Transfer learning for 3d MRI images brain tumor segmentation-Final report

<u>Written by</u>:

Tomer Shimshi 203200480 tomershomsho@gmail.com

Avi Epsitein 320521461 Avi.zvi.epstein@gmail.com

<u>Abstract:</u> Brain tumor segmentation is a critical task for patient's disease management. Currently MRI's are hand annotated by a neurologist which is highly time consuming and sometimes prone to mistakes. In this article we attempt to solve the task of brain tumor segmentation. Given the Multimodal Brain Tumor Segmentation Challenge (BraTS) 2019 training dataset. To achieve this goal we have used the backbone of a pretrained resnet nammed Med3Dl net(link) this pretrained network was trained on a diverse set of medical 3d images acquired from x-ray CT and MRI to generally segment internal body parts. We use this backbone for transfer leaning on our 4 different types of MRI input images (T1,T1ce,T2,Flair) each one of these input images is inserted to the mednet and concatenated at the output. Doing this achieves good global segmentation of the different parts of the brian but is not specific enough to segment the exact tumor. Hence we have combined this system to a unet inspired network (Unets are known for their high results in segmentation problems especially for medical purposes) we then take our initial 4 3d inputs and down convolve them to the size of output of the med3Dnet and concatenate,we then upconvlove and cancatinat again. repeating this process 3 times until we reached the label size containing the segmentation of the tumor(4 segmentations :ET - Enhanced Tumor, TC - Tumor Tore, WT - Whole Tumor and background). Doing this we try to combine the best of both Resnet and Unet network architectures. The performance we were able to achieve from the validation is

Dice sore of: **WT:74.1% TC: 35.3% ET:24.6%**

For comparison the state of the art networks trained with much more resources than ours is:

Dice score of ET- 0.81,WT- 0.91 and TC-0.85

Therefore we believe that using our network architecture could receive the same if not better results with the sufficient resources. Moreover we have demonstrated the power of transfer learning and combining 2 state of the art networks for a difficult segmentation task.

Introduction:

Brain tumor is one of the most serious brain diseases, among which the malignant gliomas are the most frequently occurring type. The gliomas can be simply divided into two categories according to the severity: the aggressive one (i.e. HGG) with the average life expectancy of nearly 2 years and the moderate one (i.e. LGG) with the life expectancy of several years. Due to the considerably high mortality rate, it is of great importance for the early diagnosis of the gliomas, which largely improves the treatment probabilities especially for the LGG. At present, the most possible ways to treat gliomas are surgery, chemotherapy and radiotherapy. For any of the treatment strategies, accurate imaging and segmentation of the lesion areas are indispensable before and after treatment so as to evaluate the effectiveness of the specific strategy.

Among all the existing imaging instruments, MRI has been the first choice for brain tumor analysis for its high resolution, high contrast and present no known health threats. In the current clinical routine, manual segmentation of large amounts of MRI images is a common practice which turns out to be remarkably time-consuming and prone to make mistakes for the raters. So, it would be of tremendous potential value to propose an automatic segmentation method.

Our project attempts to solve this task with the results of fully segmenting the brain tumor to 4 segmentations ET WT CT given 4 different MRI images T1 T1ce T2 and flair

Related work:

Many researchers have proposed several effective methods based on deep learning or machine learning methods to solve the problem. Among those proposed methods, Zikic et al. [1]used a shallow CNN network to classify 2D image patches which were captured from the MRI data volumes in a sliding window fashion. Zhao et al. [2] converted the 3D tumor segmentation task to 2D segmentation in triplanes and introduced multi-scales by cropping different patch sizes. Havaei et al. [3] proposed a cascaded convolutional network, which can capture local and global information simultaneously.iek et al. [4] extended the traditional 2D U-net segmentation network to a 3D implementation which makes the volume segmentation to a voxel-wise fashion.Kamnitsas et al. [5] proposed a dual pathway 3D convolution network named DeepMedic to incorporate multi-scale contextual information, and used the 3D fully connected CRF as the postprocess method to refine the segmentation result. Chen et al. [6] improved DeepMedic by first cropping 3D patches from multiple layers selected from the original DeepMedic and then merging those patches to learn more information in the network, besides, deep supervision was introduced in the network to better propagate the gradient.Ma.et al [7]employed a feature representations learning strategy to effectively explore both local and contextual information from multimodal images for tissue segmentation by using modality specific random forests as the feature learning kernels.

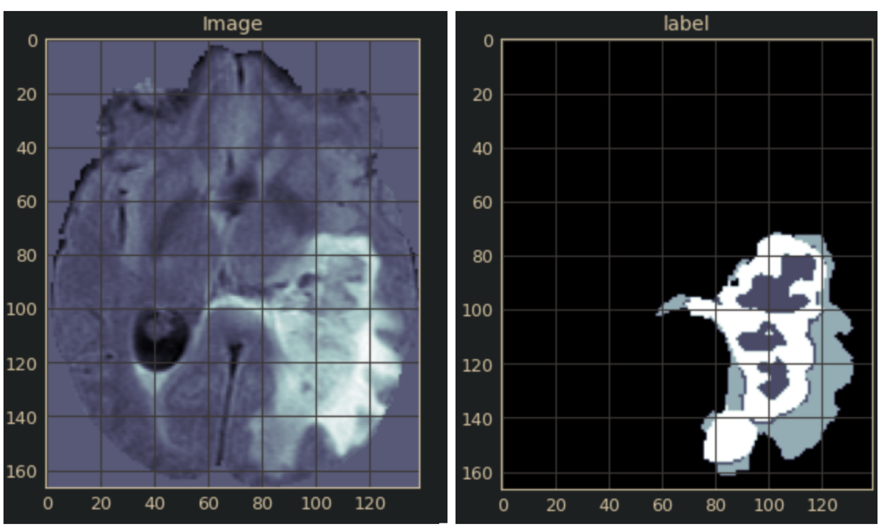
Moreover med 3d used a resnet encoder decoder to train a general purpose 3D medical segmentation system by training on large amount of diverse images received from MRI's CT's and x rays of different parts of the human body parts encoding them all throw the backbone and decoding them to 8 different segmentation tasks to later use this back bone for transfer learning on the Lung Nodule Dataset.

