Brain MR images segmentation using 3D CNN with features recalibration mechanism for segmented CT generation

Imene Mecheter [a,*], Maysam Abbod [a], Habib Zaidi [b,c,d,e], Abbes Amira [f,g]

[a] Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UK
[b] Division of Nuclear Medicine and Molecular Imaging, Department of Medical Imaging, Geneva University Hospital, Geneva, Switzerland
[c] Geneva University Neurocenter, Geneva University, Geneva, Switzerland
[d] Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[e] Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark
[f] Department of Computer Science, University of Sharjah, Sharjah, United Arab Emirates
[g] Institute of Artificial Intelligence, De Montfort University, Leicester, UK

ARTICLE INFO

ABSTRACT

The segmentation of MR (magnetic resonance) images is a simple approach to create Pseudo CT images which are useful for many medical imaging analysis applications. One of the main challenges of this process is the bone segmentation of brain MR images. Deep convolutional neural networks (CNNs) have been widely and efficiently applied to perform MR images segmentation. The aim of this work is to propose a novel excitation-based CNN by recalibrating the network features adaptively to enhance the bone segmentation by segmenting the brain MR images into three tissue classes: bone, soft tissue, and air. The proposed method combines two types of features excitation mechanisms namely: (1) spatial squeeze and channel excitation block (cSE) and (2) channel squeeze and spatial excitation block (sSE). The two blocks are combined sequentially and integrated seamlessly into a 3D convolutional encoder decoder network. The novelty of this work emerges in the combination of the two excitation blocks sequentially to improve the segmentation performance and reduce the model complexity. The proposed approach is evaluated through a comparison with computed tomography (CT) images as ground truth and validated with other methods in the literature that applied deep CNN approaches to perform MR image segmentation for PET attenuation correction. Brain MR and CT datasets which consist of 50 patients are used to evaluate the proposed method. The segmentation performance of the three brain classes is evaluated using precision, recall, dice similarity coefficient (DSC), and Jaccard index. The presented method improves the bone tissue segmentation compared to the baseline model and other methods in the literature where the DSC is improved from 0.6278 ± 0.0006 to 0.6437 ± 0.0006 with an improvement percentage of 2.53% for bone class. The proposed excitation-based segmentation network architecture demonstrates promising and competitive results compared with other methods in the literature and reduces the model complexity thanks to the sequential combination of the two excitation blocks.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

The increase use of magnetic resonance (MR) imaging in clinical routine thanks to its advantages over computed tomography (CT) images requires methods for creating pseudo CT images by obtaining the CT numbers from MR images. The synthesis of pseudo CT images is an essential step in the integrated process of many neurological images analysis such as positron emission tomography (PET) attenuation correction using MR images (MRAC) [1–3].

Commonly, the PET attenuation is corrected using the attenuation coefficients from CT images by rescaling the Hounsfield units to linear attenuation coefficients maps. However, the CT scan exposes the patient to radiation dose and generates images with low soft tissue contrast [4]. Recently, various learning based methods using deep learning have been proposed to learn the complex mapping from the tissue details of MR images to CT images in the same patients [5–12]. Another way to generate pseudo CT images is to segment MR images into different tissue classes. It has been proven that the MR image segmentation based method for PET quantification is a simple and robust approach which has been applied in commercial PET/MR scanners [13].

---

* Corresponding author.
  E-mail address: imene.mecheter@brunel.ac.uk (I. Mecheter).

Deep convolutional neural network (CNN) has been applied successfully to perform segmentation tasks in different medical imaging applications [14–18]. Various fully convolutional segmentation network architectures have been proposed, yet few of them emphasis on the characterization of spatial-wise and channel-wise features and patterns.

In this research, an adaptive features recalibration mechanism is applied using squeeze and excitation (SE) block that can be embedded at any stage of any network architecture. This mechanism has been introduced firstly for 2D images classification application by [19] then proposed for 2D images segmentation application by [20]. The 3D squeeze and excitation block is introduced in this work to segment 3D MR images. Additionally, the concept of combining the channel and spatial excitation blocks sequentially is proposed to recalibrate the features more efficiently and reduce the model complexity.

The main objective of this research is to improve the MR images segmentation method to generate segmented CT images which can be converted to pseudo CT. The proposed approach aims to integrate the SE block with CNN architecture to influence the features extraction process. Another objective is studying the impact of each type of the SE blocks on the segmentation results to find the optimal network architecture that improves the segmentation performance and reduces the model complexity as well as the required computing resources. The main contributions herein are designing a 3D encoder decoder convolutional network with 3D SE blocks and the combination of two SE blocks sequentially to firstly excite the useful channels then focus the spatial attention on certain regions with informative details. Moreover, the proposed CNN is trained using the dice similarity coefficient as an objective function to overcome the issue of the voxels unbalancing. The proposed network architecture is evaluated using clinical MR and CT brain datasets to enhance the bone segmentation in brain MR images with minimal increase of model complexity.

The paper is structured as follows. Section 2 reviews briefly the related work on MR images segmentation using deep learning and excitation mechanisms. The data description and the applied methodology are defined in Section 3. The experimental setup is presented in Section 4.The conducted experiments with the segmentation results and the limitations are discussed in Section 5. Finally, the conclusion and future work are given in Section 6.

2. Related work

Deep learning has been applied slightly to address the problem of segmentation based MR attenuation correction for brain PET images and has shown its superiority compared to the conventional machine learning techniques. For instance, deep learning guided segmentation methods perform better than Dixon-based soft-tissue and air segmentation using machine learning [14].

Fully CNN has been applied to generate segmented MR images using segmented CT images as ground truth to create pseudo CT images. Liu et al. [21] have applied the vanilla implementation of Segnet architecture using a dataset of 40 T1-weighted volumes. The main drawback of this architecture is the use of cross entropy loss as an objective function which does not take into account the class unbalancing issue. The number of bone voxels is much less than the air and soft tissue voxels in the brain volume; Hence, the classifier will be biased towards the majority class. There is a need to consider using an objective function which deals with the class unbalancing issue while training the network. This method shows promising results, yet the results are not cross validated and the segmentation results need further validation and testing.

Later, Jang et al. [22] have applied the pretrained network with T1-weighted images to segment ultrashort echo time (UTE) MR sequences. The combination of the segmented UTE images with Dixon images are used to generate the pseudo CT images. They have also enhanced the segmentation results using the conditional random field technique as a post processing step. Although this method shows good results, it consists of complicated aspects such as the use of UTE sequences and the application of a post processing step which increases the method complexity.

In addition to fully CNN, there are other networks architectures that have been applied to learn the mapping from MR to CT images to generate pseudo CT with continuous values such as generative adversarial network [5–9], U-Net [10], residual U-Net [11], and HighRes3DNet [12].

Focusing on the applied CNNs for MR images segmentation, there are some studies that aim at improving the network architecture by proposing features recalibration mechanisms to boost the features and improve the segmentation results. Hu et al. [19] were the first who have introduced the features recalibration approach using channel excitation block for ILSVRC 2017 classification competition. Since then, this mechanism has been widely and successfully applied to solve different applications such as breast density categories classification [23], action detection [24], and sea ice images classification [25]. The spatial squeeze and channel excitation also has been applied for various medical images segmentation applications such as brain tumor segmentation [26], prostate zonal segmentation [27] and pancreas segmentation using CT images [28]. Moreover, this excitation mechanism has been proposed as a feature fusion module using public datasets by [29,30]. Roy et al. [20] were the first who have introduced the spatial excitation mechanism and combined it with channel excitation using different fully CNNs architectures. The two types of excitations have been used successfully for white matter hyperintensities segmentation using parallel combination [31].

The channel and spatial excitation mechanisms have shown substantial improvement in the segmentation of medical images. However, to the best of our knowledge, no literature study has applied this technique to improve the bone segmentation in MR images for pseudo CT images generation. Moreover, this is the first study that combines the channel and spatial excitation mechanisms sequentially to improve the segmentation and reduce the model complexity.

3. Materials and methods

3.1. Data acquisition

The objective of this study is to build an efficient segmentation method using the most conventional MR sequences such as T1-weighted and T2-weighted. These sequences can be acquired more frequently in the clinical routine in contrast to other sophisticated MR sequences (e.g., ZTE and UTE) which have the drawbacks of the long acquisition time and the challenging implementation [32,33].

The brain imaging datasets consist of 50 patients that have acquired CT and MR images after giving informed consent. The patients have been diagnosed with different clinical conditions such as epilepsy, brain tumor, and neurodegenerative diseases.

