Universitat de Girona

Medical Image Segmentation and Applications

Final Project Report

Tissue Segmentation Project

Student:

Ahmed Gouda

Supervisors:

Prof. Arnau Oliver
Prof. Xavier Lladó

1 Introduction

Brain tissues segmentation using any medical diagnostic tool can be challenging due to the varying nature of the used pathology images. This project is based on using a database from MICCAI Grand Challenge of MR Brain Image Segmentation[1]. This challenge aims to apply automatic segmentation of T1 brain image into Cerebrospinal fluid (CSF), Gray matter (GM) and White matter (WM). The provided database has some major challenges which are addressed in the following points:

- The database is generated from different MRI scanners. Consequently, the images of the database are not consequent, for instance they have a different size, shape, noise and contrast intensity.
- From analyzing the data visually, there are some regions of the brain are not labeled, like the cerebral outer layer and some regions towards the spinal cords, as shown in figure 1.

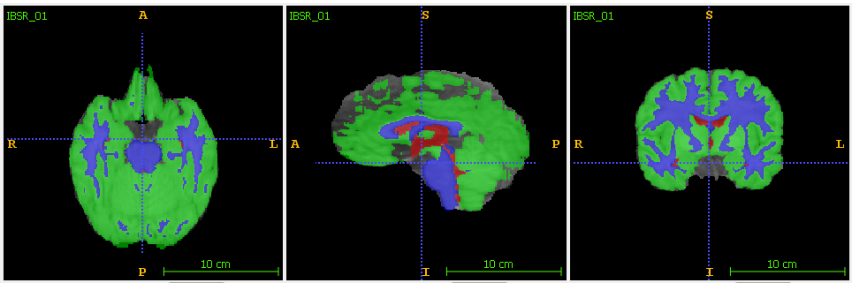

Figure 1: Some unlabeled brian regions

- The number of the training images of the provided training set are too few, while the Deep Learning approach requires a huge training database to achieve high accuracy.
- Applying 3D volume segmentation for the whole brain MRI image is computationally expensive, especially when this image is with high resolution.

2 New Labels Generation

The non labeled regions is considered one of the main problems of the Mr Brain database. Since these regions are not classified as brain tissue, it is supposed

to be classified as a background. In fact these unlabeled regions are have a similar range of intensities of labeled regions. Thus, the can be classified as CSF, WM or GM. Moreover, there areas size are very small and not balanced with the background region. In order to solve this problem, these regions are included as a fifth label to make the training model capable to classify these regions only, as shown in figure 2.

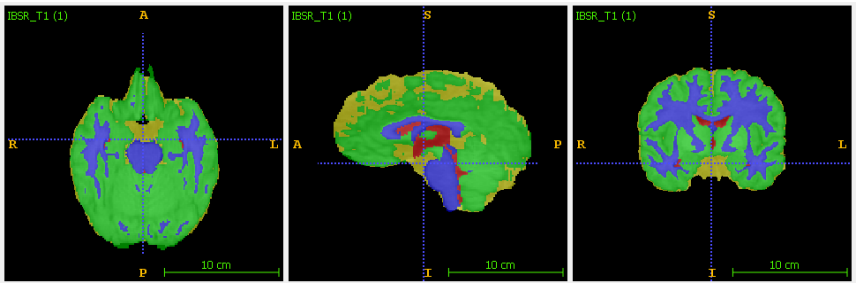

Figure 2: The new generated labels

3 Pre-processing

The pre-processing is the first stage in the implemented algorithm, and is is divided into two sub stages, which are intensity normalization and 3D image blocking. Since the given database in this project are generated from different scanners, intensity normalization is very important stage to normalize the variation in intensity values between the given images. There are two different normalization methods are used. The first method is mean and standard deviation normalization to the region of interest(ROI) of the brain as shown in the following equations.

$$I_{zm} = I_{in} - mean(roi(I_{in})) \tag{1}$$

$$I_{out} = \frac{I_{zm}}{std(roi(I_{in}))} \tag{2}$$

The second normalization method is re-scaling all the intensities values of the image between 0 and one according to the following equation.

$$\begin{cases} I_{out} = \frac{I_{in}+min(I_{in})}{max(I_{in})-min(I_{in})} & if \quad min(I_{in}) < 0 \\ I_{out} = \frac{I_{in}-min(I_{in})}{max(I_{in})} & otherwise \end{cases} \tag{3}$$

In the second prepossessing stage, the input 3D image is into divided into smaller overlapped 3D voxels. This blocking technique address the challenges

of GPU memory restriction and limited training data, and it makes the model capable to train small 3D volume without GPU memory restriction problems.