After seeing the results of med 3D network and the state of the art unet segmentation success we propose to combine a pre trained resnet(Med3D) transfer learning with a Unet network, to tackle this challenge. Inputting 4 3 Dimages at a time. We believe that this combination of general segmentation of the resnet to fine segmentation of the unet will be able to speed up training and receive better results in the long run.

Data:

Brats 2019 is are multimodal scans in NIFTI file format each input to the system is comprised of 4 3D MRI scans native T1,post-contrast T1-weighted, T2-weighted and Fluid Attenuated Inversion Recovery (T2-FLAIR). All the imaging data have been segmented manually, by one to four raters, following the same annotation protocol, and their annotations were approved by experienced neuro-radiologists. Annotations comprise the GD-enhancing tumor (ET — label 4), the peritumoral edema (ED — label 2), and the necrotic and non-enhancing tumor core (NCR/NET — label 1),The data is distributed after their pre-processing, i.e. co-registered to the same anatomical template, interpolated to the same resolution (1 mm^3) and skull-stripped.meaning the voxels are interpolated and mapped to xyz courdents.

After transforms of center crop and random rotation flip we acquired a data set of 334 labeled 3D images for training and another 43 images for validation.



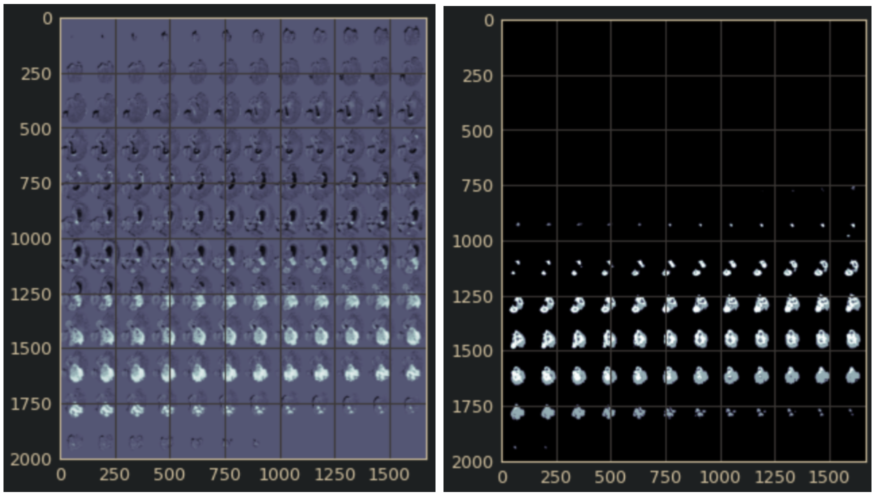


Figure 1: brats raw data visualization

Methods:
Based on the thorough analysis of the brain segmentation task we propose a combination of transfer learning from a pre-trained 3d medical image resnet as a backbone with a concatenated Unet.

Our initial approach was to use the Med3D network to achieve the goal of the task. As you can see in figure 2 the pretrained network used a seg8 resnet consisting of an encoder and 8 decoders, for each training segmentation problem in order to make a universal encoder of internal body parts for medical imaging.

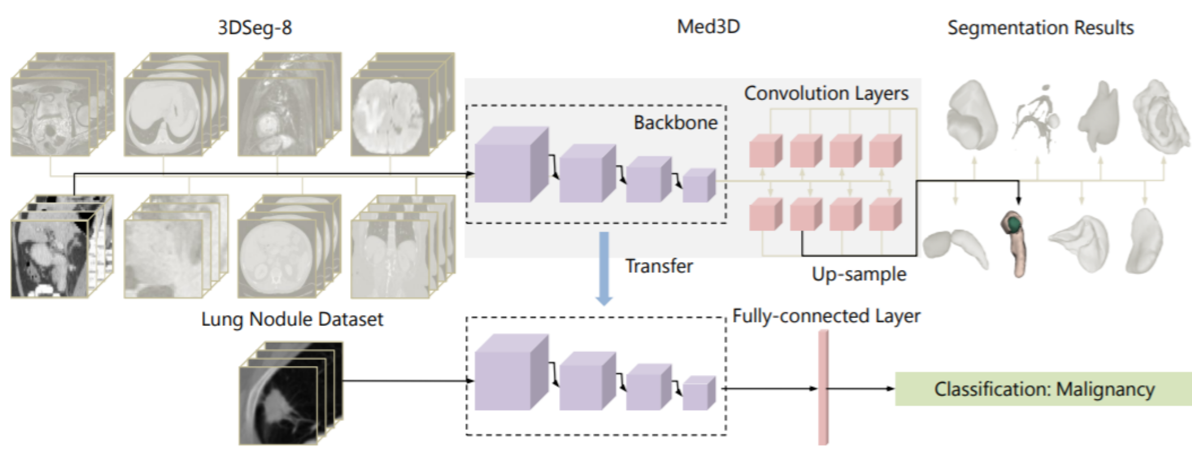


Figure 2 :Med3D method

Here you can see how Med3D is able to generalize segmentation of internal body parts. After using this model on our brats 2019 dataset we were able to achieve general segmentation of the patient's brain structure by feeding each image input (T1,T1ce,T2,FLAIR) to the system but we were not able to get exact segments of the tumors.

