Treatment response assessment of glioma tumors using deep learning based on magnetic resonance imaging

Zahra Papi
Isfahan University of Medical Sciences
Sina Fathi
Iran University of Medical Sciences
Fatemeh Dalvand
Shahid Beheshti University
Mahsa Vali
Isfahan University of Technology
Ali Yousefi
University of Isfahan
Mohammad Hemmatyar Tabatabaei
Abadeh University
Alireza Amouheidari
Isfahan Milad Hospital
Iraj Abedi (i.abedi@med.mui.ac.ir)
Isfahan University of Medical Sciences

Research Article

Keywords: Magnetic resonance imaging, Deep learning, Treatment response, Transfer learning, Gliomas, Radiotherapy

Posted Date: March 28th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2733829/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Neuroinformatics on July 17th, 2023. See the published version at https://doi.org/10.1007/s12021-023-09640-8.

Abstract

Background: Glioma is the most common primary intracranial neoplasm in adults. Radiotherapy is a treatment approach in glioma patients, and Magnetic Resonance Imaging (MRI) is a beneficial diagnostic tool in treatment planning. Treatment response assessment in glioma patients is usually based on the Response Assessment in Neuro Oncology (RANO) criteria. The limitation of assessment based on RANO is two-dimensional (2D) manual measurements. Deep learning (DL) has great potential in neuro-oncology to improve the accuracy of response assessment.

Method: In the current research, firstly, the BraTS 2018 Challenge dataset included 210 HGG and 75 LGG were applied to train a designed U-Net network for automatic tumor and intra-tumoral segmentation, followed by training of the designed classifier with transfer learning for determining grading HGG and LGG. Then, designed networks were employed for the segmentation and classification of local MRI images of 49 glioma patients pre and post-radiotherapy. The results of tumor segmentation and its intra-tumoral regions were utilized to determine the volume of different regions and treatment response assessment.

Results: Treatment response assessment demonstrated that radiotherapy is effective on the whole tumor and enhancing region with p-value ≤ 0.05 with a 95% confidence level, while it did not affect necrosis and peri-tumoral edema regions.

Conclusion: This work demonstrated the potential of using deep learning in MRI images to provide a beneficial tool in the automated treatment response assessment so that the patient can obtain the best treatment.

Introduction

Gliomas are the most common types of glial-based primary tumors in adults, which are specified with several malignancy grades, namely grade I (pilocytic astrocytoma, gangliocytoma, and ganglioglioma), grade II (astrocytoma, oligodendroglioma, and oligoastrocytoma), grade III (anaplastic astrocytoma, anaplastic oligodendroglioma, and anaplastic ependymoma), and grade IV (glioblastoma) according to World Health Organization (WHO) (Ahmed, Oborski, Hwang, Lieberman, & Mountz, 2014; Ranjbarzadeh et al., 2021; Tiwari, Srivastava, & Pant, 2020).

Grades I and II of gliomas usually are known as benign or low-grade brain tumors; however, grades III and IV are more aggressive and specified as malignant or high-grade tumors (Ghaffari, Sowmya, & Oliver, 2020). Despite the recent advances in therapeutic and diagnostic methods for malignant gliomas, the median survival is less than five years for patients with anaplastic glioma (grade III) and 15 months for patients with glioblastoma (GBM) (Ahmed et al., 2014). Gliomas represent about 80% of all malignant brain and central nervous system (CNS) tumors. More than 50% of these gliomas are diagnosed as GBM (Barnholtz-Sloan, Ostrom, & Cote, 2018). In the United States, the annual incidence of all brain

tumors, gliomas, and GBMs per 100,000 population between 2010 to 2014 were 22.64, 5.74, and 3.20, respectively (Barnholtz-Sloan et al., 2018).

The neuroimaging data provide precious information related to the shape, size, location, and metabolism of brain tumors for clinicians, which can be used to assess the status of the tumor before and after therapeutic intervention (Magadza & Viriri, 2021). Magnetic resonance imaging (MRI) sequences such as T1-weighted (T1w), T2-weighted (T2w), T1-contrasted (T1c), and fluid attenuation inversion recovery (FLAIR) provide substantial contrast for various brain tissues, which can be used to discriminate the different parts of tumor and normal tissues (Saha & Panda, 2018; Wadhwa, Bhardwaj, & Singh Verma, 2019).

Brain tumor segmentation for locating and assessing the tumor region and tumor classification for recognizing its grade are essential steps for choosing the proper treatment plan and effectiveness of treatment. As a result, physicians usually perform these procedures manually before starting the treatment plan; in the meantime, manual segmentation and classification of tumors are laborious and time-consuming tasks and different specialists may have varying diagnoses, which may reduce treatment effectiveness. (Ranjbarzadeh et al., 2021).

