ML Methods for Brain Tumor Segmentation and Survival Prediction

Elliott Leow, Jennifer Ye, Jay Kim

Machine Learning, CS 601.475/675

1 Introduction

Brain tumors are among the most life-threatening forms of cancer, boasting an overall survival rate of 33.4%. Despite the numerous advances in the fields of medicine and patient care, these tumors remain extremely dangerous and require intensive treatment planning. As a result, the ability to predict a patient's survival time remains a crucial part in ensuring informed and effective decision-making for healthcare providers. However, this task is vastly complex due to the wide variability in factors found with each individual patient as well as the unpredictable nature of the brain tumors themselves. For this reason, we turn to Machine Learning methods with the hopes of tackling this daunting task and produce accurate survival rate predictions.

The Brain Tumor Segmentation (BraTS) challenge is an annually held competition in which researchers, students, and healthcare professionals can attempt to find solutions to this issue. Every year, they offer a dataset that contains brain tumor imaging information to be used for medical image analysis and prediction tasks. The focus of this project will be on the BraTS 2020 Challenge dataset, which contains pre-operative MRI scans from a variety of sources, narrowing in on gliomas which are a prominent variant of brain tumors. Our approach in particular follows an ML pipeline that integrates tumor segmentation, dimensionality reduction, and finally predictive modeling. Initially, the multi-modal MRI scans are segmented for the purpose of extracting meaningful spatial features. These higher dimensional results are then passed through a 3D autoencoder structure in order to capture the most prominent latent features, as well as reduce the overall complexity of the data. Finally, a series of methods are run on this latent space representation to enable prediction of patient survival time.

Our approach aims to explore the complex relationships between segmented tumor information and features and survival outcomes by utilizing the aforementioned methods. By leveraging the BraTS 2020 dataset, we will discuss how our model provides insights into the capabilities of Machine Learning architectures for survival prediction.

2 Methods

To motivate our implemented Machine Learning methods, we first must discuss the data presented in the BraTS 2020 set. The dataset contains brain scans taken using the Magnetic Resonance Imaging (MRI) method which create detailed images of the brain, which consequently captures a tumor as well. These images were stored in an NIfTI format, widely used for 3-dimensional medical imaging.

There are four modalities or types of MRI images found in the dataset as shown in Figure 1. These include T1 which captures the basic structure of the brain, T1 contrast which is taken post dye-injection which highlights tumor areas, T2 which highlights fluids, and finally FLAIR which focus specifically on fluid produced by tumor swelling. The variety of modalities in this dataset allow for stronger feature extraction later on. Due to the 3D nature of the inputted data, the following methodologies require functionality that allows for the use of this higher level of complexity.