For this reason we decided to connect this transferred network to an untrained Unet which has been proven to be very effective on medical segmentation tasks. Unets are built by double convolving the input then downsampling arriving at a bottleneck which is then upsampled by convtranspose concatenated and again double convolved. In figure (3) you can see the classic 2D unet:

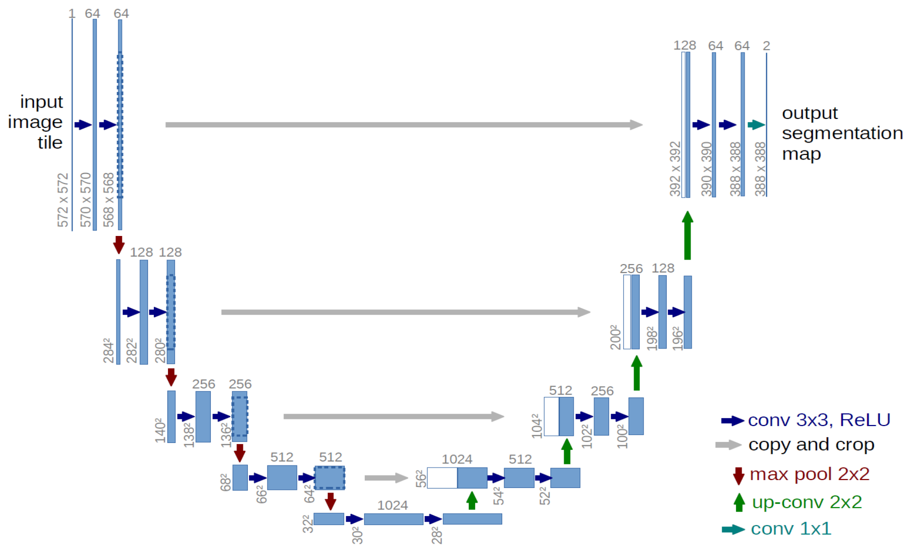


Figure 3: classic 2D unet

Inspired by these 2 networks we took our 4 3d images and ran them through the pretrained Med3D resnet and obtained a bottleneck encoder. In parallel we applied a semi unet for 3D images were we down sampled and convolved our original 4 3d images 2 times to match the size of the output from the med3D we then upsampled with convtranspose concatenated and again double convolved.repeating this 3 times until we reached the size of our original mask label. Using the same technique of the original 2D unet with different depths.

In figure 4 we can see the breakdown of our model:

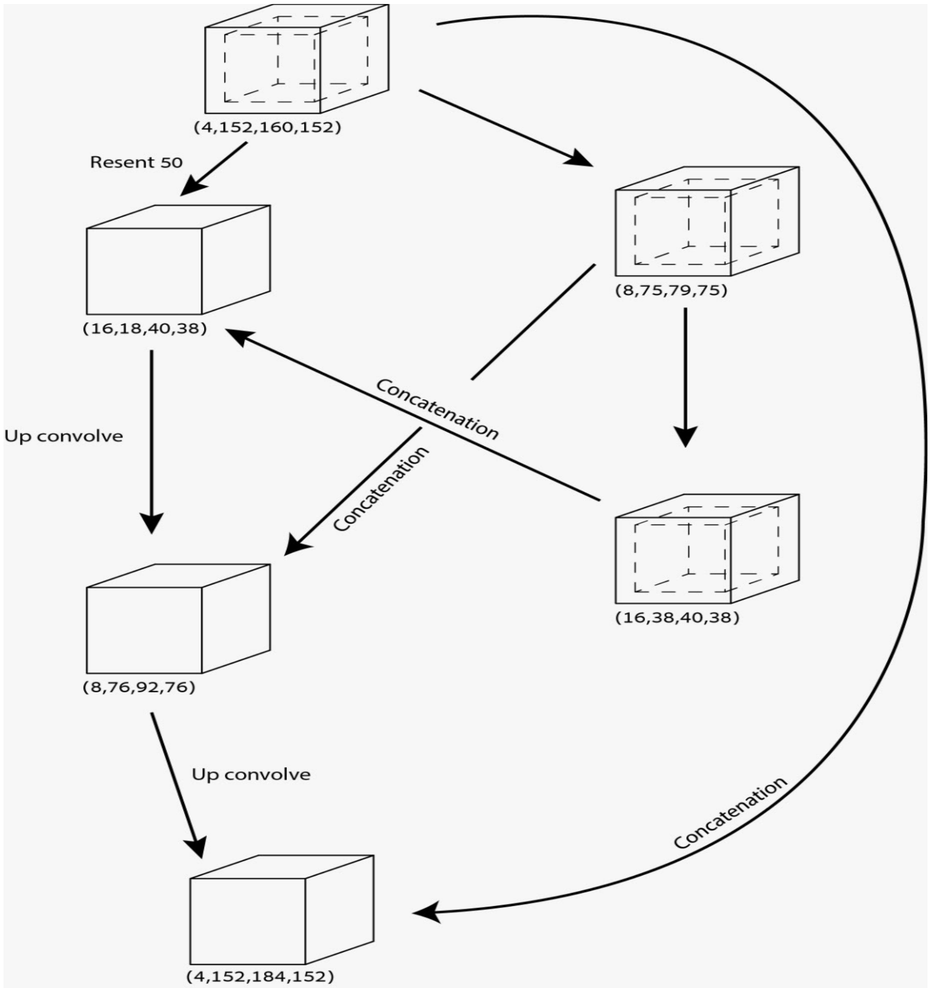


Figure 4: our model

As in Unet 2d we interpolated the images to match the concatenation if needed.

We believe that using the pretrained resnet will be able to extract general structures of the brain while the 3d unet will be able to zoom in on the specific structure we want to segment.(ie the tumor segmentation) .

