Segmentation of 2D Brain MR Images using Deep Neural Architectures

A summer internship report
Submitted by
Angad Ripudaman Singh Bajwa (Roll No-107118014)
B. Tech, 4th Semester
Electrical and Electronics Engineering,
National Institute of Technology, Tiruchirappalli

Under the Supervision of
Dr. Rajeev Srivastava
Professor and Head
Department of Computer Science and Engineering
IIT (BHU), Varanasi

TABLE OF CONTENTS

1.Abstract

Brain tumour segmentation is an essential task in medical image processing. Early diagnosis of brain tumours plays a crucial role in improving treatment possibilities and increases the survival rate of the patients. Manual segmentation of the brain tumours for cancer diagnosis, from large number of MRI images, is both a difficult and time-consuming task. There is a need for automatic brain tumour image segmentation. The purpose of this project is to provide an automatic brain tumour segmentation method of MRI images to help locate the tumour accurately and quickly.

2.Introduction

A cancerous or non-cancerous mass or growth of abnormal cells in the brain is known as brain tumor. Tumors may start in the brain, or cancer elsewhere in the body may spread to the brain. While brain tumours are not very common, they are one of the most lethal cancers. Four standard MRI modalities used in brain image study are T1-lighted MRI (T1), T2-lighted MRI (T2), T1-lighted MRI (T1-Gd) and Fluid Attenuated Inversion Recovery (FLAIR).

Generally, T1 images are used for distinguishing healthy tissues and in this project, I use T1 brain images for study.

Often automatic brain tumour segmentation methods use hand-crafted features such as edges, corners, histogram of gradient, local binary pattern. These features are extracted and then given to the classifier. The training procedure of the classifier is not affected by the nature of those features.
Semantic segmentation is commonly used in medical imaging to identify the precise location and shape of structures in the body, and is essential to the proper assessment of medical disorders and their treatment.
Recently, deep convolutional neural networks (CNNs) have led to substantial improvements for numerous computer vision tasks like object detection, image classification, and semantic segmentation [1].
In this project I predict the brain tumour location using one such CNN architecture- U Net. I study three types of U-Net architectures. One is the simple U-Net, the second is a U-Net with additional Res Net connections and lastly a M- Net, a new unique approach to semantic segmentation. All previous works on this dataset are for classification of tumour types. None of the previous works performed on this dataset Ire intended for segmentation purposes.

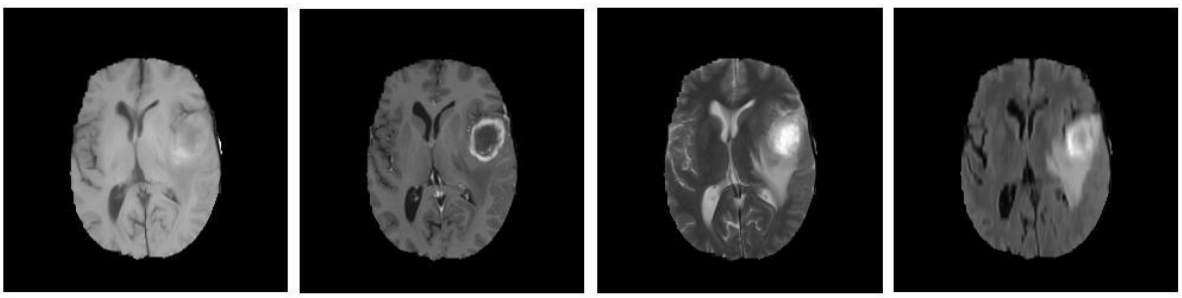

Figure 1: Four different MRI modalities showing a high grade glioma, each enhancing different subregions of the tumor. From left; T1, T1-Gd, T2, and FLAIR[2]

Challenges

Automatic segmentation of gliomas is a very challenging problem. Tumour bearing brain MRI data is a 3D/2D data where tumour shapes, size and location can vary greatly from patient to patient. Also, tumour boundaries are usually unclear and irregular with discontinuities, posing great challenge especially against traditional edge-based methods. In addition to these, brain tumour MRI data obtained from clinical scans or synthetic databases11 are inherently complex and require large device memory to perform the tumour segmentation task.

Problem Statement- To come up with an accurate method for brain image segmentation which can aid detection of brain tumours.