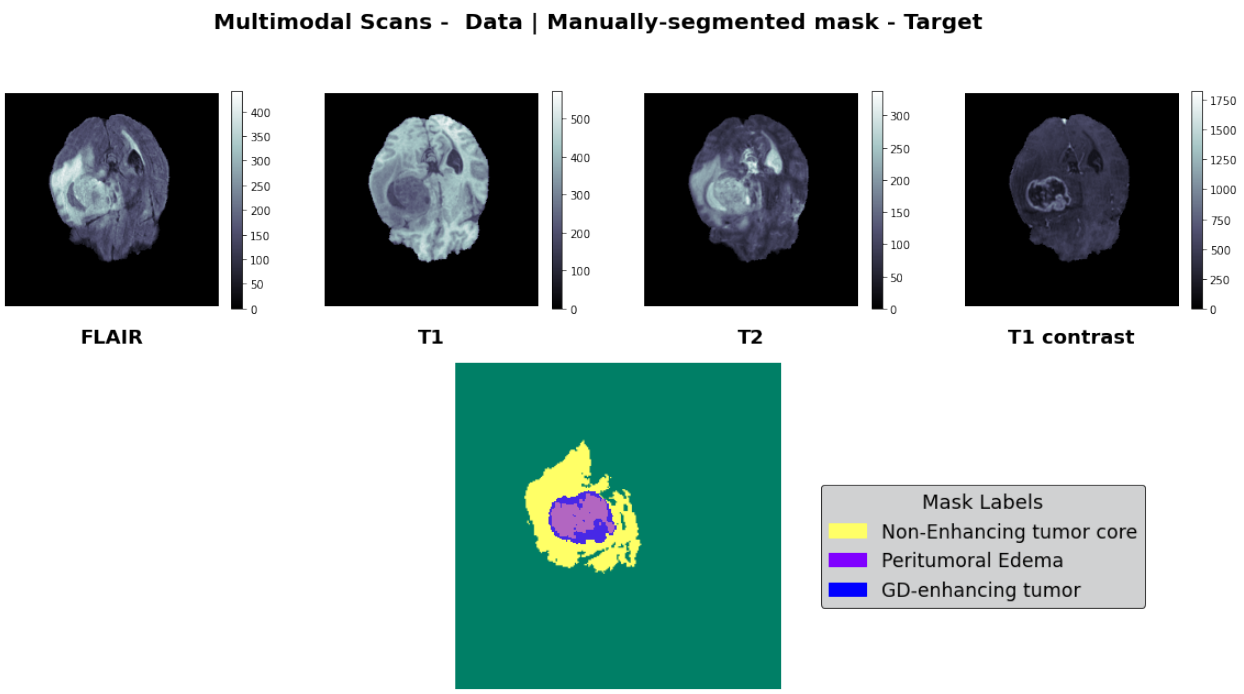

Figure 1: Displays the multimodal MRI images found in the BraTS dataset, T1, T2, T1-contrast, and FLAIR

2.1 Segmentation

The next step in our pipeline was brain tumor segmentation, which is a process in which different parts of a brain tumor are identified and labeled. Given the 3D images as shown in Figure 2 from the BraTS dataset, the brain's structure is mapped, and regions such as tumorous regions, healthy tissues, or other features are highlighted.

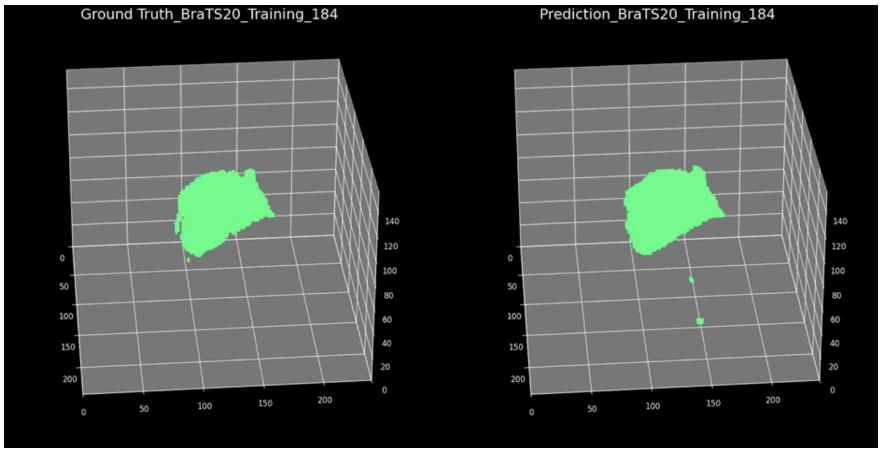

Figure 2: Visualization of the mapped out 3D volume from the BraTS 2020 dataset

The following methodology was derived from a BraTS challenge user, utilizing the popular architecture UNet based on a Convolutional Neural Network (CNN) structure specifically used for image segmentation.

For this project we used the extended 3D-UNet version to handle our previously specified requirements. 3D-Unet, following the general flow of a traditional CNN, consists of an encoder, the bottleneck, and a decoder. After inputting our 3D images into the architecture, the encoder then extracts features through a series of convolutional layers. In our case, we utilized 3 convolutional layers, and then utilized max pooling to reduce the dimensionality of the data. Through the layers, the spatial dimensions of our brain feature map continued to decrease, however the number of features themselves continued to increase which captures increasingly complex and high-level features.

2.2 Supervised Autoencoder

Autoencoders are a type of neural network architecture designed for efficient data compression and reconstruction. The compressed representation, known as the latent space, captures the most salient features as learned by the autoencoder as shown in Figure 3.

