Multi-Task Neural Network for Brain Tumor Segmentation and Survival Rate Prediction on BraTS Dataset

Cecchin Andrea
andrea.cecchin.5@studenti.unipd.it

Dugo Alberto
alberto.dugo@studenti.unipd.it

Soldà Matteo
matteo.solda.1@studenti.unipd.it

Abstract

This paper presents a unified deep learning framework based on the U-Net architecture for the simultaneous tasks of brain tumor segmentation and survival prognosis. Utilizing the publicly available BraTS Challenge dataset for training and evaluation, the proposed model leverages multimodal MRI data to generate precise segmentations of tumor regions and to predict patient survival duration. By addressing these two critical challenges in clinical oncology within a single architecture, the approach demonstrates the potential of integrated deep learning solutions for comprehensive and clinically relevant analysis of brain tumors.

1. Introduction

CBTRUS[5] maintains and regularly updates a registry of malignant and non-malignant brain tumors. In global data, the incidence of malignant brain tumors is approximately 3.5 per 100,000 people. This incidence increases to 3.9 for the male population and 3.1 for the female population. This indicates that approximately 173,699 men and 148,032 women are diagnosed with brain tumors each year, for a total of 321,731 individuals. The average annual mortality in the United States alone between 2017 and 2021 was 4.41 per 100,000 people. This represents a death of 17,411 people per year, or 48 deaths per day.

Brain Tumor Segmentation Challenge is an annual competition, firstly launched by the Perelman School of Medicine[4] in 2012, focused on evaluating the performance of algorithms in tasks related to the segmentation of brain tumors in multimodal MRI scans. In particular, our work focused on the following tasks proposed by the competition:

- Segmentation of Gliomas: it is the task of recognizing a tumor region, if present. It requires the segmentation of the three typical regions which characterize a brain tumor:
  - Enhancing Tumor: it is the region of a brain tumor characterized by the presence of a disrupted blood-brain barrier.
  - Peritumoral Edema: it is the swelling of tissue surrounding a brain tumor, caused by the accumulation of fluid.
  - Necrotic and Non-Enhancing Tumor Core: it is the central region of a brain tumor that contains dead tissue.
- Prediction of Overall Survival: this task involves predicting the survival rate of patients based on their MRI scans.

Being able to correctly identify and classify tumor areas is a highly important task, because it increase the survival rate of patients. Despite that, those tasks are extremely challenging. The difficulty of these tasks, and the reason why tumor segmentations are performed by human experts, lies in the challenge of visualizing certain tumor areas in MRI scans.

Indeed, while peritumoral edema is easily visible even without image preprocessing or contrast modification, the Enhancing Tumor is only visible in post-contrast scans, where it appears brighter. Even more difficult is the segmentation of the Necrotic and Non-Enhancing Tumor Core, which does not show contrast enhancement in MRI scans.

Modeling and training a deep neural network to perform these tasks would decrease the time required to accurately identify tumor regions in MRI scans, potentially increasing the accuracy of human-made segmentations. Additionally, predicting survival rates could be useful in defining an irreversible clinical condition, where tumor segmentations and surgery attempts would be futile.

2. Related Works

The most cited work about tumor segmentation made through Deep Learning Networks is the one wirtten by Havaei et Al.[2] that presents a Convolutional Neural Network that exploits local features and some global contextual features at the same time. This specific model is implemented through a fully connected layer which allows a 40 fold speed up and a 2-phase training procedure. The overall dice score of this model is of 0.86 (complete regions), 0.86 (core regions) and 0.77 (enhancing tumor regions). There is also an article made by Gordillo et Al.[1] that explore the state of the art of MRI brain tumor segmentation, comparing various models.

3. Dataset

For this project, we used the official dataset[3] provided for the BraTS Challenge held in 2020, obtained from the Kaggle dataset repository. The data it contains, consisting of the combination of 19 different smaller datasets created using various clinical protocols and scanners, include MRI scans from 494 patients.

For each patient, four different three-dimensional brain scans are provided. T1-weighted MRI is a type of scan that emphasizes the anatomical details of the brain. It provides good contrast between gray matter and white matter, making brain structures clearly visible. T1 contrast-enhancing, on the other hand, is a T1 image acquired after the administration of a gadolinium-based contrast agent. The contrast accumulates in areas with high vascular permeability, such as tumors, which absorb the contrast and appear brighter than the surrounding tissues. The T2 scan highlights water-rich areas, such as edema, while T2-FLAIR, or simply FLAIR (Fluid Attenuated Inversion Recovery), is a variant of T2 in which the cerebrospinal fluid signal is suppressed, allowing better visualization of brain lesions.

