

Segmentation of Brain Tumors Using Three-Dimensional Convolutional Neural Network on MRI Images 3D MedImg-CNN

Ahmed Kharrat, University of Sfax, MIRACL Laboratory ISIMS, Tunisia*

 <https://orcid.org/0000-0001-8439-8877>

Mahmoud Neji, University of Sfax, MIRACL Laboratory ISIMS, Tunisia

ABSTRACT

In MR images containing glioblastomas, we regard the issue of fully automated brain tumor segmentation. We suggest a 3D MedImg-CNN (3 Dimensional Convolutional Neural Network) approach that achieves high results while being highly efficient, a combination that current techniques have struggled to achieve. Our 3D MedImg-CNN is directly formed on the raw image modalities and thus learns directly from the data a characteristic representation. We suggest a new-cascaded architecture of two pathways that each provides a model of details in tumors. Completely leveraging our model's convolutionary nature also helps us to segment in one minute a full cerebral picture. The efficiency of the suggested 3D MedImg-CNN with CNN segmentation system is determined using the dice similarity coefficient (DSC). The Experiments carried out on the 2013, 2015 and 2017 BraTS datasets, demonstrate that the proposed approach is one of the dominant in literature as it is one of the most effective. Keywords Brain Tumor, Convolutional Neural Networks, Deep Learning, Segmentation

1. INTRODUCTION

The central nervous system's management center the brain is responsible for executing all activities through the human body. A mass or growth of abnormal cells in the brain is called a brain tumor, some brain tumors are noncancerous (benign), and some brain tumors are cancerous (malignant), recently brain tumor became the second cause of deaths in young adults and children suffering from cancer. Central Brain Tumor Registry of the United States (CBTRUS) stated that there are 64,530 diagnosed new cases of central nervous system and initial stage of brain tumors is diagnosed since 2011. The number has exceeded 600,000 of people who live with the disease (Kharrat et al., 2015; Abraham et al., 2017). Assessing these tumors using Magnetic resonance imaging (MRI) is a widely used imaging technique, but it produces a large amount of data that prevents manual segmentation in a reasonable time, besides having significant variation from various experts without the global 3D brain structure (Kharrat & Néji, 2019). So, automatic and reliable segmentation methods are required, which led to the growing number of machine learning studies based on neuroimaging data that are aiming to both develop diagnostic tools that help brain MRI classification and automatic volume segmentation, and understand the mechanics of diseases, including the neurodegenerative ones. The goal of brain tumor segmentation is to detect the area of the brain based on texture from information in MRI images. Segmentation methods typically look for active tumor tissue (vascularized or not),

DOI: 10.4018/IJCINI.20211001.0a4

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

necrotic tissue and edema (swelling near a tumor) by exploiting multiple magnetic resonance imaging (MRI) modalities, such as T1, T2, T1-Contrasted (T1C) and Flair. Recently, Convolutional neural networks (CNNs) (Schmidhuber, 2015) are a type of deep artificial neural networks widely used in the field of computer vision. They have been applied to many tasks, including image classification (Kharrat & Néji, 2018; Krizhevsky et al., 2012; Paul Justin et al., 2017; Schmidhuber, 2015; Simonyan & Zisserman, 2016), super-resolution (Kim et al., 2016) and semantic segmentation (Shelhamer et al., 2017). Recent publications report their usage in medical image segmentation and classification (). For instance, a novel brain tumor segmentation method is developed by Wu Deng et al. (2019) by integrating fully convolutional neural networks (FCNN) and dense micro-block difference feature (DMDF) into a unified framework so as to obtain segmentation results with appearance and spatial consistency. Firstly, these authors proposed a local feature to describe the rotation invariant property of the texture. In order to deal with the change of rotation and scale in texture image, Fisher vector encoding method is used to analyze the texture feature, which can combine with the scale information without increasing the dimension of the local feature. The obtained local features have strong robustness to rotation and gray intensity variation. Then, the non-quantifiable local feature is fused to the FCNN to perform fine boundary segmentation. Since brain tumors occupy a small portion of the image, deconvolutional layers are designed with skip connections to obtain a high quality feature map. Jijun Tong et al. (2019) Tong introduced an automatic brain tumor segmentation method using kernel sparse coding and texture feature from Fluid Attenuated Inversion Recovery (FLAIR). Initially, MRI images are pre-processed to reduce the noise and enhance the contrast. Then, sparse coding is carried out on the first and the second order statistical eigenvector extracted from the raw MRIs. The kernel dictionary learning is used to extract the non-linear features to construct two adaptive dictionaries for healthy and pathologically tissues respectively. After that, a kernel-clustering algorithm based on dictionary learning is developed to code the voxels, then the linear discrimination method is used to classify the target pixels. In the end, the flood-fill operation is used to improve the quality of segmentation. In another research, Muhammad Sajjad et al. (2019) developed a Convolutional Neural Network (CNN) based multi-grade brain tumor classification system. Firstly, tumor regions from an MR image are segmented based on a deep learning technique. Secondly, extensive data augmentation is utilized to effectively train the proposed system. Finally, a pre-trained CNN model is fine-tuned for brain tumor grade classification. An automated brain tumor segmentation algorithm using deep convolutional neural network (DCNN) is presented by Saddam Hussain et al. (2017). A patch based approach along with an inception module is used for training the deep network by extracting two co-centric patches of different sizes from the input images. Recent developments in deep neural networks such as dropout, batch normalization, non-linear activation and inception module are used to build a new Linear nexus architecture. The module overcomes the over-fitting problem arising due to scarcity of data using dropout regularization. Images are normalized and bias field corrected in the pre-processing step and then extracted patches are passed through a DCNN, which assigns an output label to the central pixel of each patch. A two-phase weighted training method is introduced and evaluated using BRATS 2013 and BRATS 2015 datasets. A similar approach was used by Guotai Wang et al. (2018), these author proposed a novel deep learning-based framework for interactive segmentation by incorporating CNNs into a bounding box and scribble-based segmentation pipeline. They proposed image-specific fine tuning to make a CNN model adaptive to a specific test image, which can be either unsupervised or supervised. They also proposed a weighted loss function considering network and interaction-based uncertainty for the fine tuning. They applied this framework to two applications: 2-D segmentation of multiple organs from fetal magnetic resonance (MR) slices, where only two types of these organs were annotated for training and 3-D segmentation of brain tumor core (excluding edema) and whole brain tumor (including edema) from different MR sequences, where only the tumor core in one MR sequence was annotated for training. On the other hand, Kamnitsas et al. (2017) introduces a 3D CNN architecture designed for various segmentation tasks involving MR images of brains. The authors benchmark their approach on the BraTS (Menze et al., 2015) and ISLES (Maier et al., 2017) challenges.