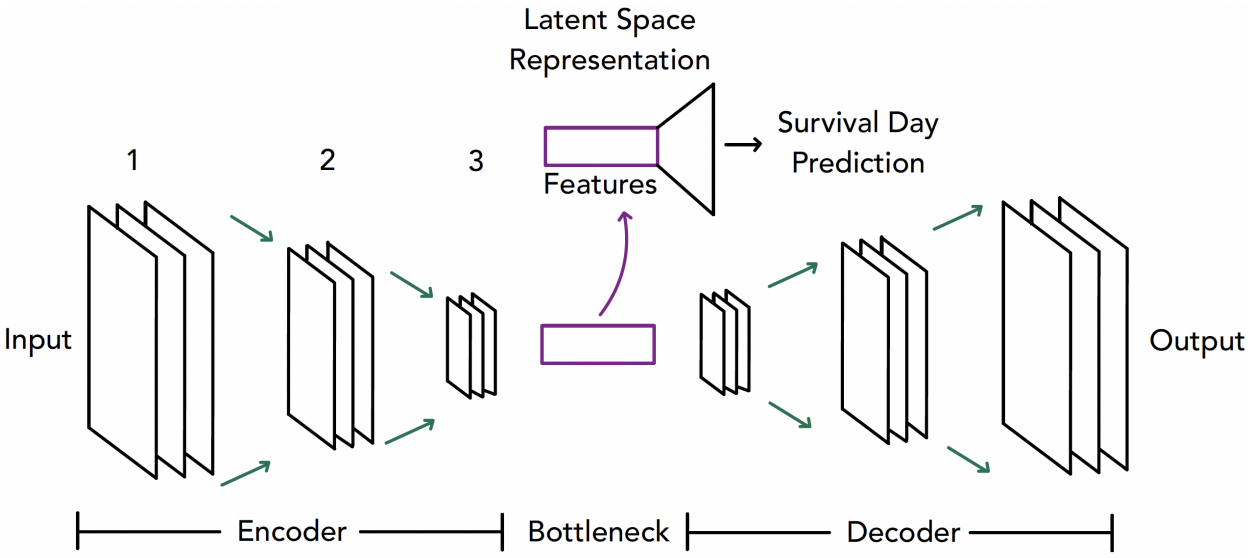

Figure 3: Basic autoencoder structure, consisting of an encoding stage, decoding stage, and bottleneck structure which represents the latent space.

Data compression is achieved by passing information through progressively smaller convolution layers, with the ultimate goal that this condensed form can accurately reconstruct the original input. Additionally, skip connections are incorporated to preserve intermediate features, ensuring that important information from earlier layers is retained and contributes to the reconstruction process.

The autoencoder accepts the 4-channel input described above. Noise was applied to the brain scan before entering the autoencoder so that the autoencoder creates a useful representation of the scan in order to reconstruct it. Then three convolution layers with group normalization are applied with ReLU nonlinearities. Each convolution layer reduces spatial dimension and allows the network to extract more complex features. After these layers, the feature map is flattened and passed through a fully-connected layer to produce latent features.

From the latent representation of the scans, a regression head was implemented that attempts to predict survival days. The Mean Squared Error loss from this task was added to the total loss and helps the latent

space become relevant in predicting survival days.

The decoder mirrors the encoding layers and produce a reconstructed scan of the brain. The Mean Squared Error loss was calculated by comparing the reconstructed image to the original image without noise and added to the total loss.

The total loss is the sum of the survival days prediction loss and the reconstruction loss.

2.3 Prediction

After generating the latent features from the autoencoder, we used them to perform regression on survival days. We compared multiple regression strategies, including Linear Regression and Neural Networks with and without PCA, and found that the **Neural Network with PCA**, trained over 100 epochs, produced the best results. The neural network comprised four hidden layers with units decreasing from 512 to 64, each using ReLU activation, along with Batch Normalization and Dropout to stabilize training and reduce overfitting. The model used the Adam optimizer and was trained with Mean Squared Error loss over 100 epochs with a batch size of 32.