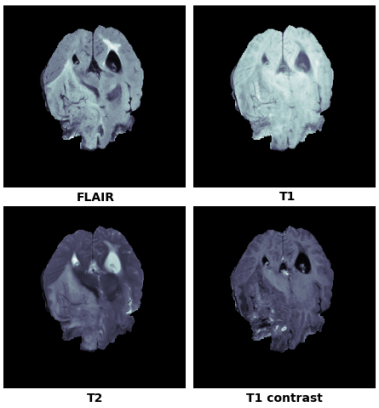

Figure 1. Example of an MRI scan with FLAIR, T1, T2 and T1 contrast-enhancing.

The images provided in the dataset have been preprocessed to facilitate their use. Specifically, all scans have been realigned and resampled to the same voxel size of 1 mm^3. Additionally, all non-brain tissues, such as the skull and scalp, have been removed.

For all patients in the training set, totaling 369, manual segmentations of tumor areas were created, verified, and validated by expert neuroradiologists. These segmentations were used during the training and validation phases of the model as the "ground truth" for the segmentations produced by our neural network.

Figure 2. A colored version of the manual segmentation provided for a patient in the training set.

To complete the dataset, a list of all patients who did not survive the tumor was provided, including their age at the time of death and the number of days from the MRI scans to that day.

4. Method

Our model implements a customized 3D U-Net architecture specifically designed for volumetric brain tumor segmentation in multimodal MRI scans.

4.1. Architecture

We developed an improved 3D U-Net variant with an exceptionally lightweight structure, optimized for efficient tumor subregion delineation. The network is

designed to be as compact as possible while preserving high segmentation accuracy. It comprises a total of 340,596 parameters, corresponding to approximately 1.30 MB of memory during inference (assuming 32-bit floating-point precision). The architecture features three resolution levels in the encoder path and a corresponding symmetric decoder path, connected via skip connections to preserve spatial information. The encoder is composed of sequential blocks that capture hierarchical features through $3 \times 3 \times 3$ convolutions, followed by Instance Normalization and LeakyReLU activations. Downsampling between resolution levels is implemented using 3D max pooling with a stride of 2. Each encoder block adopts a residual-like structure, with dual convolutional layers that facilitate improved gradient flow during training. The decoder path restores spatial resolution using 3D transposed convolutions (kernel size 2, stride 2) for upsampling. Features from the encoder are concatenated with the upsampled representations via skip connections, effectively combining low-level spatial details with high-level semantic features. Dynamic padding is applied to ensure precise spatial alignment during concatenation. To mitigate overfitting, a dropout layer with a rate of 0.3 is inserted prior to the final $1 \times 1 \times 1$ convolutional layer, which maps feature representations to the output segmentation classes. Model weights are initialized using Kaiming initialization, which is well-suited for networks employing LeakyReLU activations. The network produces probability maps for four target classes (background and three tumor subregions), while preserving the original spatial dimensions of the input volume. This architecture achieves a strong balance between computational efficiency and segmentation performance, making it well-suited for clinical deployment, especially in resource-constrained environments.

4.2. Input Preparation

Our model processes the full 3D volumetric MRI scans from the BraTS 2020 dataset. Each input sample consists of four co-registered modalities (T1, T1ce, T2, and FLAIR) with dimensions $4 \times 240 \times 240 \times 155$, where the first dimension represents the input channels. Prior to feeding the data into our network, we applied intensity normalization independently for each modality using z-score standardization, ensuring each modality had zero mean and unit variance within the brain region.

To address memory constraints while preserving the volumetric context essential for accurate segmentation, we implemented patch-based processing during both training and inference.

4.3. Loss Function

5. Experiments

All text must be in a two-column format.

6. Conclusions

All text must be in a two-column format.

References

[1] Nelly Gordillo, Eduard Montseny, and Pilar Sobrevilla. State of the art survey on mri brain tumor segmentation. Magnetic Resonance Imaging, 31(8):1426–1438, 2013.

[2] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. Medical Image Analysis, 35:18–31, 2017.

[3] Kaggle. Kaggle brats2020 dataset.

[4] Perelman School of Medicine University of Pennsylvania. Brain tumor segmentatation (brats) challenge.

[5] CBTRUS Central Brain Tumor Registry of the United States. Cbtrus fact sheet.