Enhancing Brain Tumor Segmentation through Deep Learning: A Comprehensive Analysis and Predictive Framework

Srinidhi Kannan, Varshini Balaji, Suryamritha M, Rimjhim Padam Singh
Department of Computer Science and Engineering, Amrita School of Computing, Bengaluru,
Amrita Vishwa Vidyapeetham, India
bl.en.u4aie21121@bl.students.amrita.edu, bl.en.u4aie21139@bl.students.amrita.edu, bl.en.u4aie21126@bl.students.amrita.edu, ps_rimjhim@blr.amrita.edu

***Abstract*—Brain tumor segmentation plays a major role in the diagnosis and treatment of patients. It has improved accuracy and efficiency in medical imaging. They extract important features from the Magnetic Resonance Imaging scan. This paper uses the Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) 2020 dataset. Various models are trained to analyze which can give accurate segmentation. The findings contribute to enhancing clinical and research applications by providing precise tumor segmentation. This helps in personalized patient care, ensuring treatment can be tailored to the specific characteristics of each tumor.In terms of overall performance, Unetr, Unet++, and Attention Unet performance are extremely good, particularly in Dice coefficient and mean IOU, making them the top three models for brain tumor segmentation.**

***Index Terms*—tumor, segmentation, medical imaging, BRATS,Unet, MRI scan,Dice coefficient,mean IOU**

I. Introduction

Brain tumor segmentation plays a significant role in medical research and diagnosis. It is used for partitioning different tumor areas on multi-modality images [1] processing and collecting the findings from an MRI scan is considered a difficult undertaking in this line of work [2]. Traditional semantic segmentation techniques are problem-specific, therefore there is no generic segmentation technique that works for all pics. However, deep learning (DL) techniques can overcome this difficulty. Convolutional neural networks (CNNs) have recently gained popularity as a model for image segmentation [3]. Using multi-scale features is one of the key components of effective segmentation [4]. Hence DeL is often used in the field of medicine to improve disease detection [5].

Brain tumors need precise diagnosis and personalized treatment. It can be malignant or benign.This can unusual cell growth in the brain or surrounding tissues. There are various symptoms such as headaches, seizures, vision or speech changes, and movement difficulties. The treatment should be done very cautiously with proper detection and planning. The treatment depends on the use of MRI, however manual interpretation can bring inconsistencies and mistakes.Accurate segmentation is essential for clinical and research purposes. The patient should be treated with the utmost care and consideration hence for the diagnosis to be efficient the brain images should be analyzed rightly.

Medical analysis has been revolutionized by CNNs, especially in areas such as the segmentation of tumors in the brain using MRI scans.CNNs excel in s excel in learning and extracting hierarchical features directly from images.

Various CNN models are compared to analyze which outperforms in the segmentation task. Deep learning-based approach outperforms traditional approaches. Data-driven insights are very crucial to understanding the tumor characteristics.The dataset used here is BRATS 2020.It uses MRI scans from multiple hospitals to study brain tumors, specifically gliomas, which vary greatly in appearance and behavior.

Overall the paper compares eight models and a detailed analysis of their performance is done. The models are Unet ,Unet++, Unet2d, Unet3d, Unet-fpn, Unet_Vgg16, Attention Unet, Unetr. Section II presents a literature review, Section III details methodology and contributions, Section IV covers results and analysis, and Section V concludes, summarizing findings and suggesting future research directions.

The key contributions of this paper are :

- Contributes to comparing and analyzing different models' efficiency in brain segmentation
- Enhancing the performance to increase the efficiency by accurate segemention

II. Literature Review

DL models are used in brain segmentation to enhance accuracy and address data imbalances. Various techniques such as ensemble learning, attention mechanisms, and integration of multi-modal data are explored for comprehensive tumor analysis. Somasundaram et al. [6] examines new developments in image segmentation and classification employing effective processing of tumor-infected human brain MRI, focusing on gliomas, and automatically responding to data imbalances. It examines the current state of tumor segmentation and detection with DL models.

Saouli et al. [8] explore a novel DL model for brain tumor segmentation by combining three Incremental Deep

CNN models. The models are intended to segment Glioblastomas, which vary in size, shape, and contrast. They employ Ensemble Learning for apt training and a unique training technique that considers affecting hyperparameters to speed up the process.