3.Theoretical Background and Literature Review

U- Net

The Unet architecture was introduced in the paper" U-Net: Convolutional Networks for Bio-medical Image Segmentation" in 2015. It was developed by Olaf Ronneberger for Bio-Medical Image Segmentation [3]. The Unet architecture contains two paths, first path is the contraction path (also called encoder part) which is used to capture the context of the input image. The encoder is a stack of convolutional and max pooling layers. The second path is a symmetric expanding path (also known as the decoder part) to enable precise localization using transposed convolutions. Basically, it is an end to end FCN connection but instead this only contains Convolutional
layers and does not contain any Dense layer because of which such architecture it cannot accept image of any size. In this type of

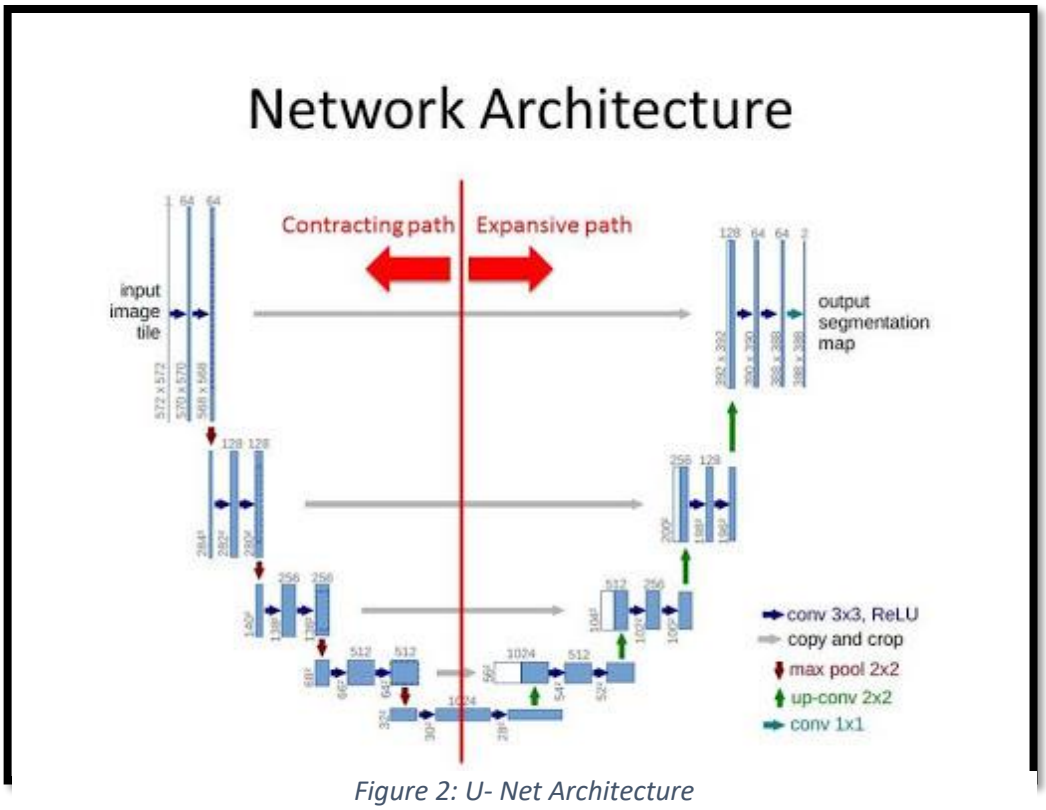

Figure 2: U- Net Architecture

architecture, 2 consecutive Convolutions layers are applied before max-pooling. This is done because when pooling is applied, lot of information is lost as the dimensions of the data are halved. So, convolution layers are stacked before each pooling, so it can build up better representations of the data without quickly losing all the spatial information[4].

U Net with skip (both short and long) connections: To the U-Net architecture I have added skip connections between the first convolution layer and the max pooling layer on each level. The main purpose of this was to increase accuracy and uniform distribution of parameters in the layers.

M Net

M-net has 4 pathways of 2D filters: two main encoding and decoding paths, and two side paths which gives the architecture functionality of deep-supervision. Each pathway has 4 steps. In the encoding path, each step has a cascade of 2D convolution filters of size 3x3 and maxpooling of 2x2, which reduces the size of input by half and allows network to learn contextual information. Skip connections are introduced to enable the network to learn better features. The decoding layer is identical to encoding layers with one exception: maxpooling is replaced by upsampling layer to double the size of input and recover an output image of original size. Similarly, skip connections are also implemented between corresponding encoding and decoding layers to ensures that the network has sufficient information to derive fine grain labelling of an image without the need for any post-processing. The left leg operates on _s with 4 maxpooling layers of size 2x2and the outputs are given as input to the corresponding encoding layers. The right leg upsamples the output of each of the decoding layers to the original size of _s. Finally, the output of the decoding layer and the right leg is processed by a1x1 convolution layer with L channels, where L is the number of structures of interest including background.
The advantage of the M-net is that all filters are 2D filters which allows end-to-end training of the network with considerably low memory requirement (~5GB)[5].

