Optimizing brain tumor segmentation using Transfer Learning

Aadithya Kandeth*
aadithya.kandeth@ufl.edu, UFID: 6980-2791

Sai Nikhil Dondapati*
sa.dondapati@ufl.edu, UFID: 2228-6439

Shaanya Singh*
shaanyasingh@ufl.edu, UFID: 3476-2752

*Department of Computer and Information Science and Engineering, University of Florida

Abstract Deep Learning for medical imaging is a fast-growing field with widespread applications across. Computer Aided Diagnosis is one of the main approaches which assists doctors in making correct decisions regarding their patients' health. In this project we plan to approach the field of Neurology with the help of Deep Learning and find the most effective models that will give us the best possible accuracy on our dataset. We chose to use transfer learning as it is hard to find a large brain MRI dataset of which all are correctly annotated. Transfer Learning allows us to make our models more accurate even in the absence of larger datasets. Transfer Learning helps us with the generalization process by giving us more flexibility to transfer knowledge between different domains. In this project, we demonstrate that transfer Learning is a great way to train models for segmentation tasks efficiently.

1 Introduction

Recent trends in Deep Learning have paved the way for medical analysis and computer-aided diagnosis in the biomedical field. They have been major contributors in the field of tumor classification/detection heavily relying on MRI images as inputs. MRI can provide a better contrast for tissues and hence makes it easier for neural networks to learn these changes in patterns and provide appropriate results. In addition to using simple neural networks, we have decided to apply a special class of deep learning, called transfer learning into our problem. Transfer learning allows us to use models that have already been trained on different images and apply it into our use-case. In this project, we used architectures such as U-net, VGG-16, ResNet50, InceptionNet, and InceptionResNet on the BraTS dataset. We then fine-tuned the pre-trained models of the different architectures for Brain Tumor Segmentation. We compare the results of these techniques to find out which pretrained model performs best. We used accuracy and IoU score as metrics to compare the results.

2 Literature Review

There has been extensive research in using deep learning to segment and classify biomedical images to diagnose diseases early. Multiple proposed solutions have been developed to segment tumors in MRI scans. Transfer learning has emerged as one of the most effective approaches for segmenting tumors and multiple papers highlight the benefits of transfer learning over standard machine learning algorithms. For example, [1] discusses the use of transfer learning for segmenting tumors in MRI images. The authors performed this study by finding studies that used

transfer learning to complete MR brain imaging tasks. More than 433 papers were evaluated by comparing data such as label availability and machine learning techniques. Among them, 129 articles were located that used transfer learning for MR brain imaging and the paper with the best results was found using different evaluation metrics. In another work [2], researchers develop a system that could classify cancers in MRI images using a newly curated dataset. This used discriminative learning and deep residual CNN based architecture for the segmentation, and the achieved accuracy was over 97%. These metrics suggested that the proposed architecture was a viable solution for the brain image segmentation problem. U-Net was developed specifically for biomedical imaging. Multiple research papers show that a U-NET model is highly effective in segmenting MRI images. For example, [3] describes a simplified U-Net based implementation. The reasoning behind this was that this approach would require lesser data to train the network. Additionally, no further data augmentation step would be required. The approach was tested on the BITE dataset and was able to perform better than the traditional benchmark techniques to achieve an IOU score of 89%. The data also used three perspective planes instead of the original 3D volumes for simplified training. [4] also leveraged a multiple U-Net based neural network architecture along with deep supervision and stochastic weight averaging to automate image segmentation on the BraTS dataset. Two model ensembles were trained on the data and the performance of each were considered while merging label maps for the patients. The complex training schemes and neural network architectures were designed with the trade-off of much higher training times. Some other neural network architectures that can be used to solve the brain tumor segmentation problem include VGG-16 and Resnet50. [5] used a CNN to automatically interpret MRI scans more precisely. This was aided by SegResnet to speed up the tumor annotation process. The highest accuracy achieved was 81%, 87%, and 91% for segmenting enhancing tumors, whole tumors and tumor cores respectively.

3 Research Methodology

3.1 Description

In this study, we implemented the state-of-the-art 3D models commonly used for Brain Tumor Segmentation. We developed U-net, VGG-16, and Resnet-50 models on the BraTS dataset. We also used transfer learning approach to train VGG-16 and Resnet-50 and compared the results with the results obtained without transfer learning approach. All the comparisons in this study are made using IoU score and accuracy. Owing to the limited pretrained 3D models, and very long training time of 3D models, we extract 2D images and masks by slicing 3D MRI volumes. Using the extracted 2D images, we trained Inception-Resnet, InceptionNet, Resnet-50, and VGG-16. We make use of transfer learning approach for developing 2D models. All the models developed in this study as part of evaluating the performance of transfer learning approach are built upon pretrained models that have been trained on ImageNet dataset. We compare the results obtained via different implementations as described above.