Their approach comprises a CNN with 3D filters and a conditional random field smoothing the output of the CNN. The authors propose dividing the input images into regions in order to address the high memory demand of 3D CNNs. Notable in Kamnitsas is the usage of an architecture consisting of two pathways. The first receives the subregion of the original image that is to be segmented, while the second receives a larger region that is downsampled to a lower resolution before being fed to the network. This enables the network to still be able to learn global features of the images.

Havaei et al., (2017) presented a 2D patch-wise architecture using local and global CNN pathways, which exploits local and global contextual features around a pixel to segment brain tumors. The pre-processing follows three steps. First, the 1% highest and lowest intensities are removed. Then, they apply an N4ITK bias correction to T1 and T1C modalities. The data is then normalized within each input channel by subtracting the channel's mean and dividing by the channel's standard deviation. Finally, they explore a cascade architecture in which the output of a basic CNN is treated as an additional source of information for a subsequent CNN. Like Kamnitsas, the first CNN receives a larger portion of the original image than the second, with the purpose of learning both global context and local details.

To conclude, the variety of CNN-based medical image segmentation methods is largely due to different attempts at addressing difficulties specific to medical images. These are chiefly the memory demands of storing a high number of 3D feature maps, the scarcity of available data and the high imbalance of classes. In dealing with the first issue, most researchers have turned to dividing images into a small number of regions and stitching together the outputs of different regions (Chen et al., 2017; Kamnitsas et al., 2017; Zhen et al., 2017) and/or using downscaled images (Cicek et al., 2016). Data augmentation is often used to address the scarcity of data (Chen et al., 2017; Cicek et al., 2016; Deng et al., 2019; Hussain et al., 2017; Kamnitsas et al., 2017; Pei et al., 2017; Sajjad et al., 2019; Tong et al., 2019; Wang et al., 2018; Zhen et al., 2017), (Drozdal et al., 2016). As for class imbalance, reported methods include using weighted loss functions (Cicek et al., 2016; Zhen et al., 2017) overlap metrics such as the dice similarity (Havaei et al., 2017; Pei et al., 2017) or deep supervision (Dou et al., 2016; Tong et al., 2019).

Recent research has shown that deep learning methods have performed well on supervised machine learning and image segmentation tasks (Wang et al., 2016; Wang et al., 2017; Wang & Zatarain, 2017). The purpose of this study is to apply deep learning methods to segment brain tumor. In this paper, we propose a successful and very efficient CNN architecture for brain tumor segmentation. Two main contributions were presented: Combining multiple segmentation maps created at different scales and using element-wise summation to forward feature maps from two stages of the network.

The remainder of this paper is structured as follows. We present our proposed methodology in section 2. Section 3 is devoted to experimental setup. In Section 4, we present the results achieved and compare with other existing approaches. Conclusions are finally drawn in Section 5.

2. MATERIALS AND METHODS

The proposed methodology, shown in Figure 1, is applied on multimodal MRI sequences and exploits the inherent pattern recognition capability of CNN to classify tumor pixels. A patch based approach is used for pixel classification, where pre-processed images are passed through a CNN and post-processed to obtain a segmented image highlighting the tumor area.

Figure 1. Generic flow diagram of our proposed method.

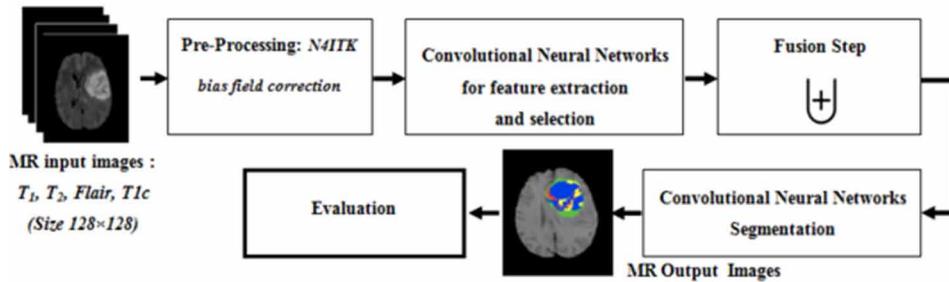
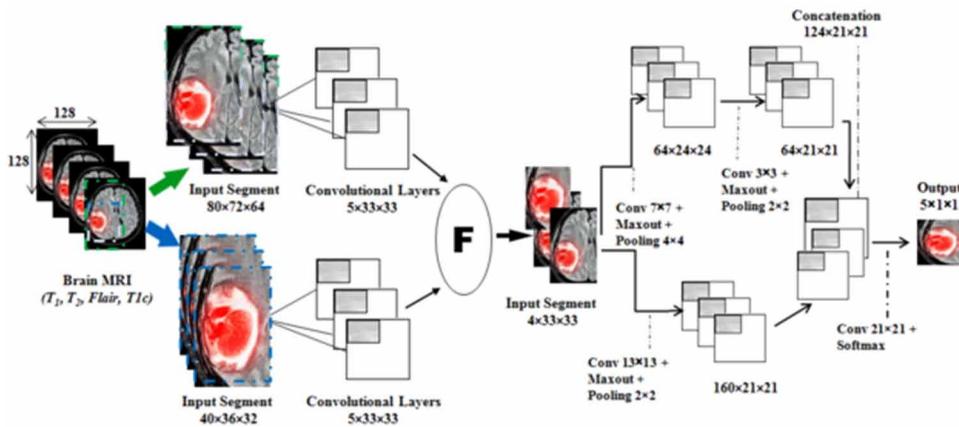


Figure 2. Architecture 3D CNN for Brain segmentation with two pathways.