Considering the reasons mentioned above, automatic brain tumor segmentation could prepare valuable morphological information for clinicians about different tumor parts, including core, enhanced, and whole tumor regions, leading to timely and proper diagnosis and treatment of brain tumors (Ranjbarzadeh et al., 2021; Saman & Jamjala Narayanan, 2019). In recent decades, machine learning (ML) advances have led to increasing interest in automatic medical image analysis. However, traditional ML approaches mainly require prior knowledge and a manual feature extraction process, which can be time-consuming. As a new subcategory of ML, deep learning recently showed major benefits in overcoming the limitations of traditional ML approaches (Fathi, Ahmadi, & Dehnad, 2022). More complex and high-level features can be extracted automatically and then given to a deep learning-based classification or segmentation algorithm, which means that the feature extraction and classification/segmentation steps are merged in deep learning (Fathi et al., 2022; Havaei et al., 2017). Among various deep learning approaches, convolutional neural network (CNN) architectures have shown superior results, especially in detecting and analyzing neurological diseases (Fathi et al., 2022; Valliani, Ranti, & Oermann, 2019).

In this study, a CNN-based architecture called U-Net has first applied to the BraTS 2018 dataset for automatic brain tumor segmentation to determine the different parts of the tumor, namely core, enhanced, and whole tumor regions. Then, another CNN architecture called VGG16 was trained to classify the images into low and high grades (LGG/HGG) tumors. The transfer learning parameter initialization method was used to decrease the learning time and enhance the performance of the model. Evaluating the applicability of the model was done by gathering and using a local dataset to assess the trained models. During the final step, the volume of the tumor regions before and after radiotherapy on cases in the local dataset was measured and statistically compared to assess the efficiency of the treatment.

To the authors' best knowledge, the assessment of treatment responses pre and post-radiotherapy in glioma patients has been mentioned in none of the reviewed studies. Hence, due to the importance of assessing treatment response in these patients, we evaluated that in glioma patients who underwent 3D conformal radiotherapy about 3 to 8 months after. Hence main contributions of this study are described as follows:

1. A multimodal approach including four primary MRI sequences, namely T1w, T2w, T1c, and FLAIR, was utilized to enhance the accuracy of segmentation and classification procedures.
2. The segmentation process was done on both LGG and HGG cases.
3. The final segmentation and classification models were applied to a local dataset to assess their practical applicability.
4. The proposed models have been conducted on the local dataset before and after radiotherapy intervention aiming to measure the effectiveness of treatment statistically.

1.1 Related works

In this section, we glance at similar studies and briefly review their automatic brain segmentation method. Brain tumor segmentation (BraTS) challenges, which have been held annually since 2012, have inspired the application of ML approaches in this field. The most popular methods used in the early years of the BraTS challenge were based on traditional ML approaches such as Random Forest (RF), logistic regression, Markov Random Field (MRF), and Conditional Random Field (CRF) (Ghaffari et al., 2020). Despite the unsatisfying performance of the traditional methods, they opened new ways for automatic brain tumor segmentation. Later advances in the computational power of computers and the ML approaches made deep learning-based methods more popular in this field. Concerning this, convolutional neural networks (CNN) were used in brain tumor segmentation for the first time in 2014.

Urban et al. in 2014 used a simple 3D-CNN architecture with three convolutional layers as a voxel-based classification method to classify edema, non-enhancing tumor, enhancing tumor, necrosis, air, and normal tissue in the multimodal images and obtained fair results for whole, core and enhancing tumor segmentation (Urban, Bendszus, Hamprecht, & Kleesiek, 2014). In 2015, Havaei et al. proposed a 2D-CNN architecture called InputCascadeCNN model for brain tumor segmentation. They achieved Dice scores of 0.88, 0.79, and 0.73 for the whole tumor, tumor core, and enhancing tumor, respectively (Havaei, Dutil, Pal, Larochelle, & Jodoin, 2015). In 2016, a new CNN-based architecture called DeepMedic was introduced by Kamnitsas et al., which comprised 11 3D convolution layers with residual connections to obtain a more efficient model. They reported Dice scores of 0.89, 0.76, and 0.72 for whole, core, and enhancing tumor regions, respectively (Kamnitsas et al., 2016).

In recent years, the number of studies using CNN-based models, especially U-Net architecture, has increased dramatically. Chen et al. applied a novel separable 3D U-Net architecture on BraTS2018 and reached the Dice scores of 0.69, 0.84, and 0.78 for enhancing tumor, whole tumor, and tumor core, respectively (Chen, Liu, Peng, Sun, & Qiao, 2019). In another submitted study to BraTS2018, Feng et

al. (Fang & He, 2018) used three 3D U-Nets with different hyperparameters and combined them via simple averaging of the probability of all classes obtained by each model. They reported Dice scores of 0.9, 0.83, and 0.78 for the whole tumor, tumor core, and enhancing tumor, respectively (Fang & He, 2018). Similar to the previous study, Caver et al. also used three different U-Nets to segment brain tumors, with the difference that each model has been employed for segmenting one region of interest. The Dice scores for the whole tumor, tumor core, and enhancing tumor were 0.87, 0.76, and 0.72, respectively (Caver, Chang, Zong, Dai, & Wen, 2018).