The MR images have been acquired using a 3 T MAGNETOM Skyra (Siemens Healthcare, Erlangen, Germany) with a 64-channel head coil. The MR scans are 3D T1-weighted (magnetization-prepared rapid gradient-echo (MP-RAGE)) sequence that have been acquired using the following parameters: TE = 2.3 ms, TR = 1900 ms, T1 = 970 ms, flip angle 8°, NEX = 1, and matrix dimension of $255 \times 255 \times 250$ with voxel size of $0.86 \times 0.86 \times 1$ mm.

The CT scans have been acquired using two different Siemens scanners where 15 patients underwent the Biograph mCT scanner (Siemens Healthcare, Erlangen, Germany) and 35 patients underwent the Biograph 64 True Point scanners (Siemens Healthcare, Erlangen, Germany). The matrix size of the CT images is $512 \times 512 \times 150$ with voxel size of $0.97 \times 0.97 \times 1.5$ mm. The research protocol has been approved by the research ethics committee of the college of engineering, design and physical sciences at Brunel University London. An example of a single slice of T1-weighted MR image, CT image, and MR overlaid on CT are illustrated in Fig. 1 in transverse, coronal, and sagittal views.

3.2. Data preprocessing

The processing of each patient volume of both T1-weighted and CT images is performed in 2D fashion by firstly identifying the region of interest and selecting 48 slices from each volume. The process of slices selection is performed manually by removing the noisy slices that do not include any details of the brain. The removed slices reside either at the beginning or at the end of each volume. Secondly, the number of pixels of the background are reduced by cropping each slice into $256 \times 256 \times 48$ matrix dimension. Due to the usage of different scanners, some of the volumes are resampled to 300×300 matrix using bilinear interpolation then cropped into $256 \times 256 \times 48$ to obtain the same dimensions for all patients. Local contrast normalization technique is applied on MR images only to downscale the range of the pixel's values. CT images are used as ground truth for this supervised learning approach where each MR volume is co-registered with its corresponding CT volume to overcome the temporal gap (within 2–3 days) between the acquisition of MR and CT images. The registration process combines the rigid Euler transformation followed by the non-rigid B-spline transformation using Elastix tool [34].

The labelling 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 labelled as bone, lower than -500 HU are labelled as air, and other pixels are labelled as soft tissue.

3.3. Methods

3.3.1. Encoder Decoder convolutional network

The baseline architecture follows the well-known 2D U-Net architecture [35] that has been applied widely in various medical applications. This network is a fully convolutional architecture that takes the shape of an encoder decoder path with the inclusion of additional paths from the high-resolution to the low-resolution features. The additional paths increase the resolution of the output and improve the localization. The U-Net architecture is amended by converting 2D layers to 3D layers, decreasing the number of down-sampling layers which cause loss of some features, and increasing the number of convolutional layers at each depth. It consists of 21 3D convolutional layers, each three convolutional layers are followed by the batch normalization layer, the rectified-linear unit (ReLU) activation function, and one 3D max-pooling layer with stride size of 2. The number of filters of the first layer is 64 and doubles in size as the depth of the network
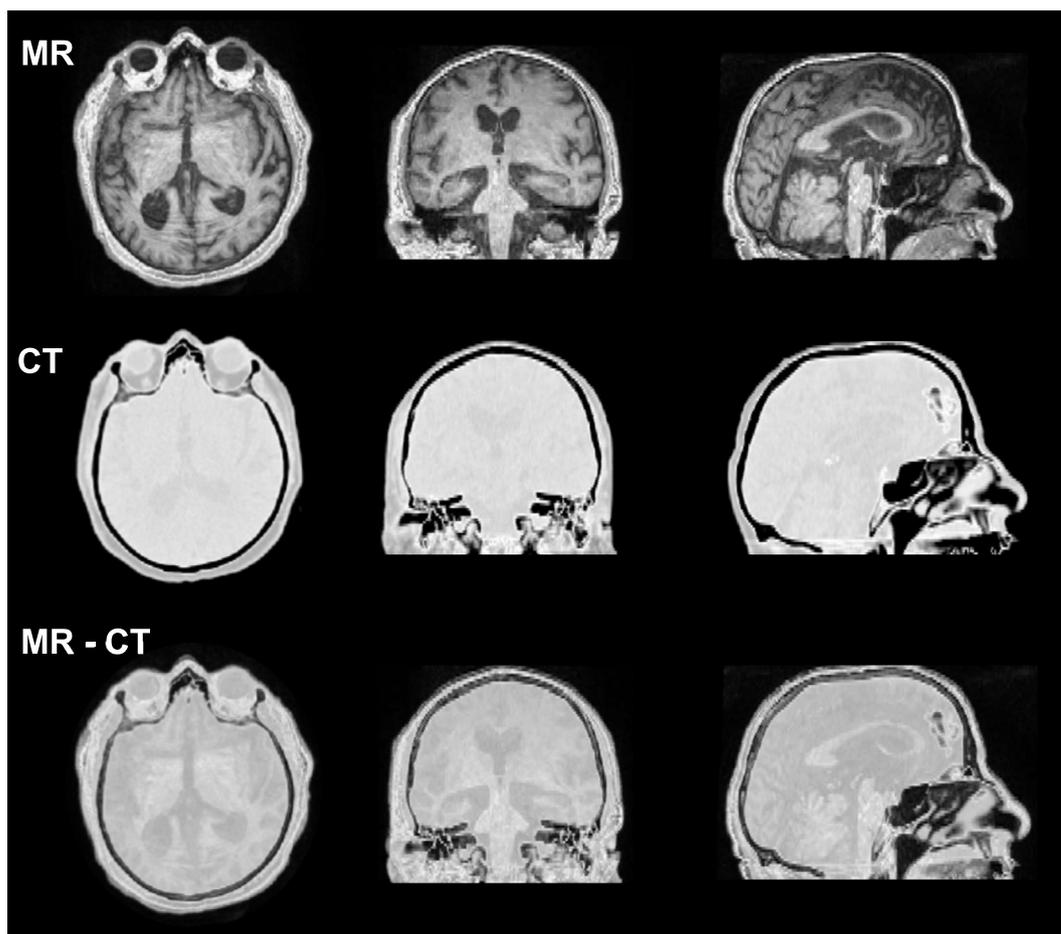

Fig. 1. An illustration of a single slice of MR, CT, and MR overlaid on CT in transverse, coronal, and sagittal views.

increases. The decoder path is the mirror of the encoder path with replacing the maxpooling layers with deconvolution layers. The use of the deconvolution layer is another modification to the U-Net architecture to allow the network to learn the weights of the upsampling process. The last layer of the network is a fully connected layer that consists of three neurons that represent the number of classes with a multiclass sigmoid activation function. The CNN hyper-parameters are selected based on a grid search procedure. Fig. 2 illustrates the baseline proposed architecture.

Another modification is the application of the dice similarity coefficient (DSC) as an objective function to train the network to overcome the problem of classes unbalancing which leads to a biased classification towards the majority class. The cross-entropy loss function that is applied in U-Net does not address this problem. The dice coefficient loss for multi-class segmentation is defined as:

$$L_{dice} = 1 - \sum_{c=1}^{C} \frac{2\sum_{h=1}^{H}\sum_{w=1}^{W}\sum_{d=1}^{D} p(c,h,w,d)gt(c,h,w,d) + s}{\sum_{h=1}^{H}\sum_{w=1}^{W}\sum_{d=1}^{D} p(c,h,w,d)^2 gt(c,h,w,d)^2 + s} \tag{1}$$

where h is the index of the height, w is the index of the width, d is the index of the channels, p is the probability of each voxel (output of the segmentation), gt is the one hot encoded value of the ground truth for class c, C is the number of tissue classes, and s is a smoothing variable to avoid division by zero.

3.3.2. Features recalibration mechanism

The down-sampling layers of the convolutional network cause loss of useful features by taking the maximum or average values of each patch of the features map. This drawback initiates the exploration of the features recalibration mechanism that aims at capturing the meaningful and distinguishable features and suppressing the less useful ones. The recalibration is represented by a generic and resilience component that consists of few blocks and can be integrated within any convolutional network architecture. The recalibration mechanism is the resultant of squeezing the features along the spatial or the channel domain then exciting the same set of features spatially or channel wise. This technique identifies a relationship between the number of channels, the spatial locations, and the features maps.