Given this we explored the results of different pretrained resnets that were supplied by the Med3D network.

Experiments:

Initially we tried to use the pre-trained med 3d without the added U-net which resulted in poor results. so we have decided to up scale the image coming out of the forward pass of our network to the original size of the image so we could visualize the data coming out of the system we then saw that we get improved results after that we performed varies tests trying to improve the network and make it into a more U-net like network after that we converged into the final net topology. At the final stage we tested 3 different net topologies to determine which gives us the best results :

U-net with pre trained Resnet -10 without changing the the pretrained weights:

We got the following dice scores : WT:74.1% TC: 33.8% ET:13%

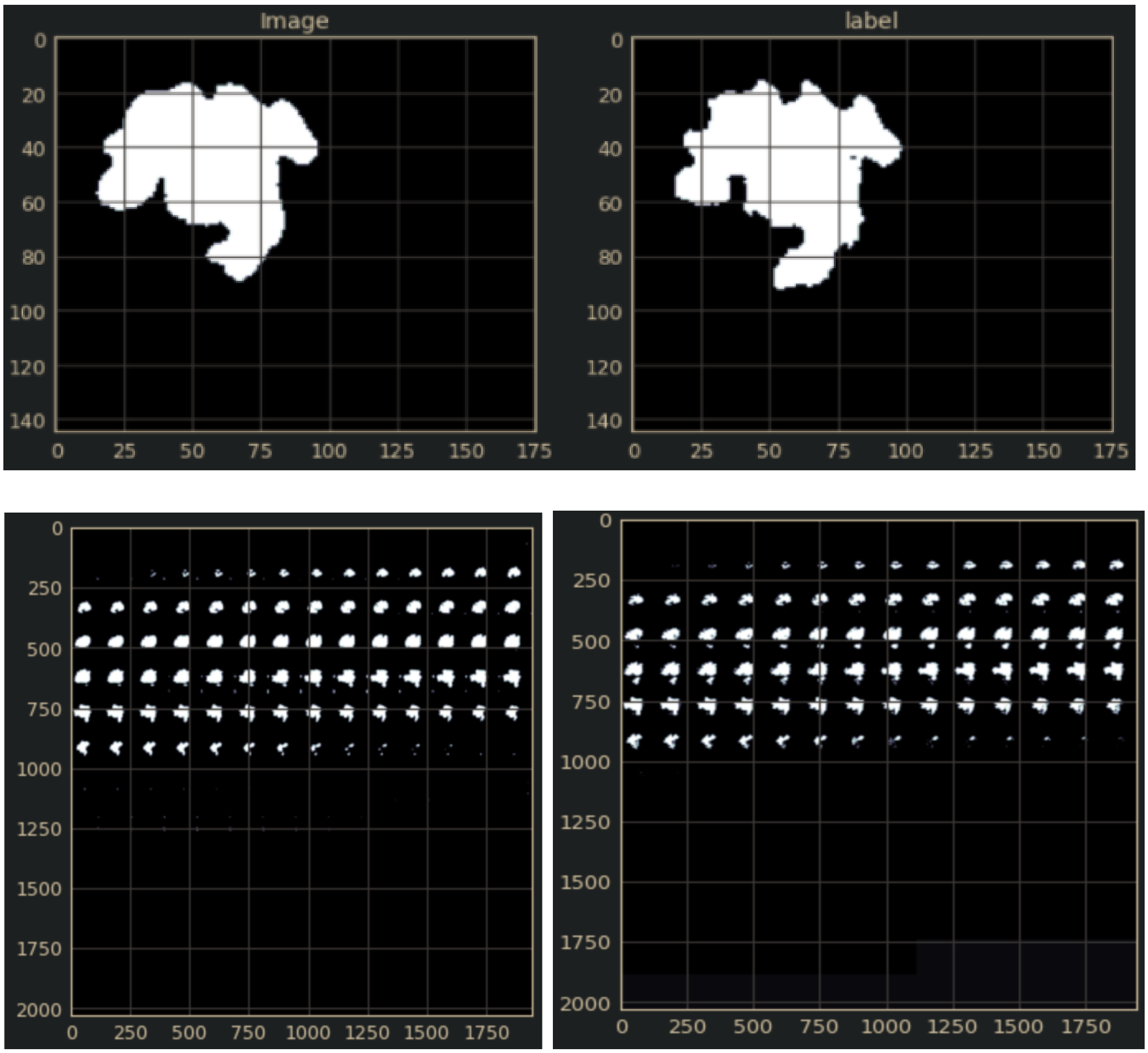
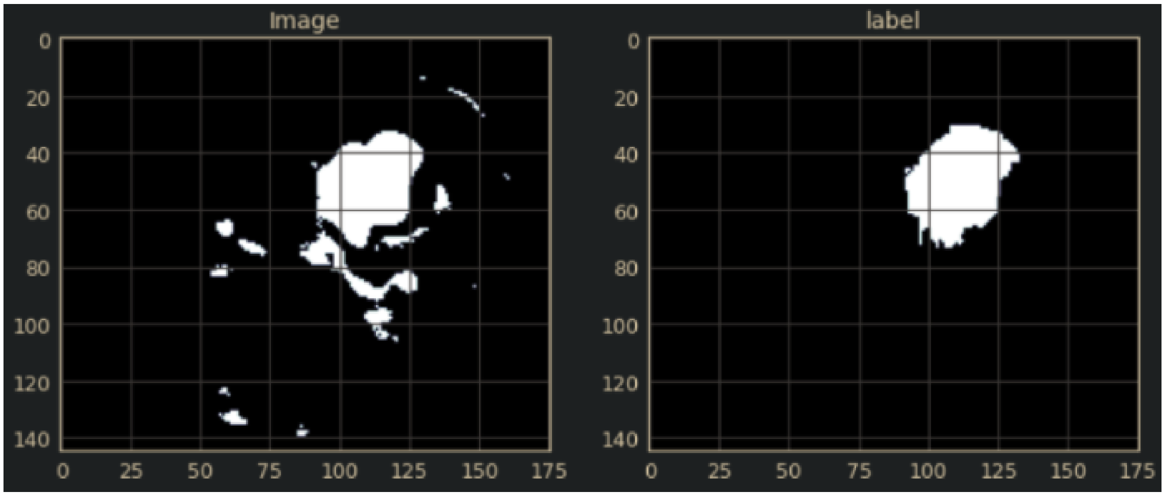
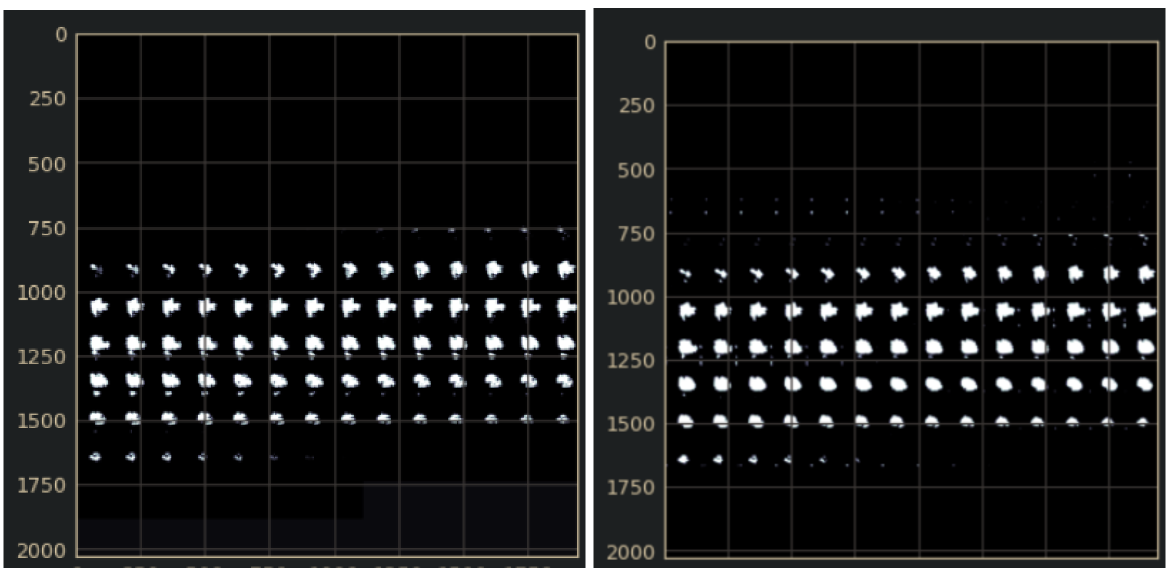


