

# Exploring Brain Tumor Segmentation and Patient Survival: An Interpretable Model Approach

Valerio Ponzi<sup>1,2</sup>, Giorgio De Magistris<sup>1</sup>

<sup>1</sup>Department of Computer, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25, Roma, 00185, Italy

<sup>2</sup>Institute for Systems Analysis and Computer Science, Italian National Research Council, Via dei Taurini 19, Roma, 00185, Italy

## Abstract

Detecting and delineating brain tumors from MRI images using artificial intelligence presents a complex challenge in medical AI. Recent progress has seen a variety of techniques employed to assist medical professionals in this task. Despite the effectiveness of machine learning algorithms in segmenting tumors, their lack of transparency in decision-making hinders trust and validation. In our project, we constructed an interpretable U-Net Model specifically tailored for brain tumor segmentation, leveraging both the Gradient-weighted Class Activation Mapping (Grad-CAM) Algorithm and the SHapley Additive exPlanations (SHAP) library. We relied on the BraTS2020 benchmark dataset for training and evaluation purposes. The U-Net model we employed yielded promising results. We then utilized Grad-CAM to visualize the crucial features attended to by the model within an image. Additionally, we enhanced interpretability by utilizing the SHAP library to elucidate the predictions made by various models (including Random Forest, KNN, SVC, and MLP) utilized for predicting patient survival days.

## Keywords

Brain Tumor, U-Net, Segmentation, Explainable Artificial Intelligence

## 1. Introduction

Brain tumors represent a significant challenge in healthcare, affecting millions of individuals worldwide with their life-threatening implications. Accurate delineation of these tumors is paramount for effective treatment strategies and ongoing monitoring of disease progression. Over the past few years, deep learning techniques have emerged as promising tools for brain tumor segmentation, with the U-Net architecture gaining popularity for its ability to capture intricate details within medical images. However, the inherent opacity of deep learning models presents a hurdle, as it limits their interpretability and makes it difficult for clinicians to comprehend the rationale behind their decisions. Explainable Artificial Intelligence (XAI) has garnered increasing importance, particularly in the medical domain, where precise tumor segmentation plays a crucial role. Tumor segmentation involves the identification and localization of tumors within medical imaging data, such as MRI scans, CT scans, or X-rays. This process is indispensable in cancer diagnosis, treatment planning, and progress tracking. XAI holds significance in tumor segmentation for several reasons. Firstly, AI models often operate as "black boxes," meaning their decision-making processes are not readily transparent [1].

In the context of medical imaging, the lack of trans-

parency inherent in deep learning models poses a significant challenge. This opacity is problematic as doctors need to understand how the model arrives at its conclusions to make informed decisions about patient care [2, 3]. Additionally, explainable artificial intelligence (XAI) plays a crucial role in mitigating biases within AI models. Biases may arise if the model is trained on data that doesn't adequately represent the population it will serve, leading to incorrect or skewed predictions. XAI can help identify and rectify these biases, thereby enhancing the model's reliability. Moreover, XAI fosters trust in AI systems by elucidating the decision-making process [4], thereby increasing the willingness of doctors and patients to rely on these models [5, 6].

In our project, we utilize the Gradient-weighted Class Activation Mapping (Grad-CAM) technique to imbue our segmentation UNET model with explainability. Grad-CAM generates heatmaps highlighting the crucial regions of input images that the model focuses on when making predictions. By visualizing these heatmaps, we gain insights into the features guiding the model's decisions, facilitating better understanding of its behavior.

Furthermore, our project incorporates the SHAP (SHapley Additive exPlanations) approach, particularly relevant for tasks like predicting patient survival based on medical imaging data in datasets like BRATS. SHAP values elucidate the contributions of individual features to the model's output, shedding light on the mechanisms underlying predictions. This transparency is vital in the medical context, where accurate predictions profoundly impact patient outcomes. Our system is trained on BRATS2020 data using a standard UNET model for

ICYRIME 2023: International Conference of Yearly Reports on Informatics, Mathematics, and Engineering. Naples, July 28-31, 2023

✉ ponzi@diag.uniroma1.it (V. Ponzi);

demagistris@diag.uniroma1.it (G. D. Magistris)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License

Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

segmentation, augmented by Grad-CAM for heatmap visualization and SHAP for survival prediction analysis. The segmentation accuracy stands at an impressive 99 percent, with specific dice scores for necrotic, edema, and enhancing regions. This comprehensive approach not only yields accurate predictions but also enhances interpretability, trust, and confidence in AI-assisted medical decision-making.