The architecture proposed in Figure 2, takes as input patches of multiple modalities and predicts the class of center pixel in respective patches. The BraTS dataset (Menze et al., 2015) lacks resolution in third dimension; so, to extract 2D patches, axial view is used. In the first convolution layer, input is the extracted patches from the original MR images (size 128×128) that correspond to various anchors used. The produced feature maps are then taken as input by the cascading layers. A network of six convolution layers is implemented to learn feature maps with various kernel sizes. Rectified linear units (RLUs) activation is used for non-linear representation since it gives a better representation (Nwankpa et al., 2018; Qiu & Cai, 2017).

To reduce the input dimensionality that is going into the next layers, three max-pooling layers are used. The max-pooling layer selects the max value and discards the rest therefore summarizing the data in a small rectangle. In this way, the irrelevant information is discarded and the next convolution layer only receives the summarized important data. Pooling layers have more beneficial effects like invariance to lightning conditions and position. To further reduce the complexity, a maxout layer is used, which reduces the number of feature maps by reducing the dimensions in third axis. It is used after convolution layer and selects two adjacent feature maps at maximum; consequently the number of maps produced by convolution layer is reduced to half. This resulted in a small improvement in performance in terms of memory and computing time used. There has been a reduction in the memory used as well as in terms of clipping time.

The two-pathway convolution layer architecture produces the input for the fusion step of the network. The concatenated input is then fed to the second part of the network and the output layer i.e. softmax activation predicts class probabilities, which are accounted for in the loss function.

The choice of numbers of kernels and layers has been validated empirically on the used databases, such as BraTS 2013, 2015 and 2017. The number of layers has been fixed on an experimental basis. Several layers of convolutions were used and the best results were obtained by the architecture presented in Figure 2.

These steps are discussed in detail in the following subsections.

2.1 Preprocessing

The MR images when extracted from volumetric data have artifacts due to different acquisition techniques and systems (Collewet et al., 2004). Especially in T1 and T1c modality, the same type of tissues has different intensities across the dataset. N4ITK bias field correction is applied using 3D slicer toolkit (Fedorov et al., 2012; Tustison et al., 2010), to T1 and T1c modalities. Image normalization is performed to ensure zero mean and unit variance. Finally, patches are normalized with respect to mean and variance. Since fusion architecture is used for the neural network, two types of patches are extracted: one having $80 \times 72 \times 64$ pixels and the other having $40 \times 36 \times 32$ pixels co-centric with the fusion step.

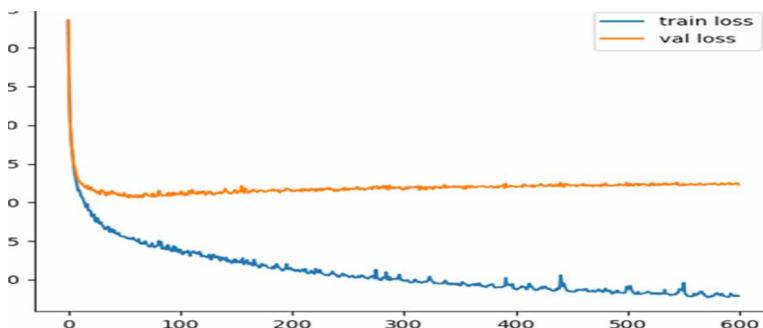
The choice of features maps and batch size is made empirically and it's limited by the quality of materials namely graphic memory, the processor and central memory. These constraints have a bad impact on the execution time as the run time can last several days without convergence.

2.2 Convolutional Neural Network (CNN) for Feature Extraction and Selection

CNN has an advantage over other classifiers as kernels used in convolution layers have same weights for all inputs, which detect same characteristics that makes them invariant by translation (Tustison et al., 2010). In stochastic optimization, gradient calculation is determined by the training distribution. Both patch wise training and fully convolutional training was made to produce any distribution. Whole medical image fully convolutional training is identical to patch wise training where each batch consists of all the receptive fields of the units below the loss for an image (or collection of images).

Usually, a non-linear activation function is used to convert features into class probabilities. Although inherently classifiers, CNNs can address segmentation tasks by throwing them to voxel-wise classification. The network processes a 3D patch around each voxel of an input image. It is trained to predict whether the central voxel is a pathology or normal brain tissue, depending on the content of the surrounding 3D patch. During training, kernel parameters are optimized using gradient descent, with the goal of minimizing the error between predictions and true labels. Figure 3. illustrates the gradient descent for error minimization.

Figure 3. The gradient descent for error minimization



One of the limitations in the above framework is that the segmentation of each voxel is processed only by computing the new representation of its neighbored patches. It is intuitive that the context is more likely to lead to better results. However, a straight-forward increase in the size of the 3D input patch would increase the memory requirement and computational burden.

Our proposed solution is to perform a parallel image processing at multiple scales. Our network architecture consists of two parallel convolutional pathways, where both have receptive fields of the same size. The entrance to the second path, however, is a patch extracted from a subsampled version of an image, thus allowing it to a larger area around each voxel. This architectural design is shown in Figure 2.

Another significant feature of our architecture is its full convolutional nature, which allows its effective application on larger parts of the image. By supplying as input segments of an image larger than the receiver field of the neurons of the final layer, the network can efficiently process the larger input and provide as output predictions for several neighboring voxels. As a result of (Dou et al., 2016; Drozdal et al., 2016), we also use this feature during training, building our training batches by extracting image segments larger in size than the network's receptive field.

2.3 Fusion Step

An earlier version of our system 3D MedImg-CNN is shown in Figure 2. Our method uses a two-pathway architecture, in which each pathway is responsible for learning about either the local details or the broader context of tissue appearances (for example, whether or not he is close to the skull). The tracks are connected by their concatenate feature cards immediately before the output layer. The fusion step used the result of the two pathway convolution layer. The result of step fusion is a patch size $4 \times 33 \times 33$ features. This patch is an input of the two pathway of the final step.

2.4 Convolutional Neural Network (CNN) for Classification and Segmentation

Finally, a prediction of the class label is made by stacking a final output layer, which is totally convolutional to the last convolutional hidden layer. The number of feature cards in this layer corresponds to the number of class tags and uses the so-called non-linearity softmax.