Channel Excitation Mechanism. This type of excitation uses the channels interdependencies by squeezing the features maps along the spatial domain then exciting along the channel domain to recalibrate and reweight the features that flow to the next layers. This type of recalibration is called spatial squeeze and channel excitation (cSE). It squeezes the global spatial features into a channel descriptor. It is more adapted for classification applications as it enables the extraction of global context information and improves the channel interdependencies. The components of this mechanism are a global average pooling to shrink the features maps spatially followed by two fully connected layers to reduce the model

3D convolution − Batch normalization − ReLU
3D maxpooling
3D Deconvolution
⋯⋯ Copy & Concatenate

Fig. 2. The architecture of the convolutional encoder decoder network that follows the U-Net.

complexity then the ReLU activation function to add nonlinearity and introduce generalization. The last layer is a sigmoid activation function to rescale the features to the interval [0,1]. The resultant vector is then multiplied by the original features maps for excitement and recalibration. Considering a feature map U with c channels:

$$U = [u_1, u_2, \ldots, u_c], \tag{2}$$

the spatial squeezing is calculated by vector z with k elements:

$$z_k = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_k(i,j), \tag{3}$$

where H and W are the height and width of the feature map. By adding the two fully connected layers and the activation functions, z transformed to:

$$s_k = \sigma(W_2(\delta(W_1 z_k))) \tag{4}$$

where σ and δ refer to sigmoid and ReLU activation functions respectively, $W_1 \in R^{\frac{c}{r \times c}}$ and $W_2 \in R^{\frac{c}{r}}$. The parameter r is the reduction ratio that compresses the c channels to $\frac{c}{r}$ channels that can be selected empirically, W_1 and W_2 are the weights of the first and second fully connected layers. The vector s is multiplied by the features map U to excite and recalibrate the features and produce the final vector:

$$X_{cSE} = [u_1.s_1, u_2.s_2, \ldots, u_c.s_c] \tag{5}$$

The architecture of this block is illustrated in Fig. 3.

Spatial Excitation Mechanism. The channel squeeze and spatial excitation mechanism (sSE) squeezes the features maps on the channel wise to produce a scalar value then excites spatially. This mechanism preserves the global semantic information and enables the representation capabilities of feature maps. It is a useful mechanism for image segmentation as it enhances the salient spatial locations. The architecture of this mechanism is shown in Fig. 4 where the spatial squeeze part is represented by the convolution operation using convolutional layer with size $1 \times 1 \times 1$:

$$q = w \times u_{(i,j)} \tag{6}$$

where w is the weight and u is a single spatial location (i,j) at the features map. The output vector q combines all channels for a spatial location then passes through the sigmoid activation function to rescale the features in the scale [0,1]. The resultant is multiplied by the feature maps u to excite and adaptively recalibrate the features. The spatially excited feature map is calculated as:

$$X_{sSE} = [\sigma(q_1).u_1, \sigma(q_2).u_2, \ldots, \sigma(q_{(H,W)}).u_{(H,W)}] \tag{7}$$

where H and W are the height and width of the feature map.

Combined Spatial and Channel Excitation mechanisms (scSE). The exploitation of the two excitation mechanisms channel and spatial wise is conducted by combining both of the blocks. This combination helps to achieve the full potential of recalibrating the features and capturing the useful information. The combination has two different configurations: parallel scheme and sequential scheme.

Parallel scheme. The parallel combination has been proposed on 2D domain by [20] where the input feature map passes through the cSE and sSE blocks concurrently. Then the output is aggregated using different strategies such as max-out, addition, multiplication, and concatenation. This parallel combination is applied herein on the 3D domain using the concatenation aggregation method as it provides the best performance. Fig. 5 depicts the architecture of the parallel combination on 3D domain.

Sequential scheme. The sequential combination of the cSE and sSE blocks is introduced in this work to decrease the model complexity and the required computing resources. Additionally, the rationale behind this combination structure is to create an empha-
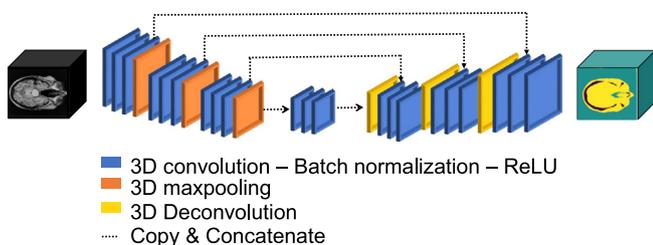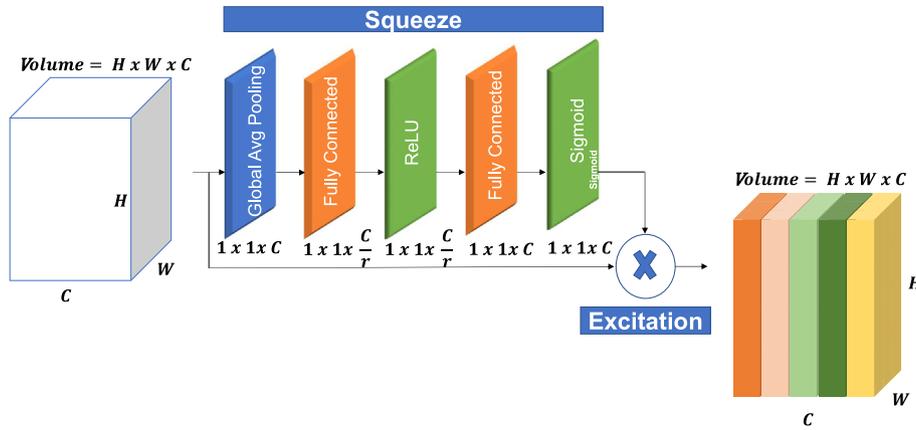
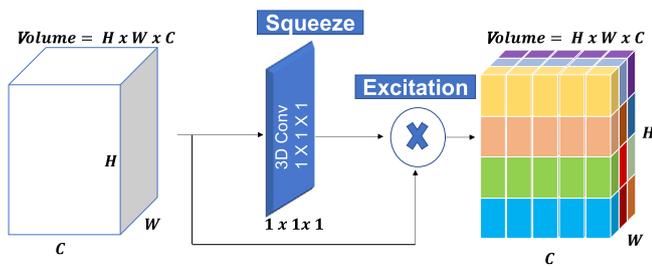**Fig. 3.** The architecture of the spatial squeeze and channel excitation block (cSE).

Fig. 4. The architecture of the channel squeeze and spatial excitation block (sSE).
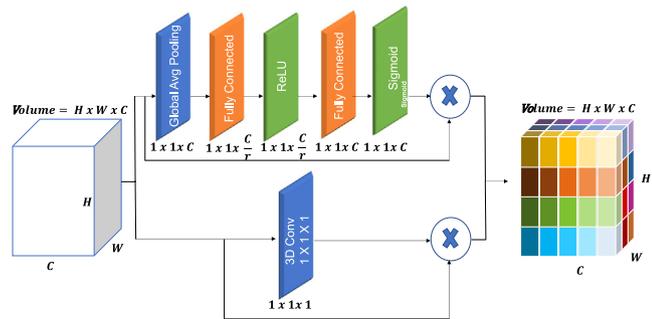

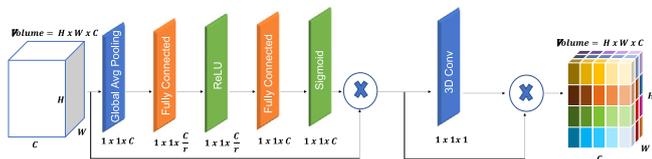**Fig. 5.** The architecture of the parallel combination of channel and spatial excitation blocks.

Fig. 6. The architecture of the sequential combination of channel then spatial excitation blocks (cSE then sSE).

sis on the recalibration mechanism by passing only the excited useful features to the next excitation mechanism. The order of the combination matters where each SE block receives different excited features either channel or spatial wise. It has been shown that the spatial excitation is more efficient than channel excitation to perform segmentation tasks [20]. Thus, performing the spatial excitation at the end of the excitation process (cSE then sSE) tends to emphasis on the spatial locations that consist of useful inter channel information that enable the distinguish of the most useful

spatial locations of the feature maps. Fig. 6 illustrates the architecture of the sequential combination of SE blocks which starts with channel then spatial excitation. Fig. 7 shows the excitation based network architecture.

4. Experimental setup

4.1. Input and training setup

The volume size of each patient is $256 \times 256 \times 48$ where 48 is the number of selected slices per patient. This huge number of pixels overwhelm the graphical processing unit (GPU) memory and creates limitations when designing the network architecture such as selecting the number of filters per convolutional layer. To overcome this situation, each volume is divided into overlapped patches to minimize the input size and increase the size of the training datasets. The application of overlapped patches preserves the spatial contextual information for each volume. Each volume is divided into 9 overlapped patches with the size of $128 \times 128 \times 48$.

