Post-Treatment Brain Tumor Segmentation using MONAI SegResNet Model

William W. Winters
MSDS 692: Data Science Practicum 1
Anderson College of Business and Computing
Regis University, Denver, CO, USA
wwinters@regis.edu

Abstract

Determining the effectiveness of brain tumor treatments from post-treatment MRI images is an area where little research has been conducted. Current studies primarily focus on tumor detection, where MRI images clearly distinguish between healthy and tumorous tissue. The MR Images used in these studies clearly delineate between healthy and tumor images. For these cases, a Convolutional neural network (CNN) can be used to effectively classify these images as tumor-positive or negative. However, post-treatment imaging presents a challenge due to the obscuration of tumor cavities by edema and disrupted blood-brain barriers (BBB). As a result, CNNs trained on well-defined tumor images often perform poorly on post-treatment images, leading to incorrect classifications.

This study aims to address this limitation by training a **SegResNet** from MONAI on a BraTS-style dataset for image segmentation. The trained model will be applied to a post-treatment dataset of glioma patients to segment the images and determine the presence or absence of tumor tissue. Accurate segmentation will enable the assessment of residual tumor tissue, which is crucial for developing informed treatment plans. By focusing on this specific area, this research seeks to improve the accuracy of post-treatment brain tumor diagnosis and follow-up care.

I. Introduction/Background

Despite a wealth of research on detecting brain tumors using MR Images, there is a notable gap in the literature on assessing the effectiveness of treatment through post-treatment imaging. Current studies primarily focus on tumor detection where MR Images clearly distinguish between healthy and tumorous tissue. Using these images, a Convolutional neural network (CNN) can be used to effectively classify them as tumor-positive or negative. However, post-treatment imaging presents a challenge due to the murkiness of tumor cavities caused by edema, necrotic tissue, and disrupted blood-brain barriers. CNNs trained on well-defined tumor images often perform poorly on post-treatment images, leading to incorrect classifications. The main issue with CNN classification is that it does not distinguish between the different types of tissue displayed in an MR Image. A better solution is to apply segmentation to the post-treatment image since it does label the different tissue types identified in the image. One of the key advantages of segmentation is the ability to delineate tumors into specific areas such as necrotic core (NCR), enhancing tumor (ET), and peritumoral edema (ED) [1]).

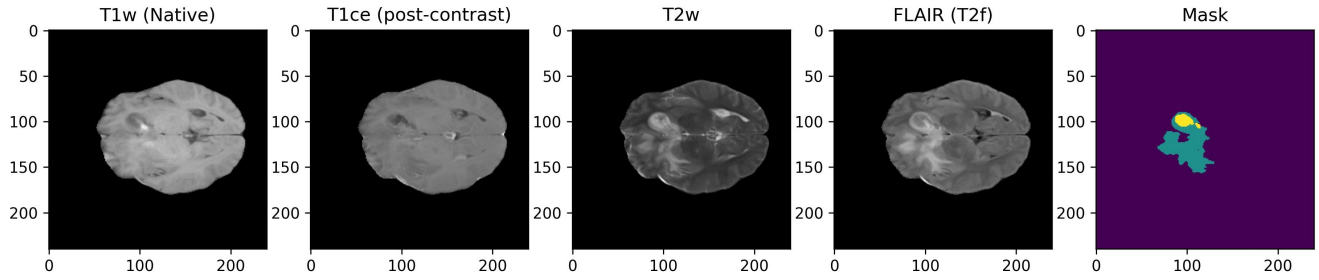
When a brain MRI is conducted, four common modalities are used. Each one highlights different anatomical features of the brain and tumor. The table below outlines them

TABLE I
Brain MRI Modalities (Sequences)

SEQUENCE	KEY USE
T1w (Native)	Reveals the structure and composition of various tissue types in the brain
T1ce (post-contrast)	A contrast agent is injected into the patient which improves the visibility (intensity) of abnormalities
T2	Highlights fluid content within brain tissue
FLAIR (T2f)	Suppresses the effects of cerebrospinal fluid in the brain to better delineate the peritumoral edema and infiltrative tumor edges [1]

The figure below illustrates how the four modalities are combined and color-coded by segmentation to assist with diagnosis. These images are from a post-treatment *glioma* patient.