3.2 Dataset Preparation

We extensively make use of BraTS dataset for this project. BraTS dataset has multimodal data with native (t1), post-contrast t1-weighted (t1ce), t2-weighted (t2), and FLAIR scans. We disregard the t1 channel since it does not contain as much relevant information as t1ce. The scans in this dataset were obtained from 19 institutions that used different clinical protocols and medical scanners for the acquisition. The decision to go with this dataset was made based on the amount

of data available and the significant improvements it offered over the earlier datasets. The dataset has GD-enhancing tumor, peritumoral edema, the necrotic tumor, and non-enhancing tumor core. We extract patches of 128*128*128 from the brain MRI volumes. Only the t1ce, t2 and flair channels are used for the input. The input dimension is therefore 128*128*128*3. In this study, we classify each pixel as one of the 3 tumor types or as non-tumorous. Therefore, the output dimension of our models is 128*128*128*4.

3.3 3D Modeling

U-Net (Figure 1) is one of the most effective networks that has been introduced for biomedical image segmentation. U-Net consists of many encoder and decoding layers. These layers altogether help us to get a precise output with high resolution. Encoder is the first half of the whole architecture, made up of convolution layers and maxpooling layers which provide downsampling to encode the input information and extract relevant features. Decode is the second half of the whole architecture, that contains convolution layers, concatenation and upsampling layers which help us decode the encoded information from the first half and generate a mask.

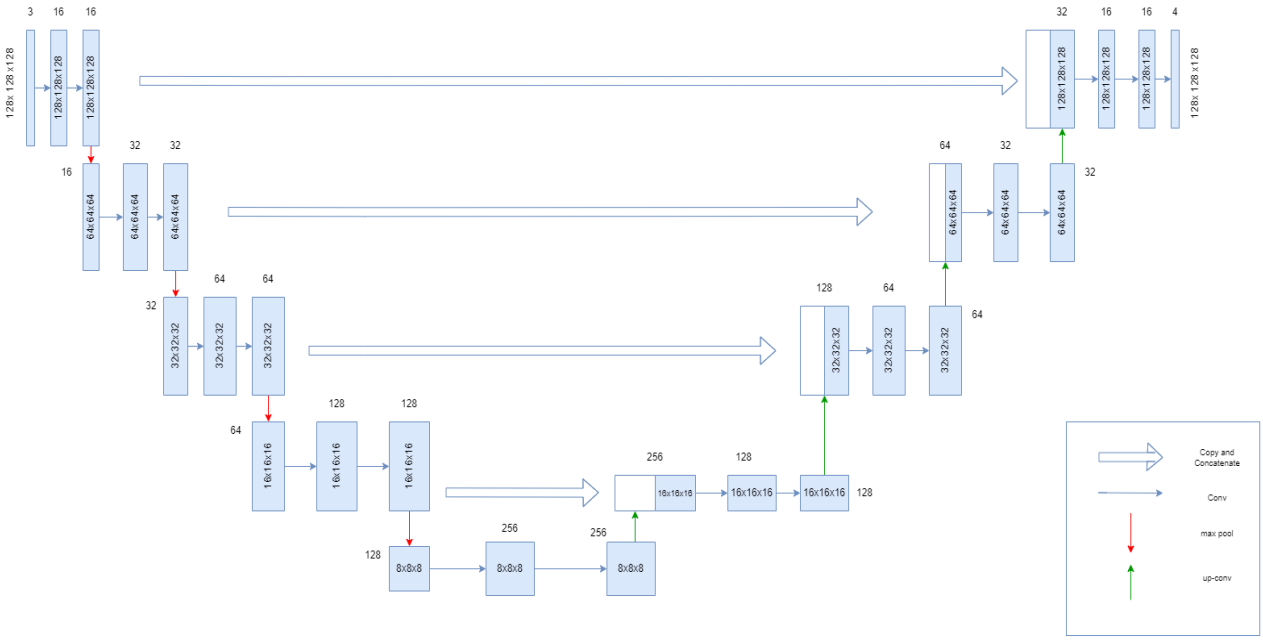

Figure 1. U-Net architecture

VGG-16 network (Figure 2) consists of dense layers, convolution layers, maxpooling layers and upsampling layers. ResNet (Figure 3) is a convolutional neural network that consists of multiple stacked residual layers. The architecture was created to solve the problem of vanishing gradients using the concept of skip connections. We use U-Net, VGG-16, and ResNet-50 in this study.