## 2. Related Works

Brain tumors are among the most perilous types of tumors globally, with gliomas emerging as the predominant primary brain tumors. Gliomas stem from the aberrant proliferation of glial cells in the brain and spinal cord, exhibiting varying degrees of malignancy and histological classifications. Individuals diagnosed with glioblastoma, the most aggressive form of glioma, typically face a survival prognosis of fewer than 14 months on average. Medical professionals frequently utilize Magnetic Resonance Imaging (MRI), a non-invasive technique, to diagnose brain tumors because of its capability to generate a wide variety of tissue contrasts in each imaging mode [7]. However, analyzing and segmenting structural MRI images of brain tumors is a challenging and time-consuming task that typically requires the expertise of professional neuroradiologists. Therefore, an automated and dependable brain tumor segmentation method would greatly facilitate the diagnosis and treatment of brain tumors.

[8] An alternative approach is suggested to focus solely on a small region of the image rather than processing the entire image, reducing computational time and addressing overfitting issues in a Cascade Deep Learning model. Additionally, a Cascade Convolutional Neural Network (C-ConvNet/C-CNN) is introduced, which extracts both local and global features through separate pathways. Moreover, to enhance the accuracy of brain tumor segmentation beyond existing models, a new Distance-Wise Attention (DWA) mechanism is employed.

In another work [9], a novel design relying on a 3D U-Net model was developed, incorporating numerous skip connections alongside cost-effective pre-trained 3D MobileNetV2 blocks and attention modules. These pre-trained MobileNetV2 blocks aid the architecture by offering fewer parameters, ensuring a manageable model size within our computational capacity, and facilitating faster convergence. Furthermore, additional skip connections were introduced between the encoder and decoder blocks to facilitate the transfer of extracted features, while attention modules were employed to filter out irrelevant features transmitted through the skip connections.

Further existing works on interpretable CNNs were examined during the execution of our project. One of

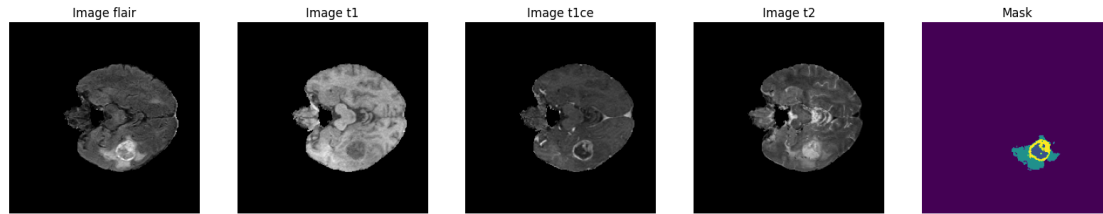
the most interesting ones was [10]. The author implemented a prototypical part network (ProtoPNet), which dissects images by identifying prototypical parts and amalgamating evidence from these prototypes to derive a final classification. The operational principle of this ProtoPNet involves comparing the latent features of  $f(x)$  with the learned prototypes. Specifically, for each class  $k$ , the network seeks evidence for  $x$  belonging to class  $k$  by assessing its latent patch representations against every learned prototype  $p(j)$  associated with class  $k$ .

In another study focusing on interpretable machine learning, researchers introduced a method for "Classification of Mass Lesions in Digital Mammography" [11]. They employed a pixel-wise annotation technique to precisely segment affected lesions, and the outcomes were subsequently depicted using GradCam and GradCam++ heatmaps. The findings demonstrated that pixel-wise annotation improved the segmentation and localization of the affected area, with the generated heatmaps maintaining focus on the impacted part of the skin, rather than encompassing all image pixels.