3D CNN's perform pixel classification without taking into account the local dependencies of labels, one can model the dependencies of labels by taking into account estimates of wise pixel probability of a first CNN as an additional entry to a second 3D CNN, in formation of a new cascade architecture. Our final network is composed of two parallel pathways. The first pathway uses two layers with different sizes and number of filters 80×72 and 40×36 in order to include different strokes and edges at the first level of convolution. After these two multi-resolution convolution, we obtain the two pathways which will be processed before and after the fusion step. The final layer deep network exhibits significantly more accurate segmentation performance. Initial learning rate is set at 0.01 and is gradually reduced during training, and also constant impulse equal to 0.8. The training time requires for final system convergence is about one day using an NVIDIA Tesla K10 GPU with 8 GB of memory. Segmentation 3D brain tumor with four modalities requires 16 seconds.

3. EXPERIMENTAL SETUP

3.1 Dataset

For testing and evaluating of our proposed system 3D MedImg-CNN, we use the main objective of the annual BraTS challenge: segmenting tumor regions in brain MRI. More concretely, the network is trained using the BraTS 2013, 2015 and 2017 training sets (Bakas et al., 2017). They contain four modalities i.e., T1, T1-Contrasted (T1C), T2 and Flair. BraTS 2013 comprises of 30 training images (20 brains with high-grade (HG) and 10 brains with low-grade (LG) tumors) and 10 brains with high-grade tumors for testing. The data are rather sparse and preprocessing steps like skull stripping

have been performed to improve the data representation. In 2013 dataset, two more data subsets are provided i.e. leaderboard and challenge data. These two subsets comprise of 65 MR images. Manual segmentation is available for training data only. BraTS 2015 contains a total of 274 images, 220 were classified as high-grade gliomas (HG), while the remaining 54 were classified as low-grade gliomas (LG), with no inclusion of images depicting healthy brains. The classes involved in the segmentation task are: (1) necrosis, (2) edema, (3) non-enhancing and (4) enhancing tumor. Whilst the BraTS 2017 obtained significant difference to the previous year's training datasets, as the annotation was manually interpolated by one to four raters and all the segmentations were approved by experts. The annotated training dataset consists of 210 high-grade gliomas (HGG) and 75 low-grade gliomas (LGG) samples (Bakas et al., 2017). For each patient, there are four different modalities for MRI data as described: T1 weighted, post-contrast T1-weighted, T2-weighted and FLAIR. The MRI were acquired from 19 institutions using various protocols and scanners with different field strengths. For each tumor, there are four kinds of labels: edema (label 2), necrosis and non-enhancing tumor (label 1) and active/enhancing tumor (label 4).

All of the images have a size of $240 \times 240 \times 155$, which can be cropped to a region of the size $160 \times 144 \times 128$, while still containing the entire brain. For some of the experiments, these cropped images were further down sampled to a size of $80 \times 72 \times 64$ for training and test.

Since our GPU memory space is limited to 8 GB of memory, we divided the images in two in terms of size and based on the experience and the ground truth this step did not influence the final results.

3.2 Implementation Details

The algorithm is implemented in Theano with CUDA/GPU and CuDNN acceleration library in python. Hyper-parameters are tuned using grid search and the parameters on which model performed best on validation data are selected. Parameters such as learning rate and momentum are varied during training. Momentum is initially set to 0.6 and is gradually increased to 0.8. Learning rate, on the other hand, is initially set to 0.01 and then is gradually decreased to 0.1×10^{-3} . Hence, a dropout value of 0.5 is used in the network to avoid over-fitting.

Segmenting brain tumor is an unbalanced classification problem where, most of the pixels are of healthy tissues.

We propose in a second step an unsupervised classification phase that precedes the segmentation phase such as Deep Belief Network (DBN) (Latha & Kavitha, 2018) in order to eliminate normal images that do not include tumors.

Experiments have been performed on BraTS 2013 dataset which has two types of tumors, HG and LG glioma, divided into four tumor classes. There are 30 volumetric images in 2013 dataset containing slices varying in the range of 150 to 220. The dataset is divided randomly into training and testing sets with 80:20 ratios. The dataset also contains synthetic data with low variance in intensity values of a similar class that is comparatively easy to classify. Therefore, only real patient data are used for evaluating the model. Evaluation metrics are determined for three tumor regions namely a) the complete tumor area (all four tumor labels), b) the core tumor area, and c) the enhancing tumor region.

The following describes the general training set-up used throughout the experiments: Of the 274 MRI volumes in the BraTS 2015 data set, 220 are used for training, while the remaining 54 are reserved for the validation and test sets, with 27 images each. BraTS 2017 dataset, consists of 285 training and 46 validation and test sets cases. The following validation (BraTS 2017) and test (BraTS 2015 and 2017) set were predicted via the five networks obtained by the fivefold cross-validation as an ensemble. The Adam optimizer is used to optimize the network settings. Training takes place on an NVIDIA Tesla k10 for 600 epochs (around 120 hours). The trained network takes roughly 4 seconds to segment one $80 \times 72 \times 64$ sized image.

3.3 Evaluation Parameters

The experimental results are evaluated based on one metric, namely dice similarity coefficient (DSC). Dice score is calculated by overlapping predicted labels with actual labels and the intersection of two contributors determine the dice score. Dice score is calculated for three categories i.e. the whole tumor, enhancing tumor, and core tumor and is given by,

$$DSC = \frac{2 \times (|L \cap P|)}{|L| + |P|} \quad (1)$$

Where L and P stand for actual labels for tumor region and predicted tumor regions respectively.

4. RESULTS AND DISCUSSION

A comparative analysis is presented in Table 1 to evaluate the effectiveness of the proposed model 3D MedImg-CNN. In fact, Table 1 details the results obtained in our proposed model. It was found that, when each modality is used on its own, T1C produces the best results on every class but edema, for which Flair, followed by T2, produces better predictions. Using this as a starting point, the network was then trained on Flair and T1C images together, which lead to dice scores very close to the ones achieved by a network that had access to all four modalities as input channels. Since T2 is the next best-performing modality, the network was then trained on images in the Flair, T1C and T2 modalities. Comparing the results of this run with the performance of the network when all modalities are available shows that, interestingly, the network achieves a much higher dice score on necrotic regions when the T1 modality is discarded. This suggests that some benefit could be achieved by training different networks on different combinations of MRI modalities.

Table 1. Segmentation results on BraTS 2015 training data.

Type	DSC			
	Necrosis (1)	Edema (2)	Enhancing (3)	Non-enhancing (4)
Flair	0.21	0.82	0.31	0.39
T1	0.35	0.53	0.26	0.48
T1C	0.56	0.67	0.52	0.83
T2	0.38	0.71	0.31	0.46

Table 2. Segmentation results on BraTS 2013 training data compared with the state-of-the-art methods.