The deep network is trained from scratch using four folds cross-validation where the size of each fold is 10 volumes. For each trial, the training and validation datasets consist of 30 and 10 patients, respectively. The testing dataset is fixed and consists of 10 patients used to evaluate the results of each trial. The layers weights are initialized using Glorot Uniform initialization scheme and updated using Adam optimizer with a moving learning rate that starts with 0.0001 then reduces by a factor of 0.75 on plateau mode when the training accuracy is not increasing for 5 continuous epochs. The batch size is 2 and the training stops after 100 epochs or when the validation loss stops decreasing. The only hyperparameter that is changed while training the parallel combination is the batch size that is set to 1 to solve the memory allocation issue of the utilized GPU.

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 the implementation of CNN architecture. The deep CNN is trained using Tesla V100 GPU with 16 GB RAM.

4.2. Evaluation metrics

The segmentation results are evaluated by comparing the segmented MR image with the ground truth CT image. The calculated evaluation metrics are precision (PRE), recall (REC), dice similarity coefficient (DSC), and Jaccard index (JAC). Moreover, the segmentation performance is assessed with 95% confidence intervals.
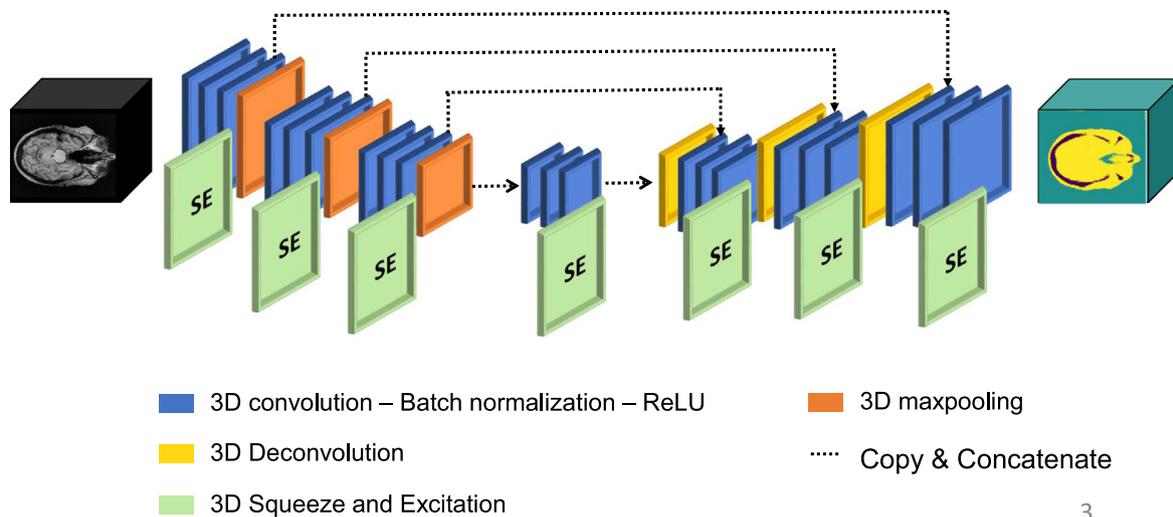
Fig. 7. The architecture of the convolutional encoder decoder network with multiple squeeze and excitation blocks.

5. Results and discussion

5.1. Segmentation results

The impact of each type of the excitation blocks is studied by placing the SE block at the same location using the same network architecture and comparing the evaluation metrics which are generated from the confusion matrix. The average of the evaluation metrics from the obtained results using four folds cross-validation for bone, soft tissue, and air classes are shown in Tables 1–3, respectively which represent the pixel-wise comparison of the testing datasets for various methods. Table 4 demonstrates the evaluation results per patient using the cSE then sSE (sequential) method without the use of validation set where the training and testing sets consist of 40 and 10 patients, respectively. The results show that the addition of any type of SE improves the segmentation results of the bone class. The spatial excitation outperforms the channel excitation as it enables the global contextual information on the spatial domain. The two types of the sequential combination (cSE-sSE and sSE-cSE) outperform the parallel combination. The performance of the spatial excitation approach for bone segmentation is very similar to the sequential combination of spatial then channel excitation (cSE-sSE).

Another important observation is that placing the cSE block before the sSE block in the sequential configuration shows superiority in terms of segmentation evaluation metrics. Starting with recalibrating the features in channel wise helps to emphasis on the most useful inter-channel features that will be excited spatially in the next stage. This helps the model to identify the spatial locations of the most useful channels.

The box plots of different evaluation metrics which are shown in Fig. 8 indicate a high variability on the segmentation results of the bone class with high standard deviation. Additionally, the intrasubject variability percentage values for each tissue class are shown in Table 5. The intrasubject variability is calculated by dividing the standard deviation over the mean of the DSC obtained from the four folds cross-validation experiments. The results reveal that the intrasubject variability of the air and soft tissue classes are negligible. However, there are some variations on the bone class. For some patients the bone class variability is very low such as patient 9 (0.81%), yet it is higher in patient 7 (51.43%). One reason of this high variability is the inhomogeneity of the datasets that include patients with different neurological clinical diagnosis such as epilepsy, brain tumor, and other neurological diseases. In fact, different neurological diseases can impact bone health differently [36]. Hence, the segmentation of each neurological condition should be addressed separately where a model should be trained and tested with datasets that include patients with the same clinical diagnosis. For instance, the DSC value of the bone class of patient 7 in Table 4 is very low (DSC = 0.2947) compared to patient 9 (DSC = 0.8744). The proposed model has the capability to segment the bone class accurately for some patients while it fails with others. The variability of the results of bone segmentation does not provide any guarantee that this method is robust for patients with different anatomical structures. Moreover, there are few outliers values of the air and soft tissue classes in terms of precision, recall, DSC, and Jaccard score. These outliers belong to the data of patient 10 which records low values compared to other patients as shown in Table 4.

The proposed method is compared with other segmentation studies that have applied deep convolutional network to segment MR images for PET attenuation correction [21,22]. The method proposed by [21] follows the SegNet architecture while a pretrained SegNet with conditional random field (CRF) technique is applied by [22]. These methods are re-implemented using our datasets and the segmentation results of the bone, soft tissue, and air

Table 1
The average of four folds cross validation for the evaluation metrics of the bone class from the confusion matrix of the testing datasets using different configurations of SE blocks.

Model	Bone			
	PRE	REC	DSC	JAC
Baseline	0.7197 ± 0.0006	0.5575 ± 0.0006	0.6278 ± 0.0006	0.4578 ± 0.0007
+ cSE	0.7366 ± 0.0006	0.5538 ± 0.0006	0.6319 ± 0.0006	0.4620 ± 0.0007
+ sSE	0.7228 ± 0.0006	0.5875 ± 0.0006	0.6474 ± 0.0006	0.4789 ± 0.0007
+ cSE and sSE (parallel)	0.7219 ± 0.0006	0.5065 ± 0.0007	0.5928 ± 0.0006	0.4222 ± 0.0006
+ cSE then sSE (sequential)	0.7373 ± 0.0006	0.5722 ± 0.0006	0.6437 ± 0.0006	0.4750 ± 0.0007
+ sSE then cSE (sequential)	0.7223 ± 0.0006	0.5631 ± 0.0006	0.6326 ± 0.0006	0.4628 ± 0.0007

Table 2

The average of four folds cross validation for the evaluation metrics of the soft tissue class from the confusion matrix of the testing datasets using different configurations of SE blocks.

Model	Soft tissue			
	PRE	REC	DSC	JAC
Baseline	0.8932 ± 0.0001	0.9283 ± 0.0001	0.9104 ± 0.0001	0.8355 ± 0.0002
+ cSE	0.8941 ± 0.0001	0.9297 ± 0.0001	0.9115 ± 0.0001	0.8375 ± 0.0002
+ sSE	0.8972 ± 0.0001	0.9274 ± 0.0001	0.9120 ± 0.0001	0.8382 ± 0.0002
+ cSE and sSE (parallel)	0.8959 ± 0.0001	0.9310 ± 0.0001	0.9131 ± 0.0001	0.8400 ± 0.0002
+ cSE then sSE (sequential)	0.8956 ± 0.0001	0.9311 ± 0.0001	0.9130 ± 0.0001	0.8399 ± 0.0002
+ sSE then cSE (sequential)	0.8933 ± 0.0001	0.9293 ± 0.0001	0.9109 ± 0.0001	0.8365 ± 0.0002

Table 3

The average of four folds cross validation for the evaluation metrics of the air class from the confusion matrix of the testing datasets using different configurations of SE blocks.