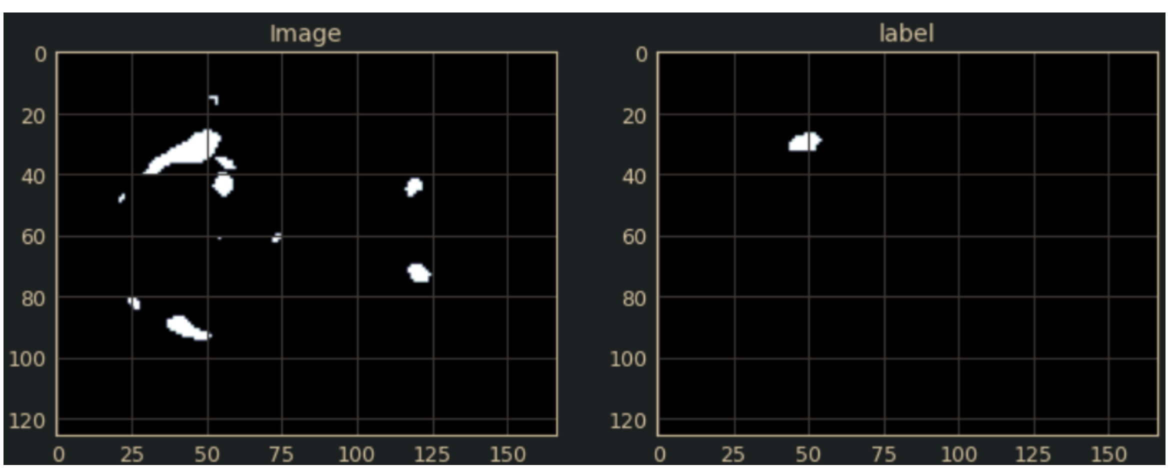
Figure 5:WT segmentation on the left is output image on the right the correct label

U-net with pre trained Resnet -10 with changing the the pretrained weights:

We got the following dice scores : WT:74.1% TC: 35.3% ET:24.6%







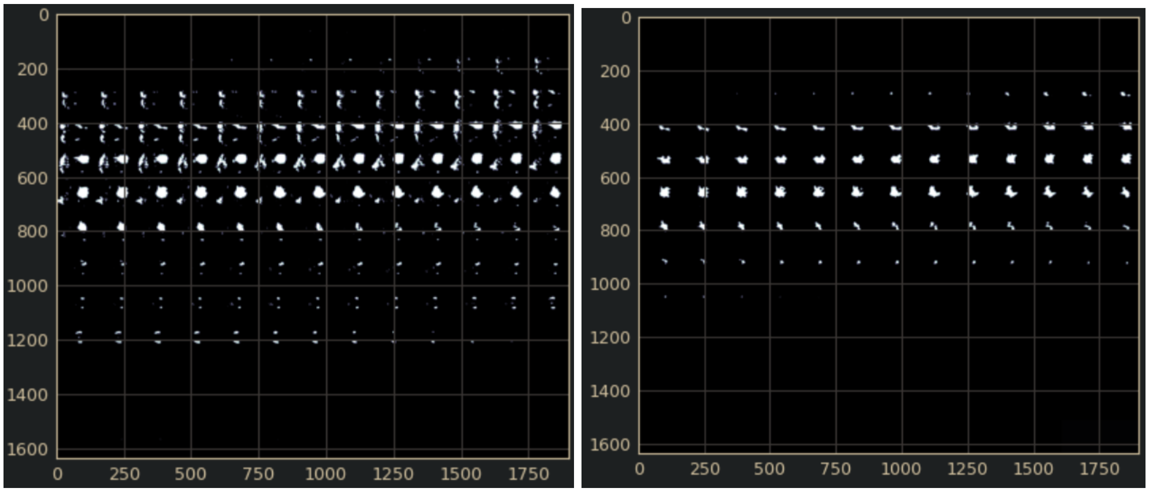


Figure 6: U-net with pre trained Resnet -10 with grad on top we can see TC and ET on the bottom

U-net with pre trained Resnet -50 without changing the the pretrained weights:

We got the following dice scores : WT:35.1% TC:16.1% ET:7.92%

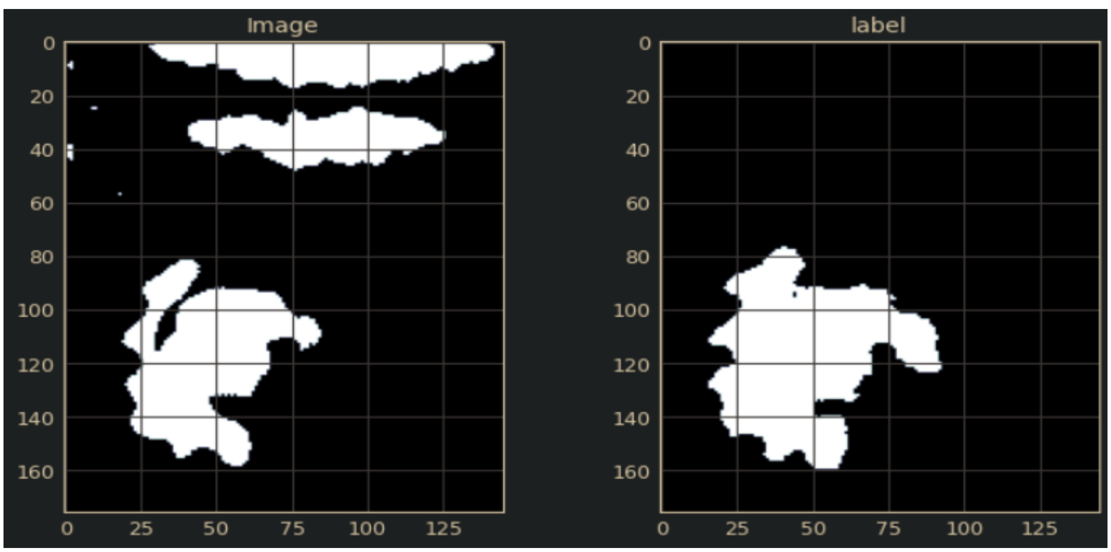


Figure 7: U-net with pre trained Resnet -50 no grad WT

U-net with pre trained Resnet -18 with changing the pretrained weights:

We got the following dice scores : WT:21.8% TC:2.58% ET:0.277%

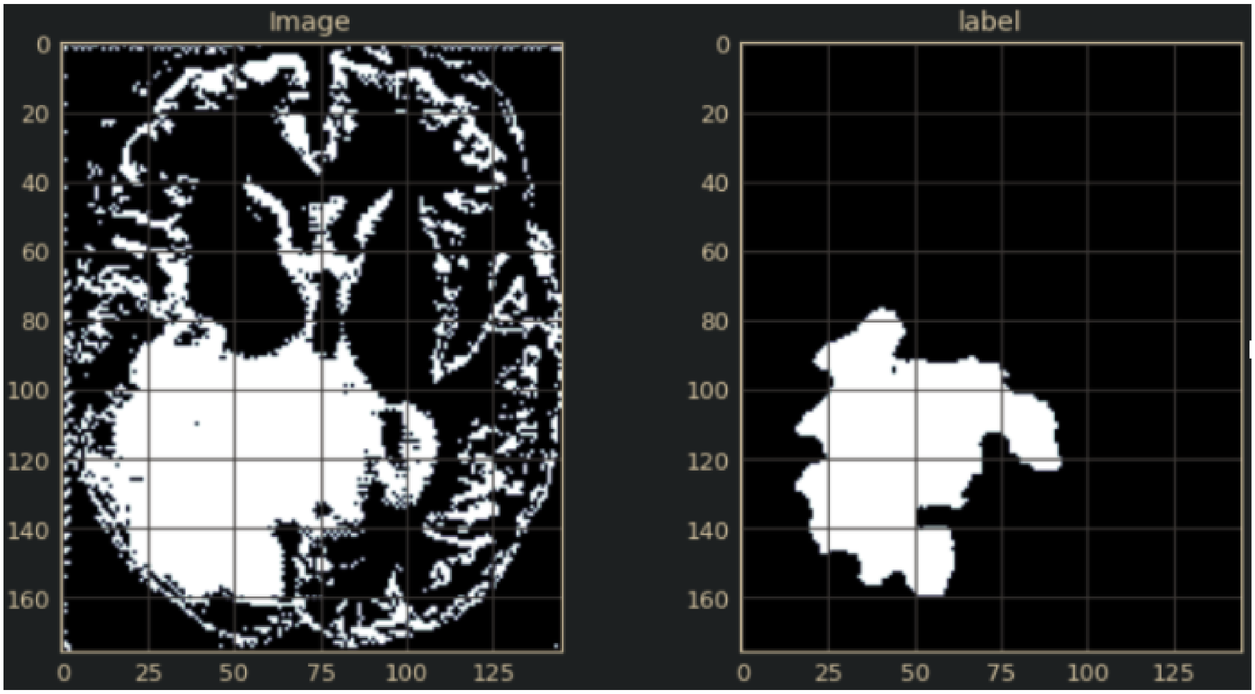


Figure 7: U-net with pre trained Resnet -18 with grad WT

Hyper parameters:

Adam optimizer, learning rate:0.001 and using a scheduler step at the beginning of each epoch for a total of 10 epochs

Loss:

One of the issues was deciding the train and validation loss as for training the dice loss will only be low if the union of same pixel values are the same. However in training we did not want the output to be binary because we wanted the gradients to have a smoother effect hence we have used the BCEWithLogitsLoss for our output of our model and in test evaluation we used the dice score link.

loss function for image segmentation tasks is based on the Dice coefficient, which is essentially a measure of overlap between two samples. This measure ranges from 0 to 1 where a Dice coefficient of 1 denotes perfect and complete overlap. The Dice coefficient was originally developed for binary data, and can be calculated as link:

$$Dice \ = 2 * |(A \cap B)| \div (|A| \ + \ |B|)$$

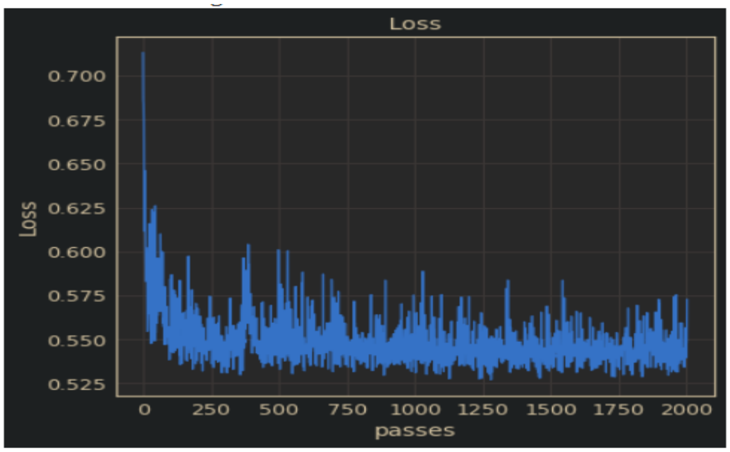
Where |A∩B| represents the common elements between sets A and B, and |A| represents the number of elements in set A (and likewise for set B).

Another issue was the batch size because each data input image is an mri from different machines and years has a different size. We encountered problems when trying to resize the images. Resizing the images resulted in skewed labels and inputs therefore we decided to run the network each time with batch size of size 1.

<u>loss graph: (as a combination of all 3 dice loss)</u>

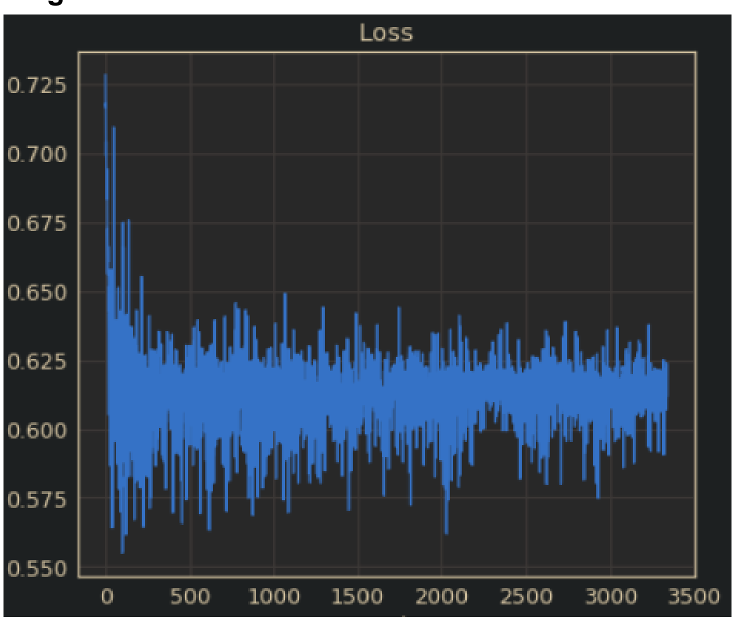
<u>Visualizing the results:</u>

The Loss values for theU-net with pre trained Resnet -10 without changing the the pretrained weights:



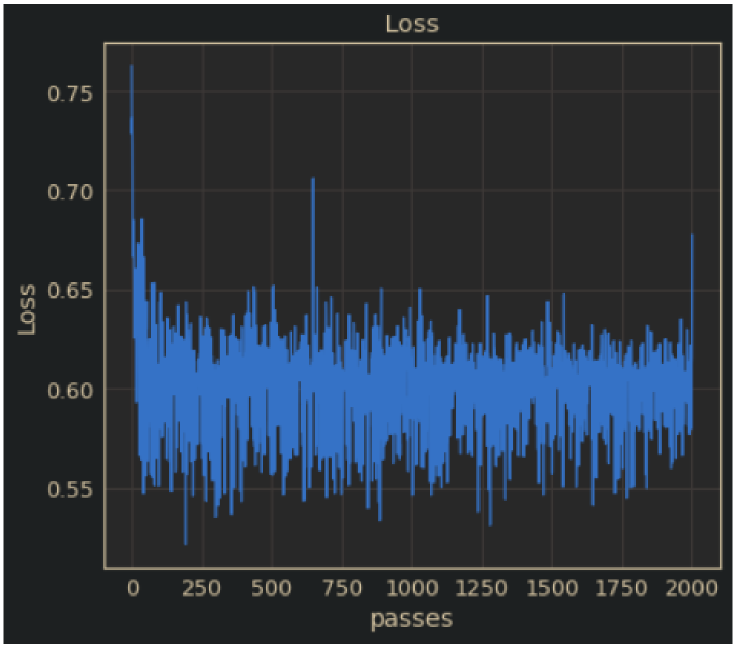
As we can see the loss started at 0.7 then quickly converged to around 0.55 and remained at approximately this value