Kermi et al. proposed a 2D U-Net architecture to segment the whole and other intra-tumor regions. In order to address the class imbalance issue, they have used novel loss functions called Weighted Cross Entropy (WCE) and Generalized Dice Loss (GDL). They achieved the Dice scores of 0.86, 0.80, and 0.76 on validation data for the whole tumor, tumor core, and enhancing tumor, respectively (Kermi, Mahmoudi, & Khadir, 2018). In Naser et al.'s study, the closest one to our study, a multi-task deep learning-based method has been proposed to segment and classify grades II and III of gliomas. The segmentation and classification tasks were carried out by U-Net and VGG16 architectures, respectively. The reported Dice score and tumor detection accuracy were 0.84 and 0.92, respectively (Naser & Deen, 2020).

Materials And Methods

The proposed model for segmentation and classification procedures is fully automated and is based on two different 2D-CNNs. The main steps of the proposed model are described as follows: preprocessing of the 3D-MRI data, training, and creation of the segmentation model by using a U-net architecture, and using a VGG16 architecture along with transfer learning to classify the images into HGG and LGG tumors. Next, the images of the local dataset from pre and post-radiotherapy were preprocessed. The automatic segmentation was applied to the local dataset to measure the volume of the tumor and its regions. Finally, the effectiveness of treatment was statistically evaluated by comparing the volume of the tumor before and after radiotherapy. The flowchart of the current study process is shown in Figure 1.

2.1 BraTS datasets

The brain tumor segmentation (BraTS) challenge, which contains a multimodal MRI image dataset, has been held since 2012 annually. The BraTS 2018 training dataset consists of 210 HGG and 75 LGG scans. These image sets consist of four MRI sequences, namely T1w, T1c, T2w, and FLAIR. All images have a volume dimension of 240× 240 × 155 and have been manually annotated by expert neuro-radiologists into four types of intra-tumoral regions, namely necrotic core, non-enhancing (1), edema (2), and active/enhancing (4) tumor.

2.2 Segmentation

2.2.1. Preprocessing

In the first step, the background of all images with unimportant information was removed; hence, the size of the images was reduced to 190×190×135. Next, data normalization have applied to each sequence of the MRI images by subtracting the mean and dividing by the standard deviation of the intensities within the slices. Ground truth and four MRI scanning sequences (T1W, T2W, T1c, and FLAIR) were saved in a 4-dimensional matrix as input.

2.2.2. Data Augmentation

Data augmentation is the process of creating synthetic data from original data aiming to increase the amount of training set and improve the generalizability of the final model. To this end, the patch-based method was performed so that four patches with a size of 160 ×160 were randomly extracted from each slice of images.

2.2.3. Network Architecture and Training

The proposed segmentation model was based on U-Net architecture. As shown in Figure 2, it consists of an encoding (left side) and decoding path (right side). The contracting path consists of five convolutional blocks, each containing two 3×3 convolutional layers with a stride of one and a 2×2 max pooling layer after each block except the last one. Hence, the size of feature maps reduces from 160×160 to 10×10. The utilized activation function was rectified linear unit (ReLU).

The expanding path includes five convolution blocks, in which every block starts with an up convolutional layer with a filter size of 2×2 and stride value of two, aiming to double the size of feature maps in both directions and decrease the number of feature maps by two. So the size of feature maps increased from 10 × 10 to 160 ×160. The number of feature maps was reduced by half from each block to the next, in order to maintain symmetry. The decoding path is followed by a concatenation layer and two convolutional layers with a filter size of 3×3. At the end of expanding path, a 1×1 convolution with a softmax activation function is used to map the multi-channel feature maps to the desired four classes. A dropout with the value of 0.2 was utilized between all two convolutional layers to prevent overfitting in all blocks. Unlike the original U-Net architecture, we used zero padding to keep the output dimension for all the convolutional layers of both the down-sampling and up-sampling paths.

Selecting the proper loss function in multi-class segmentation problems with class imbalance in the foreground and background brain tumor data is essential. This also improved model accuracy. Accordingly, we used a hybrid loss function that combines Weighted Cross Entropy (WCE) and Generalized Dice Loss (GDL). Four patches with the size of 160×160×4 are extracted randomly from each subject and used for training the model. The stochastic gradient-based (SGD) algorithm was used as a parameter optimization algorithm. Other hyperparameters such as momentum, initial learning rate, and epoch numbers were 0.9, 0.01, and 5, respectively, in which the learning rate exponentially decayed with the factor of 0.0001. The proposed method was tested and evaluated quantitatively on both BraTS 2018 and local datasets.

2.3. Classification

2.3.1. preprocessing

The preprocessing steps for the classification phase are described as follows: Considering the different number of slices with tumor regions in images of each patient, 32 slices around the slice with the largest tumor region were extracted to increase the accuracy and reduce the computational costs in each modality. The proposed automated segmentation model was conducted on images to reduce the effect of background pixels in the non-tumor area. Next, the intensity of non-tumor pixels was multiplied by 0.2 to reduce the effect of background pixels. The size of images was reduced to 224×224, and the normalization process was accomplished by subtracting the mean intensity of pixels divided by the standard deviation as well. Three MRI sequences, namely T2W, T1c, and FLAIR, were used as three-channel input images so that the size of each image was 224×224×3. A data augmentation technique has been applied on LGG samples with horizontal rotation in favor of balancing classes in the training set. Generally, the training and test set comprised 6890 and 1824 slices, respectively.