3 Results

3.1 Prediction results

The training set had 369 patient samples that contained a labeled segmentation of the brain tumor and corresponding survival days of the patient. In addition to the MSE metric we previously discussed, we also calculated the Spearman correlation for our results. It describes the strength and direction of a monotonic relationship between two variables, in our case, between our y_{pred} and y_{true} for SD prediction. The higher and closer to 1 the Spearman correlation is, the stronger the relationship.

	MSE	Spearman correlation
Survival Days	159197.84	0.4865
Patient Age	182.3997	0.5351

Table 3.1: Model performance on predicting survival days and age

As shown in Table 3.1, our Patient Age score which we used as a standard for comparison was much better than the Survival Days, which is to be expected as the Survival Days prediction exhibits a much higher level of complexity. The Spearman correlations were comparable, though the Patient Age still showed a slightly higher correlation strength. We compared this loss against previous attempted methods: a standard AutoEncoder without the added supervision on survival days and using just age to predict survival days.

3.2 Comparison

	MSE
Our Proposed Method	159197.84
Patel J. et al.	152467.00
Agravat R.R. et al.	116083.48
Soltaninejad M. et al.	109564.00

Table 3.2: Our model performance compared to other published models

4 Dicussion

The imbalance in our dataset, which predominantly comprises samples with shorter survival times (600 days or fewer), complicates accurate prediction for patients with longer life expectancy and early-stage tumors. This is not surprising, as these patients are underrepresented in our sample, and their tumors are typically smaller and more challenging to distinguish from normal tissues. Addressing this imbalance is crucial for improving model performance in predicting outcomes for patients with nascent tumors. Overcoming this challenge should be a focus in future tumor research within the machine learning domain, especially if we aim to advance early tumor detection capabilities.

In addition, the complexity of our model was limited by the amount of computational power we could obtain as many well-performing models used significantly higher amounts of compute.

Future improvements on the model would involve making the segmentation of the tumor more detailed. Doing this would make the extraction of higher level features other than voxel statistics easier. In addition, the structure of the autoencoder could be further optimized and expanded by adding further layers and possibly investivating other autoencoder architectures such as variational autoencoders.

5 References

1. Brain tumor segmentation (brats) challenge 2020: Scope. Brain Tumor Segmentation (BraTS) Challenge 2020: Scope — CBICA — Perelman School of Medicine at the University of Pennsylvania. (2020, May 18). https://www.med.upenn.edu/cbica/brats2020/

2. 3D-unet medical image segmentation for tensorflow: Nvidia NGC. NVIDIA NGC Catalog. (n.d.). https://catalog.ngc.nvidia.com/orgs/nvidia/resources/unet3d$_$medical$_$for$_$tensorflow

3. Agravat, R. R., & Raval, M. S. (1970, January 1). 3D semantic segmentation of brain tumor for overall survival prediction. SpringerLink. https://link.springer.com/chapter/10.1007/978-3-030-72087-2$_$19

4. Patel, J., Chang, K., Hoebel, K., Gidwani, M., Arun, N., Gupta, S., Aggarwal, M., Singh, P., Rosen, B. R., Gerstner, E. R., & Kalpathy-Cramer, J. (1970, January 1). Segmentation, survival prediction, and uncertainty estimation of gliomas from multimodal 3D MRI using selective kernel networks. SpringerLink. https://link.springer.com/chapter/10.1007/978-3-030-72087-2$_$20

5. Polomarco. (2020, November 17). Brats20_3dunet_3dautoencoder. Kaggle. https://www.kaggle.com/code/polomarco/brats20-3dunet-3dautoencoder/notebook

6. Soltaninejad, M., Pridmore, T., & Pound, M. (1970, January 1). Efficient MRI brain tumor segmentation using multi-resolution encoder-decoder networks. SpringerLink. https://link.springer.com/chapter/10.1007/978-3-030-72087-2$_$3

7. Huang, H., Zhang, W., Fang, Y., Hong, J., Su, S., Lai, X. (2021, August 18). Overall survival prediction for gliomas using a novel compound approach. Frontiers in oncology. https://pmc.ncbi.nlm.nih.gov/articles/PMC8416476/#B42