Method	DSC		
	Complete tumor area	Core tumor area	Enhancing tumor region
3D MedImg-CNN	0.92	0.83	0.86
Havaei et al. (2017)	0.88	0.79	0.73
Hussain et al. (2017)	0.80	0.67	0.85
Fidon et al. (2017)	0.88	0.77	0.72
Kayalibay et al. (2017)	0.87	0.78	0.71

Figure 4. Comparison of our trained network with the state-of-the-art methods on BraTS 2013

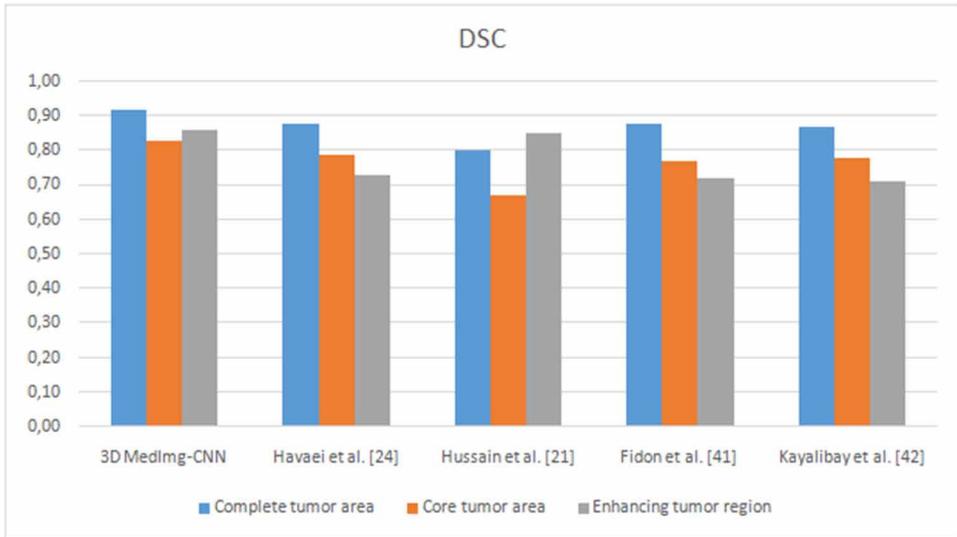


Table 3. Segmentation results on BraTS 2015 training data compared with the state-of-the-art method.

Method	DSC			
	Necrosis	Edema	Enhancing	Non-enhancing
3D MedIm-CNN	0.52	0.88	0.68	0.9
Kayalibay et al. (2017)	0.49	0.84	0.5	0.8

Figure 5. Comparison of our trained network with the state-of-the-art methods on BraTS 2015.

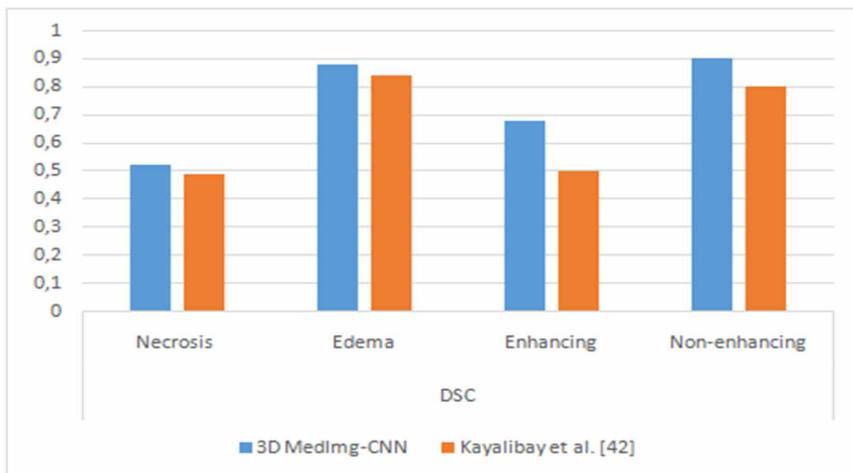


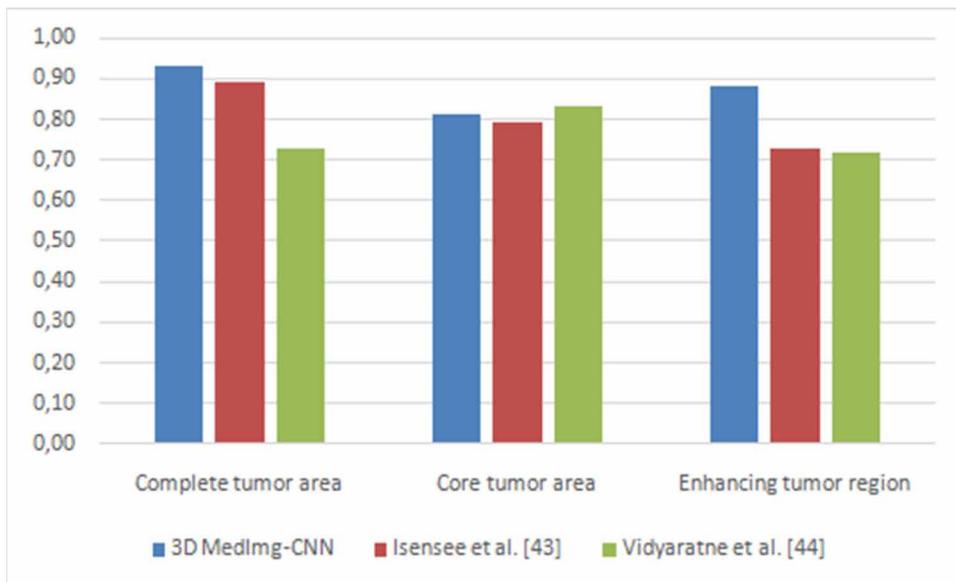
Table 4. Segmentation results on BraTS 2017 dataset.

Dataset	DSC		
	Complete tumor area	Core tumor area	Enhancing tumor region
BraTS 2017 Training data	0.91	0.79	0.89
BraTS 2017 Test data	0.93	0.81	0.88

Table 5. Segmentation results on BraTS 2017 test data compared with the state-of-the-art methods.