4 U-Net Architecture

Recently, Convolutional Neural Networks(CNN) methods have been widely developed to applied to whole brain segmentation. The straightforward strategy of performing whole brain segmentation is to fit all brain volume to a 3D CNN based segmentation network. In this project, the traditional 3D U-Net was utilized for multi-class segmentation. The U-net architecture is divided into two main parts which are encoder and decoder, as shown in figure 3. In the encoder part, the input image is down-sampled into the latent space using convolution and max pooling layers. It generates a compressed version of the input image with a lot less information. U-NET model elements are copied and cropped from the encoder to the decoder. In the decoder part, the compressed image is up-sampled and concatenated to the same level decoding layers in a symmetric path of the encoder layers in order to reconstruct the image structure prior to segmentation labels.

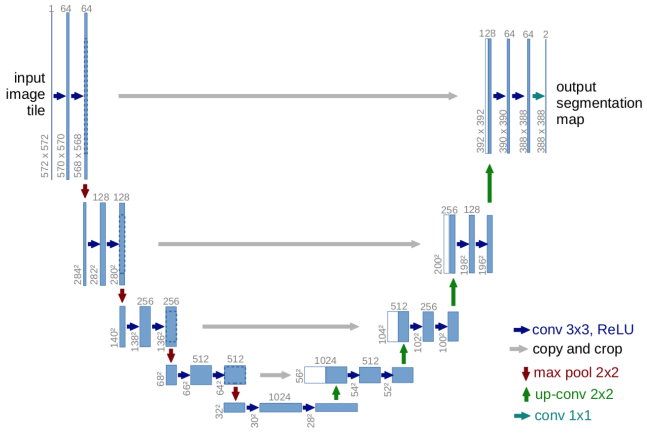

Figure 3: Basic U-Net architecture

In our case, the 3D U-Net is applied to segment different classes, which are the background class and the three brain tissue classes in the 4 labeled images, and additional fifth labels in the case of five labels images. Batch Normalization and Dropout techniques are applied by to the UNet architecture in order to make it more robust and to decrease the training over-fitting. These techniques are explained in this section.

4.1 Batch Normalization

Bach Normalization is considered one of the most important techniques in Deep Learning. It makes the hyper-parameter search problem much easier. Therefore, it speeds up the learning operation, and it makes the network much more robust. In this technique, the input of each layer is normalized by adjusting and scaling its values after applying activation function in the previous layer.

In the Unet architecture, the Batch normalization is applied in both of encoder and decoder parts, after the Relu activation layer. Figures 4 and 5 shows the effect of applying Batch Normalization using the same network hyper-parameters. As shown in figure 5, applying batch normalization increase the learning speed of the the parameters and provides more stable learning accuracy along epochs in comparison to figure 4.

Figure 4: Loss and accuracy values without applying Batch Normalization and Dropout

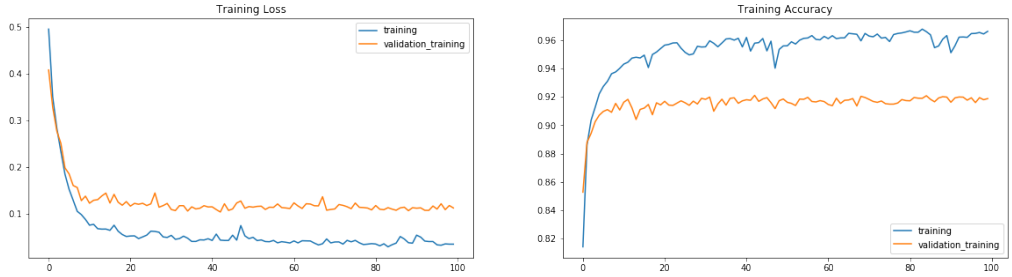

Figure 5: Loss and accuracy values without applying Batch Normalization only

4.2 Dropout

Over-fitting is considered another neural network issue. It occurs when the accuracy increases with the training data, while the accuracy of the validation set decreases or does not change, as shown in figure 5. In this condition, the model fails to reach the generalisation condition. Therefore, Dropout is considered one of effective regularization techniques to reduce training over-fitting by randomly dropping out nodes during training. This technique very computationally cheap and remarkably effective method to reduce over-fitting and enhance generalization error in for different deep neural networks architectures.

