Brain Magnetic Reasonance Imaging Segmentation using Convolutional Auto Encoder Network for PET Attenuation Correction

***Abstract*—MR images segmentation is one of the most robust MR based attenuation correction methods which have been adopted in clinical routine for PET quantification. However, the segmentation of 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 was applied to segment the brain into three tissue classes: air, soft tissue, and bone. Geometric and intensity augmentation were applied to increase the training datasets and overcome the domain shift limitation.**

The dice similarity coefficients of air, soft tissue, and bone tissues are 0. 96 $\pm$ 0.01, 0.86 $\pm$ 0.02, and 0.63 $\pm$ 0.06 respectively. Despite the small datasets used in this work, the results are promising and outperformed other machine learning based segmentation methods. The results achieved showed the feasibility of deep learning with data augmentation to perform accurate and robust MR images segmentation.

I. 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 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 have 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 doors for researchers to explore new research questions and propose new complicated methods to correct PET attenuation using MR images without the need to use CT images.

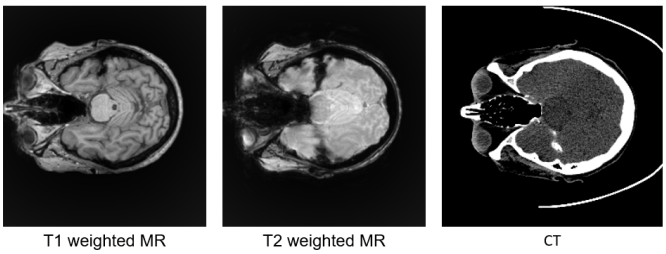

Fig. 1. Example of training datasets: T1-w and T2-w MR images with their corresponding CT image.

Different methods have been applied to address the PET attenuation correction problem using MR imaging which can be categorized into: segmentation based methods, atlas based methods 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 have 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].

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 the two MR protocols where T1 weighted (T1-w) MR image looks darker than T2-weighted (T2-w) in most tissue classes.

In this work, 2D deep learning network was applied to perform brain MR images segmentation into three tissue classes using deep convolutional auto encoder architecture. The applied deep network followed the SegNet architecture. Two data augmentation techniques which are geometric augmentation and intensity augmentation were 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 the intensity augmentation. Despite the small available datasets, the obtained results are promising compared to other MR based attenuation correction methods [15], [16], [17]. The structure of the paper is as follows: section 2 presents briefly the related work. Section 3 illustrates the proposed methodology. Section 4 represents the obtained results and evaluation. The discussion is presented in section 5 and finally the conclusion is covered in section 6.

II. Related Work

Deep learning has been proposed for several MR based attenuation correction [18], [19], [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 mean 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 consisted of 30 patients in addition to 14 new patients which used to train the pretrained model. The achieved dice coefficient is 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 consisted of 50 patients and recorded a cortical bone dice coefficient value of 0.77.

III. Proposed Methodology

A. 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 weighted MR images. The age range of the patients is 64.6 ± 11.7 years. The patients data showed 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 images scans used for this study are 3D T1- weighted magnetization prepared rapid gradient-echo, MP-RAGE (TE/TR/TI, 2.3 ms/1900 ms/ 970 ms, flip angle 8 ; NEX = 1, voxel size $0.8\times0.8\times0.8$ mm3) and 3D T2-weighted turbo spin-echo, TSE (TE/TR, 100 ms/6200 ms,NEX = 2; voxel size $0.4\times0.4\times4$ mm3). Afterward, 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.

B. Data preprocessing

Each MR volume of both T1-weighted and T2-weighted images was converted into 2D slices. 25 volumes have 149 slices and only 4 volumes have 112 slices. The dimension of each slice is 512 X 512. The background was removed by cropping each slice into 256 X 256 image. Then, each image was normalized using local contrast normalization technique. The same processing steps applied to CT volumes. The total number of 2D images for the whole dataset is 2982 slices. Among them, only 2328 were 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 majority of the image is background.

C. Ground truth generation

CT images were used as ground truth for the supervised training process. MR images were firstly co-registered with CT images to determine a common coordinate system that enables the pixel-based comparison of images. Each MR slice was 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]. Afterward, simple pixel intensity-based thresholding was applied to create the CT ground truth. The CT image was segmented into three tissue classes which are air, bone, and soft tissue. Hounsfield values greater than 600 HU were classified as bone, less than -500 HU were classified as air, and others were labeled soft tissue.

D. Deep network model

1) Segmentation Network Architecture: The deep segmentation network followed the convolutional auto encoder (CAE) architecture which connects an encoder with a decoder. The encoder extracts high level features from raw images and the decoder deconvolves the extracted features and reproduces the original format 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 3X3 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 Figure 2.