Despite their success, these models have two limitations:
(1) their optimal depth is apriori unknown, requiring extensive architecture search or inefficient ensemble of models of varying depths
(2) their skip connections impose an unnecessarily restrictive fusion scheme, forcing aggregation only at the same scale feature maps of the encoder and decoder sub-networks.

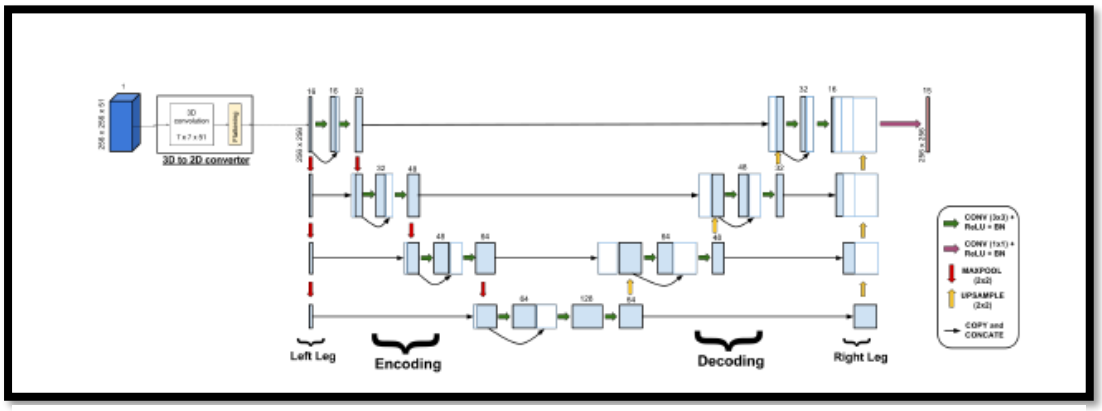

Figure 3: M- Net Architecture [5]

4. Results and Analysis

a. Dataset Description

This brain tumour T1-lighted CE-MRI image-dataset consists of 3064 slices. There are 1047 coronal images. Coronal images are those which are captured from the back of the head. Axial images, those taken from above the skull, are 990 in number. The dataset also contains1027 sagittal images that are captured from the side of the skull. This dataset has a label for each image, identifying the type of the tumour. These 3064images belong to 233 patients. The dataset includes three types of tumours- 708 Meningiomas, 1426 Gliomas, and 930 Pituitary tumours, which are publicly available in: (http://dx.doi.org/10.6084/m9.figshare.1512427).

The size of each image is 512X512 pixels[6].

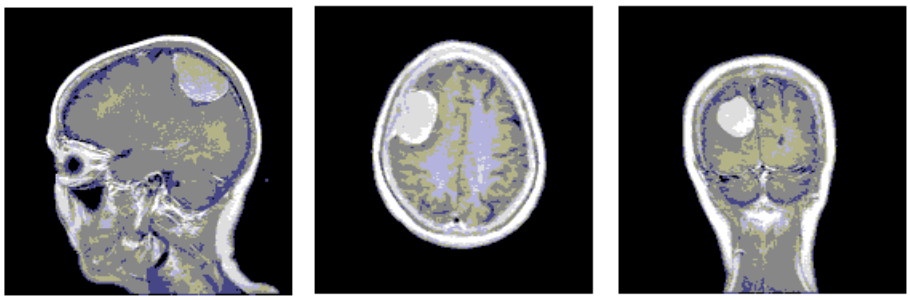

Figure 4:Brain MRI slices captured from different directions [6]

a) Axial view *b) Coronal view* *c) Sagittal view*

b. Performance Metrics

I evaluate the segmentation results using the Dice coefficient. The Dice coefficient, also called the overlap index, is a metric used for validation in medical image segmentation. The pair-wise overlap of the segmentations is calculated by:

$$\text{Dice Coefficient} = \frac{2*TP}{2*TP+FP+FN} \quad \text{…….. (1)}$$

where TP is true positive results or correctly segmented tumour pixels, FP is false positive, and FN is the false negative results of the segmentation. False positive results are obtained when a pixel which is non-tumorous is classified as tumorous. Also, FN refer to the number of pixels that are tumorous and are falsely labelled as non-tumorous[7].