The concept of utilizing GradCam for visual interpretation and explanation of model results originated from a related study, which utilized GradCam for visual explanations across a wide range of CNN-based models. This approach combines Grad-CAM with fine-grained visualizations to produce high-resolution, class-discriminative visualizations. It was applied to various off-the-shelf image classification, captioning, and visual question-answering (VQA) models, including those based on ResNet architectures [12].

The domain of explainable Artificial Intelligence (xAI) is relatively new but evolving rapidly, with the introduction of numerous libraries designed to elucidate the outputs of opaque deep learning models. One such notable library is the SHAP (SHapley Additive exPlanations) library. SHAP assigns an importance value to each feature for a specific prediction. In their work [13], SHAP was applied to the brats dataset. For each input feature, SHAP calculates the importance value, offering various calculation methods, including two model-agnostic methods applicable regardless of the trained network type, and four specific model methods, one of which is DeepExplainer.

In this study, DeepExplainer was utilized to determine the importance values for a given combination of 3D MRI voxel and age values. DeepExplainer efficiently approximates SHAP values for a deep neural network model by recursively propagating DeepLIFT multipliers, thereby deriving an effective linearization technique from the SHAP values. By inputting an example data point into DeepExplainer, importance values for every pixel in the 3D voxel, as well as for the age value, are determined. These important values can then be visually represented by integrating them into a background image, which



**Figure 1:** These images show a set of brain scans from the dataset and the corresponding mask. From left to right: Flair, T1, T1ce, the T2, and the mask that emphasizes important regions.

represents the input.

### 3. Dataset

In this project, we utilized the BRATS2020 Dataset, a commonly employed medical imaging dataset utilized for both brain tumor segmentation and classification tasks. It represents an enhanced iteration of the BRATS2015 dataset and is made available by the Multimodal Brain Tumor Segmentation Challenge (BRATS).

The dataset comprises MRI scans of the brain obtained from patients diagnosed with diverse types of brain tumors, including gliomas, meningiomas, and pituitary adenomas. These scans encompass four distinct modalities: T1-weighted (T1), T1-weighted contrast-enhanced (T1ce), T2-weighted (T2), and fluid-attenuated inversion recovery (FLAIR) images <sup>1</sup>. Accompanying each MRI scan is a ground truth segmentation map delineating the tumor's location and extent. Containing a total of 369 MRI scans, the BRATS2020 dataset designates 335 scans for training and 34 for testing. These scans were sourced from various medical institutions and meticulously annotated by multiple experts. Furthermore, the dataset provides additional patient-related information such as age, gender, and tumor subtype. Widely recognized as a benchmark dataset, BRATS2020 serves as a standard for assessing the efficacy of algorithms in brain tumor segmentation and classification. Researchers leverage this dataset to innovate and validate new approaches for automating these processes, aiming to enhance the accuracy and efficiency of diagnosis and treatment for individuals afflicted with brain tumors.

#### 3.1. Data Preprocessing

The dataset utilized in our project comprises MRI scans of patients with various types of brain tumors, encompassing four modalities: T1-weighted (T1), T1-weighted contrast-enhanced (T1ce), T2-weighted (T2), and Flair Images, alongside corresponding ground truth segmentation masks. Each MRI image contains 155 slices, of which we selected slices ranging from 22 to 100, capturing the

most pertinent tumor data after rigorous experimentation.

During data preprocessing, we identified irregular patterns in file 355, prompting its removal from the dataset to ensure the integrity of our results and model training. Additionally, we standardized the image size to 128\*128 for training purposes. Our training data comprises stacked Flair and T1 images, while the model receives segmentation masks as labels, which are subsequently one-hot encoded for compatibility.

## 4. Our Methodology

In this project, our approach is delineated into the following steps:

- Employing the Unet model for intricate segmentation tasks, adept at delineating tumor contours within MRI scans.
- Enacting an array of Machine Learning Algorithms to prognosticate patient survival, harnessing extracted features from segmented tumor regions alongside ancillary clinical data.
- Deploying SHAP (SHapley Additive exPlanations) to furnish comprehensive elucidations of predictions rendered by machine learning models, facilitating an enhanced understanding of the underlying determinants contributing to patient survival prognostications.