Model	Air			
	PRE	REC	DSC	JAC
Baseline	0.9630 ± 0.0001	0.9636 ± 0.0001	0.9633 ± 0.0001	0.9291 ± 0.0001
+ cSE	0.9590 ± 0.0001	0.9622 ± 0.0001	0.9606 ± 0.0001	0.9242 ± 0.0001
+ sSE	0.9639 ± 0.0001	0.9636 ± 0.0001	0.9638 ± 0.0001	0.9301 ± 0.0001
+ cSE and sSE (parallel)	0.9618 ± 0.0001	0.9717 ± 0.0001	0.9667 ± 0.0001	0.9356 ± 0.0001
+ cSE then sSE (sequential)	0.9631 ± 0.0001	0.9633 ± 0.0001	0.9632 ± 0.0001	0.9290 ± 0.0001
+ sSE then cSE (sequential)	0.9636 ± 0.0001	0.9628 ± 0.0001	0.9632 ± 0.0001	0.9290 ± 0.0001

Table 4

The evaluation metrics of the brain segmentation of the three tissue classes of each patient in the testing datasets using the sequential configuration of cSE then sSE blocks.

Model	Bone				Soft tissue				Air			
	PRE	REC	DSC	JAC	PRE	REC	DSC	JAC	PRE	REC	DSC	JAC
Patient 1	0.8209	0.5072	0.6270	0.4566	0.9197	0.9606	0.9397	0.8863	0.9731	0.9921	0.9825	0.9656
Patient 2	0.8791	0.6871	0.7713	0.6278	0.9325	0.9488	0.9406	0.8878	0.9564	0.9488	0.9406	0.8878
Patient 3	0.6654	0.3154	0.4280	0.2722	0.9033	0.9368	0.9198	0.8514	0.9513	0.9703	0.9607	0.9244
Patient 4	0.6882	0.3086	0.4261	0.2708	0.8518	0.9491	0.8978	0.8146	0.9772	0.9814	0.9793	0.9594
Patient 5	0.9173	0.7404	0.8194	0.6940	0.9350	0.9667	0.9506	0.9058	0.9808	0.9914	0.9861	0.9725
Patient 6	0.4966	0.5041	0.5003	0.3336	0.9187	0.8914	0.9048	0.8262	0.9701	0.9910	0.9804	0.9616
Patient 7	0.3977	0.2341	0.2947	0.1728	0.9394	0.9176	0.9283	0.8663	0.9514	0.9937	0.9721	0.9457
Patient 8	0.6312	0.6080	0.6194	0.4486	0.9385	0.9280	0.9333	0.8749	0.9815	0.9921	0.9868	0.9738
Patient 9	0.9098	0.8417	0.8744	0.7769	0.9644	0.9725	0.9684	0.9388	0.9869	0.9926	0.9897	0.9796
Patient 10	0.6549	0.5020	0.5683	0.3970	0.6758	0.9084	0.7750	0.6327	0.9208	0.7516	0.8277	0.7060
Mean	0.7061	0.5249	0.5929	0.4450	0.8979	0.9380	0.9158	0.8485	0.9650	0.9605	0.9606	0.9277
SD	0.1759	0.1990	0.1878	0.1985	0.0835	0.0265	0.0536	0.0841	0.0201	0.0748	0.0490	0.0827

classes are depicted in Tables 6–8, respectively. It is worth mentioning that the application of CRF [22] as a post processing technique has not shown any enhancement in the segmentation results compared to the use of SegNet solely [21].

The segmentation results of some 2D slices are illustrated in Fig. 9 which includes the input MR images, CT images as ground truth, the segmentation results of the proposed network, and the segmentation results of other studies.

5.2. Position of SE block

The position of SE block has an impact on the performance of the CNN. Different configurations are tested to find the optimal place of the SE component by calculating the DSC. The sequential combination of spatial and channel excitation is placed after each 3 convolutional layers at the encoder part, the bottleneck part, the decoder part, and after the classifier layer. The different configurations are illustrated in Fig. 10 and the DSC of different configurations is recorded in Table 9.

5.3. Combining mechanism

The comparison between the sequential and the parallel combination of channel and spatial excitation components is evaluated by calculating the DSC of each configuration. One drawback of

the parallel combination is the need to train large number of parameters which require a large GPU memory. Due to this limitation, two different architectures are tested to minimize the number of parameters that can be trained with the available GPU memory. The first experiment is conducted by inserting the csSE block into a network that consists of only ten convolutional layers and two downsampling layers using the batch size of 2. The second experiment is conducted by inserting the csSE block into the baseline architecture (shown in Fig. 7) with the batch size of 1. The segmentation results are recorded using the DSC as shown in Table 10. The sequential combination not only reduces the number of training parameters but also shows superiority of the segmentation results in terms of DSC using two different architectures. The DSC of the brain class shows substantial improvement compared to other classes which are slightly improved. The sequential combination mechanism is capable of recalibrating the features in channel wise then spatial wise significantly by helping the network to recalibrate the useful spatial regions of the most dominant channels.

5.4. The relation between SE block and convolution layers

The design of placing the SE block after n convolutional layers is studied and evaluated by conducting some experiments. In theory, the addition of one SE block after each convolutional layer is the optimal design especially with shallow convolutional networks
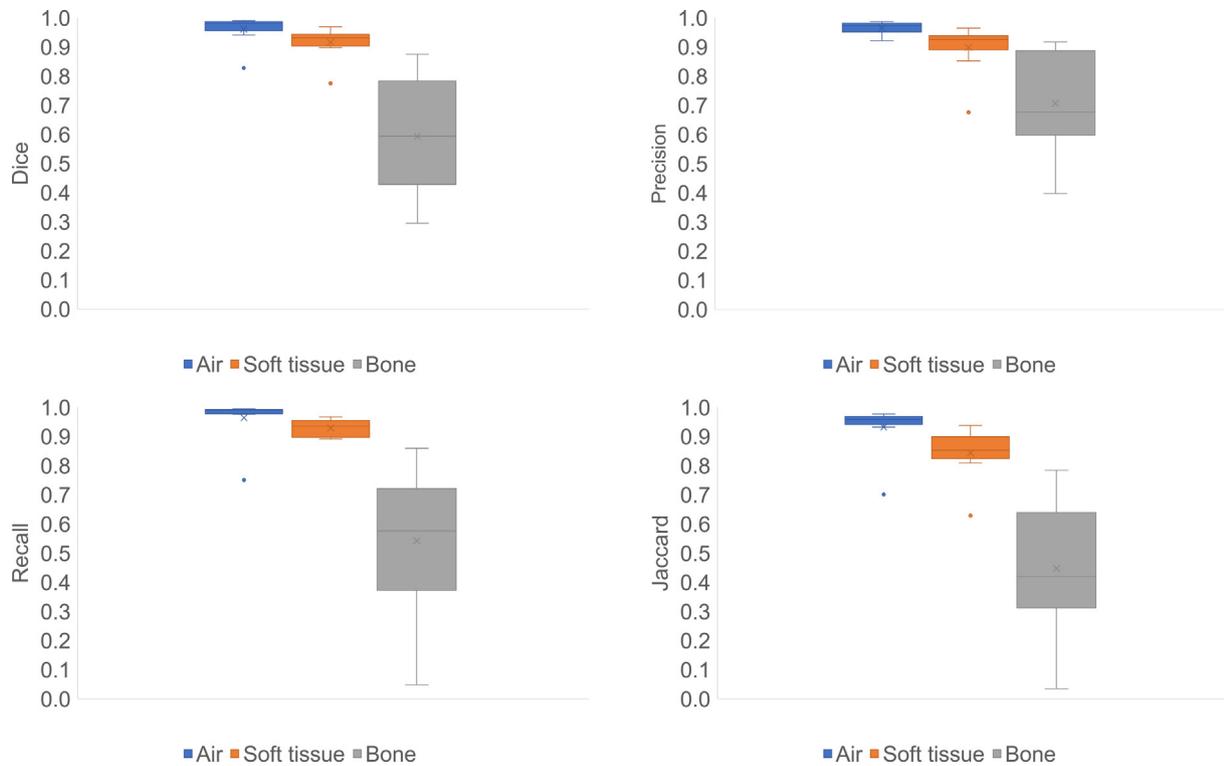
Fig. 8. The box plot of the evaluation metrics of the three tissue classes of the testing datasets using the sequential configuration of cSE then sSE blocks.

Table 5
The intrasubject variability based on the calculated DSC of the three tissue classes which are segmented using the sequential combination method (cSE-sSE).

Patient	Intrasubject variability		
	Bone %	Soft tissue %	Air %
Patient 1	5.72	0.36	0.22
Patient 2	5.12	0.46	1.60
Patient 3	8.86	0.34	0.12
Patient 4	14.04	0.83	0.18
Patient 5	3.53	0.57	0.14
Patient 6	3.78	0.57	0.20
Patient 7	51.43	0.40	0.17
Patient 8	10.29	0.56	0.09
Patient 9	0.81	0.19	0.20
Patient 10	2.01	0.18	0.21

that have few trainable parameters. However, due to GPU memory limitation, it is not possible to add one SE block after each convolutional layer in deep networks as proposed in this study where 21 convolutional layers are used to enable the extraction of dominant features. The tradeoff between the number of convolutional layers and the number of SE blocks leads to include one SE block after each three convolutional layers. This design produces the optimal results experimentally as shown in Table 11. The results show that the inclusion of more SE blocks can substitute the addition of more convolutional layers to the model. Thus, using less convolutional layers with the addition of one SE block after each convolutional layer is able to generate better features rather than using more convolutional layers with less SE blocks.