In the U-Net encoder part, Dropout is applied after each max pooling layer, while in the decoder part it is applied before each activation layer. Figure 6 shows the effect of applying dropout using the same network hyperparameters.The gap between the training and the validation accuracies is significantly decrease by comparing it to the accuracies in figure 5.

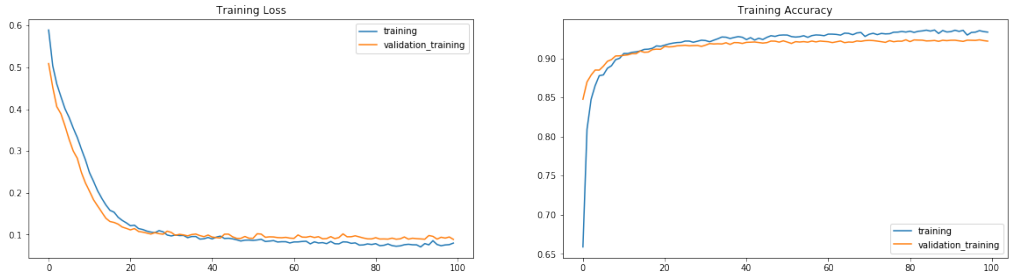

Figure 6: Loss and accuracy values after applying Batch Normalization and Dropout

5 Loss Functions

Our main aim in the Deep Learning is to minimize the error between the predicted labels and the real labels. This error optimization problem can be represented by an objective function which is often referred as a cost function or a loss function. In the provided code, the cross entropy loss function is used. One of the main drawbacks of this loss function that the network is highly sensitive to classes who are appearing more frequently than the class which appears infrequently. Since, we have a class imbalance problem between different class of brain tissue and the background (for example the number of pixels of each tissue and the background are not equal for each 3D MRI image). Therefore, the multi dice loss function is more robust solution in this case as

it calculates the overlab between each class and the corresponding label which is rescaled from 0 to 1, according to the following formula.

$$DC_i = \frac{2\,|A_i \bigcap B_i|}{|A_i| + |B_i|} \quad i = 1, ., ., N \tag{4}$$

$$DiceLoss = 1 - \frac{\sum_N^{i=1} DC_i}{N} \tag{5}$$

6 Data Augmentation

Data augmentation is one of strategies that is based on increasing the number of training images of the available training database, without including new data. It requires applying some transformation and image processing techniques to generate a new training data from the training set and include while the training stage. This technique increases the robustness of the training model and it contributes to achieve generalisation state.
In this project, a new database generated from the training set by applying different zooming factors and cropping the boundaries to the original image size. Then, Gamma correction is applies with random Gamma values greater and lower than 1, in order to change the intensity distribution of the image as shown in figure 7. For the associated label images, only zooming and cropping operation is applied.

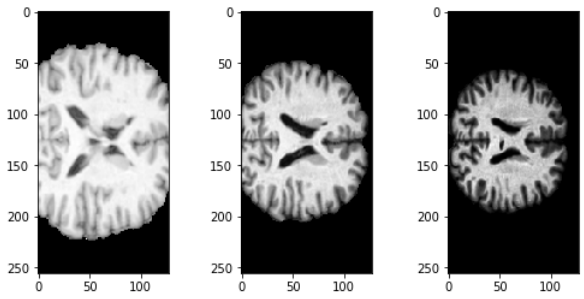

Figure 7: Data augmentation with different zooming and Gamma correction operations

7 Experimental Works and Results

For the experimental works, I modified the medical imaging segmentation python notebook wich is developed by Sergi[2]. I added batch normalization and dropout for the existing U-Net architecture. In addition, I rebuilt a

deeper U-Net model with four encoder and decoder layers. Furthermore, another 3D segmentation model based on U-Net architecture is which is model called "Modified UNet"[3]. This model is applied for train tumor segmentation (BraTS 2018 challenge).

One third of the provided training set is used as a validation set during the training. The data augmentation is applied only on the training part. However the provided validation set is used for evaluating the selected model model weights after the training process.