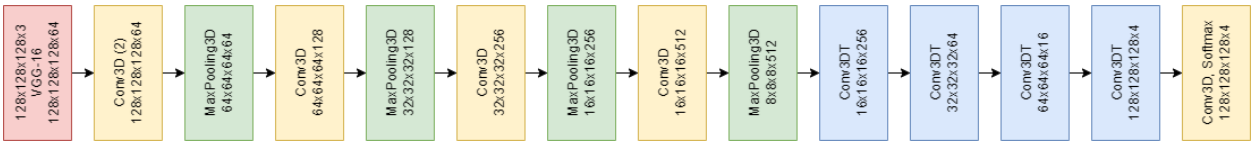

Figure 2. VGG-16 architecture for 3D inputs.

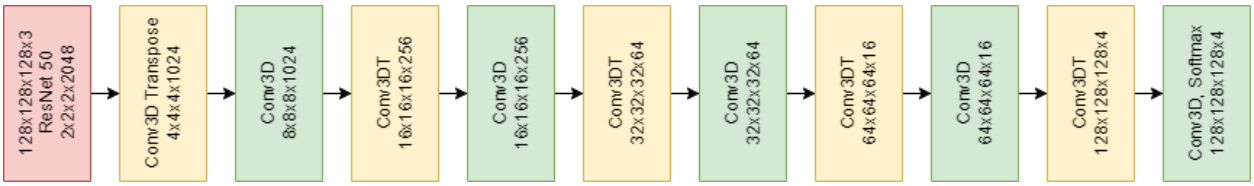

Figure 3. ResNet-50 architecture for 3D inputs.

3.4 2D Modeling

In order to boost the training speed for comparing several transfer learning approaches, we extract 2D images and masks from the MRI volumes by slicing the input. We build Inception-ResNet, Inception Net, VGG-16, and ResNet-50 on the extracted 2D data. After the models are trained and predictions are made, we perform post-processing to reconstruct the 3D mask back. As part of our transfer learning approach, we make use of models pretrained on ImageNet dataset.