Dice Loss= 1.0 – Dice Coefficient

Binary Cross Entropy loss

$$\text{Loss} = -\frac{1}{\text{output size}} \sum_{i=1}^{\text{output size}} y_i \cdot \log \hat{y}_i + (1 - y_i) \cdot \log (1 - \hat{y}_i) \quad \text{……. (2)}$$

Where y is the target value (in our case 1 or 0, tumor or non-tumor pixel) and yhat is the predicted value for that pixel. Output size is the image size in our 512X512.

I used two loss functions, one with only Binary Cross Entropy and the other a sum of Binary Cross Entropy and Dice Loss.

c.Results

I split the dataset into a training, validation and test set in the ratio of 80:10:10. Thereby having 2451 training images, 307 validation images and 306 test images. I implement the above methodology using Keras with a Tensorflow backend and the ADAM optimizer (learning rate=0.0001). The model is trained for 50 epochs.

Due to limited RAM and GPU I split the training into 2 sections. The first training was up to 30 epochs after which I saved the model checkpoint and continued training for 20 more epochs.
I use Google Colab to implement the model. Colab uses Intel(R) Xeon(R) CPU @ 2.20GHz, 12 GB of RAM and NVIDIA Tesla K80 TPU.

On comparing the three models, I find that the simple U- Net performs the best with a dice score of 0.672.

Table 1: Result Analysis and Comparison

Model	epochs	Loss_function	Dice_coef
U-Net	50	0.0577790886	0.672770553
U-Net (with skip connections)	50	0.0536935590	0.6101779341
M-Net	50	0.0764972120	0.6455184817