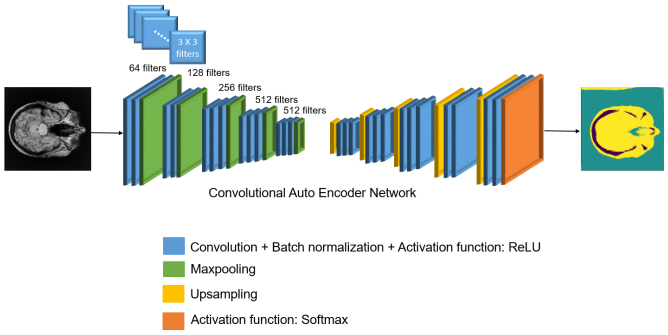

Fig. 2. An illustration of the SegNet architecture

2) Network Training: The deep network was trained from scratch using 12 T1-weighted MR datasets and 11 T2-weighted MR datasets. In the testing phase, 3 patients from T1-weighted datasets and 3 patients from T2-weighted datasets were selected for testing and prediction. A grid search strategy was followed to fine tune the network parameters. The network weights were initialized using he normal scheme and updated using Adam optimizer with a fixed learning rate 0.001 where beta 1 and beta 2 were set to 0.9 and 0.999 respectively. At each step, the network read 10 mini batches and calculated the multiclass cross-entropy loss for 50 epochs. In order to create randomization, the training data were shuffled after each epoch.

E. Computing Environment

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

IV. Experimental Results and Evaluation

The network was firstly trained and tested with T1-weighted 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. Afterward, the best hyper parameters were used to train the network without validation dataset.

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

The third experiment was conducted by combining both T1-weighted and T2-weighted 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 by using different MR protocols which in terms increases the size of training datasets while applying domain adaptation. The resultant trained model will be able to segment both T1-weighted and T2-weighted MR images.

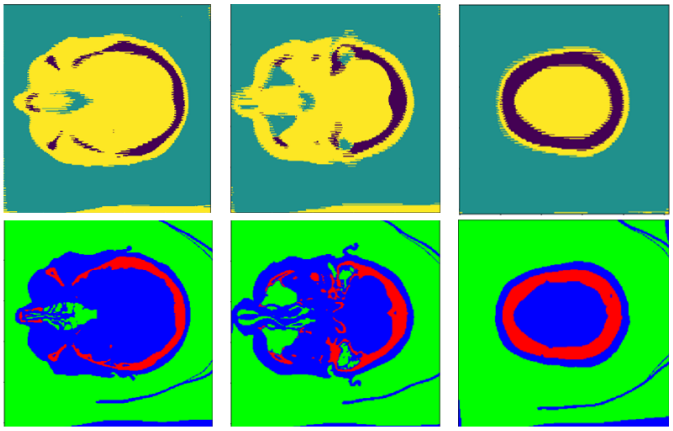

Fig. 3. Example of segmented T1-weighted slices (above) with their corresponding CT slices as ground truth (bottom).

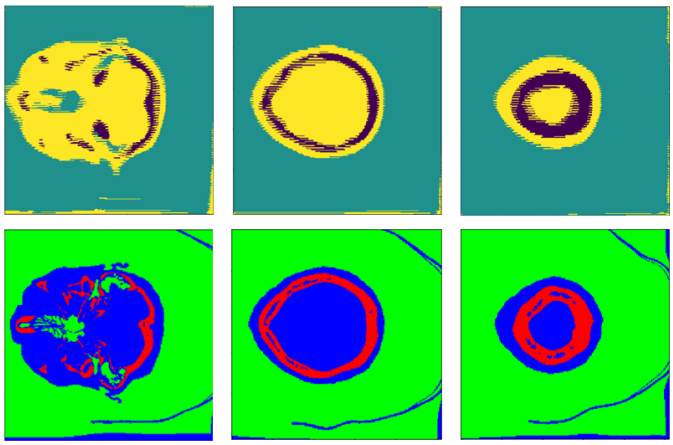

Fig. 4. Example of segmented T2-weighted slices (above) with their corresponding CT slices as ground truth (bottom).

The segmentation results of both T1 and T2 weighted images are illustrated in Figure 3 and 4 with their corresponding ground truth. The evaluation metrics are recorded in Table I.

Additionally, data augmentation was 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 II shows the mean dice similarity coefficient while applying data augmentation for training datasets. Data augmentation was not applied to testing datasets. Tables III,

TABLE I
Dice similarity coefficient of the conducted experiments with different MR protocols

MR images 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

TABLE II
DICE SIMILARITY COEFFICIENT OF THE CONDUCTED EXPERIMENTS WITH GEOMETRIC DATA AUGMENTATION

MR images 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 III
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 IV
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 V
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

IV, V illustrate the segmentation evaluation metrics after the application of data augmentation.

V. DISCUSSION

The results showed promising segmentation results by comparing the baseline results (T1-w model) with other models which trained with T2-w images and combination of T1-w and T2-w images. Although the same model and the same training parameters were applied to train both T1-w and T2-w images, the intensity values of T2-w images generated a more accurate model. Additionally, the combination of T1-w and T2-w images for training and testing the model increased the segmentation accuracy as well as it increased 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 showed 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 was not improved. Moreover, the application of data augmentation with combined T1-w and T2-w images did not improve the segmentation performance. One reason of that is the need to generate more variations of data since this training datasets contain a mixture of two MR protocols with different intensity values. Advanced types of data augmentation should be applied to these datasets to improve the segmentation.