Method	DSC		
	Complete tumor area	Core tumor area	Enhancing tumor region
3D MedImg-CNN	0.93	0.81	0.88
Isensee et al. (2017)	0.89	0.79	0.73
Vidyaratne et al. (2018)	0.73	0.83	0.72

Figure 6. Comparison of our tested network with the state-of-the-art methods on BraTS 2017.



The Table 2 and 3 show the comparison between our proposed model 3D MedImg-CNN and the currently published state-of-the-art methods from BraTS 2013 and BraTS 2015 respectively (see Figure 4 and Figure 5). It is evident from these Tables that our implemented architecture outperforms state-of-the-art methods (Havaei et al., 2017; Hussain et al., 2017), (Fidon et al., 2017; Kayalibay et al., 2017) in terms of dice similarity coefficient (DSC).

Results from the BraTS 2017 dataset segmentation performance presented in Table 4 shows that our network 3D MedImg-CNN is capable of accurately segmenting large tumor regions as well as fine grained details. Let us mention how the thin wall of the enhancing region in the uppermost part

of the tumor as mentioned in Figure 6 was segmented with voxel-level accuracy, whilst the manual ground truth label was spilled into the bordering edema region. Moreover, the small spot of enhancing tumor which is surrounded by edema in the ground truth segmentation can, upon closer inspection of the raw data, be identified as a blood vessel that has been erroneously included in the enhancing tumor region by the annotator. The former non-enhancing tumor label, which was integrated into the necrosis label for the BraTS 2017 challenge, is often not well defined in the training data. As a result, our algorithm learns an accurate prediction in segmentation task of this label based on the context rather than on image evidence, which can sometimes be related to where to place it.

Quantitatively, using the training and validation data of the BraTS 2017 database, the Dice coefficients for complete, core and enhancing tumor segmentation are of 0.93, 0.81 and 0.88 respectively. As seen from this result, our methods outperform other approaches on this dataset according to the state-of-the-art methods (Table 5).

Qualitative segmentation results using the trained neural networks are shown in Figure 7 and Figure 8.

Figure 7. Model outputs for brain MRI, depicted alongside the ground truth. Colors correspond to: necrosis (green), non-enhanced (red) and enhanced tumor (orange) and edema (yellow).

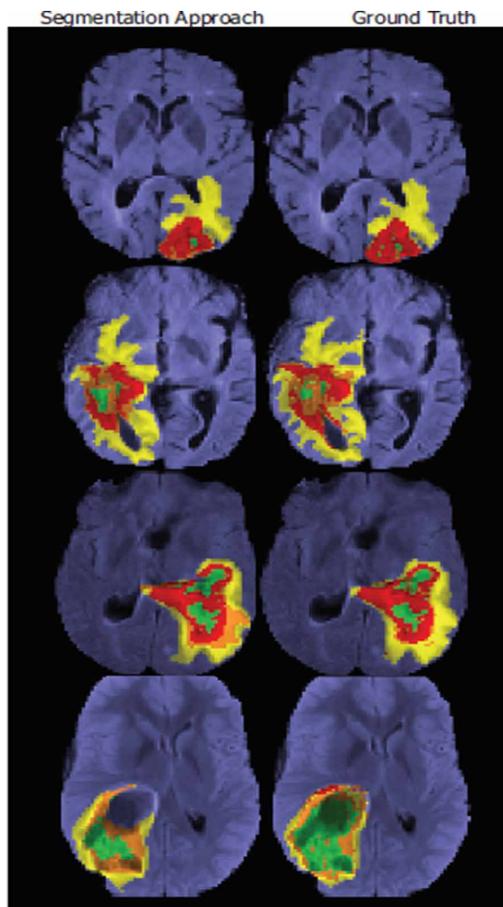
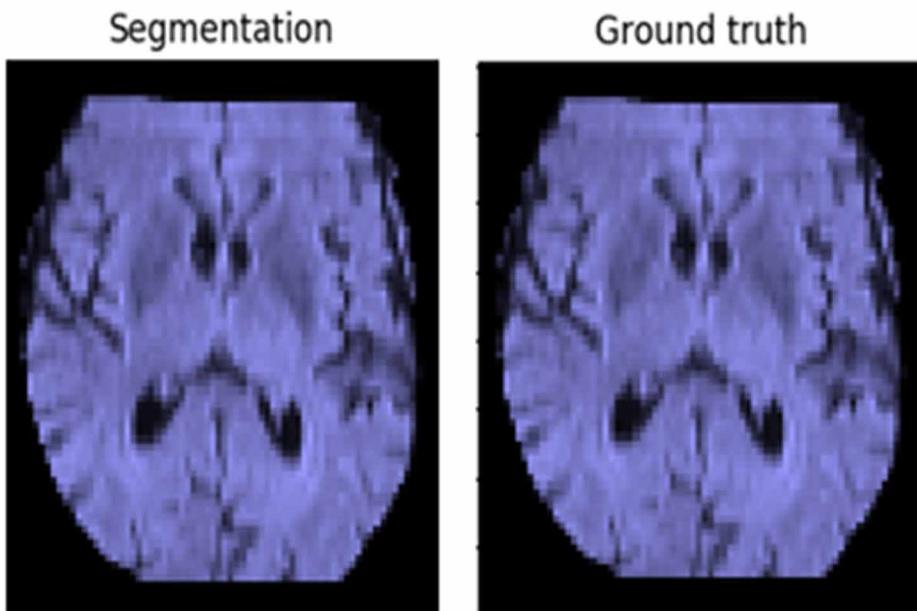


Figure 7 shows visual segmentation produced by our model from BraTS 2013, BraTS 2015 and BraTS 2017 respectively. The larger receiver field in the two-path process allows the model to have more contextual information about the tumor and thus provides better segmentations. In addition, with its two pathways, the model is flexible enough to recognize the fine details of the tumor rather than making a very smooth segmentation as in a one trajectory process. By allowing a second phase of training and learning by the true class distribution, the model corrects most of the classification errors produced in the first phase.

Figure 8. Detection of normal Brain, depicted alongside the ground truth.



Based on experiments results shown in Figure 8, we conclude that our system can easily detect the normal brain without tumor compared with ground truth image. The dice similarity coefficient (DSC) value is 0.98 similar in the ground truth image.