2.3.2. Network Architecture and Training

In this study, a VGG16 architecture and transfer learning method were used to classify HGG and LGG. The VGG16 is a predefined 2D-CNN-based architecture with 16 layers comprising five convolutional blocks with ReLu activation function, max-pooling layers at the end of each block, and three fully connected layers (see Figure 3). Lower layers in the pre-trained CNNs contain the generic and low-level features, while the top layers contain the more specific and rich features. Hence, the parameters of earlier layers can be frozen due to low-level features. In the current study, only the last three layers were left trainable in favor of classification. The optimization algorithm and loss function used in the current study were SGD and categorical cross entropy, respectively. Other hyperparameters such as learning rate, batch size, and epochs number were 0.001, 16, and 8, respectively. The model was built-up using Keras and Tensor flow as the backend. The data was split into training, validation, and test sets with a ratio of 60:20:20.

2.4. Local dataset

In the current study, a local dataset of MRI scans obtained from 49 patients with HGG and LGG glioma tumors (30 males with median age 46 (28-63) years old, and 19 females with median age 49 (31-66) years old) were retrospectively collected in order to validate of proposed segmentation and classification models. Patients were longitudinally scanned as part of their routine clinical follow-up. The time interval between pre and post-therapy scans varied from 3 to 8 months, and the dataset was collected from 2018-2020 from Isfahan Milad Hospital and General Hospital. 46.9% of patients were diagnosed with HGG, and 53.1% with LGG. The patients who died or were unable to participate in the study and patients with pilocytic astrocytoma (grade I) due to its high prevalence in children were excluded from the study as exclusion criteria. While LGG (grade II) and HGG (grade III, IV) were included in this study because of their high prevalence in adults.

2.4.1. Data labeling

The gold standards for proposed segmentation and classification models in the local dataset were the radiologist's diagnosis and histopathological results. The ground truth for the MRI images of patients was based on manual segmentation, initially performed by an MRI expert and rechecked by a radiologist. The manual segmentation was performed at the slice (2D) level using ITK-SNAP commercial software.

2.4.2. Data preprocessing Measuring the volume of different tumor regions

Data preprocessing steps for the local dataset describe as follow: 1) Realignment of T2WI, T1WI, and FLAIR to T1c images performed using CaPTK for longitudinal scans. 2) All images from different time points were registered to the baseline scan. 3) Bias field correction performed using intensity inhomogeneity correction algorithm implemented in CaPTK and signal intensity normalization images. 4) Intensity normalization was performed in order to overcome the high heterogeneity in the images.

The automatic segmentation process was applied to the local dataset, pre and post-radiation therapy to determine the whole tumor volume and its different regions. The total number of pixels in each tumor region with dimensions of 1 mm^3 was considered to determine the volume.

Results

3.1. Segmentation results

As mentioned before, a multimodal approach based on four MRI sequences has been used for automatic segmentation in this study. You can see the result of the segmentation model on two sample subjects and compare the results with ground truth in Figure 4.

As shown in Figure 4, four left columns display four sequences of MR images, and the following two columns show the manual annotation of neuro-radiologists and the proposed model results, respectively. The split ratio used in this study was 80:20 for the training and validation dataset, respectively. Performance metrics were used to evaluate segmentation results, including mean dice similarity coefficient (DSC), sensitivity, specificity, and Hausdorff distance. The result of which for enhanced tumor (ET), whole tumor (WT), and necrotic tumor (NT) regions are given in Table 1. The DSC was the main parameter for evaluating the segmentation accuracy on the clinical dataset. The performance results of 0.76, 0.71, and 0.70 for DSC in the WT, NT, and ET respectively.

Table 1. Quantitative evaluation of segmentation results

Performance metrics	Necrotic tumor (NT)	Whole tumor (WT)	Enhanced tumor (ET)
Hausdorff distance	7.20	8.82	4.43
Dice	0.71	0.76	0.70
Specificity	0.998	0.996	0.999
Sensitivity	0.86	0.91	0.80

3.2. Classification results

The classification performance of the proposed model is assessed based on both BraTS 2018 and local datasets. Figure 5 shows the results of the confusion matrix for the two mentioned datasets. Table 2 also reports the performance of classifying the tumor grades (LGG vs. HGG).

Table 2. Results of classification

dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
BraTS	99.1	98.93	99.91	99.69	99.71
Clinical	89.8	91.30	88.46	87.5	92

3.3. Tumor volume results

In order to evaluate the effectiveness of radiotherapy treatment, we have measured and compared the tumor region volumes before and after therapy. The results were statistically analyzed with SPSS v25 software. The paired sample t-test and Wilcoxon signed ranks were used as statistical tests with a significance level of 0.05 in this study.