From experimental works, is mean and standard deviation normalization provides better accuracy than 0 and 1 re-scaling method. Therefore, it is used in the all experiments.In addition, some hyper-parameters tuning is applied such as patch-size, the overlap size of the 3D blocks, different loss functions, different types of optimizer. The highest archived accuracy is obtained with the following hyper-parameters:

- Optimizer: Adam with learning rate (0.003)
- Loss Function: Dice Loss
- Dropout: 0.5
- Patch Size: 128
- Block Size: 32, 32, 32
- Block Sampling Step Size: 16, 16, 16
- Number of Training Epochs: 100

The following table shows a quantitative analysis for the provided validation set in terms of Dice Similarity Coefficient (DSC) between the predicted labels and the associated ground truth. Overall, CSF tissue shows the lowest DSC value in comparison to WM and GM. Applying U-Net segmentation for 5 labeled image data-set has a slight improvement more than the 4 labeled data-set. Moreover, despite of using the 4 Layers UNet model has more order of complexity than 3 Layers model, it shows a small improvement in dice score.

The best experimental results was obtained from applying 4 Layers U-Net and Modified U-Net architectures with the 5 labels data. The mean DSC accuracy for the both models are almost similar, while the GM and WM have

a small improvement in Modified U-Net. Choosing the best model between 4 Layers U-Net and Modified U-Net model is a trade-off, since the results of the 4 Layers U-Net model has a higher variance the more than the Modified U-Net. In terms of the networks complexity, 4 Layers U-Net architecture has around 9 millions trainable parameters, Modified U-Net has around 2 millions. Therefore, applying Modified U-Net in this segmentation problem is much way efficient.

DSC of the Validation Set					
	3 Layers UNet (4 Labels)	4 Layers UNet (4 Labels)	3 Layers UNet (5 Labels)	4 Layers UNet (5 Labels)	Modified UNet (5 Labels)
CSF Mean	90.53	90.67	90.57	**90.98**	90.94
GM Mean	93.60	93.89	94.17	94.43	**94.63**
WM Mean	92.56	92.82	93.30	93.70	**93.97**
CSF Min	88.20	87.95	88.05	88.88	**89.38**
GM Min	92.98	93.06	93.59	93.47	**93.97**
WM Min	89.20	89.61	91.87	92.56	**92.72**
CSF Max	92.04	91.86	92.06	**92.35**	92.16
GM Max	94.36	94.42	94.55	**95.23**	94.91
WM Max	95.26	95.59	95.56	**96.07**	95.75

Figure 8 shows the accuracy and the loss values for the training and the validation set along 100 epochs, in the experiment of using Modified U-Net with 5 labels data.

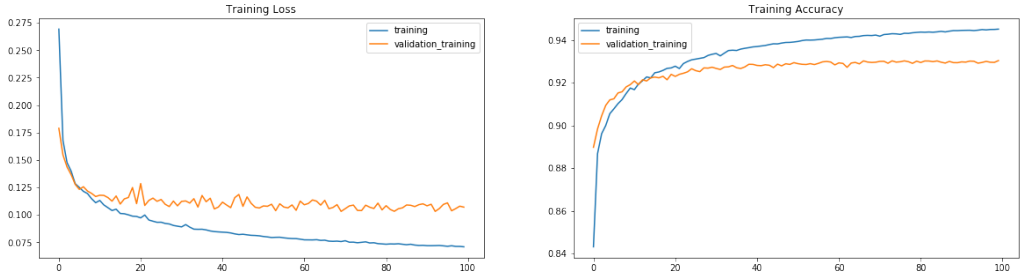

Figure 8: Loss and the Accuracy Modified U-Net with 5 labels data experiment

The following figures 9 and 10 shows a qualitative analysis in the case of applying Modified U-Net on 5 labels data . As shown in figure 10, the predicted and the ground truth labels are visually similar. Furthermore, the Model has successfully segmented the cerebral outer layer of the brain.