The proposed algorithm 3D MedImg-CNN performs well in specifying tumor region as is evident from the lack of false positives in detections. It also detects enhancing tumor better than most state-of-the-art techniques and gives comparable results on other metrics.

It is observed that the trained model faces difficulty, predicting minority classes. But by increasing training data, this problem can be dealt with. Two variations of the proposed architecture are also tested where a number of features in the fully connected layer is varied. It has been observed that too many features in the fully connected layers lead to over-fitting and if features are reduced enormously, the model does not learn significantly leading to under fitting. It is also observed that fully connected layers are time-consuming compared to convolutions and thus a trade-off between segmentation time and accuracy is achieved in the fully connected layer.

5. CONCLUSION

Brain tumor segmentation has a very important role in diagnostic procedures. With accurate segmentation, clinical diagnostic not only becomes easy, but also the chances of subject's survival increase tremendously. In this paper, a 3D MedImg-CNN architecture for brain tumor segmentation is presented. This algorithm incorporates both global and local features since context is important when it comes to tumor segmentation task. The use of max-pooling, max-out, and drop-out complement the learning process, improving training and testing speed by reducing features in the fully connected layer as well as reducing a number of parameters, which in turn reduce the chances of over-fitting. Evaluation results show that the proposed network architecture is promising and performs particularly well in detecting an enhancing tumor as well as specifying tumor to actual tumor region only. This approach can be applied to the segmentation of any type of region of interest (ROI) in medical images.

Our architecture illustrates promising performance, with capabilities for delicate segmentations. The difficulties are observed in the segmentation of the lesions of particularly low size. The separation of the lesions in various categories, for example according to their size and their treatment by different classifiers could simplify the task for every learner and help to limit the problem.

ACKNOWLEDGMENT

The authors would like to thank Xin Yao (<http://www.cs.bham.ac.uk/xin/>) for his validation of our architecture: **3D MedImg-CNN** for Brain segmentation with two pathways and all his pertinent remarks during the Workshop Internet of Things (IoT) Workshop (<http://devfest.ieee.tn/>).

REFERENCES

- Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., & Davatzikos, C. (2017). *Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection*. The Cancer Imaging Archive. doi:10.7937/K9/TCIA.2017.KLXWJJ1Q
- Chen, H., Dou, Q., Lequan, Y., Qin, J., & Heng Pheng, A. (2017). VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. *NeuroImage*, 161, 135–146. PMID:28445774
- Cicek, O., Abdulkadir, A., Lienkamp, S., Brox, T., & Ronneberger, O. (2016). *3d u-net: Learning dense volumetric segmentation from sparse annotation*. CoRR, abs/1606.06650
- Collewet, G., Strzelecki, M., & Mariette, F. (2004). Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magnetic Resonance Imaging*, 22(1), 81–91. doi:10.1016/j.mri.2003.09.001 PMID:14972397
- Deng, W., Shi, Q., Luo, K., Yang, Y., & Ning, N. (2019). Brain Tumor Segmentation Based on Improved Convolutional Neural Network in Combination with Non-quantifiable Local Texture Feature. *Journal of Medical Systems*, 43(6), 152. doi:10.1007/s10916-019-1289-2 PMID:31016467
- Dou, Q., Chen, H., Jin, Y., Yu, L., Qin, J., & Heng, P.A. (2016). *3D deeply supervised network for automatic liver segmentation from CT volumes*. CoRR, abs/1607.00582
- Drozdal, M., Vorontsov, E., Chartrand, G., Cadoury, S., & Pal, C. (2016). *The importance of skip connections in biomedical image segmentation*. CoRR, abs/1608.04117
- Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J. V., Pieper, S., & Kikinis, R. (2012). 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*, 30(9), 1323–1341. doi:10.1016/j.mri.2012.05.001 PMID:22770690
- Fidon, L., Li, W., Garcia-Peraza-Herrera, L. C., Ekanayake, J., Kitchen, N., Ourselin, S., & Vercauteren, T. (2017). Scalable multimodal convolutional networks for brain tumour segmentation. *Proc. MICCAI*, 285-293. doi:10.1007/978-3-319-66179-7_33
- Gao Xiaohong, W., Hui, R., & Tian, Z. (2017). Classification of CT brain images based on deep learning networks. *Computer Methods and Programs in Biomedicine*, 138, 49–56. doi:10.1016/j.cmpb.2016.10.007 PMID:27886714
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P., & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35, 18–31. doi:10.1016/j.media.2016.05.004 PMID:27310171
- Hussain, S., Anwar, S. M., & Majid, M. (2017). Brain tumor segmentation using cascaded deep convolutional neural network. *Engineering in Medicine and Biology Society (EMBC), 39th Annual International Conference of the IEEE*, 1998-2001.
- Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., & Maier-Hein, K. H. (2017). Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. In *International MICCAI Brainlesion Workshop*. Springer.
- Kamnitsas, K., Ledig, C., Newcombe, V. F., Joanna Simpson, P., Andrew Kane, D., David Menon, K., Daniel, R., & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36, 61–78. doi:10.1016/j.media.2016.10.004 PMID:27865153
- Kayalibay, B., Jensen, G., & van der Smagt, P. (2017). *CNN-Based segmentation of medical imaging data*. CoRR, abs/1701.03056
- Kharrat, A., Ben Halima, M., & Ben Ayed, M. (2015). MRI brain tumor classification using Support Vector Machines and meta-heuristic method. *Proceedings of the 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 446-451. doi:10.1109/ISDA.2015.7489271