The results of paired t-test showed that the radiotherapy had been significantly effective in decreasing the volume of the whole tumor. However, the volume differences of other tumor regions (NT and edema) were not statistically significant pre and post-radiotherapy. Meanwhile, the non-parametric Wilcoxon's test was significant, with a 95% confidence level for evaluating the mean difference between ET volumes pre and post-therapy.

Discussion

In this study a deep learning-based model has been introduced to segment glioma tumors and their intratumoral regions via U-Net automatically, then we classified HGG and LGG using the results of the tumor segmentation phase. The volume of different tumor regions was determined using the segmentation results from pre and post-radiotherapy. The most critical finding was the positive effect of radiation therapy in glioma patients on the WT and the ET volumes; however, it had no significant effect on the peritumoral edema and NT volume.

A patch-based data augmentation method was performed in the segmentation stage, in which four random areas from each image slice were extracted for both LGG and HGG data. In the classification stage, the data augmentation was performed only by rotating the LGG data. The use of data augmentation methods has increased the variety of data for training, reducing prediction errors in the test data and improving the model's generalizability. However, using extra data augmentation methods would lead to increase training time and computational cost.

Considering the expected effect of radiation therapy on the treatment of brain tumors and the effects of stopping the growth and death of the cells, the results of this study, which indicate the reduction of WT volume, as well as the reduction absorptionof contrast agent in the ET region, are consistent with the radiology principles. The effects of radiotherapy on tumor cells, which grow faster than healthy cells, leading to single-stranded and double-stranded DNA breaks, accelerating apoptosis. Consequently, the removal of malignant cells leads to the reduction of the WT volume (Kim et al., 2019).

Results obtained in this study using DL approach represent no significant difference in the NT and edema volume between pre and post-radiotherapy, which was not unexpected. However, in general, the peritumoral edema is expected to decrease due to the treatment response and the reduced number of malignant cells. It is noteworthy that a variety of mechanisms influence the origin and persistence of peritumoral edema. Both vasogenic and cytotoxic edema is involved in tumor-induced edema. Considering the unique characteristics of the brain and central nervous system (CNS), it should be kept in mind that the elimination of peritumoral edema caused by the tumor may require a long time. At the same time, the patient's MRI images were performed 3-8 months after radiotherapy. Although the peritumoral edema induced by malignant cells can be diminished by radiotherapy, it can trigger peritumoral edema due to different mechanisms observed post-radiotherapy on MRI.

Necrotic regions are clearly visible, especially in grade IV post-radiotherapy, and the presence of necrotic areas can lead to misdiagnosis. Necrotic areas can be caused by cancer cell destruction post-radiotherapy, especially in hypoxic areas. On the other hand, radiation-induced necrosis is a potential long term CNS complication and also distinguishing radiation-induced necrosis from tumor recurrence is especially challenging in neuro-oncology.

Reviews of studies revealed that recently proposed CNN-based architectures, especially the U-net, could obtain desirable results. Chen et al. (Chen et al., 2019), Fang et al. (Fang & He, 2018), Caver et al. (Caver et al., 2018), and Kermi et al. (Kermi et al., 2018), in their studies, applied U-net-based architectures on BraTS2018 for automatic brain tumor segmentation and obtained promising results. Although their results were slightly better than ours, their analysis was only limited to tumor segmentation. None of the abovementioned studies reported using tumor segmentation and classification to perform a treatment response assessment.

Mazaheri et al. (Mazaheri et al., 2022) recently reviewed studies that used DL to assess treatment responses in different cancers, including brain cancers. A combination of hand-crafted features and MRI images (or fMRI) has been used in most brain cancer-related studies to predict survival time and quantify

tumor response (Han et al., 2020; Lao et al., 2017; Nie et al., 2019). Similarly to our study, Kickingereder et al. (Kickingereder et al., 2019) deployed a DL-based model to segment the brain tumor into enhanced and non-enhanced regions and to quantify tumor response. However, unlike our study, they did not use the BraTS dataset to create the model. The review done by Tandel et al. (Tandel et al., 2019) concluded that most of the studies focused on tumor segmentation and tissue classification, while there is a great deal of potential for further research into tumor grading. Furthermore, the overall training time of our proposed model was almost 12 hours, which is much lower than studies that reported the training time.

Conclusion

Using automatic segmentation and classification methods from medical images has an important role in helping physicians to diagnose medical lesions. Automatic methods of tumor segmentation and intramural as well as tumor classification to determine the grade of glioma tumor can help physicians in accurate diagnosis and optimal treatment planning. Different methods of deep learning can be helpful in the treatment response assessment in order to more accurately determine the criteria of treatment response, including RNAO and determining the 3D volume of different tumor regions.