Fig. 1. Post-treatment Glioma MR Image



Each sequence (modality) highlights different structures of the tumor cavity. These structures are then combined into the mask image and color-coded to aid with diagnosis. In the "Mask" image, the blue area surrounded by yellow is the core tumor (CT); yellow indicates disruptions to the blood-brain barrier and may indicate necrotic or scar tissue. Green is edema (swelling). While the displayed images are 2D, the MR Images are often stored as 3D or 4D images and include additional planes. The table below explains the viewing planes.

TABLE II
MR IMAGE VIEW PLANES

PLANE	**EXPLAINATION**
Axial	*Represents "slices" of the body*
Sagittal	Images that are taken perpendicular to the axial plane
Coronal	Images that are perpendicular to the sagittal plane

These planes provide sufficient information for the segmentation model to form a 3D or 4D understanding of the region of interest (ROI).

II. PROBLEM STATEMENT

The use of convolutional neural networks (CNNs) for brain tumor detection has been explored in various studies, often relying on a specific labeled dataset comprising clearly defined healthy and tumorous MR images. However, a significant gap exists in these studies: they primarily focus on detecting brain tumors but provide limited support for radiologists in identifying tumor tissue in post-treatment images. In addition, the author attempted to use a CNN model trained on the images described above to classify unlabeled post-treatment MR Images of various patients without success. The model incorrectly classified tumor cavities as tumors, failing to distingush between the different tissue types within the cavity.

This limitation highlights the need for an alternative approach. Further research suggests that image segmentation could be a viable solution, enabling the mapping and color-coding of different tissue types within the cavity. This capability would, in turn, facilitate the detection of brain tumor tissue in post-treatment images.

III. RELATED WORK

Over the past few years, there has been significant growth in research on brain tumor quantification using magnetic resonance imaging (MRI) technology. Many of these studies have focused on developing state-of-the-art algorithms for tumor segmentation, often adapting techniques originally designed for other medical applications. While these methods have shown considerable accuracy in segmenting brain tumors,

the field has progressed to develop algorithms specifically tailored for this purpose. A key benchmark for these developments has been the work of the Multimodal Brain Tumor Image Segmentation (**BraTS**) organization, which has established a standard for evaluating brain tumor segmentation algorithms [2].

A. MONAI

Building on this foundation, the Medical Open Network for Artificial Intelligence (MONAI) has developed a suite of PyTorch-based models and methods to facilitate the segmentation of brain tumor images. According to MONAI, the application of artificial intelligence (AI) in healthcare holds great promise for improving the detection and diagnosis of diseases, including brain tumors. By automating and enhancing the analysis of medical images, AI can help enable more positive treatment outcomes [3]. Many of MONAI's models are built on top of the U-Net architecture, which was designed to address the limitations of traditional CNNs in biomedical image processing. Traditional CNNs used for image segmentation would have to examine each pixel in an image and predict its class label, a computationally and time intensive process. In contrast, U-Net takes a different approach: instead of predicting the class label of each individual pixel, it makes predictions based on regions of interest [4]. This approach enables more efficient and effective image segmentation.

While MONAI has made significant progress in developing U-Net-based models for brain tumor segmentation, other techniques are also worth exploring. For instance, Jin et al. [5] proposed applying a transformer model, typically used in natural language processing (NLP) applications, to this task. Their approach utilized a modified Swin Transformer and a patch acquisition strategy, which yielded acceptable results in improving segmentation performance. The authors further investigated this by suggesting a hybrid combination of Convolutional neural networks (CNNs) and transformers for image segmentation [5]. Another approach found in the literature is to use transfer learning. This method trains a model on labeled data and the model is used to extract features from the target dataset which are then classified using another CNN [6]. A common thread among the techniques described so far is that they all require labeled data for training. In other words, the training dataset must be manually segmented, typically by a qualified radiologist. Although pre-trained models are available, they too rely on labeled data. An alternative approach that has been explored in the literature is self-supervised learning (SSL), which may offer a solution to this problem.