- Kharrat, A., & Néji, M. (2017). A Hybrid Feature Selection for MRI Brain Tumor Classification. In A. Abraham, A. Haqiq, A. Muda, & N. Gandhi (Eds.), *Innovations in Bio-Inspired Computing and Applications (IBICA). Advances in Intelligent Systems and Computing* (Vol. 735, pp. 329–338). Springer. doi:10.1007/978-3-319-76354-5_30
- Kharrat, A., & Néji, M. (2018). Classification of brain tumors using personalized deep belief networks on MRImages: PDBN-MRI. *Proc. SPIE 11041, Eleventh International Conference on Machine Vision (ICMV), 110412M.*, 1-9. doi:10.1117/12.2522848
- Kharrat, A., & Néji, M. (2019). Feature selection based on hybrid optimization for magnetic resonance imaging brain tumor classification and segmentation. *Applied Medical Informatics*, 41(1), 9–23.
- Kim, J., Lee, J. K., & Lee, K. M. (2016). *Accurate image super-resolution using very deep convolutional networks*. CoRR, abs/1511.04587
- Korolev, S., Safiullin, A., Belyaev, M., & Dodonova, Y. (2017). *Residual and Plain Convolutional Neural Networks for 3d Brain MRI classification*. CoRR, abs/1701.06643
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems*, 1097-1105.
- Latha, M., & Kavitha, G. (2018). Detection of Schizophrenia in brain MR images based on segmented ventricle region and deep belief networks. *Neural Computing & Applications*, 30, 1–12.
- Maier, O., Menze, B. H., von der Gablentz, J., Häni, L., Heinrich, M. P., Liebrand, M., Winzeck, S., Basit, A., Bentley, P., Chen, L., Christiaens, D., Dutil, F., Egger, K., Feng, C., Glocker, B., Götz, M., Haeck, T., Halme, H.-L., Havaei, M., & Reyes, M. et al. (2017). ISLES 2015 - A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Medical Image Analysis*, 35, 250–269. doi:10.1016/j.media.2016.07.009 PMID:27475911
- Menze, B., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.-A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., & Van Leemput, K. et al. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS). *IEEE Transactions on Medical Imaging*, 34(10), 1993–2024. doi:10.1109/TMI.2014.2377694 PMID:25494501
- Novikov, A., Major, D., Lenis, D., Hladuvka, J., Wimmer, M., & Katja, B. (2017). *Fully convolutional architectures for multi-class segmentation in chest radiographs*. CoRR, abs/1701.08816
- Nwankpa, C. E., Ijomah, W., Gachagan, A., & Marshall, S. (2018). *Activation Functions: Comparison of Trends in Practice and Research for Deep Learning*. arXiv:1811.03378.
- Paul Justin, S., Plassard Andrew, J., Landman Bennett, A., & Daniel, F. (2017). Deep learning for brain tumor classification. *Proc. SPIE 10137, Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 129-138.
- Pei, Y., Qin, H., Ma, G., Guo, Y., Chen, G., Xu, T., & Zha, H. (2017). Multi-scale Volumetric ConvNet with Nested Residual Connections for Segmentation of Anterior Cranial Base. In *International Workshop on Machine Learning in Medical Imaging*. Springer. doi:10.1007/978-3-319-67389-9_15
- Qiu, S., & Cai, B. (2017). *Flexible rectified linear units for improving convolutional neural networks*. arXiv preprint arXiv:1706.08098.
- Sainath, T. N., Mohamed, A. R., Kingsbury, B., & Ramabhadran, B. (2013). Deep convolutional neural networks for LVCSR. Acoustics, speech and signal processing (ICASSP), IEEE international conference on, 8614-8618.
- Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., & Baik, S. W. (2019). Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *Journal of Computational Science*, 30, 174–182. doi:10.1016/j.jocs.2018.12.003
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61(1), 85–117. doi:10.1016/j.neunet.2014.09.003 PMID:25462637

Shelhamer, E., Long, J., & Darrell, T. (2017). Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 640–651. doi:10.1109/TPAMI.2016.2572683 PMID:27244717

Simonyan, K., & Zisserman, A. (2016). *Very deep convolutional networks for large-scale image recognition*. CoRR, abs/1409.1556

Tong, J., Zhao, Y., Zhang, P., Chen, L., & Jiang, L. (2019). MRI brain tumor segmentation based on texture features and kernel sparse coding. *Biomedical Signal Processing and Control*, 47, 387–392. doi:10.1016/j.bspc.2018.06.001

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., & Yushkevich, P. A. et al.. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6), 1310–1320. doi:10.1109/TMI.2010.2046908 PMID:20378467

Vidyaratne, L., Alam, M., Shboul, Z., & Iftekharuddin, K. (2018). Deep learning and texture-based semantic label fusion for brain tumor segmentation. *Proc SPIE Int Soc Opt Eng.* , 1-16. doi:10.1117/12.2292930

Wang, G., Li, W., Zuluaga, M. A., Pratt, R., Patel, P. A., Aertsen, M., Doel, T., David, A. L., Deprest, J., Ourselin, S., & Vercauteren, T. (2018). Interactive Medical Image Segmentation Using Deep Learning with Image-Specific Fine Tuning. *IEEE Transactions on Medical Imaging*, 37(7), 1562–1573. doi:10.1109/TMI.2018.2791721 PMID:29969407

Wang, Y., Widrow, B., Zadeh, L. A., Howard, N., Wood, S., Bhavsar, V. C., Budin, G., Chan, C., Fiorini, R. A., Gavrilova, M. L., & Shell, D. F. (2016). Cognitive intelligence: Deep learning, thinking, and reasoning by brain-inspired systems. *International Journal of Cognitive Informatics and Natural Intelligence*, 10(4), 1–20. doi:10.4018/IJCINI.2016100101

Wang, Y., Zadeh, L. A., Widrow, B., Howard, N., Beaufays, F., Baciú, G., & Raskin, V. (2017). Abstract Intelligence: Embodying and Enabling Cognitive Systems by Mathematical Engineering. *International Journal of Cognitive Informatics and Natural Intelligence*, 11(1), 1–15. doi:10.4018/IJCINI.2017010101

Wang, Y., & Zatarain, O. A. (2017). A Novel Machine Learning Algorithm for Cognitive Concept Elicitation by Cognitive Robots. *International Journal of Cognitive Informatics and Natural Intelligence*, 11(3), 31–46. doi:10.4018/IJCINI.2017070103

Zhen, X., Chen, J., Zhong, Z., Hrycushko, B., Zhou, L., Jiang, S., Albuquerque, K., & Gu, X. (2017). Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: A feasibility study. *Physics in Medicine and Biology*, 62(21), 8246–8263. doi:10.1088/1361-6560/aa8d09 PMID:28914611

Ahmed Kharrat is an assistant Professor at Higher Institute of Computer Science and Multimedia of Sfax (Tunisia). He received the Bachelor Degree in 2001 in Computer Science. He completed his Master degree and PhD degree in Computer Sciences from University of Sfax. He is currently a member of the Research Multimedia, Information systems and Advanced Computing Laboratory (MIRACL). Her main research interests include Cognitive Informatics and Image Processing.