Deep Learning based Segmentation for Multi MR Imaging Protocols using Transfer Learning for PET Attenuation Correction

Imene Mecheter
Department of Electronics and Computer Engineering
Brunel University London
Uxbridge, UK
imene.mecheter@brunel.ac.uk

Abbes Amira
Institute of Artificial Intelligence
De Montfort University
Leicester, UK

Maysam Abbod
Department of Electronics and Computer Engineering
Brunel University London
Uxbridge, UK

Habib Zaidi
Division of Nuclear Medicine and Molecular Imaging
Geneva University Hospital
Geneva, Switzerland

***Abstract*—Magnetic resonance (MR) image segmentation is a robust technique used for PET attenuation correction. However, the segmentation of the brain into different tissue classes is a challenging task because of the similarity of pixel intensity values. The objective of this work is to propose a deep learning network to segment T1-weighted MR images of a dataset consists of 50 patients. Additionally, transfer learning is applied to segment another MR image protocol which is T2-weighted. The pretrained network with T1-weighted images is finetuned then tested with a dataset of 14 patients only. The Dice coefficients of air, soft tissue, and bone classes for T1-weighted MR images are 0.98, 0.92 , and 0.79 respectively. The results of transfer learning show the feasibility of finetuning a deep network trained with T1-weighted images to segment T2-weighted images.**

***Index Terms*—Magnetic Resonance Imaging, Segmentation, PET Attenuation Correction, Deep Learning, Transfer Learning**

I. Introduction

Positron emission tomography (PET) and magnetic resonance (MR) are two complemantry imaging modalities that provide physiological and morphological information of different tissues in both normal and pathological cases. Recently, the hybrid PET/MR imaging system has been commercialized and adopted in clinical domain [1]. The main challenge is the PET quantification process which requires the photon attenuating properties which cannot be derived directly from MR image intensitties. The attenuation maps are commonly derived from computed tomography (CT) images because of the direct mapping between CT intensities and attenuation coefficients. Therefore, it is necessary to explore more complicated ways to convert the MR image intensities to 511 KeV attenuation maps [2].

978-1-7281-2547-3/20/$31.00 ©2020 IEEE

Three methods have been performed to address the PET attenuation correction problem using MR imaging which are: segmentation, atlas, and emission methods [2]. The segmentation based method is considered the most robust and simple method which has been deployed in commercial scanners and introduced in clinical practice [2]. Atlas based method shows its superiority, however, it is not considered yet as a robust method to be introduced in the clinical domain [3].

Deep learning networks have been recently applied on various computer vision applications. In the medical domain, different deep network architectures have been proposed for medical images segmentation [4] such as Fully convolutional network [5], Segnet [6], and U-Net [7].

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 [8]. 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.

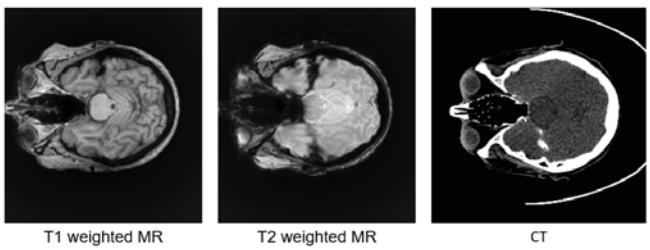

Fig. 1. Example of training datasets: T1-w and T2-w MR images with their corresponding CT image.

In this work, a deep convolutional auto encoder (CAE) network is applied to segment the brain MR image into three tissue classes. The network is trained from scratch using T1-w MR images. Afterwards, transfer learning is applied to finetune the pretrained network to segment T2-w MR images.

The structure of the paper is as follows: section 2 describes the data. Section 3 presents the proposed methodology. Section 4 illustrates the experimental results and evaluation. Finally the conclusion is covered in section 5.

II. Data

A. Data acquisition

Brain MR and PET/CT images are acquired as part of the clinical workup of patients. The dataset consists of 50 patients of T1-w MR images and 14 patients of T2-w MR images. The datasets show clinical indication of dementia (70%), epilepsy (25%) and brain tumors (5%). The age range of the patients is 64.6 $\pm$ 11.7 years.