U-NET

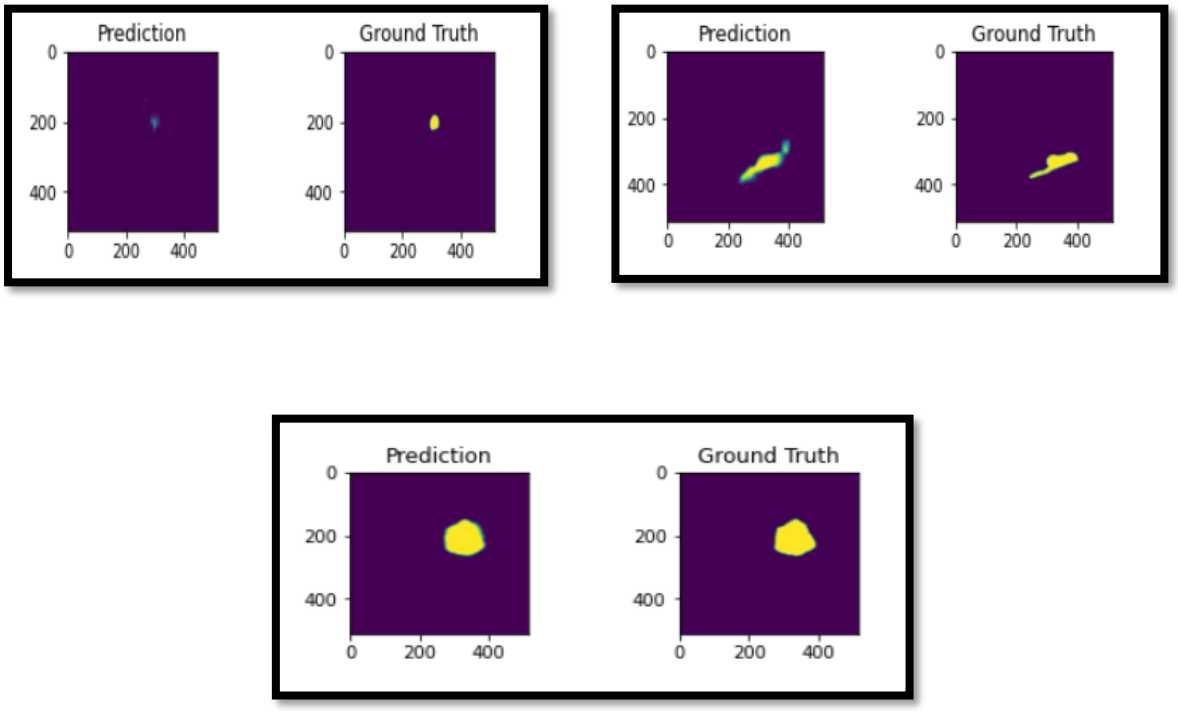

Figure 5 : U- Net: Prediction VS Ground truth for unseen data

M- Net

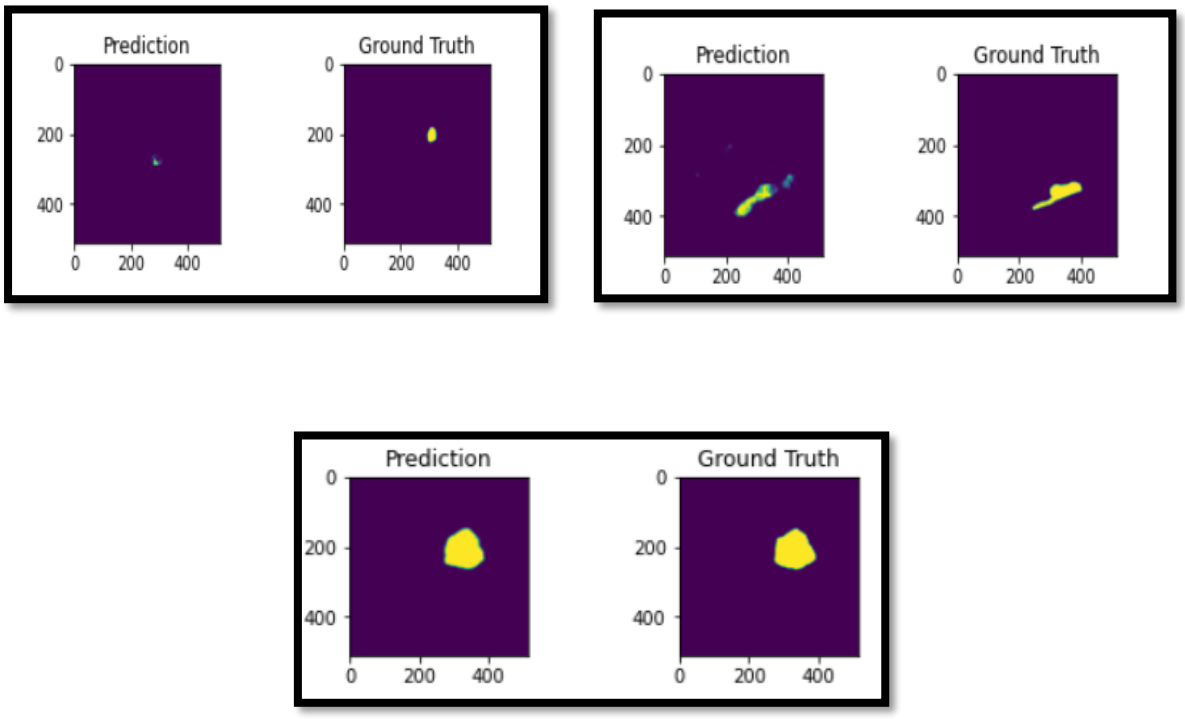

Figure 6: M- Net: Prediction VS Ground truth for unseen data

U-Net with skip connections

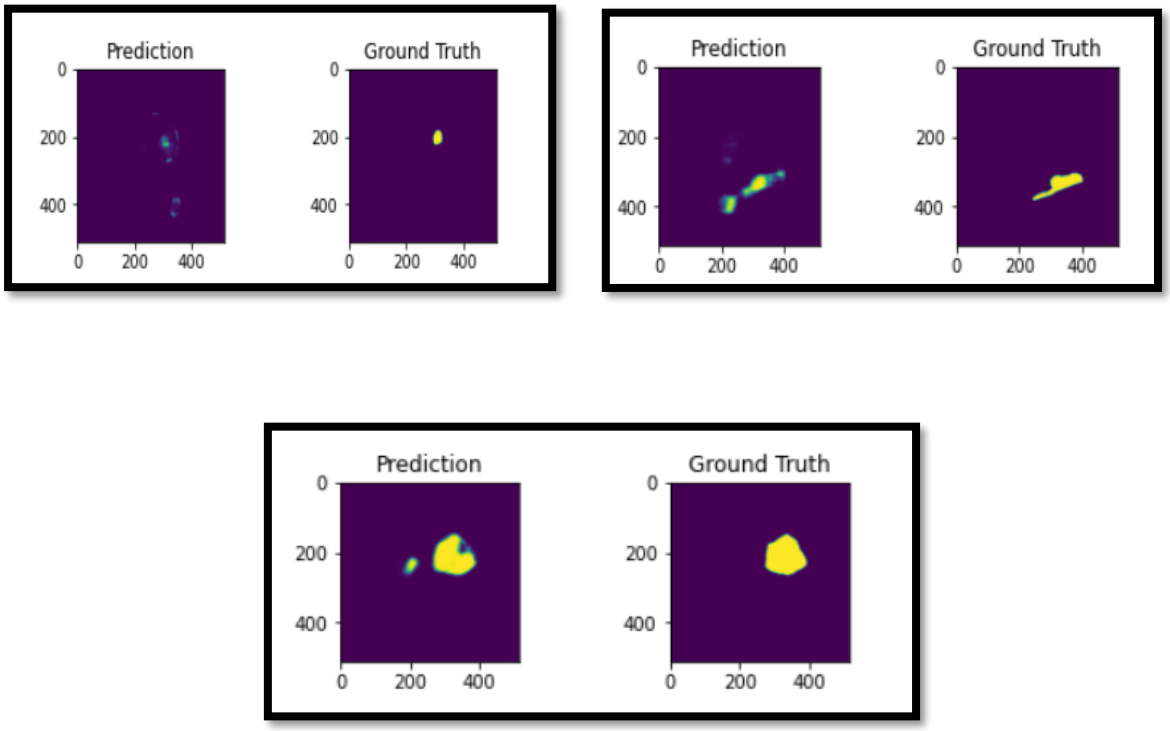

Figure 7: U- Net with skip connections: Prediction VS Ground truth for unseen data

5.Conclusion

I experimented with three different U-Net architectures and tried two loss functions to derive the best results. After testing on an unknown dataset, the simple U- Net performs the best on the given dataset. I also introduced skip connections between convolution filters and deep supervision functionality in the network which allows it to learn features better. I come up with a memory efficient way of detecting tumor's that can reduce both the time and difficulty faced in clinics. Early discovery can prevent patients from entering severe stages of the tumor.

6.References

1. Arnab Kumar Mondal, Jose Dolz and Christian Desrosiers," Few-shot 3D Multi-modal Medical Image Segmentation using Generative Adversarial Learning"
2. Ali Isin,Cem Direkoglu , MelikeSah ,"Review of MRI-based brain tumor image segmentation using deep learning methods"
3. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation,2015.
4. Malav Bateriwala and Pierrick Bourgeat," Enforcing temporal consistency in Deep Learning segmentation of brainMR images"