B. Self-Supervised Learning

Recent breakthroughs in contrastive learning methods have significantly reduced the computational resources required for self-supervised learning. A notable approach is the algorithm proposed by Caron et al. [7], which leverages contrastive learning without the need for pairwise comparisons. This method, called SwAV, simultaneously clusters data and enforces consistency between cluster assignments generated from different augmentations (or views) of the same image. The results show promise and may lead to the elimination of manually labeling image segmentation regions by qualified radiologists. Another advantage of applying SSL models to image segmentation is that it mitigates human error in the labeling process. Research has shown that factors such as high workloads and bias can lead to mis-classification of images. For example, a study by Zang et al. [8]. found that diagnostic errors occur in 3.5% to 4.5% of cases. Similarly, Kasalak et al. [9] reported approximately the same error rate as Zang in real-time radiology practices, attributing them to heavy workloads. These mislabeled training data can compromise the accuracy of a model. By leveraging SSL models and reducing reliance on labeled images, the accuracy of image segmentation in brain tumor detection may significantly improve. The use of SSL in medical image segmentation is a domain that warrants further study.

IV. METHODOLOGY/APPROACH

This study conducted a intensive review of literature, selected a method to apply, trained a segmentation mode, then used the model to segment a post-treatment MR Image volume.

A. Background

Many studies utilize brain tumor MR Images from Kaggle.com, chosen for their accessibility and widespread citation in current literature as publicly available source of brain MRI datasets. However, concerns regarding the accuracy of image labels have been raised within the research community. To ensure data reliability, the author limited his analysis to images sourced from credible, clinically validated medical imaging repositories. The shift to high-quality, clinically acquire MR Images introduced new challenges, since such images are typically stored in the NIfTI format as 3D volumetric data. Consequently, additional research was required to understand and process this file format effectively. All of the images used in this study can trace their origins to the Cancer Imaging Archive (TCIA) the only difference between the downstream and origin volumes is how the modalities are represented. Images directly acquired from TCIA are not processed and each modality is stored in its own NIfTI volume. MSD images, on the other hand, have been processed and the modalities have been stacked into a single NIfTI volume. After identifying data sources a file storage strategy was developed and implemented. Over 140GB of NIfTI volumes were downloaded from TCIA and MSD. These were stored in two different repositories.

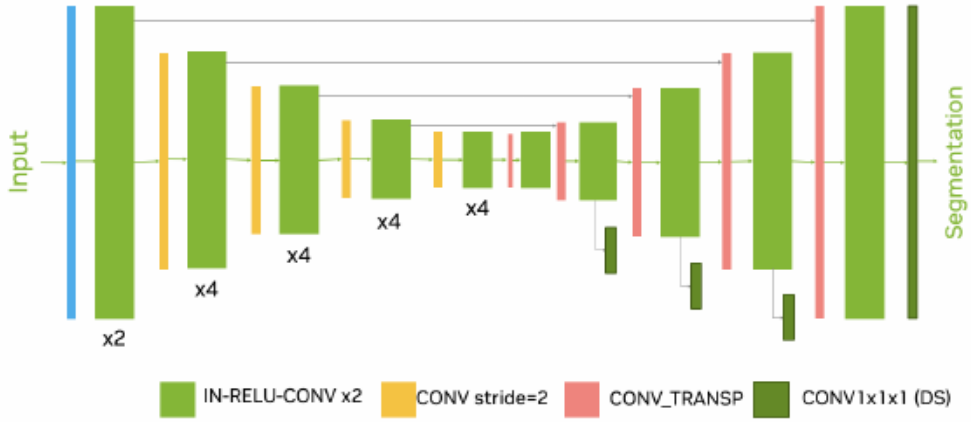
B. Model Selection

A review of literature was conducted to help select a model to train and preform brain tumor MR Image segmentation. MONAI has a collection of modern models that are optimized for image segmentation that are based on the N-Net architecture and compatible with PyTorch's framework. In addition, the MONAI site has image segmentation tutorials using its models and processes. A 2026 study singled out MONAI's SegResNet model as providing adequate accuracy with smaller datasets [10]. In addition, it has the following characteristics which makes it a good fit for this study:

- It is a residual encoder/decoder CNN designed specifically for 2D/3D medical segmentation
- Uses less GPU resources than a transformer-style model making it a good fit for training with consumer grade GPUs
- it is a top performer in challenges like BraTS
- It is part of the vast MONAI ecosystem

The figure below visually describes the SegResNet architecture

Fig. 2. SegResNet Network Architecture. The network uses repeated ResNet blocks with Instance normalization [11]



C. Model Development

The MONAI tutorial was closely followed with a few modifications to enhance its ability to effectively train with the MSD dataset. Mainly, an early-stopping mechanism was added to stop the training loop when validation loss bottomed out for a set number of epochs. When processing NIfTI volumes there are four input channels the model has to deal with. Namely, T1w, T1ce, T2w, and FLAIR. All of these

channels are used for segmentation. In addition, there are three planes of view to consider. For this study, only the axial plan was analyzed; however, other planes could have been added to aid in computing the volume of the different tumor structures that are segmented. Under the BraTS labeling standard the following output channels can be segmented and labeled. These include:

TABLE III
BRATS STANDARD SEGMENTATION LABELS

LABEL	DESCRIPTION
0	normal brain tissue (non-tumor/background)
1	necrotic/non-enhancing tumor core (TC)
2	edema (ED)
3	not used to avoid confusion with binary encoders
4	enhancing tumor (ET) [1]

For this study the output labels were collapsed into three output classes: TC (tumor core), WT (whole tumor), and ET (enhancing tumor).

D. Measuring Model Performance

The Sørensen-Dice similarity coefficient (DCS) also known as the Dice score is a measure of similarity between two sets and is the most common metric used for image segmentation with deep learning models. As a measure of the amount of relative overlap, the DSC fits the intuitive idea of a measure of similarity. Its interpretation is simple. The value of 1 means perfect resemblance and 0 means that there is no overlap. However, it also gives the same weight to elements outside of the ROI. In other words, only the number of misplaced pixels or voxels matter, not their location. Due to this limitation it may not be ideal for every situation [18]. To measure the performance of the model used in this study, the DCS was used. It measured training performance and the model's ability to make correct predictions when compared to a *ground truth* when using the model for inference segmentation.

V. DATA DESCRIPTION

The training data used in this study were obtained from the **Medical Segmentation Decathlon (MSD)**, which includes a subset of data from the BraTS (Brain Tumor Image Segmentation) challenges (specifically, the BraTS 2016–2017 datasets) [12]. The dataset comprises 750 multi-parametric MR brain scans from patients with either glioblastoma (GBM) or lower-grade glioma (LGG). Each scan includes the following MR sequences: native T1-weighted (T1), T1-weighted post-gadolinium enhancement (T1ce), T2-weighted (T2), and T2 Fluid-Attenuated Inversion Recovery (T2-FLAIR or T2f). All samples were co-registered to a reference atlas space using the SRI24 brain structure template, resampled to isotropic voxel resolution of 1 mm^3, and skull-stripped using various methods followed by manual refinements. All annotations for the tumor sub-regions were approved by expert board-certified neuroradiologist [13]. 484 volumes and their labels comprised the training dataset and the validation dataset used contained 266 volumes. This dataset can be downloaded from the *Medical Segmentation Decathlon* Google Drive or AWS S3 bucket [14]. All brain tumor MR Image files downloaded from BraTS or the Medical Segmentation Decathlon are in NIfTI format. NIfTI stands for the *Neuroimaging Informatics Technology Initiative* and can be recognized by their file ending of `*.nii` or `*.nii.gz`. The format used by NIfTI files store metadata and images in a single package [15].