At the first step, 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- 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 $0.8\times0.8\times0.8$ mm^3) and 3D T2-w turbo spin-echo, TSE (TE/TR, 100 ms/6200 ms, NEX = 2; voxel size $0.4\times0.4\times4$ 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 $\pm$ 13.9 MBq 18F-FDG.

B. Data preprocessing

Each MR volume of both T1-w and T2-w is split into 2D slices. The number of slices per each volume is different and the dimension of each slice is 512×512. Each slice is cropped into a 256×256 image to remove the background. Then, local contrast normalization technique is applied on each image. The same processing steps are applied to CT volumes which are used as ground truth. 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. The discarded slices have just few pixels that correspond to the brain tissue while the majority of the pixels correspond to the background which is not useful for the training process.

C. Ground truth generation

Due to the unavailability of manually segmented MR images, the CT images are used as ground truth for the supervised training process as they can be segmented properly by using a simple thresholding technique. Firstly, the MR images are registered with CT images to determine a common coordinate system that enables the pixel-based comparison of images. The registration process includes the rigid Euler transformation followed by the non-rigid B-spline transformation using Elastix tool [9]. The labeling of CT images is performed by applying a simple pixel intensity-based thresholding to segment the brain into three tissue classes which are air, bone, and soft tissue. Hounsfield values which are greater than 600 HU are labeled as bone, lower than -500 HU are labeled as air, otherwise are labeled as soft tissue.

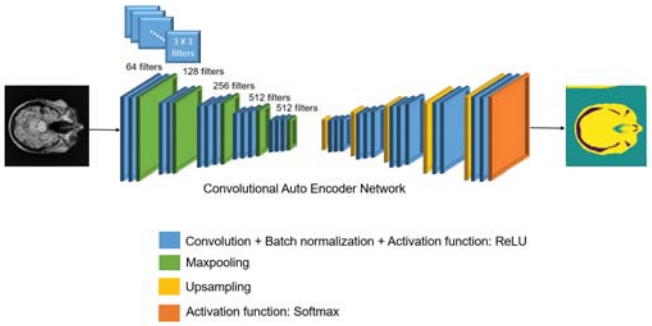

Fig. 2. An illustration of the SegNet architecture

III. Proposed Methodology

The proposed method consists of two parts. The first part is training a deep convolutional auto encoder network from scratch to segment T1-w MR images into three tissue classes. The second part is finetuning the pre-trained network with T1-w images to segment T2-w MR images.

A. Deep CAE Network to segment T1-w MR

The segmentation network architecture is based on the Segnet which follows the convolutional auto encoder architecture. SegNet is one common deep CAE model that has been applied successfully to segment the brain MR images [10], [11]. This model consists of an extracting path to capture the features and an expanding path for localization and resizing. The output of such networks preserves the same size of the input after extracting the meaningful features. The extracting path consists of 13 convolutional layers with their corresponding batch normalization layers and rectified-linear unit (ReLU) activation function in addition to 5 maxpooling layers. Each convolutional layer consists of multiple 3×3 filters. The expanding path is the mirror of the extracting path with upsampling layers to retrain the original size of the input. The final layer of the network is a multiclass softmax classifier that produces class probabilities for each pixel. Figure 2 illustrates the network architecture.

The network is trained from scratch using T1-w MR datasets using two folds cross validation. A grid search strategy is followed to find the best training parameters. The network weights are initialized using He normal scheme [12] and updated using Adam optimizer with a learning rate that starts with 0.001 then reduces when there is no improvement on the training loss for more than 10 epochs. Beta 1 and Beta 2 which are the parameters of Adam's optimizer that controls the decay rates are set to 0.9 and 0.999 respectively. The network reads 10 samples per each batch and calculates the multiclass cross-entropy loss for 300 epochs. Dropout with a factor of 0.7 is applied to reduce the overfitting.