5.5. Model complexity

The additional excitation blocks increase the number of trainable parameters of the network and increases the model complexity as well. The additional number of parameters of using the cSE block is given by:

$$I = \frac{2}{r}\sum_{s=1}^{S} N_s C_s^2 \quad (8)$$

where r is the reduction ratio that is set to 16 in all experiments, S is the number of stages/blocks in the network, N is the number of repeating the cSE block which is 1 in all experiments, and C is the number of channels in the feature map. On the other hand, the sSE block only introduces C_s additional parameters. The overall model complexity is given by:

$$X = \frac{2}{r}\sum_{s=1}^{S} N_s C_s^2 + C_s \quad (9)$$

The number of parameters of different architectures and configurations along with the increase rate of the model complexity are illustrated in Table 12. The results show that the sequential combi-

Table 6
The comparison of the bone class results using the sequential excitation method (cSE-sSE) with other techniques reported in the literature.

Model	Bone			
	PRE	REC	DSC	JAC
SegNet[21]	0.6278 ± 0.0006	0.3649 ± 0.0006	0.4616 ± 0.0007	0.3000 ± 0.0006
SegNet + CRF[22]	0.6278 ± 0.0006	0.3649 ± 0.0006	0.4616 ± 0.0007	0.3000 ± 0.0006
Proposed method	**0.7373 ± 0.0006**	**0.5722 ± 0.0006**	**0.6437 ± 0.0006**	**0.4750 ± 0.0007**

Table 7

The comparison of the soft tissue class results using the sequential excitation method (cSE-sSE) with other techniques reported in the literature.

Model	Soft tissue			
	PRE	REC	DSC	JAC
SegNet[21]	0.8696 ± 0.0002	0.8810 ± 0.0002	0.8753 ± 0.0002	0.7783 ± 0.0002
SegNet + CRF[22]	0.8696 ± 0.0002	0.8810 ± 0.0002	0.8753 ± 0.0002	0.7783 ± 0.0002
Proposed method	**0.8956 ± 0.0001**	**0.9311 ± 0.0001**	**0.9130 ± 0.0001**	**0.8399 ± 0.0002**

Table 8

The comparison of the air class results using the sequential excitation method (cSE-sSE) with other techniques reported in the literature.

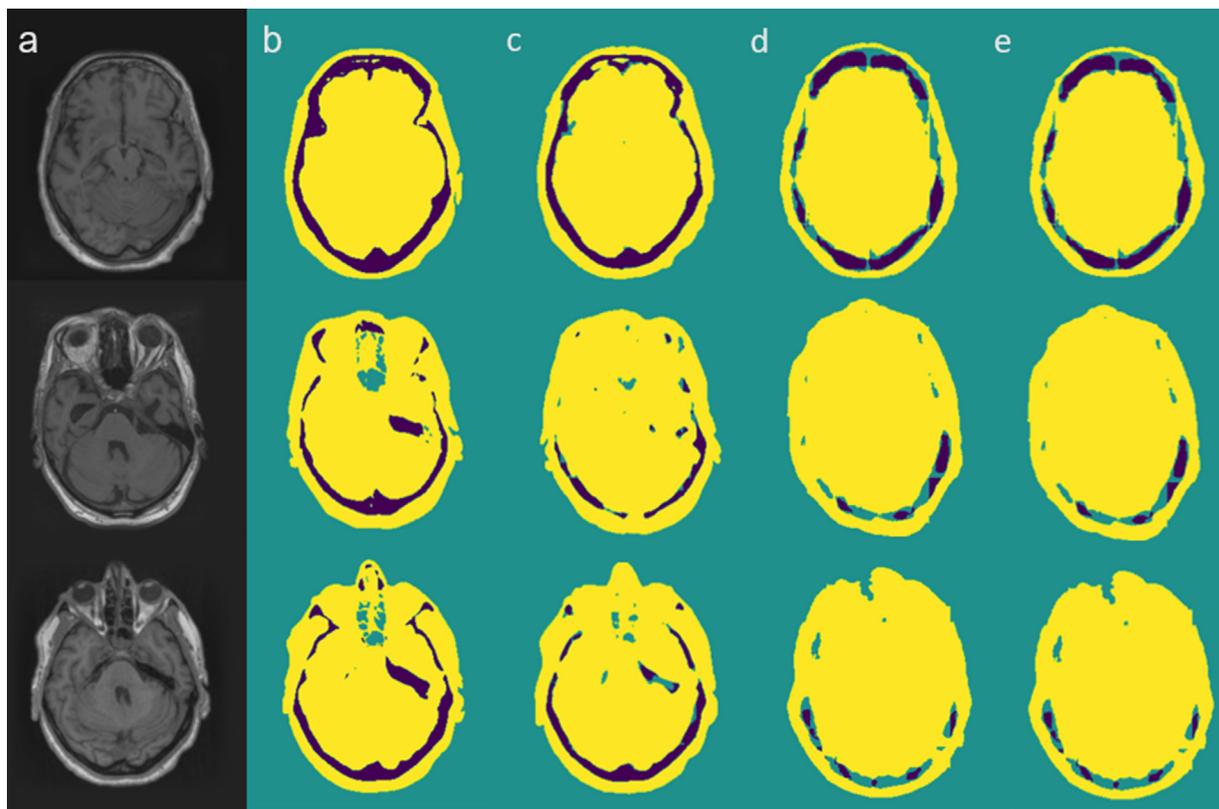| Model	Air			
	PRE	REC	DSC	JAC
SegNet[21]	0.9083 ± 0.0002	0.9510 ± 0.0001	0.9291 ± 0.0001	0.8677 ± 0.0002
SegNet + CRF[22]	0.9083 ± 0.0002	0.9510 ± 0.0001	0.9291 ± 0.0001	0.8677 ± 0.0002
Proposed method	**0.9631 ± 0.0001**	**0.9633 ± 0.0001**	**0.9632 ± 0.0001**	**0.9290 ± 0.0001**

Fig. 9. Some slices of the testing data. MR image (a), CT image(ground truth) (b), Segmented MR image with the proposed deep network (c), Segmented MR image generated by [21] (d), Segmented MR image generated by [22](e). The colors in columns b–e refer to the following classes: green is the air, yellow is the soft tissue, and purple is the bone.

nation of SE components can be added to any architecture with almost no cost. Although the parallel combination mechanism requires more computing resources with high computation cost, the runtime of the conducted experiments for both parallel and sequential (cSE-sSE) mechanisms is very similar.

5.6. Limitations

There are some limitations of the proposed work such as the obtained segmentation results rely on co-registered MR-CT pairs which makes the produced results very sensitive to any registration error [37]. Another limitation is the generation of pseudo CT images with constant values by assigning a predefined attenuation

coefficient to each class in the segmented CT images. Moreover, the segmentation of the brain into three tissue classes (air, soft tissue, and bone) might be an inadequate approximation for attenuation correction. Therefore, there is a consideration of segmenting the brain into more classes in the future work.