Cherguif et al. [9] describe a DL strategy that employs deep CNN based on the U-Net model. The approach was tested using real photos from the BRATS 2017 dataset, which comprised both HGG and LGG patients. According to the results, this technique can produce segmentation that is both effective and resilient when compared to manually determined ground truth.

Madhupriya et al. [10] employs DL techniques, CNN and Probabilistic Neural Networks (PNN), to find of tumors in brain in MRI images. They suggest an intelligent system that employs DL to segment aberrant brain tissues and compare them to normal brain tissues and tumor cells. They employ a CNN architecture with 3x3 and 7x7 overlap, as well as a cascaded architecture, with the Brats13 imagine dataset. In contrast to traditional image processing and computer vision models, the model considers both local and global information. Thillaikkarasi et al. [11] demonstrates a novel algorithm for MRI pictures called Kernel-based CNN with Multiple SVM (M-SVM), and uses 40 MRI images to compare its performance to current techniques. Pre-processing, extraction of features, classification of image using M-SVM, and tumor segmentation using a kernel-based CNN are all part of the suggested methodology. The method shows good segmentation and integration of DL techniques, achieving almost 84% accuracy. Sajid et al. [12] present a hybrid CNN that incorporates contextual and local data. Using a patch-based approach, dropout regularization and batch normalization reduce overfitting. Beyond state-of-the-art techniques, this strategy exhibits significant improvements, attaining high dice score, sensitivity, and specificity values across the whole tumor region, providing accurate segmentation and efficient management of overfitting and imbalanced data.

Raut et al. [13] Despite the drawbacks of autoencoders and K-Means segmentation, a CNN-based DL approach is suggested for the identification and segmentation of brain tumors from MRI images. This method improves classification accuracy through convolution layers, max pooling, and activation functions such as ReLU. Yogananda et al. [14] investigates the segmentation of brain tumors by automated DL networks, demonstrating its ability to appropriately represent the subcomponents of the tumor. The network performs admirably when using the Dice coefficient and cross-validation, providing a quick fix for clinical workflow integration. Notwithstanding its benefits, its drawbacks include a lack of available data, computational restrictions, and the possibility of performance disparities in clinical settings. Naser et al. [15] employ DL to segment and grade LGG brain tumors from MRI images, showing encouraging results in terms of segmentation and grading accuracy. The study highlights the promise of deep learning for non-invasive, automated tumor segmentation, detection, and grading in LGG, despite obstacles such data validation and class imbalance. Akil et al. [16] presents a Deep CNN model that incorporates Occipito-Temporal pathway-inspired selective attention processes for the automatic segmentation of Glioblastoma brain tumors. The model shows promise for research and clinical applications, as well as computational efficiency, with noteworthy performance on BRATS-2018. Magadza et al. [17] investigates DL strategies for brain tumor segmentation, including ensemble approaches, UNet models, hierarchical segmentation, and cascaded architectures. Although these techniques are promising for assessing tumor features in order to plan treatments, they are not as robust as expert performance. Specific restrictions and difficulties, such computational needs and generalizability, are not sufficiently explored in this article. DL techniques nevertheless provide a great deal of progress in medical image analysis. Aboussaleh et al. [18] minimizes operator intervention through data augmentation and FLAIR image integration in a novel CNN-based technique [18] for the simultaneous prediction and segmentation of brain cancers. Despite attaining a Dice similarity coefficient of 82.35% and good accuracy, the study is limited by the specificity of the dataset and the absence of resource analysis. To address the issues of interpretability and generalizability, more research is required. Liu et al. [19] The paper investigates DL methods for brain tumor segmentation, emphasizing multi-modality integration, network design, and imbalanced segmentation. While highlighting the advantages of multi-modality data, it also points out the constraints in terms of investigating novel topologies and computing difficulties. Huang et al. [20] Glioma MRI image synthesis is enhanced by CoCa-GAN, a 3D Context-aware Generative Adversarial Network, which produces intricate synthetic images. Nevertheless, it is deficient in comprehensive segmentation comparisons, quantitative evaluation metrics, and computing efficiency.