Declarations

Acknowledgements

This study was funded by Isfahan University of Medical Sciences and ethically confirmed by Iran National Committee for Ethics in Biomedical Research [code No. IR.MUI.MED.REC.1398.697]

Statements and Declarations

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Ahmed, R., Oborski, M. J., Hwang, M., Lieberman, F. S., & Mountz, J. M. (2014). Malignant gliomas: current perspectives in diagnosis, treatment, and early response assessment using advanced quantitative imaging methods. *Cancer management and research, 6*, 149.
2. Barnholtz-Sloan, J. S., Ostrom, Q. T., & Cote, D. (2018). Epidemiology of brain tumors. *Neurologic clinics, 36*(3), 395-419.
3. Caver, E., Chang, L., Zong, W., Dai, Z., & Wen, N. (2018). *Automatic brain tumor segmentation using a U-net neural network.* Paper presented at the Pre-conference proceedings of the 7th MICCAI BraTS challenge.
4. Chen, W., Liu, B., Peng, S., Sun, J., & Qiao, X. (2019, 2019//). *S3D-UNet: Separable 3D U-Net for Brain Tumor Segmentation.* Paper presented at the Brainlesion: Glioma, Multiple Sclerosis, Stroke and

Traumatic Brain Injuries, Cham.

5. Fang, L., & He, H. (2018). *Three pathways U-Net for brain tumor segmentation.* Paper presented at the Pre-conference proceedings of the 7th medical image computing and computer-assisted interventions (MICCAI) BraTS Challenge.
6. Fathi, S., Ahmadi, M., & Dehnad, A. (2022). Early diagnosis of Alzheimer's disease based on deep learning: A systematic review. *Computers in biology and medicine, 146*, 105634. doi:https://doi.org/10.1016/j.compbiomed.2022.105634
7. Ghaffari, M., Sowmya, A., & Oliver, R. (2020). Automated Brain Tumor Segmentation Using Multimodal Brain Scans: A Survey Based on Models Submitted to the BraTS 2012–2018 Challenges. *IEEE Reviews in Biomedical Engineering, 13*, 156-168. doi:10.1109/RBME.2019.2946868
8. Han, W., Qin, L., Bay, C., Chen, X., Yu, K.-H., Miskin, N., . . . Young, G. (2020). Deep transfer learning and radiomics feature prediction of survival of patients with high-grade gliomas. *American Journal of Neuroradiology, 41*(1), 40-48.
9. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., . . . Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical image analysis, 35*, 18-31.
10. Havaei, M., Dutil, F., Pal, C., Larochelle, H., & Jodoin, P.-M. (2015). *A convolutional neural network approach to brain tumor segmentation.* Paper presented at the BrainLes 2015.
11. Kamnitsas, K., Ferrante, E., Parisot, S., Ledig, C., Nori, A. V., Criminisi, A., . . . Glocker, B. (2016). *DeepMedic for brain tumor segmentation.* Paper presented at the International workshop on Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries.
12. Kermi, A., Mahmoudi, I., & Khadir, M. T. (2018). *Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes.* Paper presented at the International MICCAI Brainlesion Workshop.
13. Kickingereder, P., Isensee, F., Tursunova, I., Petersen, J., Neuberger, U., Bonekamp, D., . . . Foltyn, M. (2019). Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. *The Lancet Oncology, 20*(5), 728-740.
14. Kim, W., Lee, S., Seo, D., Kim, D., Kim, K., Kim, E., . . . Youn, B. (2019). Cellular stress responses in radiotherapy. *Cells, 8*(9), 1105.
15. Lao, J., Chen, Y., Li, Z.-C., Li, Q., Zhang, J., Liu, J., & Zhai, G. (2017). A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. *Scientific Reports, 7*(1), 1-8.
16. Magadza, T., & Viriri, S. (2021). Deep learning for brain tumor segmentation: a survey of state-of-the-art. *Journal of Imaging, 7*(2), 19.
17. Mazaheri, Y., Thakur, S. B., Bitencourt, A. G., Lo Gullo, R., Hötker, A. M., Bates, D. D., & Akin, O. (2022). Evaluation of cancer outcome assessment using MRI: A review of deep learning methods. *BJR| Open*(1), 20210072.
18. Naser, M. A., & Deen, M. J. (2020). Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Computers in Biology and Medicine, 121*, 103758.

19. Nie, D., Lu, J., Zhang, H., Adeli, E., Wang, J., Yu, Z., . . . Shen, D. (2019). Multi-channel 3D deep feature learning for survival time prediction of brain tumor patients using multi-modal neuroimages. *Scientific Reports, 9*(1), 1-14.
20. Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S., Anari, S., Naseri, M., & Bendechache, M. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Scientific Reports, 11*(1), 1-17.
21. Saha, M., & Panda, C. (2018). A review on various image segmentation techniques for brain tumor detection. *Sci Res*, 21-30.
22. Saman, S., & Jamjala Narayanan, S. (2019). Survey on brain tumor segmentation and feature extraction of MR images. *International Journal of Multimedia Information Retrieval, 8*(2), 79-99. doi:10.1007/s13735-018-0162-2
23. Tandel, G. S., Biswas, M., Kakde, O. G., Tiwari, A., Suri, H. S., Turk, M., . . . Khanna, N. (2019). A review on a deep learning perspective in brain cancer classification. *Cancers, 11*(1), 111.
24. Tiwari, A., Srivastava, S., & Pant, M. (2020). Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019. *Pattern Recognition Letters, 131*, 244-260.
25. Urban, G., Bendszus, M., Hamprecht, F. A., & Kleesiek, J. (2014). *Multi-modal Brain Tumor Segmentation using Deep Convolutional Neural Networks*.
26. Valliani, A. A.-A., Ranti, D., & Oermann, E. K. (2019). Deep Learning and Neurology: A Systematic Review. *Neurology and Therapy, 8*(2), 351-365. doi:10.1007/s40120-019-00153-8
27. Wadhwa, A., Bhardwaj, A., & Singh Verma, V. (2019). A review on brain tumor segmentation of MRI images. *Magnetic Resonance Imaging, 61*, 247-259. doi:https://doi.org/10.1016/j.mri.2019.05.043