6. Conclusion

This work presents an improved approach that segments the brain into three tissue classes using an excited-based fully CNN by embedding squeeze and excitation blocks. These blocks recalibrate the features by capturing the most useful and distinguishable ones in channel and spatial wise in sequential structure. The pro-
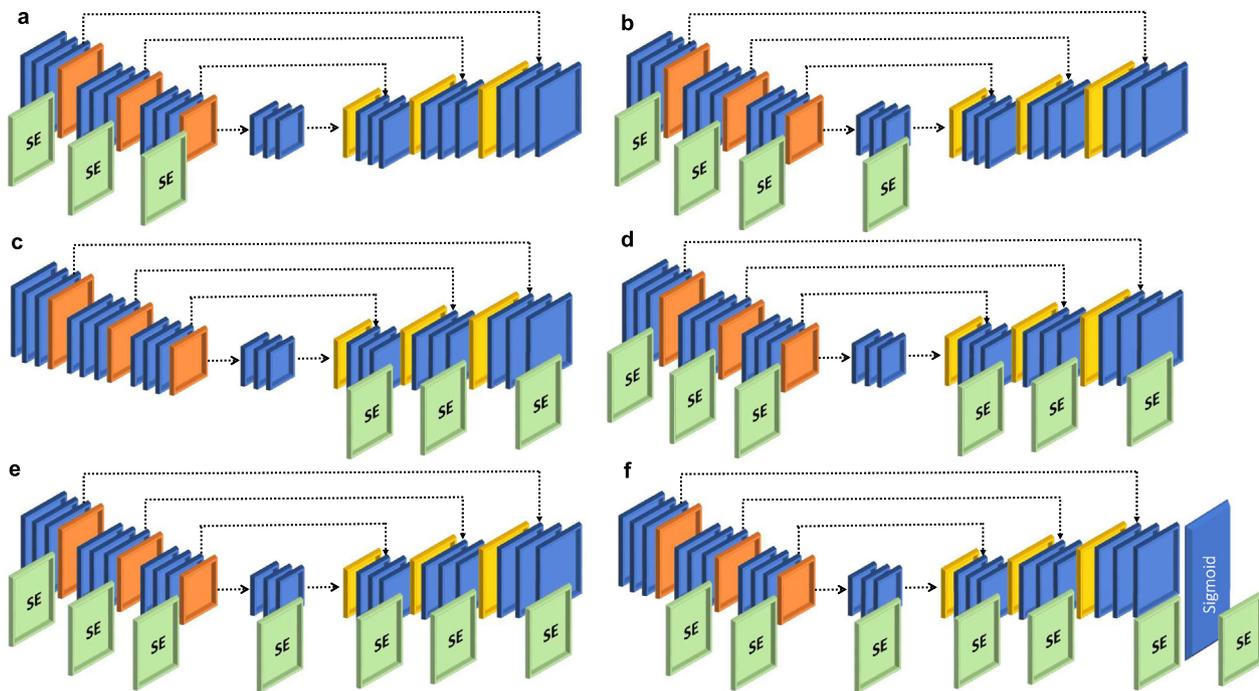
Fig. 10. Different network architectures (a–f) of placing the SE block at different locations.

Table 9

The dice similarity coefficient of the three tissue classes which are segmented using different models with different positions of the csSE block.

Model	DSC (bone)	DSC(soft tissue)	DSC (air)
without csSE block	0.6179	0.9070	0.9623
csSE at the encoder	0.6258	0.9078	0.9629
csSE at the encoder + bottleneck layer	0.6393	0.9089	0.9624
csSE at the decoder	0.6273	0.9074	0.9621
csSE at the encoder + decoder	0.6393	0.9055	0.9624
csSE at the encoder + decoder + bottleneck layer	**0.6484**	**0.9123**	**0.9633**
csSE at the encoder + decoder + bottleneck layer + classifier layer	0.6429	0.9087	0.9616

Table 12

The number of parameters and the complexity increase rate of the model using different excitation mechanisms.

Model	Number of parameters	Complexity increase rate
Baseline	37, 834, 115	
Baseline with cSE	37, 942, 659	0.3%
Baseline with sSE	37, 835, 530	0.004%
Baseline with cSE and sSE combined in parallel	55, 438, 154	47%
Baseline with cSE and sSE combined sequential	37, 889, 802	0.15%

Table 10

The comparison of the segmentation results which are generated using different combination mechanisms of SE blocks.

Combination mechanism	Conv layers	Downsampling layers	Batch size	DSC (bone)	DSC(soft tissue)	DSC(air)
Parallel	21	3	1	0.6069	0.9086	0.9640
Sequential (cSE-sSE)	21	3	1	**0.6236**	**0.9094**	**0.9660**
Parallel	10	2	2	0.6368	0.9104	0.9635
Sequential (cSE-sSE)	10	2	2	**0.6402**	**0.9111**	**0.9634**

Table 11

The results of studying the relation between the number of SE blocks and the convolutional layers.

SE type	Conv	SE	DSC (bone)	DSC(soft tissue)	DSC(air)
cSE	7	7	0.6249	0.9097	0.9610
cSE	14	7	0.6108	0.9077	0.9601
cSE	21	7	0.6302	0.9090	0.9614
cSE-sSE	7	7	0.6394	0.9095	0.9620
cSE-sSE	14	7	0.6192	0.9080	0.9629
cSE-sSE	21	7	0.6484	0.9123	0.9633

posed approach shows its efficiency to improve the bone class segmentation while minimizing the model complexity and the required computing resources. The results also show that the features recalibration mechanism enhances the segmentation results by comparing the segmented MR images with the ground truth CT images and other proposed methods in the literature. The experiments are conducted using the brain datasets that consist of patients with different conditions and diagnosis. The variability of the obtained segmentation results does not provide any guarantee that the proposed method is robust for patients with anatomic variability.

CRediT authorship contribution statement

Imene Mecheter: Conceptualization, Methodology, Software, Writing – original draft. **Maysam Abbod:** Critical revision. **Habib Zaidi:** Writing – review & editing, Data analysis and interpretation. **Abbes Amira:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

[1] Y. Chen, H. An, Attenuation Correction of PET/MR Imaging, Magnetic Resonance Imaging Clinics of North America 25 (2) (2017) 245–255 (May 2017).

[2] D. Izquierdo-Garcia, C. Catana, MR Imaging-Guided Attenuation Correction of PET Data in PET/MR Imaging, PET Clinics 11 (2) (2016) 129–149.

[3] Z. Chen, S.D. Jamadar, S. Li, F. Sforazzini, J. Baran, N. Ferris, N.J. Shah, G.F. Egan, From simultaneous to synergistic MR-PET brain imaging: A review of hybrid MR-PET imaging methodologies, Human Brain Mapping 0 (0).

[4] I. Mecheter, L. Alic, M. Abbod, A. Amira, J. Ji, Mr image-based attenuation correction of brain pet imaging: Review of literature on machine learning approaches for segmentation, Journal of Digital Imaging 2020 (2020) 1–18.

[5] H. Yang, X. Lu, S.-H. Wang, Z. Lu, J. Yao, Y. Jiang, P. Qian, Synthesizing Multi-Contrast MR Images Via Novel 3D Conditional Variational Auto-Encoding GAN, Mobile Networks and Applications 26 (1) (2021) 415–424.

[6] A. Abu-Srhan, I. Almallahi, M.A.M. Abushariah, W. Mahafza, O.S. Al-Kadi, Paired-unpaired Unsupervised Attention Guided GAN with transfer learning for bidirectional brain MR-CT synthesis, Computers in Biology and Medicine 136 (2021) (Sep. 2021) 104763.

[7] L. Tao, J. Fisher, E. Anaya, X. Li, C.S. Levin, Pseudo CT Image Synthesis and Bone Segmentation From MR Images Using Adversarial Networks With Residual Blocks for MR-Based Attenuation Correction of Brain PET Data, IEEE Transactions on Radiation and Plasma Medical Sciences 5 (2) (2021) 193–201.

[8] M. Hemsley, B. Chugh, M. Ruschin, Y. Lee, C.-L. Tseng, G. Stanisz, A. Lau, Deep Generative Model for Synthetic-CT Generation with Uncertainty Predictions, in: A.L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M.A. Zuluaga, S.K. Zhou, D. Racoceanu, L. Joskowicz (Eds.), Medical Image Computing and Computer Assisted Intervention –, 2020, Springer International Publishing, Cham, 2020, pp. 834–844.

[9] A. Ranjan, D. Lalwani, R. Misra, GAN for synthesizing CT from T2-weighted MRI data towards MR-guided radiation treatment, Magnetic Resonance Materials in Physics, Biology and Medicine (2021).

[10] R. Farjam, H. Nagar, X.K. Zhou, D. Ouellette, S.C. Formenti, J.K. DeWyngaert, Deep learning-based synthetic CT generation for MR-only radiotherapy of prostate cancer patients with 0.35T MRI linear accelerator, Journal of Applied Clinical Medical Physics n/a (n/a).

[11] Y. Chen, C. Ying, M.M. Binkley, M.R. Juttukonda, S. Flores, R. Laforest, T.L.S. Benzinger, H. An, Deep learning-based T1-enhanced selection of linear attenuation coefficients (DL-TESLA) for PET/MR attenuation correction in dementia neuroimaging, Magnetic Resonance in Medicine n/a (n/a).

[12] K. Kläser, T. Varsavsky, P. Markiewicz, T. Vercauteren, A. Hammers, D. Atkinson, K. Thielemans, B. Hutton, M.J. Cardoso, S. Ourselin, Imitation learning for improved 3D PET/MR attenuation correction, Medical Image Analysis 71 (2021) (Jul. 2021) 102079.

[13] A. Mehranian, H. Arabi, H. Zaidi, Vision 20/20: Magnetic resonance imaging–guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities, Medical Physics 43 (3) (2016) 1130–1155.