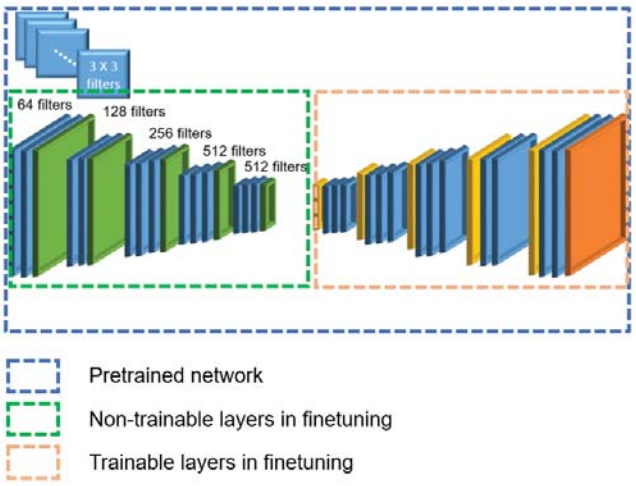

Fig. 3. An illustration of the transfer learning process to segment T2-w MR images

B. Transfer learning to segment T2-w MR

The pretrained model with T1-w MR images is used as the base model to train a small dataset of T2-w MR images. The base model is finetuned by freezing the extracting path while re-training the expanding path. The process of freezing the parameters of specific layers preserves the extracted features from T1-w MR images. The learning rate is set to a smaller value (1e-4) than the learning rate of the pretrained model to avoid destroying the learnt features. The required number of epochs to finetune the model is only 30 epochs. Figure 3 illustrates the transfer learning process.

C. Computing Environment

MATLAB and Python programming languages are used to build the proposed method. The data preprocessing is performed using MATLAB while Python libraries: Keras and TensorFlow are used for deep learning. The deep network model is trained using Tesla V100 GPU with 16 GB RAM as part of "Raad2" GPU cluster.

IV. Experimental Results and Evaluation

The first conducted experiment is the training and testing of the Segnet network to segment T1-w MR images into three different tissue classes: air, soft tissue, and bone. Then, the second experiment is the application of transfer learning by finetuning the pretrained model with T1-w images to fit the T2-w images.

A. Results of the segmentation of T1-w MR images

The segmentation results of T1-w MR images are illustrated in Figure 4. The first row corresponds to three input slices, the second row shows their corresponding CT images as ground truth, and the last row represents the results of the segmentation. Table I shows the segmentation evaluation metrics of the testing dataset from T1-w MR images.

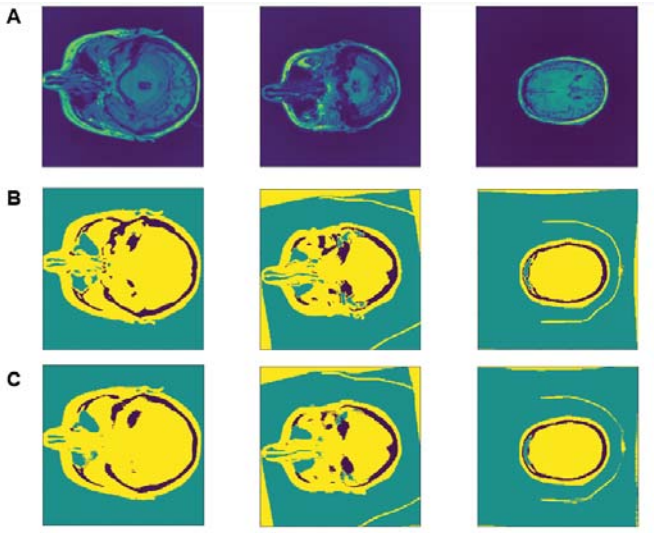

Fig. 4. Example of three MR T1-w slices (A) with their corresponding CT slices as ground truth (B) and the automatic segmented slices using deep learning(C). The colors in row B and C refer to the following classes: green is the air class, yellow represents the soft tissue class, and the purple represents the bone class.

TABLE I
Evaluation metrics of the segmentation results of brain T1-w MR images

Metric	Air	Soft tissue	Bone
Precision	0.98	0.93	0.80
Recall	0.98	0.93	0.77
Dice	0.98	0.92	0.79
Jaccard Index	0.96	0.86	0.65

The air and soft tissue classes achieve higher segmentation dice than the bone class. 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.

B. Results validation

The obtained segmentation results are validated with two other approaches that applied deep learning to segment the brain MR images for PET attenuation correction. Following the network architecture and training options adapted from [10], a 2D segnet network is trained using our datasets. The second approach [13] applies the same segnet network architecture with the addition of conditional random field as a post processing setp. The segmentation results of each approach are illustrated in table II.