Figures

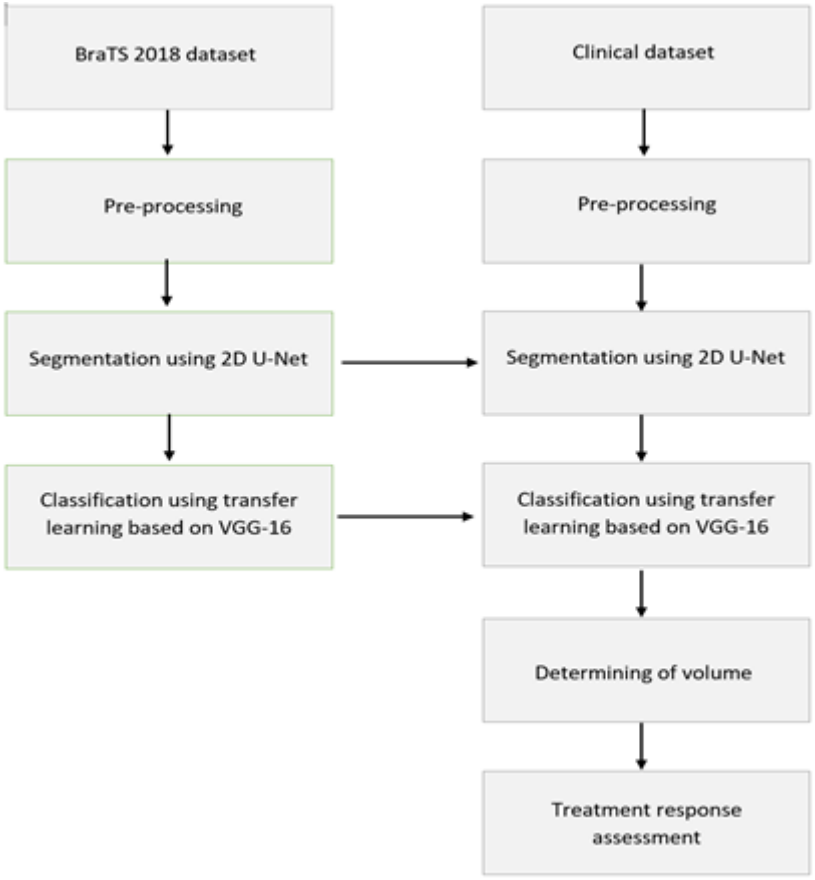

Figure 1

Flowchart of our study process

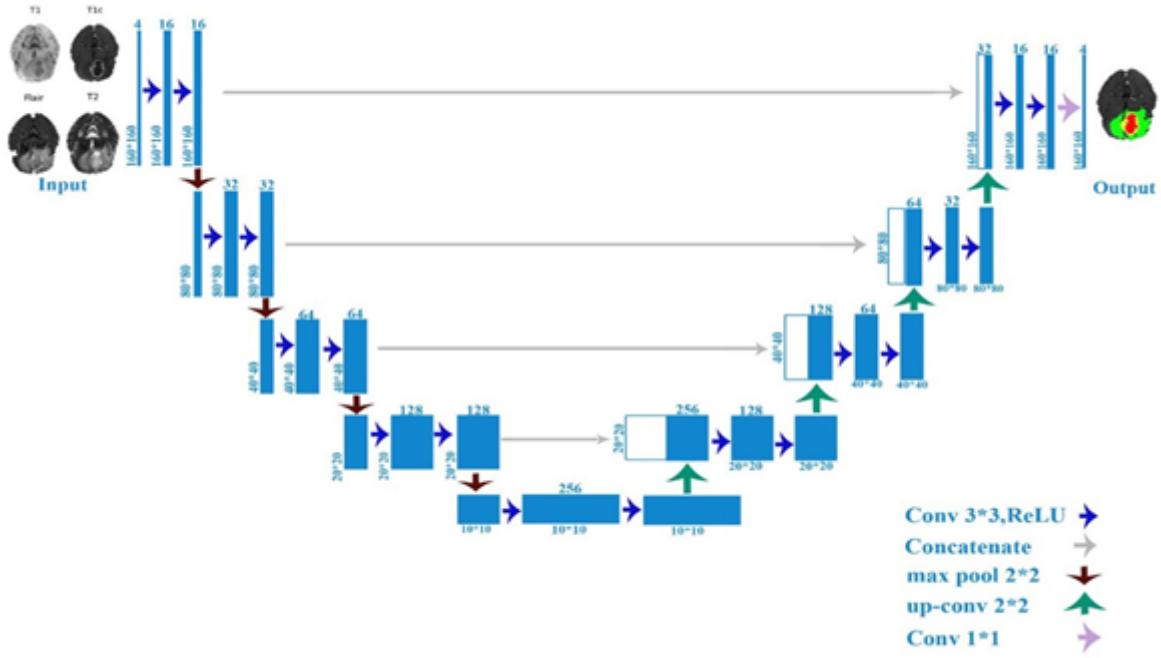

Figure 2

The proposed U-Net architecture used for automatic segmentation

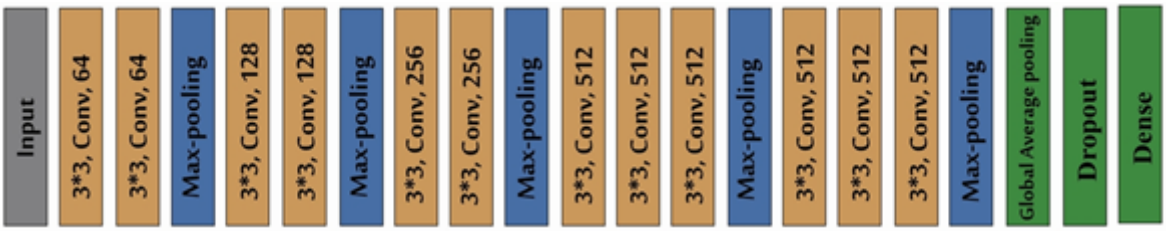

Figure 3

The proposed VGG16 architecture used for the classification

