

Bone	Precision	Recall	Dice/F1-score	Jaccard
Patient T1-w 01	0.74	0.65	0.69	0.53
Patient T1-w 02	0.66	0.60	0.63	0.46
Patient T1-w 03	0.66	0.55	0.60	0.43
Patient T2-w 04	0.60	0.61	0.60	0.43
Patient T2-w 05	0.69	0.60	0.64	0.47
Patient T2-w 06	0.61	0.64	0.62	0.45

5 Discussion

The evaluation metrics values show promising segmentation results by comparing the baseline results (T1-w model) with other models which are trained with T2-w images only and combined T1-w and T2-w images. Although the same model and the same training parameters are applied to train both T1-w and T2-w images, the intensity values of T2-w images generate a more accurate model. Additionally, the combination of T1-w and T2-w images for training and testing the model increases the segmentation accuracy as well as it increases the size of the training datasets.

Data augmentation is a technique which can be applied to overcome the limitations of small datasets and domain shift issues. The obtained results with data augmentation show an improvement of segmentation accuracy. A clear point to notice is that the improvement of segmentation accuracy includes both bone and soft tissue classes; however, the air class segmentation is not improved. Moreover, the application of data augmentation with combined T1-w and T2-w images does not improve the segmentation performance. One reason of that is the need to generate more variations of data since this training dataset contains a mixture of two MR protocols with different intensity values. Advanced types of data augmentation should be applied to improve the segmentation.

By comparing the evaluation metrics between the different tissue classes, it is clear that the air tissue class achieves higher segmentation accuracy followed by the soft tissue and finally the bone tissue. The challenging segmentation of bone tissue originates from the imbalance tissue classes as such the total number of air and soft tissue pixels per each slice is much higher than the bone pixels.

Compared to other T1-w MR based attenuation correction literatures, this work uses the smallest training and testing datasets. Our dataset consists of 12 patients for training and 3 patients for testing while [16] used 40 patients for training and 10 patients for testing. However, the results achieved are very promising as they can be improved by using more datasets. Moreover, the used datasets include different cases of anatomical abnormalities which ensures the robustness of the model with different diseases.

The training time varies from each experiment since the total number of training datasets is different. The approximate required training time of T1-w and T2-w with data augmentation is only one hour and half. The labeling of a single slice takes only 6 ms.

6 Conclusion and Future Work

In this study, an MR based attenuation correction method using deep learning has been applied to segment the brain images into three tissue classes: air, soft tissue and bone. The proposed method shows that training with different MR protocols and data augmentation techniques improves the segmentation accuracy and hence improves the PET attenuation correction and quantification. In the future work, the network will be trained and tested with more data to ensure the robustness of the model.

Acknowledgment. The authors extend a special thanks to Texas A&M at Qatar's Advanced Scientific Computing (TASC) center who provided the access and technical support to High Performance Computing cluster "Raad2".

References

1. Fei, B., Yang, X., Nye, J.A., Aarsvold, J.N., Raghunath, N., Cervo, M., Stark, R., Meltzer, C.C., Votaw, J.R.: MR/PET quantification tools: registration, segmentation, classification, and MR-based attenuation correction. Med. Phys. **39**(10), 6443–6454. https://doi.org/10.1118/1.4754796
2. Rundo, L., Militello, C., Tangherloni, A., Russo, G., Vitabile, S., Gilardi, M.C., Mauri, G.: NeXt for neuro-radiosurgery: a fully automatic approach for necrosis extraction in brain tumor MRI using an unsupervised machine learning technique. Int. J. Imaging Syst. Technol. **28**(1), 21–37 (2018). https://doi.org/10.1002/ima.22253
3. Gunawardena, K., Rajapakse, R., Kodikara, N.: Applying convolutional neural networks for pre-detection of Alzheimer's disease from structural MRI data. In: 2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), pp. 1–7. IEEE (2017)
4. Valverde, S., Cabezas, M., Roura, E., González-Villà, S., Pareto, D., Vilanova, J.C., Ramio-Torrenta, L., Rovira, À., Oliver, A., Lladó, X.: Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. NeuroImage **155**, 159–168 (2017)
5. Praveen, G., Agrawal, A., Sundaram, P., Sardesai, S.: Ischemic stroke lesion segmentation using stacked sparse autoencoder. Comput. Biol. Med. (2018)
6. Chen, Z., Jamadar, S.D., Li, S., Sforazzini, F., Baran, J., Ferris, N., Shah, N.J., Egan, G.F.: From simultaneous to synergistic mr-pet brain imaging: a review of hybrid MR-PET imaging methodologies. Hum. Brain Mapping **39**(12), 5126–5144 (2018)
7. Mehranian, A., Arabi, H., Zaidi, H.: Vision 20/20: magnetic resonance imaging-guided attenuation correction in PET/MRI: challenges, solutions, and opportunities. Med. Phys. **43**(3), 1130–1155 (2016). https://doi.org/10.1118/1.4941014
8. Shi, K., Fürst, S., Sun, L., Lukas, M., Navab, N., Förster, S., Ziegler, S.I.: Individual refinement of attenuation correction maps for hybrid PET/MR based on multi-resolution regional learning. Comput. Med. Imaging Graph. **60**, 50–57 (2017)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
10. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
11. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
12. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2017)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241. Springer, Heidelberg (2015)

14. Tommasi, T., Lanzi, M., Russo, P., Caputo, B.: Learning the roots of visual domain shift. In: European Conference on Computer Vision, pp. 475–482. Springer, Heidelberg (2016)
15. Liu, F., Jang, H., Kijowski, R., Bradshaw, T., McMillan, A.B.: Deep learning MR imaging-based attenuation correction for PET/MR imaging. Radiology **286**(2), 676–684 (2017)
16. Arabi, H., Zeng, G., Zheng, G., Zaidi, H.: Novel adversarial semantic structure deep learning for MRI-guided attenuation correction in brain PET/MRI. Eur. J. Nucl. Med. Mol. Imaging 1–14 (2019)
17. Jang, H., Liu, F., Zhao, G., Bradshaw, T., McMillan, A.B.: Technical note: deep learning based MRAC using rapid ultrashort echo time imaging. Med. Phys. **45**(8), 3697–3704 (2018). https://doi.org/10.1002/mp.12964
18. Han, X.: MR-based synthetic ct generation using a deep convolutional neural network method. Med. Phys. **44**(4), 1408–1419 (2017)
19. Leynes, A.P., Yang, J., Wiesinger, F., Kaushik, S.S., Shanbhag, D.D., Seo, Y., Hope, T.A., Larson, P.E.: Zero-Echo-Time and Dixon Deep Pseudo-CT (ZeDD CT): direct generation of pseudo-CT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. J. Nuclear Med. **59**(5), 852–858 (2018). https://doi.org/10.2967/jnumed.117.198051
20. Nie, D., Trullo, R., Lian, J., Petitjean, C., Ruan, S., Wang, Q., Shen, D.: Medical image synthesis with context-aware generative adversarial networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 417–425. Springer, Heidelberg (2017)
21. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging **29**(1), 196–205 (2010)
22. 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, pp. 1026–1034 (2015)