### 4.1. UNET for Tumour Segmentation

The U-Net architecture is a convolutional neural network initially devised for biomedical image segmentation but widely applicable across various computer vision segmentation tasks. It comprises a contracting path and an expanding path, forming a "U" shape. The contracting path functions akin to a traditional convolutional neural network, capturing image context through successive convolutional and max-pooling layers. These layers reduce spatial resolution while augmenting channel depth to extract broader image features. Conversely, the expanding path reconstructs spatial resolution and yields

the final segmentation map. It employs convolutional and up-sampling layers to enhance spatial resolution while reducing channel depth. Up-sampling methods like bilinear interpolation or transposed convolution are commonly used.

Moreover, the U-Net incorporates skip connections between corresponding layers of the contracting and expanding paths. These connections enable the network to circumvent spatial information loss from pooling operations and merge local and global image features effectively. Skip connections concatenate feature maps from corresponding layers, followed by a 1x1 convolutional layer to decrease channel depth. The concatenated feature maps then feed subsequent convolutional and up-sampling layers in the expanding path. Training the U-Net model involves end-to-end optimization using pixel-wise cross-entropy loss. This loss function compares predicted segmentation maps with ground truth maps, guiding parameter adjustments of convolutional filters to minimize the loss and generate accurate segmentation maps [14].

Our UNET architecture operates on input images with dimensions of (128, 128, 2). Initially, a convolutional layer with 64 filters, a 3x3 kernel size, and "same" padding is employed, followed by batch normalization and ReLU activation. Subsequently, the encoder phase comprises multiple down-sampling blocks, each featuring two 3x3 convolutional layers with 64 filters, followed by batch normalization and ReLU activation. After each block, the filter count doubles, and spatial resolution is halved via max-pooling. The bottleneck layer is characterized by four 3x3 convolutional layers with 1024 filters, alongside batch normalization and ReLU activation. Conversely, the decoder phase involves up-sampling blocks, consisting of 2x2 transpose convolutional layers with 512 filters. These layers are concatenated with corresponding feature maps from the encoder part, followed by two 3x3 convolutional layers with 64 filters, batch normalization, and ReLU activation. Finally, a 1x1 convolutional layer with 4 filters is employed in the final layer, succeeded by a softmax activation function to yield a probability distribution across the four segmentation classes.

## 4.2. GradCam Algorithm

GradCam, short for Gradient-weighted Class Activation Mapping, serves as a valuable tool in enhancing the interpretability of complex neural network models, particularly in medical imaging tasks such as brain tumor segmentation using MRI images. To utilize GradCam effectively for brain tumor segmentation, we first embark on training a UNET Model using BRATS Data, a well-known dataset extensively used in the field for brain tumor segmentation tasks. This initial step is crucial as it lays the foundation for the subsequent interpretability

analysis.

Once the UNET Model is trained, the next phase involves the practical application of GradCam for explanation generation. We start by selecting an input MRI image containing a brain tumor, which serves as the subject for interpretation. This MRI image is then fed through the trained UNET Model to obtain the output segmentation map, which delineates the tumor region within the image.

To delve deeper into understanding the model's decision-making process, we employ the concept of gradient computation. Specifically, we compute the gradients of the output segmentation map concerning the feature maps of the last convolutional layer. These gradients provide valuable information regarding the importance of different regions within the input image in influencing the model's segmentation decision. Leveraging these gradients, we proceed to compute the GradCam heatmap for the input MRI image. This heatmap effectively highlights the regions within the image that exert the most significant influence on the segmentation decision made by the UNET Model. By overlaying this heatmap onto the input MRI image, we create a visually intuitive representation that facilitates the interpretation of the model's decision-making process.

Through this approach, we gain valuable insights into the inner workings of the UNET Model for brain tumor segmentation. By visualizing the regions of the input image that contribute most significantly to the model's predictions, we enhance the interpretability of our model, thereby fostering greater trust and understanding among stakeholders in the medical domain.

In summary, by integrating GradCam into our workflow for brain tumor segmentation, we not only improve the transparency and interpretability of our model but also empower clinicians and researchers with actionable insights into the diagnostic process, ultimately leading to more informed decision-making and better patient outcomes [12].

## 4.3. SHAP Explanations

SHAP (SHapley Additive exPlanations) is a model-agnostic method for interpreting the predictions of machine learning models. It can help to identify which features in the input data contributed the most to a particular prediction [15].