Despite substantial progress, the subject of brain tumour segmentation still has major research gaps. These include limitations in model generalizability across varied patient groups and imaging modalities, a lack of real-time processing capabilities for clinical usage, and difficulties in achieving both high precision and recall at once. Furthermore, many models struggle to handle variations in tumor shape, size, and location efficiently. Our study fills these gaps by creating a comprehensive segmentation framework that includes multi-modal imaging data, powerful custom image generating algorithms, and hybrid DL structures. This strategy improves model resilience, provides greater generalization across multiple clinical circumstances, and increases tumor segmentation precision and recall, all of which contribute to more accurate and dependable clinical outcomes.

III. Methodology

A. Dataset Description

The BRATS 2020 dataset consists of MRI images including T1, T1c, T2, and FLAIR for recording the features of the brain tumors. T1 images give native scans that involve different slice thickness, while T1c comprises of contrast-enhanced views with constant voxel value. T2 images show the features of the tumor through the axial scan and the different thickness of the slices. FLAIR images are axial, coronal or sagittal and offer information about the presence and location of the tumor. Pretreatment segmentation classes include "NOT tumor," "NECROTIC/CORE," "EDEMA," and "ENHANCING" for determining specific tumor areas. The extensive dataset is useful for detailed analysis and designing of the brain tumor segmentation models.

B. Data Preprocessing

During the data preprocessing step, open-source Neuroimaging Python libraries are used to manage the loading of MRI images along with their segmentation masks from the BRATS2020 dataset. This includes accessing several image modalities like T1, T1c, T2, and FLAIR, together with their segmentation masks. After loading data, various visualizations are used to understand the structure of the dataset and its characteristics. The raw images are presented along with the masks, which allow the researchers to decipher the spatial relationship of the tumor areas between the two modalities. These visualization techniques are beneficial since they form the basis for the rest of the modeling approach and provide a first glimpse into the nature of the data before model creation begins.

C. Data Generation

During the data generation phase, custom data generators are developed to handle large datasets while ensuring equality in the number of batches of images and their corresponding masks. It also reduces bias that results from large variation in the distribution of data in classes or instances. The primary activity of the data generation phase is to read MRI images and the corresponding masks from the dataset. Then we resize these images and masks to host the expected dimensions by the model for input. Thus, resizing it makes the data fit in the model architecture and can be easily incorporated in the training process. This preventative measure maximizes efficiency and assures that the model will be trained on a sample set that is balanced and definitively diverse, thus improving the accuracy of the model and its ability to generalize effectively.

D. Dataset Splitting

In data partitioning, the data is split into train, val and test set for an improved evaluation of the model. This systematic approach allows for a stable performance assessment of the model while also avoiding overfitting or the use of biased values. It is in this way possible to learn more about the differences in data and increase the chances of the model generalizing on unseen data through proper data allocation.

E. Model Training

Model training refers to running the architecture with proper optimization functions and learning rates. The training data is used to modify the parameters of the model while tuning the parameters of the model and checking for over-fitting is done using the validation set.

1) UNET: UNet is a familiar architecture implemented in medical image segmentation such as brain tumor segmentation. This model comprises a contracting path for context gathering and an expansive path for accurate positioning. To overcome this issue, UNet adopts skip connections which ensures that high-resolution features are passed through the segmentation of tumor regions.

2) UNET2D: UNet2D is derived from UNet is a network that works on the 2D slices of image. It for brain tumor segmentation shows high accuracy and F1 score making it efficient. It may perform well in terms of image segmentation, but it might not be efficient in capturing the depth of the image.

3) UNET3D: UNet3D is an enhancement to the UNet ,which directly operates over 3D volumes as opposed to 2D image slices and makes use of spatial information for segmentation purposes. Despite the fact it is beneficial in increasing the sample size to define volumetric properties, it may take more time and resources when training the model.

4) UNET-FPN: UNet-FPN is an improved version of the UNet that adds a feature pyramid network to the structure, allowing feature extraction at multiple scales. This further helps the model in understanding the context, at each of the resolutions, thus achieving better segmentation.

5) UNET++: The proposed network architecture is the improved version of the UNet, called UNet++, in which the dense connections are added to help reuse the features at different scales. This leads to accurate segmentation of the tumor, particularly in cases where the shape and size of the growth is non-simple.