Since the BraTS challenge is specifically focused on brain tumor segmentation, the organizers of the Medical Segmentation Decathlon aimed to establish an international competition that addresses medical image segmentation across a broader range of anatomical structures and pathologies – not limited to a single disease [16]. The Decathlon provides multi-modal MR images of various organs, including the heart, liver, hippocampus, prostate, and others.

VI. Expected Outcomes

A **SegResNet** model will be trained to perform brain tumor segmentation on a BraTS-derived dataset. Following training, the model will achieve clinically acceptable segmentation accuracy on the validation set, as evaluated by standard metrics (e.g., Dice coefficient, Hausdorff distance). The trained model weights will be saved and applied to segment post-treatment glioma cases. This post-treatment segmentation can help quantify residual or recurrent tumor burden, enabling more objective assessment of treatment response and supporting evidence-based clinical decision-making in neuro-oncology.

VII. Timeline

TABLE IV
Project Timeline

Phase 1	Timeframe	Description
Collect and Process Images	Week 1	Identify brain MRI sources and download datasets from them. Develop directory structure to store images for easy use
Research and Review of Literature	Week 2	Limit review of literature to peer-reviewed articles and develop an annotated bibliography
Select Model and Method	Week 3	Based on the review of literature select the model and method to use for brain tumor segmentation
Code and Train Model	Week 4	Create model using Jupyter-lab, Pytorch, and Python
Evaluate Results	Week 5	Evaluate model's performance using loss and Dice scores
Fine Tune Model	Week 6	Based on loss and dice scores find tune the model
Complete Paper and Present Findings	Week 7	Pull it all together

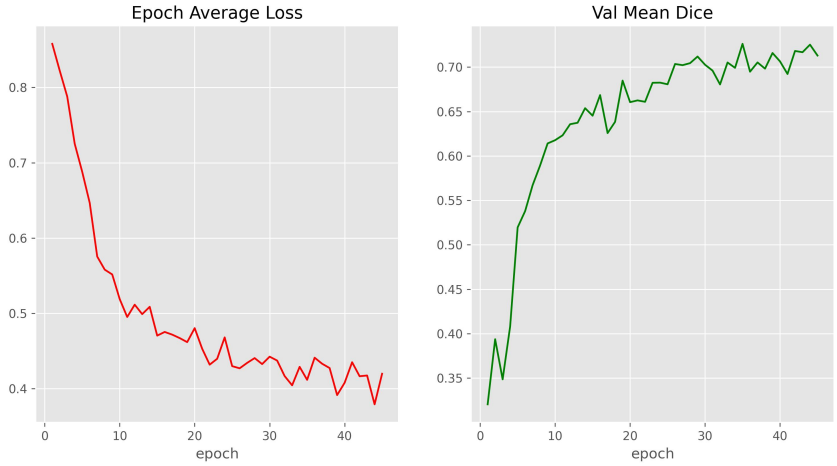
VIII. Conclusion

While substantial research exists on detecting brain tumors using MRI, there remains a gap in the literature regarding the assessment of treatment response through post-treatment imaging. In this study, I trained a segmentation model using a BraTS-style dataset and applied it to segment post-treatment MRI scans from a glioma post-treatment dataset available in the Cancer Imaging Archive (TCIA) [17]. Segmentation performance was evaluated using the standard Dice similarity coefficient (DSC).