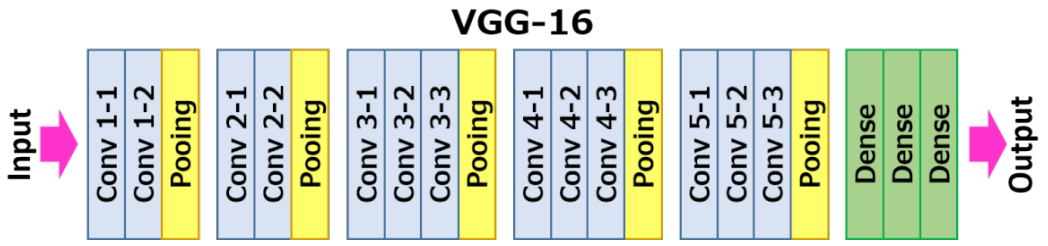

Figure 4. Standard VGG-16 architecture for 2D inputs

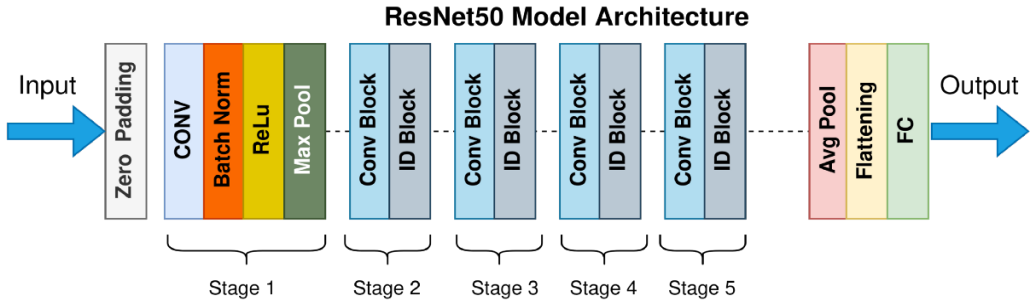

Figure 5. Standard ResNet-50 architecture for 2D inputs

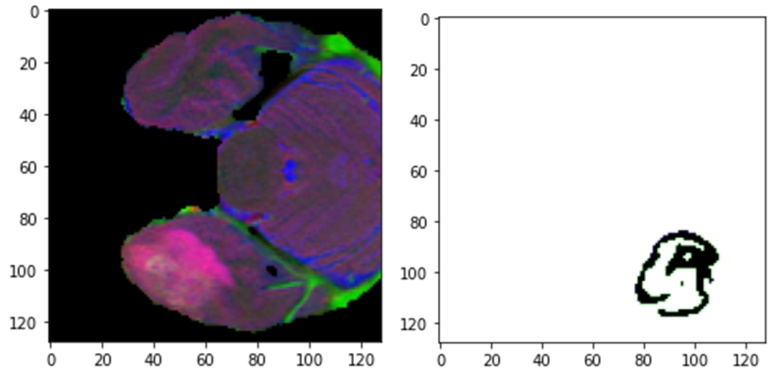

Figure 6. An example 2D image and mask obtained by slicing 3D MRI Volumes

3.5 Metrics

Because of the huge class imbalance in the dataset, we use IoU score along with accuracy as metrics. We use total loss as the sum of dice loss, categorical cross entropy loss, and focal loss.

3.5.1 IoU score (Intersection over Union)

$$IoU = \frac{Area\ of\ Intersection/Overlap}{Area\ of\ Union}$$

3.5.2 Dice Loss (Extensively used for segmentation tasks)

$$Dice = \frac{2\,|A \cap B|}{|A| + |B|}$$

$$Dice\ Loss = 1 - Dice$$

3.5.3 Cross Entropy Loss

$$Cross\ Entropy\ Loss = -Y_{act}\ ln(Y_{pred}) - (1 - Y_{act})\ ln\ (1 - Y_{pred})$$

Where Y_{act} is actual value and Y_{pred} is predicted value of Y.

If p_t is probability, then the equation for cross entropy loss can also be written as

$$CE(p, y) = CE\ (p_t) = -ln(p_t)$$

3.5.4 Categorical Focal Loss

If r is a focusing parameter greater than or equal to 0, and p_t is probability, then focal loss is

$$Focal\ loss = -(1 - p_t)^r\, ln(p_t)$$

4 Results and Analysis

4.1 Results obtained from 3D modeling

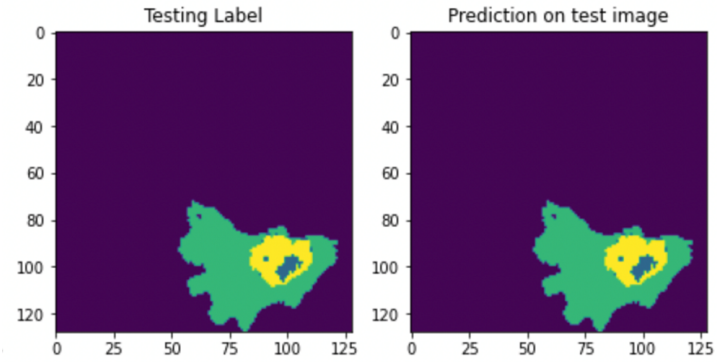

Figure 7. An example Output from 3D Modeling

Model	Approach	Accuracy	Mean IOU
U-net	No transfer learning	0.97	0.83
VGG-16	No transfer learning	0.94	0.74
VGG-16	Transfer learning	0.94	0.63
Resnet-50	Transfer learning	0.95	0.67
Baseline CNN	No transfer learning	0.75	0.24

Table 1. Performance of different 3D Models on BraTS dataset

4.2 Results obtained from 2D Modeling:

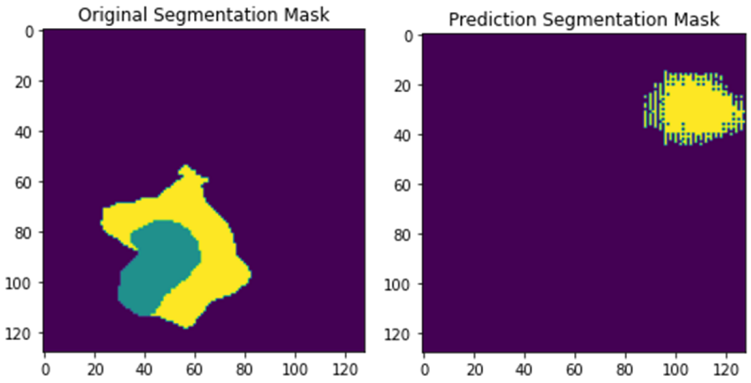

Figure 7. An example Output from 2D Modeling

Model	Approach	Accuracy	Mean IOU
Inception resnet	Transfer learning	0.87	0.53
Inception net	Transfer learning	0.76	0.61
Resnet-50	Transfer learning	0.51	0.25
VGG-16	Transfer learning	0.61	0.32

Table 2. Performance of different 2D Models on BraTS dataset

We can clearly see that 3D models outperform 2D models. We also note that the pretrained models are easy to train and they offer better generalization. In this project, selecting evaluation metrics has been very important for developing models that perform and generalize well.