[14] L. Liu, S. Chen, X. Zhu, X.-M. Zhao, F.-X. Wu, J. Wang, Deep convolutional neural network for accurate segmentation and quantification of white matter hyperintensities, Neurocomputing 384 (2020) (Apr. 2020) 231–242.

[15] Y. Yang, R. Xie, W. Jia, Z. Chen, Y. Yang, L. Xie, B. Jiang, Accurate and automatic tooth image segmentation model with deep convolutional neural networks and level set method, Neurocomputing 419 (2021) (2021) 108–125.

[16] H. Chen, Z. Qin, Y. Ding, L. Tian, Z. Qin, Brain tumor segmentation with deep convolutional symmetric neural network, Neurocomputing 392 (2020) (Jun. 2020) 305–313.

[17] J. Hu, Y. Song, L. Zhang, S. Bai, Z. Yi, Multi-scale attention U-net for segmenting clinical target volume in graves' ophthalmopathy, Neurocomputing 427 (2021) (Feb. 2021) 74–83.

[18] X. Liu, S. Wang, Y. Zhang, D. Liu, W. Hu, Automatic fluid segmentation in retinal optical coherence tomography images using attention based deep learning, Neurocomputing (2021).

[19] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7132–7141.

[20] A.G. Roy, N. Navab, C. Wachinger, Concurrent Spatial and Channel Squeeze & Excitation in Fully Convolutional Networks, arXiv:1803.02579 [cs]ArXiv: 1803.02579 (Jun. 2018).

[21] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, A.B. McMillan, Deep Learning MR Imaging–based Attenuation Correction for PET/MR Imaging, Radiology 286 (2) (2018) 676–684.

[22] H. Jang, F. Liu, G. Zhao, T. Bradshaw, A.B. McMillan, Technical Note: Deep learning based MRAC using rapid ultrashort echo time imaging, Medical Physics 8 (2018).

[23] J. Deng, Y. Ma, D.-A. Li, J. Zhao, Y. Liu, H. Zhang, Classification of breast density categories based on SE-Attention neural networks, Computer Methods and Programs in Biomedicine 193 (2020) (Sep. 2020) 105489.

[24] Y. Gao, H. Liang, B. Liu, Y. Wang, Action detection based on 3d convolution neural network with channel attention mechanism, in: IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2020, pp. 602–606.

[25] Y. Han, C. Wei, R. Zhou, Z. Hong, Y. Zhang, S. Yang, Combining 3d-cnn and squeeze-and-excitation networks for remote sensing sea ice image classification, Mathematical Problems in Engineering (2020).

[26] P. Liu, Q. Dou, Q. Wang, P.-A. Heng, An Encoder-Decoder Neural Network With 3D Squeeze-and-Excitation and Deep Supervision for Brain Tumor Segmentation, IEEE Access 8 (2020) 34029–34037, conference Name: IEEE Access (2020).

[27] L. Rundo, C. Han, Y. Nagano, J. Zhang, R. Hataya, C. Militello, A. Tangherloni, M. S. Nobile, C. Ferretti, D. Besozzi, M.C. Gilardi, S. Vitabile, G. Mauri, H. Nakayama, P. Cazzaniga, USE-Net: Incorporating Squeeze-and-Excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets, Neurocomputing 365 (2019) (Nov. 2019) 31–43.

[28] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert, Attention U-Net: Learning Where to Look for the Pancreas, arXiv:1804.03999 [cs]ArXiv: 1804.03999 (May 2018).

[29] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, N. Sang, Bisenet: Bilateral segmentation network for real-time semantic segmentation, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 325–341 (2018).

[30] Q. Cheng, C. Zhang, Z. Li, Z. Wang, Attention and Multi-layer Fusion for Real-time Semantic Segmentation, in: 2019 12th International Symposium on Computational Intelligence and Design (ISCID), vol. 1, 2019, pp. 142–145, iSSN: 2473–3547 (Dec. 2019).

[31] B. Hou, X. Xu, G. Kang, Y. Tang, C. Hu, Hybrid Attention Densely Connected Ensemble Framework for Lesion Segmentation From Magnetic Resonance Images, IEEE Access 8 (2020) 188564–188576, conference Name: IEEE Access (2020).

[32] H.T. Fabich, M. Benning, A.J. Sederman, D.J. Holland, Ultrashort echo time (ute) imaging using gradient pre-equalization and compressed sensing, Journal of Magnetic Resonance 245 (2014) 116–124.

[33] T. Fukuda, K. Wengler, D. Tank, S. Korbin, J.M. Paci, D.E. Komatsu, M. Paulus, M. Huang, E. Gould, M.E. Schweitzer, et al., Abbreviated quantitative ute imaging in anterior cruciate ligament reconstruction, BMC musculoskeletal disorders 20 (1) (2019) 1–14.

[34] S. Klein, M. Staring, K. Murphy, M.A. Viergever, J.P. Pluim, Elastix: a toolbox for intensity-based medical image registration, IEEE transactions on medical imaging 29 (1) (2010) 196–205.

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

[36] R.R. Kelly, S.J. Sidles, A.C. LaRue, Effects of Neurological Disorders on Bone Health, Frontiers Psychology (2020), Publisher: Frontiers.

[37] M.C. Florkow, F. Zijlstra, L.G. Kerkmeijer, M. Maspero, C.A. van den Berg, M. van Stralen, P.R. Seevinck, The impact of mri-ct registration errors on deep learning-based synthetic ct generation, in: Medical Imaging 2019: Image Processing, vol. 10949, International Society for Optics and Photonics, 2019, p. 1094938 (2019).

Imene Mecheter received her B.S degree in Computer Engineering in 2013 and her M.S. degree in Computing in 2017 from Qatar University. She is currently pursuing her Ph.D. degree in the department of Electronic and Electrical Engineering with Brunel University London, United Kingdom. She is also a software developer at Iberdrola company. Her current research interests include deep neural networks and medical image analysis.

Maysam Abbod received his Ph.D. in Control Engineering from University of Sheffield in 1992. From 1993 to 2006 he was with the Department of Automatic Control and Systems Engineering at the University of Sheffield as a research associate and senior research fellow. His main research interests are in intelligent systems for modelling, control, and optimisation. Developed systems were applied to industrials and biomedical modelling, and computer control of manufacturing systems.

Habib Zaidi is Professor and head of the PET Instrumentation & Neuroimaging Laboratory at Geneva University Hospital and faculty member at the medical school of Geneva University. He is also a Professor at the University of Groningen (Netherlands) and the University of Southern Denmark. His research is supported by the Swiss National Foundation, private foundations and industry (Total 8.3M US$) and centers on hybrid imaging instrumentation (PET/CT and PET/MRI), computational modelling and radiation dosimetry and deep learning. He was guest editor for 11 special issues of peer-reviewed journals and serves on the editorial board of leading journals in medical physics and medical imaging. He has been elevated to the grade of fellow of the IEEE, AIMBE and the AAPM. His academic accomplishments in the area of quantitative PET imaging have been well recognized by his peers since he is a recipient of many awards and distinctions among which the prestigious (100'000$) 2010 Kuwait Prize of Applied sciences (known as the Middle Eastern Nobel Prize). Prof. Zaidi has been an invited speaker of over 160 keynote lectures and talks at an international level, has authored over 330 peer-reviewed articles in prominent journals and is the editor of four textbooks.

Abbes Amira received his Ph.D. degree in 2001 from Queen's University Belfast, United Kingdom. Since then, he has taken many academic and consultancy positions in the United Kingdom, Europe, Asia, and the Middle East. He is currently the Dean of the College of Computing and Informatics at the University of Sharjah, UAE.

During his career to date, Prof. Amira has been successful in securing substantial funding from government agencies and industry; he has supervised more than 25 PhD students and has over 350 publications in top journals and conferences in the area of embedded systems, IoT, image and signal processing. He obtained many international awards, including the 2008 VARIAN prize offered by the Swiss Society of Radiobiology and Medical Physics, CAST award, DELL-EM Envision the future (2018), IET premium award (2017) and many best paper and recognition awards in IEEE international conferences and events. Prof. Amira has participated as guest editor and member of the editorial board in many international journals including recent special issues in IEEE IoT Journal and Elsevier Pattern Recognition. He has taken visiting professor positions at the University of Tun Hussein Onn, Malaysia and the University of Nancy, Henri Poincare, France. Prof. Amira has also conducted consultancy services for several government agencies and companies in the private sector. He is a Fellow of IET, Fellow of the Higher Education Academy, Senior member of the IEEE, and Senior member of ACM. His research interests include artificial intelligence, embedded systems, high-performance computing, big data and IoT, connected health, image and vision systems, biometric and security.