A. Training and Validation

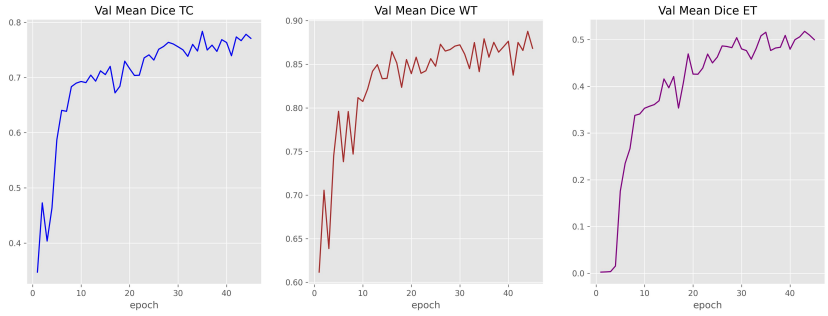
Like other CNN training scenarios training and validation datasets were used during the model's training process. Segmentation uses different metrics to gauge model performance. Validation loss and Dice coefficient are tracked and recorded. In addition, the Dice score for each region of interest (ROI) is tracked and used for evaluation. This figure below displays the average validation loss and mean Dice score for each epoch.

Fig. 3. Validation Loss and Dice Scores per Epoch



Furthermore, the Dice score for tumor core (TC), whole tumor (WT), and enhancing tumor (ET) were also recorded for each epoch and indicates which ROIs segmented accurately and which ones did not.

Fig. 4. Validation Mean Dice Score for TC, WT, and ET



Average validation Dice scores are described in the following table.

TABLE V
AVERAGE DICE SCORE DURING TRAINING

METRIC	VALUE
Mean Dice Score	0.74
TC Mean Dice Score	0.80
WT Mean Dice Score	0.89
ET Mean Dice Score	0.54

The Dice coefficient – also known as the Dice similarity coefficient (DSC) – is the most widely used metric for evaluating image segmentation performance. It quantifies the overlap between a predicted segmentation and the corresponding ground truth annotation. Ranging from 0 (no overlap) to 1 (perfect overlap), higher values indicate better agreement between the predicted and reference segmentation. As with many performance metrics in model training and evaluation, a value closer to 1 signifies superior segmentation accuracy for the region of interest (ROI) [18]. In BraTS competitions the DCS is computed for each ROI class (WT, TC, ET). For WT a score of ≥ 0.85 is considered *good*, a score of ≥ 0.80 is *good* for TC, and a score of ≥ 0.75 for ET is *good* for ET [18].

B. Inference

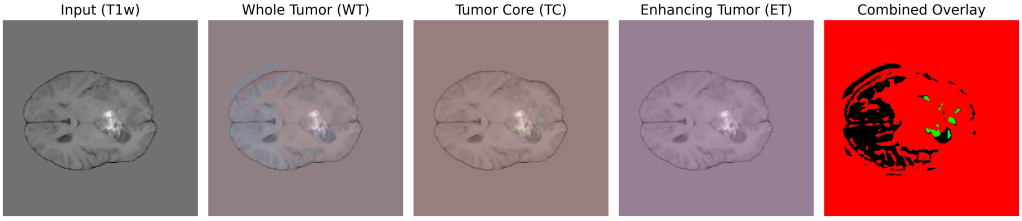
A representative image was selected at random from a post-treatment *glioma* patient dataset sourced from The Cancer Imaging Archive (TCIA). Semantic segmentation was performed on this image using the model's best-performing training weights. The ground truth segmentation—manually delineated and provided as a NIfTI file within the dataset—was used to evaluate the model's output. The Dice Similarity Coefficient (DSC) was computed between the model's predicted segmentation and the ground truth mask. Summary metrics, including the DSC, are listed in the table below.

TABLE VI
INFERENCE DICE SCORES ON POST-TREATMENT IMAGE

METRIC	VALUE
Mean Dice Score	0.49
TC Mean Dice Score	0.62
WT Mean Dice Score	0.73
ET Mean Dice Score	0.12

Typically, the ET segment yields low scores in this type of inference, likely due to its small size. For reference, a figure is provided below with the inferred segmentation overlaid on a T1n image. To assess the accuracy of the model's predictions, a ground truth is necessary, which allows for the calculation of the Dice Similarity Coefficient (DSC).