6) UNET-VGG: UNet-VGG16 utilizes the UNet structure and combines it with a pre-trained VGG16 network, utilizing VGG's convolutional layers for feature extraction. This approach helps in faster convergence of the training and also improves on the segmentation especially when data set is very limited.

7) UNETR: Recursive UNet or UNetR applies segmentation predictions in a recursive manner to fine tune the model. The above-discussed iterative refinement mechanism helps in enhancing the segmentation accuracy through gradual refinement of feature representations.

8) Attention UNET: Attention UNet also involves the use of attention mechanisms in the UNet model that helps the model to shift its attention towards specific areas in the image during segmentation. This attention mechanism enhances the segmentation results particularly challenging cases such as tumors may have nested structures or the background contains many structures.

F. Evaluation

The model training involves accuracy, mean IoU and class-specific measures which define the model performance correctly at every stage of evaluation. After training, the result analysis entails comparing these measures while also visually observing segmentation outcomes at the same time. The two perspectives allow one to assess the model's positive and

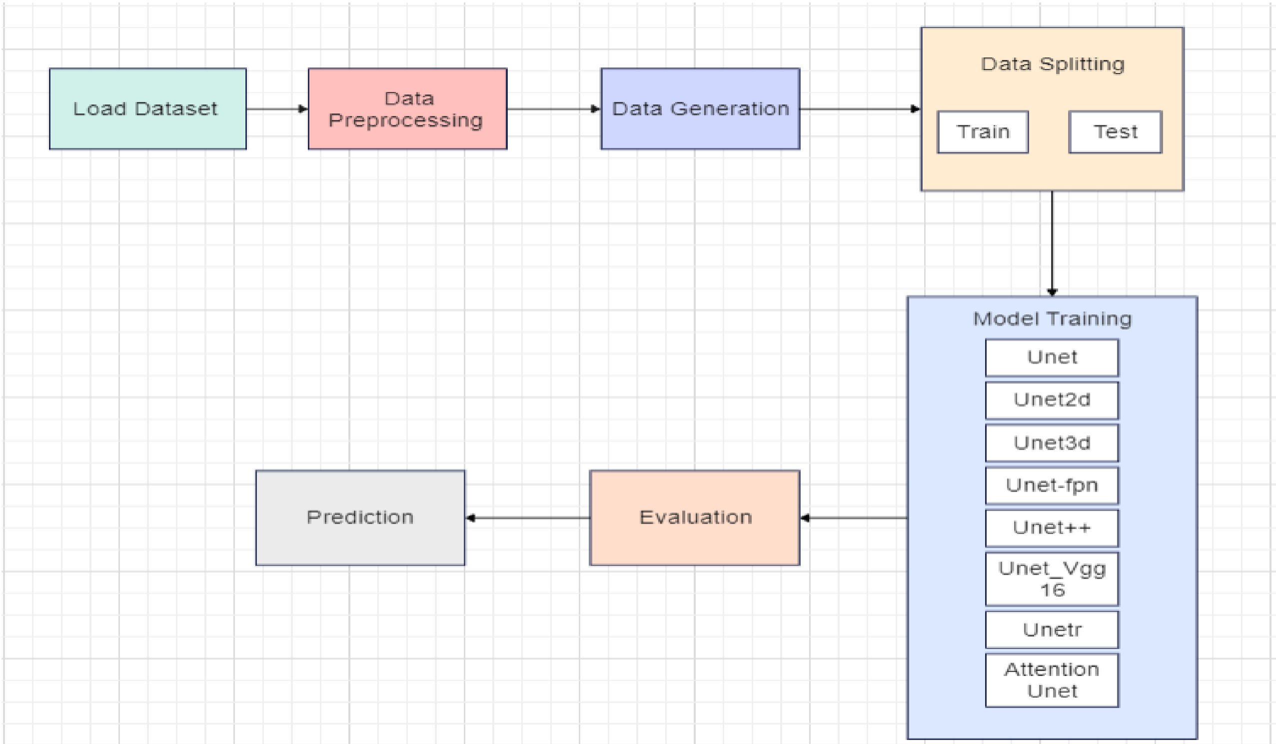

Fig. 1. Prediction Results

negative aspects while making changes in the iterative process. Through deploying a strategy of performance analysis, segmentation accuracy and robustness are improved through repeated enhancements to the architecture and training process. It also maintains the provision of improved reliability and more clinically useful segmentation results by iteratively refining the model's ability to accurately segment brain tumour areas.