In this project, we use SHAP to determine a patient survival. The patient survival data is already provided in the survival info CSV file, which contains the following columns Brats20ID, Age, Survival days, Extent of Resection. We preprocess the data and determine whether the extension of the tumor was short, medium, or large. Then the data is trained and tested using the various classification algorithms including KNN, Random Forest, etc. The



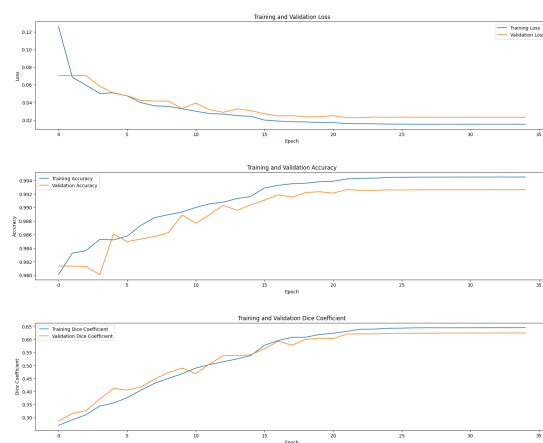
results of KNN, SVC, MLP, and Random Forest are then explained through the SHAP library. We used SHAP Kernel explainer and Tree Explainer to get the SHAP values and visualize them using the SHAP summary plot and SHAP force plot.

## 5. Results

The results section includes the results obtained on the UNET Model train and test data, Visualizations using the GradCam Algorithm, and Results obtained on survival predictions data, and its explanations using SHAP.

### 5.1. Results on UNET Model

Figure 2 shows the training and validation data results of our model. On the test data, the model achieved the following results accuracy score: 0.9912, mean you score: 0.8250, dice coefficient: 0.6590, precision: 0.9930, sensitivity: 0.9923, specificity: 0.9952, dice coefficient necrotic: 0.3912, dice coefficient edema: 0.71, dice coefficient enhancing: 0.68



**Figure 2:** Unet Model Results. The first plot represents the Loss chart, while the second is the Accuracy chart. The third is related to the Dice coefficient. This coefficient is a common metric used to evaluate the similarity between two sets of data. In the context of image segmentation, it is often used to measure the accuracy of a segmentation algorithm by comparing the segmentation output with a ground truth segmentation. [16, 17]

Our model was then validated on the test images to visualize the output segmentations made by the model. Figure 3 and figure 4 shows the results of the model. It perfectly segments all three classes namely "Neurotic/core", "Edema", and "Enhancing". The area comprising the tumour was perfectly identified by the model hence giving us perfect segmentations.

### 5.2. GradCam Heatmaps

To interpret the model predictions, we use the GradCam technique. Our GradCam visualization function builds the gradient model using Unet model inputs, the last convolution layer of the model, and model outputs. The gradient model is then provided with a test image for which it computes the gradients of the output segmentation relating to the last convolution layer. These gradients are then used to compute the Heatmap. By visualizing the heatmap generated by Grad-CAM, we can gain insights into which parts of the input image are most important for the interpretable model to make its segmentation. Figure 5 shows the original and the GradCam heatmaps generated on that MRI Image by the model, we can see that the model focuses more on the Tumour area to predict the correct segmentation mask.

### 5.3. Patient Survival Prediction

For predicting patient survival various ML algorithms were used. Initially, we used a Random Forest Classifier with 3 trees to predict the extent of survival. The survival extent was categorized into three categories, small, medium, or long. For further experiments, we used the KNN classifier. Next, we used a Support Vector classifier on the same data in the context of getting better accuracy scores. Lastly, we experimented by training and testing the model using a Multi-Layer Perceptron (MLP) classifier. The results of all those algorithms are included in the Table 1.