Fig. 5. Inference Output Labels from MR Image



C. Further Research

This study trained a segmentation model using a publicly available architecture and applied it to a post-treatment MRI, achieving reasonable segmentation accuracy—a result consistent with prior literature. However, the approach required labeled data in the form of manual segmentation for each training volume. Manually segmenting MRI scans is labor-intensive, time-consuming, and Prone to inter- and intra-rater variability. While supervised learning remains common, recent advances have increasingly shifted toward self-supervised learning (SSL) to reduce reliance on costly annotated data. Given its potential to improve efficiency and scalability, SSL in medical image segmentation warrants further investigation.

REFERENCES

[1] B. Bonato, L. Nanni, and A. Bertoldo, "'advancing precision: A comprehensive review of mri segmentation datasets from brats challenges (2012-2025)'," *Sensors*, vol. 25, no. 6, Mar 2025. [Online]. Available: https://doi.org/10.3390/s25061838

[2] Menze, Bjoern H. and Jakab, Andras and Bauer, Stefan and Kalpathy-Cramer, Jayashree and Farahani, Keyvan and Kirby, Justin and Burren, Yuliya and Porz, Nicole and Slotboom, Johannes and Wiest, Roland and Lanczi, Levente and Gerstner, Elizabeth and Weber, Marc-André and Arbel, Tal and Avants, Brian B. and Ayache, Nicholas and Buendia, Patricia and Collins, D. Louis and Cordier, Nicolas and Corso, Jason J. and Criminisi, Antonio and Das, Tilak and Delingette, Hervé and Demiralp, Çağatay and Durst, Christopher R. and Dojat, Michel and Doyle, Senan and Festa, Joana and Forbes, Florence and Geremia, Ezequiel and Glocker, Ben and Golland, Polina and Guo, Xiaotao and Hamamci, Andac and Iftekharuddin, Khan M. and Jena, Raj and John, Nigel M. and Konukoglu, Ender and Lashkari, Danial and Mariz, José António and Meier, Raphael and Pereira, Sérgio and Precup, Doina and Price, Stephen J. and Raviv, Tammy Riklin and Reza, Syed M. S. and Ryan, Michael and Sarikaya, Duygu and Schwartz, Lawrence and Shin, Hoo-Chang and Shotton, Jamie and Silva, Carlos A. and Sousa, Nuno and Subbanna, Nagesh K. and Szekely, Gabor and

Taylor, Thomas J. and Thomas, Owen M. and Tustison, Nicholas J. and Unal, Gozde and Vasseur, Flor and Wintermark, Max and Ye, Dong Hye and Zhao, Liang and Zhao, Binsheng and Zikic, Darko and Prastawa, Marcel and Reyes, Mauricio and Van Leemput, Koen, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024. [Online]. Available: https://doi.org/10.1109/TMI.2014.2377694

[3] M. J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, B. Murrey, A. Myronenko, C. Zhao, D. Yang, V. Nath, Y. He, Z. Xu, A. Hatamizadeh, A. Myronenko, W. Zhu, Y. Liu, M. Zheng, Y. Tang, I. Yang, M. Zephyr, B. Hashemian, S. Alle, M. Z. Darestani, C. Budd, M. Modat, T. Vercauteren, G. Wang, Y. Li, Y. Hu, Y. Fu, B. Gorman, H. Johnson, B. Genereaux, B. S. Erdal, V. Gupta, A. Diaz-Pinto, A. Dourson, L. Maier-Hein, P. F. Jaeger, M. Baumgartner, J. Kalpathy-Cramer, M. Flores, J. Kirby, L. A. D. Cooper, H. R. Roth, D. Xu, D. Bericat, R. Floca, S. K. Zhou, H. Shuaib, K. Farahani, K. H. Maier-Hein, S. Aylward, P. Dogra, S. Ourselin, and A. Feng, "Monai: An open-source framework for deep learning in healthcare," 2022. [Online]. Available: https://arxiv.org/abs/2211.02701

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmenation." [Online]. Available: https://arxiv.org/abs/1505.04597