G. Prediction

In the prediction process of the model, MRI volumes of brain scans are used to produce segmentation masks that underline various regions of interest like the actual tumor core, edema, and enhancing parts. Basically, these masks are generated through learned patterns and features of the input MRI data by the model. In this way, the model produces a prediction of the presence and size of the tumor based on the MRI volumes slice by slice, which can become useful for medical diagnosis and planning. Comparing the predicted masks with the actual MRI images and ground truth maps allow the determination of the reliability of the model as well as assist clinicians in understanding the results better.

IV. Results

Based on the evaluation of quantitative performance measures provided in the Table I, several conclusions can be drawn about the usefulness of each model for segmentation of brain tumors.Unet2d exhibits high accuracy (99.14%) and F1 score (99.44%), suggesting precise delineation of tumor regions. Unet++, despite slightly lower accuracy, achieves the highest mean IOU (0.8293) and a competitive Dice coefficient (0.6251), which demonstrates the excessive correspondence of the predicted and ground truth masks. Unetr stands out with the highest Dice coefficient (0.8045), emphasizing its capability for accurate segmentation. In terms of overall performance, Unetr, Unet++, and Attention Unet performance extremely good, particularly in Dice coefficient and mean IOU, making them the top three models for brain tumor segmentation.

In Figure 2, we're presented a series of images extracted from MRI scans of the brain, each serving a unique role in medical analysis. The initial image shows the raw MRI scan, offering a fundamental view of the brain's internal structure without any added annotations. Next, the "Ground Truth" image highlights specific areas within the brain tissue, indicating regions of interest or potential abnormalities, with distinct colors representing different conditions. Lastly, the "Predicted Classifications" image overlays color-coded markings on the MRI scan, indicating the model's predictions for various tissue types such as necrotic/core tissue (in red), edema or swelling (in green), and enhancing tissue (in blue). These predictions are invaluable in medical diagnosis and treatment planning, providing valuable insights into the condition of brain tissue. It's truly impressive how technology integrates imaging methods with predictive analysis to aid in understanding complex brain tissue features.

V. Conclusion Future Scope

Based on quantitative performance criteria, Unetr, Unet++, and Attention Unet excel at brain tumor segmentation. Unetr

TABLE I
MODEL PERFORMANCE METRICS

Model	Acc (%)	Prec (%)	Recall (%)	F1 (%)	Mean IOU	Loss	Dice Coeff. (%)
Unet	96.61	96.62	96.61	96.59	0.7632	0.0061	0.8029
Unet2d	99.14	98.60	99.42	99.44	0.6934	0.0278	0.5531
Unet3d	97.65	98.31	97.78	98.31	0.7012	0.0883	0.3152
Unet-fpn	95.65	92.54	93.69	94.34	0.6700	0.2000	0.8300
Unet++	98.32	98.35	98.35	98.34	0.8293	0.0178	0.6251
Unet_Vgg16	98.32	98.60	98.35	98.62	0.6260	0.0791	0.2871
Unetr	98.45	98.45	98.46	98.45	0.8940	0.0768	0.8045
Attention Unet	98.94	99.20	98.97	98.98	0.8560	0.0357	-