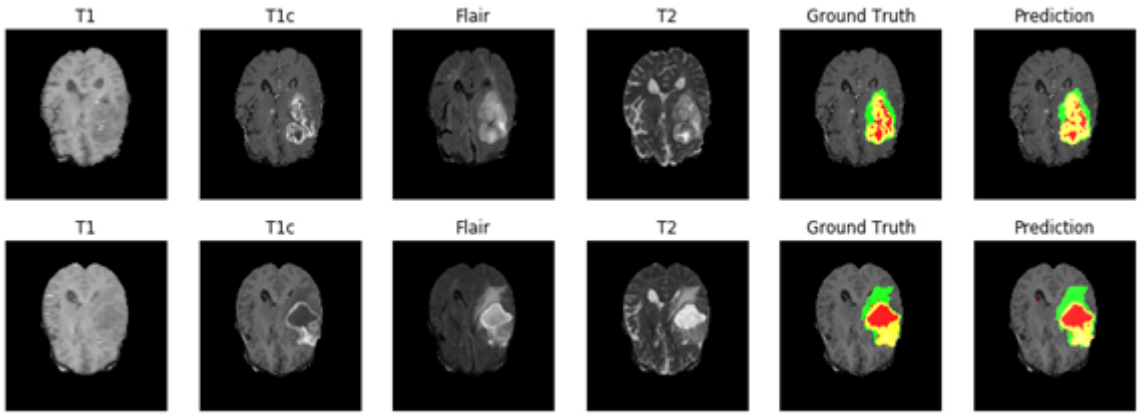

Figure 4

The results of the segmentation model on two random samples

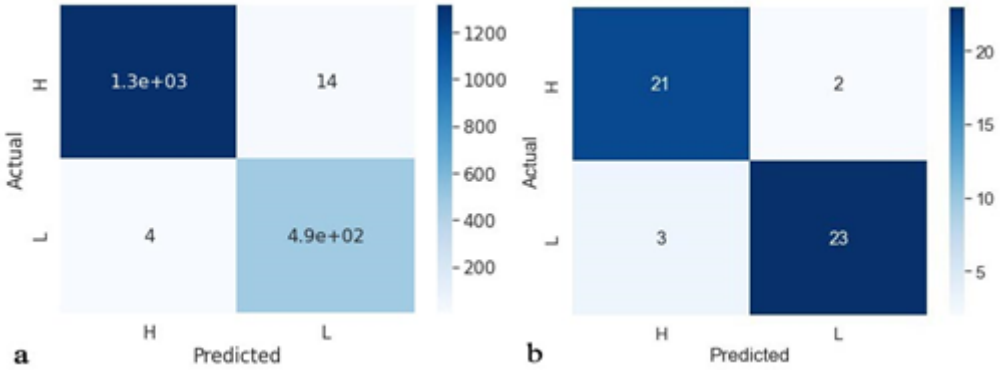

Figure 5

Confusion matrix for classification of the tumor grade (HGG vs. LGG) in - a) BraTS dataset - b) local dataset