Algorithm Name	Accuracy	Precision	F1
RFC	0.64	0.64	0.66
KNN	0.61	0.62	0.61
SVC	0.55	0.58	0.55
MLP	0.50	0.59	0.56

**Table 1**  
Survival Prediction Results.

### 5.4. SHAP for Patient Survival Predictions

For interpreting and explaining the predictions of the above-mentioned models for the task of patient survival, we used the SHAP library. SHAP provides us with the ability to explain the predictions of our machine learning and deep learning models via its various built-in explainers. For this task, we used SHAP's kernel explainer and tree explainer. The SHAP Kernel Explainer interprets machine learning models by calculating the contribution of each feature using Shapley values. It provides explanations by evaluating the impact of including or excluding features on the model's predictions, helping

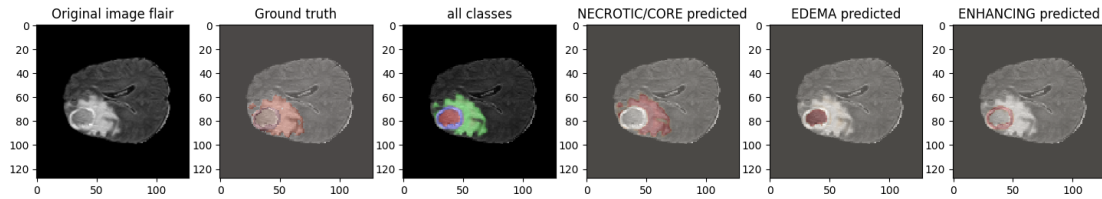


Figure 3: Segmentation Results 1

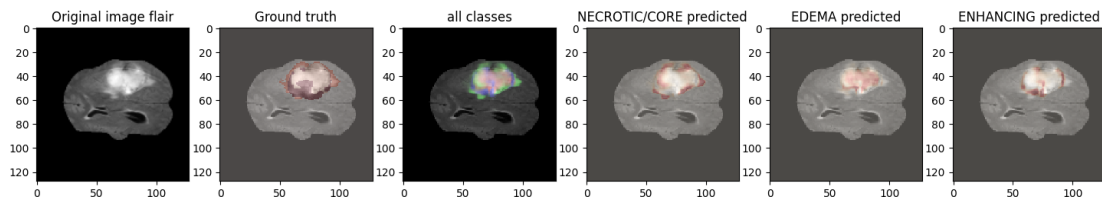


Figure 4: Segmentation Results 2

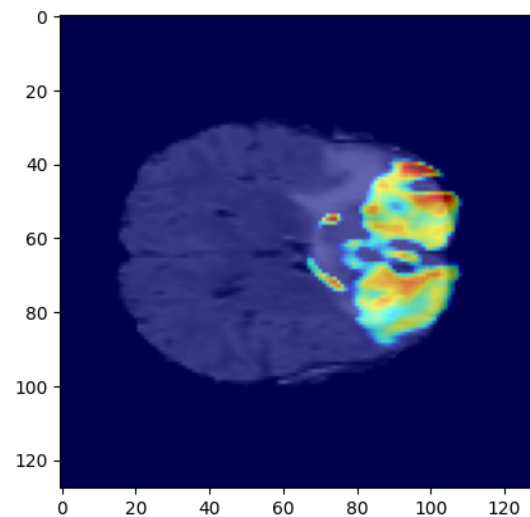


Figure 5: GradCam Heatmap Image, shows that the model focuses more on the Tumour area to predict the correct segmentation mask

understand feature importance and model behavior. The SHAP Tree Explainer is a method designed for interpreting tree-based machine learning models, such as decision trees or random forests. It computes the Shapley values by approximating the model with a set of additive tree-based models, enabling the attribution of feature contributions to individual predictions made by tree models. Figure 6 and Figure 7 display the results of our SHAP explanations.

As we can notice from the above table, the results are

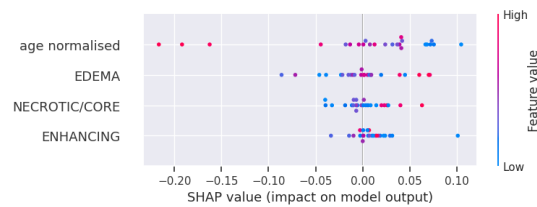


Figure 6: SHAP Summary Plot for Random Forest Classifier

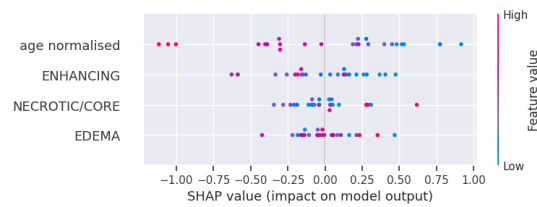


Figure 7: SHAP Summary Plot for KNN

not quite promising, with KNN and Random Forest being slightly better than our other experimented models. In future work, we would try to achieve better scores by trying various ensembles of models on the data.