C. Results of transfer learning and finetuning

The results of the segmentation of T2-w MR images are shown in Figure 5. Table III illustrates the dice similarity coefficient of the three classes. The first row in the table refers to the results of transfer learning where the pretrained model is used as it is without any finetuning. The results show that there

TABLE II
COMPARISON WITH OTHER DEEP LEARNING BASED SEGMENTATION METHODS

Method	Air	Soft tissue	Bone
Segnet [10]	0.97	0.90	0.66
Segnet+ CRF [13]	0.84	0.57	0.07
Proposed method	**0.98**	**0.92**	**0.79**

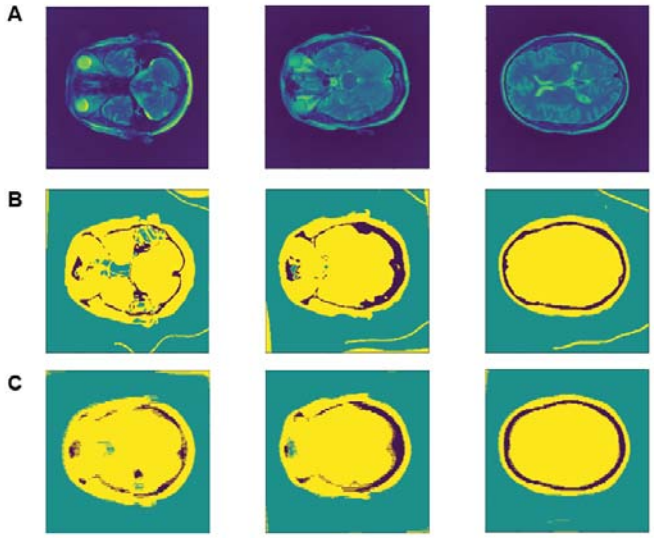

Fig. 5. The segmentation results of T2-w MR images using transfer learning. Three MR T2-w slices are shown in row (A) with their corresponding CT slices as ground truth (B) and the automatic segmented slices using transfer learning(C). The colors in row B and C refer to the following classes: green is the air class, yellow represents the soft tissue class, and the purple represents the bone class.

is an improvement in the segmentation results as we finetune more convolutional layers of the pretrained model. The best results are achieved by freezing the layers of the extracting path (encoder) and retraining the layers of the expanding path (decoder).

Similarily to T1-w MR image segmentation results, the air class achieves the highest dice value followed by the the soft tissue then the bone classes. Despite the finetuning of more convolutional layers helps to improve the segmentation accuracy of both the soft tissue and bone class, the air class segmentation does not show any improvement.

TABLE III
THE DICE SIMILARITY COEFFICIENT OF THE T2-W MR BRAIN TISSUE CLASSES: AIR, SOFT TISSUE, AND BONE USING TRANSFER LEARNING AND FINETUNING THE T1-W PRETRAINED MODEL.

	Air	Soft tissue	Bone
Transfer learning	0.97	0.85	0.51
Finetune: 2 CONV	0.97	0.86	0.50
Finetune: 4 CONV	0.97	0.87	0.58
Finetune: 6 CONV	0.97	0.87	0.60
Finetune: 8 CONV	0.97	0.87	0.60
Finetune: Decoder	**0.97**	**0.87**	**0.63**

The training time varies from training a network from scratch to finetuning a pretrained model. The approximate required time of training T1-w MR images from scratch is 6 hours. However, the finetuning of the pretrained model requires only 10 minutes. Moreover, the predication time is very short such as labeling a single slice takes only 6 ms.

V. CONCLUSION

In this study, an MR image based attenuation correction method using deep learning has been applied to segment the brain MR images into three tissue classes: air, soft tissue and bone. Transfer learning technique is applied to show the feasibility of finetuning a model trained with T1-w MR images to segment T2-w MR images. The promising results indicate that there is a potential to build a common model to segment different MR images protocols. The successful application of deep learning can improve the segmentation accuracy and hence improves the PET attenuation correction and quantification.

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 the High Performance Computing cluster "Raad2".

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep learning for brain mri segmentation: state of the art and future directions," *Journal of digital imaging*, vol. 30, no. 4, pp. 449–459, 2017.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.
[13] H. Jang, F. Liu, G. Zhao, T. Bradshaw, and A. B. McMillan, "Deep learning based mrac using rapid ultrashort echo time imaging," *Medical physics*, vol. 45, no. 8, pp. 3697–3704, 2018.