5 Conclusion

In this project, we faced some problems because of the long training times and because of the huge variance between different datasets available publicly. For this reason, we conclude that it is essential to develop pretrained repositories that can be used for different medical imaging applications. From our evaluations, we can conclude that in the case of brain tumor segmentation, 3D models clearly outperform 2D models. Generalization can also be improved by increasing the dataset size. Since we don't have many pretrained 3D models and that inter-domain transfer learning is a challenging task, this project is overall very challenging. Handling the huge imbalance in the dataset is another challenge we would like to make note of. IoU score is an appropriate metric for segmentation tasks. We can also see from the results that the U-Net model performs better as the concatenation layers in it boost the IoU score.

6 Future work

As of today, the number of pretrained models that exist for 3D data is very limited. The restricted access because of various factors like privacy concerns makes this issue a lot more prominent when it comes to medical imaging and, more specifically, MRI scans. So, a lot of these models must be built from scratch. The pretrained models that do exist are usually trained on very less data, so they don't yield good results when used for transfer learning. Another issue to try and solve in the future would be to accelerate the training process. Long training times are a major concern when it comes to evaluating multiple models. One possible way to speed up the training could be using anchor free methods with predefined priors. Anchor free algorithms are a lot faster and more generalizable for computer vision tasks like object detection. Considering that the MRI scan data has very limited availability and has significant class imbalance, we need to figure out additional methods to handle this in the future. Some of these methods could include data resampling, generating synthetic samples using SMOTE, or using GANs to produce synthetic data.

References

[1] Juan Miguel Valverde, Vandad Imani, Ali Abdollahzadeh, Riccardo De Feo, Mithilesh Prakash,,Robert Ciszek, and Jussi Tohka. Transfer Learning in Magnetic Resonance Brain Imaging: a Systematic Review. *arXiv:2102.01530v2 [eess.IV] 1 Apr 2021.*

[2] Yusuf Brima, Mossadek Hossain Kamal Tushar, Upama Kabir, and Tariqul Islam. Deep Transfer Learning for Brain Magnetic Resonance Image Multi-class Classification. *arXiv:2106.07333v2 [cs.CV] 15 Jun 2021.*

[3] Jason Walsh, Alice Othmani, Mayank Jain, Soumyabrata Dev. Using U-Net Network for Efficient Brain Tumor Segmentation in MRI Images. *arXiv:2211.01885 [eess.IV].*

[4] Theophraste Henry, Alexandre Carre, Marvin Lerousseau, Theo Estienne, Charlotte Robert, Nikos Paragios, Eric Deutsch. Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-net neural networks: a BraTS 2020 challenge solution. *arXiv:2011.01045v2 [eess.IV] 27 Nov 2020.*

[5] Cheyu Hsu, Chunhao Chang, Tom Weiwu Chen, Hsinhan Tsai, Shihchieh Ma & Weichung Wang. Brain Tumor Segmentation (BraTS) Challenge Short Paper: Improving Three-Dimensional Brain Tumor Segmentation Using SegResnet and Hybrid Boundary-Dice Loss. *Lecture Notes in Computer Science book series (LNCS,volume 12963).*

[6] Shaoguo Cui, Lei Mao, Jingfeng Jiang, Chang Liu, and Shuyu Xiong. Automatic Semantic Segmentation of Brain Gliomas from MRI Images Using a Deep Cascaded Neural Network. *Journal of Healthcare Engineering Volume 2018, Article ID 4940593, 14 pages.*

[7] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging (Volume: 34, Issue: 10, October 2015).*

[8] Yongchao Xu, T. Géraud, I. Bloch. From neonatal to adult brain MR image segmentation in a few seconds using 3D-like fully convolutional network and transfer learning. *2017 IEEE International Conference on Image Processing (ICIP).*

[9] Yuta Tokuoka, Shuji Suzuki, Yohei Sugawara. An Inductive Transfer Learning Approach using Cycle-consistent Adversarial Domain Adaptation with Application to Brain Tumor Segmentation. *arXiv:2005.04906 [cs.CV].*

[10] Yan Shen and Mingchen Gao. Brain Tumor Segmentation on MRI with Missing Modalities. *Lecture Notes in Computer Science book series (LNIP,volume 11492).*