5. Raghav Mehta and Jayanthi Sivaswamy, "M-NET: A CONVOLUTIONAL NEURAL NETWORK FOR DEEP BRAIN STRUCTURE SEGMENTATION"
6. Zahra Sobhaninia, Safiyeh Rezaei, Alireza Noroozi, Mehdi Ahmadi, Hamidreza Zarrabi, Nader Karimi, Ali Emami,Shadrokh Samavi , "Brain Tumor Segmentation Using Deep Learning by Type Specific Sorting of Images"
7. Adrian V. Dalca1;2, Evan Yu3, Polina Golland2, Bruce Fischl1, Mert R. Sabuncu3;4, and Juan Eugenio Iglesias1;2;5," Unsupervised Deep Learning for Bayesian Brain MRI Segmentation"
8. https://github.com/zhixuhao/unet
9. https://github.com/Jack-Etheredge/Brain-Tumor-Segmentation-3D-UNet-CNN
10. https://github.com/Issam28/Brain-tumor-segmentation
11. https://github.com/IntelAI/unet
12. https://github.com/polo8214/Brain-tumor-segmentation-using-deep-learning

7.Acknowledgement

I would like to thank Professor Rajeev Srivastava for guiding and mentoring me for the entirety of this project. Sir's vast experience in research inspired me to first do a thorough research on brain image segmentation and the problems faced in it. Sir also encouraged me to explore more methods, improve upon the pre-existing methods and try new datasets.

Really grateful for being given this opportunity which has further piqued my interest in research. Thank-you for your guidance and I shall remember your inputs in my future endeavors as well.

8.Appendix

The code is available on my Github profile- https://github.com/angadbajwa23/Brain-Image-Segmentation