## 6. Conclusion and Future Works

Brain tumor segmentation through machine learning has significantly assisted medical professionals in efficiently locating and resecting tumors. Various techniques have been explored to accurately segment tumors from MRI im-

ages, including the utilization of the UNET Model in this project. Recent advancements in semantic segmentation have introduced several notable models: Mask R-CNN: This CNN architecture extends the Faster R-CNN object detection model to include a mask prediction branch, allowing it to perform object detection and instance segmentation simultaneously. DeepLab V3+: Designed for semantic segmentation of images, DeepLab V3+ employs dilated convolution to capture multi-scale context without increasing the number of parameters. PSPNet: Utilizing a pyramid pooling module, PSPNet captures global context at multiple scales, facilitating accurate predictions for objects of various sizes [18, 19]. FCN: Fully Convolutional Networks perform dense pixel-wise prediction of image labels, accommodating input images of arbitrary size and producing output images of the same size with predicted labels for each pixel [20]. Segment Anything Model (SAM): Facebook’s SAM, an open-source state-of-the-art computer vision model, is designed for image segmentation tasks [21]. These models, alongside UNET, have showcased state-of-the-art results in segmentation tasks. Our objective is to collaborate with these models on our data and visualize their outcomes on the BRATS2020 dataset.

In addition to tumor segmentation, we aimed to enhance the interpretability of the UNET model by employing GradCam. However, with advancements in the eXplainable Artificial Intelligence (XAI) field, several other visual interpretation techniques have emerged. We plan to explore these techniques, including GradCam++, SmoothGradCam++, Guided GradCam, and Score-CAM, to provide more precise and insightful model interpretations.

Moreover, in the realm of patient survival prediction, current models exhibit low accuracy and struggle with generalization. To address this, our future work will involve experimenting with sequential neural network models to achieve better results. Additionally, we will focus on tuning hyperparameters and exploring different parameter sets to improve model performance on both training and test data. These efforts aim to enhance the accuracy and reliability of patient survival predictions, thus advancing the impact of medical AI in clinical settings.