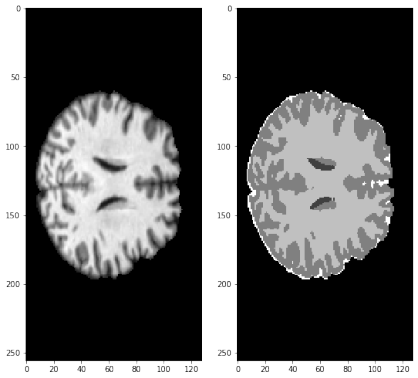

Figure 9: Registration framework

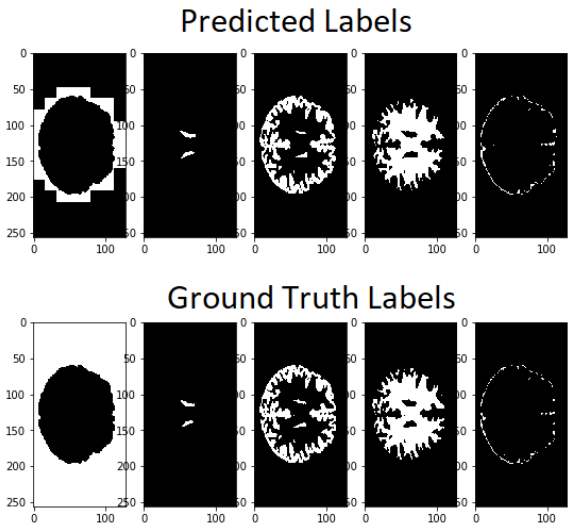

Figure 10: Registration framework

According to the previous quantitative results, the trained Modified U-Net model is used to for predicting the labels of the provided testing set. Figure 11 shows the segmentation results for the three testing images.

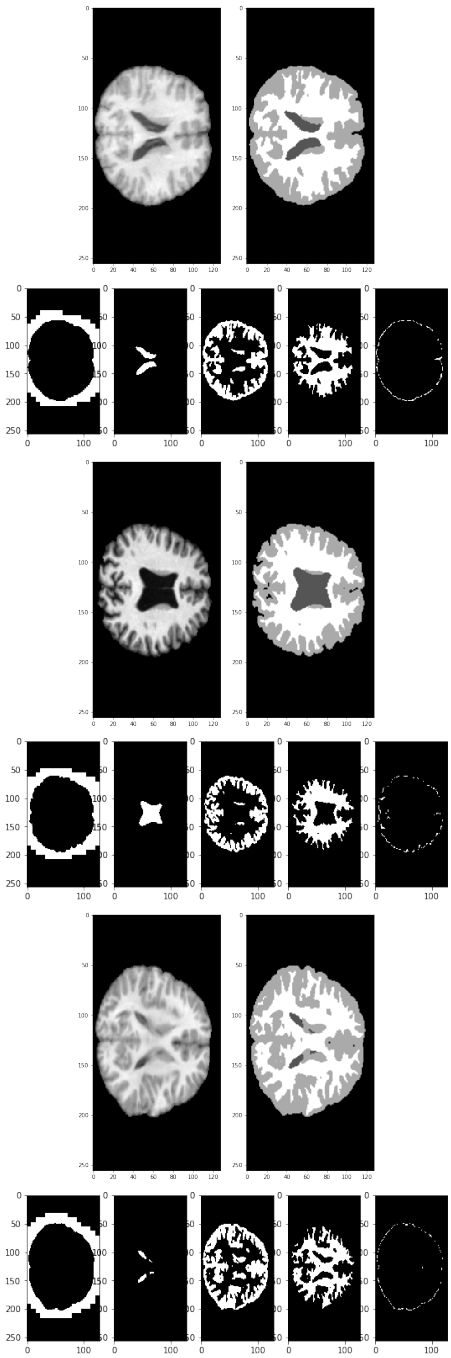

Figure 11: Predicted labels of the testing images

8 Conclusion

In conclusion, the aim of this project is to segment three different classes of brain tissues, which are CSF, WM and GM. The provided data in this project includes many challenges such as non consistencies of the data, limited numbers of training set and non-labeled regions. Moreover, computational challenges and GPU memory limitations for volume segmentation and for using complex network architectures. Slicing a high resolution 3D images into small 3D blocks can help in solving the GPU memory limitations for volume segmentation. In addition, U-Net is very powerful CNN architecture for medical image segmentation, even some state-of-art networks are designed based on modifying this architecture.

In this project, three different U-Net architectures are applied for two different number image labeling approaches. The first approaches using the original data which include the unlabeled brain regions as a background, while the other approach is based on generating an additional label for the unlabeled regions. From the experimental, using an additional labels approach provides a small enhancement in the segmentation accuracy in comparison to 4 labels approach. Although using 4 Layers U-Net and Modified U-Net architectures have close accuracy results, the Modified U-Net has less complex parameters.

9 References

References

[1] https://mrbrains13.isi.uu.nl/

[2] https://github.com/sergivalverde/MAIA_seminar.git

[3] Po-Yu Kao, Thuyen Ngo, Angela Zhang, Jefferson Chen, and B. S. Manjunath, "Brain Tumor Segmentation and Tractographic Feature Extraction from Structural MR Images for Overall Survival Prediction." arXiv preprint arXiv:1807.07716, 2018.