The Loss values for theU-net with pre trained Resnet -10 with changing the the pretrained weights:



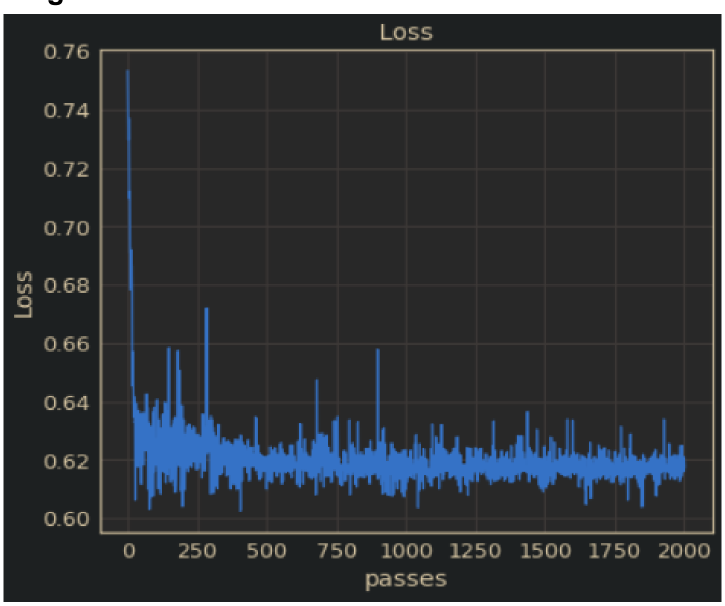
As we can see the loss started at 0.725 then quickly converged to around 0.615 and remained at approximately this value

The Loss values for theU-net with pre trained Resnet -50 without changing the the pretrained weights:



As we can see the loss started at 0.765 then quickly converged to around 0.61 and remained at approximately this value

The Loss values for theU-net with pre trained Resnet -18 with changing the the pretrained weights:



As we can see the loss started at 0.75 then quickly converged to around 0.621 and remained at approximately this value

Conclusion: in this project we have demonstrated the power of transfer learning from general 3d data to a specific brain segmentation task. We have combined 2 complex networks Resnet and unet in order to get the most out of the generality of the pretrained med 3d net and the spacifty of the 3D unet architecture.
As you can see in the results above in the WT results we generally get a good segmentation however for the TC and ET the network struggles to converge, this is because many times the segmentation is very small and hard to detect. We assume that if we could train for longer we would get better results.
In addition it is possible that because we take the bottleneck of the resnet to be with 4 channels we lose important information from the resnet. It is possible that removing this last layer and keeping the depth of the resnet large we would also get better results.
Moreover, adding the transfer model bottleneck to an early stage of a deeper Unet could also result in better results because we will be giving the network attention to the main areas of interest in an earlier stage.

The model that we currently think has the highest probability of success is: **U-net with pre trained Resnet -10 with changing the the pretrained weights**

As has been shown we have tackled a difficult segmentation task and combined various deep learning techniques to achieve good segmentation of MRI's, so that one day we will have faster and more accurate segmentation for people suffering from brain cancer.

Appendix:

Med3D: `https://github.com/Tencent/MedicalNet`

Brats2019:https://github.com/askerlee/segtran

1. Zikic, D., Ioannou, Y., Brown, M., Criminisi, A.: Segmentation of brain tumor tissues with convolutional neural networks. Proceedings MICCAI-BRATS (2014) 36–39

2. Zhao, L., Jia, K.: Deep feature learning with discrimination mechanism for brain tumor segmentation and diagnosis. In: 2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), IEEE (2015) 306–309

3. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., Larochelle, H.: Brain tumor segmentation with deep neural networks. Medical image analysis 35 (2017) 18–31

4. C¸ i¸cek, O., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d u- ¨ net: learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention, Springer (2016) 424–432

5. Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B.: Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. Medical image analysis 36 (2017) 61–78

6. Chen, S., Ding, C., Liu, M.: Dual-force convolutional neural networks for accurate brain tumor segmentation. Pattern Recognition 88 (2019) 90–100

7. Ma, C., Luo, G., Wang, K.: Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of mr images. IEEE transactions on medical imaging 37(8) (2018) 1943–1954

8. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). IEEE transactions on medical imaging 34(10) (2014) 1993–2024

9. Lin, T.Y., Goyal, P., Girshick, R., He, K., Doll´ar, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. (2017) 2980–2988

10. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. Scientific data 4 (2017) 170117 11

11. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S.M., Rozycki, M., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint arXiv:1811.02629 (2018)

12. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-lgg collection. The Cancer Imaging Archive 286 (2017)

13. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. the cancer imaging archive (2017). (2017)

14. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4itk: improved n3 bias correction. IEEE transactions on medical imaging 29(6) (2010) 1310

Our code:
https://github.com/TomerShimshi/DeepLearning-course-Final-Project-Brats-challenge

link