## References

- [1] H. Saleem, A. R. Shahid, B. Raza, Visual interpretability in 3d brain tumor segmentation network, *Computers in Biology and Medicine* 133 (2021) 104410. URL: <https://www.sciencedirect.com/science/article/pii/S0010482521002043>. doi:<https://doi.org/10.1016/j.combiomed.2021.104410>.
- [2] F. Yan, Y. Chen, Y. Xia, Z. Wang, R. Xiao, An explainable brain tumor detection framework for mri analysis, *Applied Sciences* 13 (2023). URL: <https://www.mdpi.com/2076-3417/13/6/3438>. doi:[10.3390/app13063438](https://doi.org/10.3390/app13063438).
- [3] G. De Magistris, R. Caprari, G. Castro, S. Russo, L. Iocchi, D. Nardi, C. Napoli, Vision-based holistic scene understanding for context-aware human-robot interaction 13196 LNAI (2022) 310 – 325. doi:[10.1007/978-3-031-08421-8\\_21](https://doi.org/10.1007/978-3-031-08421-8_21).
- [4] G. Borowik, M. Woźniak, A. Fornaia, R. Giunta, C. Napoli, G. Pappalardo, E. Tramontana, A software architecture assisting workflow executions on cloud resources, *International Journal of Electronics and Telecommunications* 61 (2015) 17 – 23. doi:[10.1515/e1etel-2015-0002](https://doi.org/10.1515/e1etel-2015-0002).
- [5] G. Yang, Q. Ye, J. Xia, Unbox the black-box for the medical explainable ai via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond, *Information Fusion* 77 (2022) 29–52. URL: <https://www.sciencedirect.com/science/article/pii/S1566253521001597>. doi:<https://doi.org/10.1016/j.inffus.2021.07.016>.
- [6] V. Ponzi, S. Russo, V. Bianco, C. Napoli, A. Wajda, Psychoeducative social robots for a healthier lifestyle using artificial intelligence: a case-study, volume 3118, 2021, pp. 26 – 33.
- [7] S. Khawaldeh, U. Pervaiz, A. Rafiq, R. S. Alkhaldeh, Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks, *Applied Sciences* 8 (2018). URL: <https://www.mdpi.com/2076-3417/8/1/27>. doi:[10.3390/app8010027](https://doi.org/10.3390/app8010027).
- [8] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, M. Ben-dechache, Brain tumor segmentation based on deep learning and an attention mechanism using mri multi-modalities brain images., *Scientific reports* (2021). URL: <http://hdl.handle.net/10147/631439>. doi:[10.1038/s41598-021-90428-8](https://doi.org/10.1038/s41598-021-90428-8).
- [9] J. Nodirov, A. B. Abdusalomov, T. K. Whangbo, Attention 3d u-net with multiple skip connections for segmentation of brain tumor images, *Sensors* 22 (2022). URL: <https://www.mdpi.com/1424-8220/22/17/6501>. doi:[10.3390/s22176501](https://doi.org/10.3390/s22176501).
- [10] C. Chen, O. Li, A. Barnett, J. Su, C. Rudin, This looks like that: deep learning for interpretable image recognition, *CoRR abs/1806.10574* (2018). URL: <http://arxiv.org/abs/1806.10574>. arXiv:1806.10574.
- [11] A. J. Barnett, F. R. Schwartz, C. Tao, C. Chen, Y. Ren, J. Y. Lo, C. Rudin, IAIA-BL: A case-based interpretable deep learning model for classification of mass lesions in digital mammography, *CoRR abs/2103.12308* (2021). URL: <https://arxiv.org/abs/2103.12308>. arXiv:2103.12308.

- [12] R. R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, D. Batra, Grad-cam: Why did you say that?, 2017. [arXiv: 1611.07450](https://arxiv.org/abs/1611.07450).
- [13] M. Eder, E. Moser, A. Holzinger, C. Jean-Quartier, F. Jeanquartier, Interpretable machine learning with brain image and survival data, *BioMedInformatics 2* (2022) 492–510. URL: <https://www.mdpi.com/2673-7426/2/3/31>. doi:10.3390/biomedinformatics2030031.
- [14] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, *CoRR abs/1505.04597* (2015). URL: <http://arxiv.org/abs/1505.04597>. [arXiv: 1505.04597](https://arxiv.org/abs/1505.04597).
- [15] S. M. Lundberg, S. Lee, A unified approach to interpreting model predictions, *CoRR abs/1705.07874* (2017). URL: <http://arxiv.org/abs/1705.07874>. [arXiv: 1705.07874](https://arxiv.org/abs/1705.07874).
- [16] S. Jadon, A survey of loss functions for semantic segmentation, in: *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE, 2020. URL: <https://doi.org/10.1109%2Fcibcb48159.2020.9277638>. doi:10.1109/cibcb48159.2020.9277638.
- [17] A. Alfarano, G. De Magistris, L. Mongelli, S. Russo, J. Starczewski, C. Napoli, A novel convmixer transformer based architecture for violent behavior detection 14126 *LNAI* (2023) 3 – 16. doi:10.1007/978-3-031-42508-0\_1.
- [18] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, 2017. [arXiv: 1612.01105](https://arxiv.org/abs/1612.01105).
- [19] V. Marcotrigiano, G. D. Stingi, S. Fregnan, P. Magarelli, P. Pasquale, S. Russo, G. B. Orsi, M. T. Montagna, C. Napoli, C. Napoli, An integrated control plan in primary schools: Results of a field investigation on nutritional and hygienic features in the apulia region (southern italy), *Nutrients 13* (2021). doi:10.3390/nu13093006.
- [20] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, 2015. [arXiv: 1411.4038](https://arxiv.org/abs/1411.4038).
- [21] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, R. Girshick, Segment anything, 2023. [arXiv: 2304.02643](https://arxiv.org/abs/2304.02643).