By comparing the evaluation metrics between the different tissue classes, it is clear that the air tissue class achieved 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 are much higher than the bone pixels.

Compared to other T1-w MR based attenuation correction literatures, this work used the smallest training and testing datasets; however, the results achieved are very promising as they can be improved by using more datasets. Moreover, the training datasets should include different cases of anatomical abnormalities in order to study the performance 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 weighted with data augmentation is only one hour and half. The labeling of a single slice took only 0.006 seconds.

VI. CONCLUSION

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 showed that training with different MR protocols and data augmentation techniques improved the segmentation accuracy and hence improved the PET attenuation correction and quantification.

ACKNOWLEDGMENT

REFERENCES

[1] B. Fei, X. Yang, J. A. Nye, J. N. Aarsvold, N. Raghunath, M. Cervo, R. Stark, C. C. Meltzer, and J. R. Votaw, "MR/PET quantification tools: Registration, segmentation, classification, and MR-based attenuation correction," *Medical Physics*, vol. 39, no. 10, pp. 6443–6454.

[2] L. Rundo, C. Militello, A. Tangherloni, G. Russo, S. Vitabile, M. C. Gilardi, and G. Mauri, "NeXt for neuro-radiosurgery: A fully automatic approach for necrosis extraction in brain tumor MRI using an unsupervised machine learning technique," *International Journal of Imaging Systems and Technology*, vol. 28, no. 1, pp. 21–37, Mar. 2018.

[3] K. Gunawardena, R. Rajapakse, and N. Kodikara, "Applying convolutional neural networks for pre-detection of alzheimer's disease from structural mri data," in *Mechatronics and Machine Vision in Practice (M2VIP), 2017 24th International Conference on*. IEEE, 2017, pp. 1–7.

[4] S. Valverde, M. Cabezas, E. Roura, S. González-Villà, D. Pareto, J. C. Vilanova, L. Ramio-Torrenta, À. Rovira, A. Oliver, and X. Lladó, "Improving automated multiple sclerosis lesion segmentation with a cascaded 3d convolutional neural network approach," *NeuroImage*, vol. 155, pp. 159–168, 2017.

[5] G. Praveen, A. Agrawal, P. Sundaram, and S. Sardesai, "Ischemic stroke lesion segmentation using stacked sparse autoencoder," *Computers in biology and medicine*, 2018.

[6] Z. Chen, S. D. Jamadar, S. Li, F. Sforazzini, J. Baran, N. Ferris, N. J. Shah, and G. F. Egan, "From simultaneous to synergistic mr-pet brain imaging: A review of hybrid mr-pet imaging methodologies," *Human brain mapping*, vol. 39, no. 12, pp. 5126–5144, 2018.

[7] 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, no. 3, pp. 1130–1155, Mar. 2016.

[8] K. Shi, S. Fürst, L. Sun, M. Lukas, N. Navab, S. Förster, and S. I. Ziegler, "Individual refinement of attenuation correction maps for hybrid pet/mr based on multi-resolution regional learning," *Computerized Medical Imaging and Graphics*, vol. 60, pp. 50–57, 2017.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[10] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[11] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[12] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[14] T. Tommasi, M. Lanzi, P. Russo, and B. Caputo, "Learning the roots of visual domain shift," in *European Conference on Computer Vision*. Springer, 2016, pp. 475–482.

[15] 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, no. 2, pp. 676–684, 2017.

[16] H. Arabi, G. Zeng, G. Zheng, and H. Zaidi, "Novel adversarial semantic structure deep learning for mri-guided attenuation correction in brain pet/mri," *European journal of nuclear medicine and molecular imaging*, pp. 1–14, 2019.

[17] H. Jang, F. Liu, G. Zhao, T. Bradshaw, and A. B. McMillan, "Technical Note: Deep learning based MRAC using rapid ultrashort echo time imaging," *Medical Physics*, vol. 45, no. 8, pp. 3697–3704, 2018.

[18] X. Han, "Mr-based synthetic ct generation using a deep convolutional neural network method," *Medical physics*, vol. 44, no. 4, pp. 1408–1419, 2017.

[19] A. P. Leynes, J. Yang, F. Wiesinger, S. S. Kaushik, D. D. Shanbhag, Y. Seo, T. A. Hope, and P. E. Larson, "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," *Journal of Nuclear Medicine*, vol. 59, no. 5, pp. 852–858, May 2018.

[20] D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang, and D. Shen, "Medical image synthesis with context-aware generative adversarial networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 417–425.

[21] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim, "Elastix: a toolbox for intensity-based medical image registration," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 196–205, 2010.