[5] C. Jin, N. S. E. M. Noor, T. F. Ng, M. S. M. Asaari, and H. Ibrahim, "Transformer-based architectures in mri brain tumor segmentation: A review," *Computerized Medical Imaging and Graphics*, vol. 129, 2026. [Online]. Available: https://doi-org.dml.regis.edu/10.1016/j.compmedimag.2026.102729

[6] S. Chen, K. Ma, and Y. Zheng, "Med3d: Transfer learning for 3d medical image analysis," 2019. [Online]. Available: https://arxiv.org/abs/1904.00625

[7] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, "Unsupervised learning of visual features by contrasting cluster assignments," *NeurIPS*, Jan 2021. [Online]. Available: https://arxiv.org/abs/2006.09882

[8] L. Zang, X. Wen, X. Jian, Xian-FengYang, and M. Li, "Diagnostic error and bias in the department of radiology: a pictorial essay," *Insights into imaging*, vol. 14, no. 16, 2023. [Online]. Available: https://link.springer.com/article/10.1186/s13244-023-01521-7

[9] Ömer Kasalak, H. Alnahwi, R. Toxopeus, J. P. Pennings, D. Yakar, and T. C. Kwee, "Work overload and diagnostic errors in radiology," *European Journal of Radiology*, vol. 167, p. 111032, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0720048X23003467

[10] J. Shen, S.-M. Jun, S. J. Holdsworth, G. M. Talou, J. A. Correia, and H. Abbasi, "Evaluating segresnet for single-modality meningioma segmentation on t1 contrast-enhanced mri on a new zealand clinical cohort," *Neuroscience Informatics*, vol. 6, no. 1, p. 100261, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772528626000051

[11] M. M. R. Siddiquee, D. Yang, Y. He, Daguang, and A. Myronenko, "Automated ischemic stroke lesion segmentation from 3d mri isles 2022 challenge report," 2022. [Online]. Available: https://arxiv.org/abs/2209.09546

[12] B. S, A. H, S. A, B. M, R. M, K. J, F. J, F. K, and D. C, "Segmentation labels and radiomic features for the pre-opeative scans of the tcga-lgg collection," 2017. [Online]. Available: https://www.cancerimagingarchive.net/analysis-result/brats-tcga-lgg/

[13] A. L. Simpson, M. Antonelli, S. Bakas, M. Bilello, K. Farahani, B. van Ginneken, A. Kopp-Schneider, B. A. Landman, G. Litjens, B. Menze, O. Ronneberger, R. M. Summers, P. Bilic, P. F. Christ, R. K. G. Do, M. Gollub, J. Golia-Pernicka, S. H. Heckers, W. R. Jarnagin, M. K. McHugo, S. Napel, E. Vorontsov, L. Maier-Hein, and M. J. Cardoso, "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," 2019. [Online]. Available: https://arxiv.org/abs/1902.09063

[14] M. S. Decathlon, "Medical segmentation decathlon," 2019. [Online]. Available: http://medicaldecathlon.com/

[15] N. I. of Health. Neuroimaging informatics technology initiative (nifti). [Online]. Available: https://nifti.nimh.nih.gov/index.html

[16] "The medical segmenation decathlon," *Nature Communications*, vol. 13, no. 4128. [Online]. Available: https://doi.org/10.1038/s41467-022-30695-9

[17] Y. D, G. F, G. J, I. E, L. L. J., N. A, P. K, S. J, T. N. H. M., and T. J, "University of missouri post-operative glioma dataset (mu-glioma-post) (verson 1) [data set]," 2024. [Online]. Available: https://www.cancerimagingarchive.net/collection/mu-glioma-post/

[18] O. Rainio and R. Klen, "Modified dice coefficients for evaluation of tumor segmentation from pet images: A proof-of-concept study," *Journal of Imaging Informatics in Medicine*, vol. 39, no. 1, pp. 785–793, 2026. [Online]. Available: https://doi.org/10.1007/s10278-025-01535-1