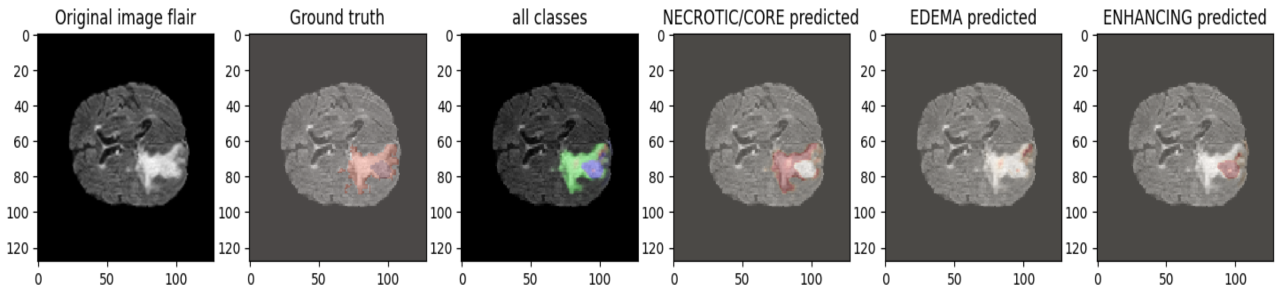

Fig. 2. Prediction Results

has the greatest Dice coefficient (0.8045), indicating a greater aptitude for accurate segmentation. Unet++ also performs remarkably well, with the greatest mean IOU (0.8293) and a competitive Dice coefficient (0.6251), demonstrating a strong agreement between anticipated and real masks. Attention Unet has a solid balance across multiple measures, making these three models the best options for accurate and dependable brain tumor segmentation.

Future studies should concentrate on resolving each of these models' specific shortcomings in order to increase their efficacy even more. Enhancing precision and recall via post-processing, loss function modification, and sophisticated data augmentation is crucial for Unet3d. Optimizing its attention mechanisms and investigating multi-scale attention frameworks could be beneficial for Attention Unet. Multi-modal imaging and transformer-based topologies could improve Unetr's segmentation quality. Furthermore, for these models to be used in clinical settings and be able to process data in real time, they will need to be tested on a variety of datasets and have their computing efficiency increased.

REFERENCES

[1] Zhu, Z., Wang, Z., Qi, G., Mazur, N., Yang, P. and Liu, Y., 2024. Brain tumor segmentation in MRI with multi-modality spatial information enhancement and boundary shape correction. Pattern Recognition, p.110553.

[2] J. N. Chandra, V. Bhavana and H. K. Krishnappa, "Brain Tumor Detection Using Threshold and Watershed Segmentation Techniques with Isotropic and Anisotropic Filters," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2018, pp. 0372-0377.

[3] Cao, F. and Bao, Q., 2020, July. A survey on image semantic segmentation methods with convolutional neural network. In 2020 international conference on communications, information system and computer engineering (CISCE) (pp. 458-462). IEEE.

[4] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.W. and Wu, J., 2020, May. Unet 3+: A full-scale connected unet for medical image segmentation. In ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 1055-1059). IEEE.

[5] Vimal Kurup, R., Sowmya, V. and Soman, K.P., 2020. Effect of data pre-processing on brain tumor classification using capsulenet. In ICICCT 2019–System Reliability, Quality Control, Safety, Maintenance and Management: Applications to Electrical, Electronics and Computer Science and Engineering (pp. 110-119). Springer Singapore.

[6] Somasundaram, S. and Gobinath, R., 2019, February. Current trends on deep learning models for brain tumor segmentation and detection–a review. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 217-221). IEEE.

[7] Reddy, D., Bhavana, V. and Krishnappa, H.K., 2018, April. Brain tumor detection using image segmentation techniques. In 2018 International conference on communication and signal processing (ICCSP) (pp. 0018-0022). IEEE.

[8] Saouli, R., Akil, M. and Kachouri, R., 2018. Fully automatic brain tumor segmentation using end-to-end incremental deep neural networks in MRI images. Computer methods and programs in biomedicine, 166, pp.39-49.

[9] Cherguif, H., Riffi, J., Mahraz, M.A., Yahyaouy, A. and Tairi, H., 2019, December. Brain tumor segmentation based on deep learning. In 2019 International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS) (pp. 1-8). IEEE.

[10] Madhupriya, G., Guru, N.M., Praveen, S. and Nivetha, B., 2019, April. Brain tumor segmentation with deep learning technique. In 2019 3rd international conference on trends in electronics and informatics (ICOEI) (pp. 758-763). IEEE.

[11] Thillaikkarasi, R. and Saravanan, S., 2019. An enhancement of deep learning algorithm for brain tumor segmentation using kernel based CNN with M-SVM. Journal of medical systems, 43, pp.1-7.

[12] Sajid, S., Hussain, S. and Sarwar, A., 2019. Brain tumor detection and segmentation in MR images using deep learning. Arabian Journal for Science and Engineering, 44, pp.9249-9261.

[13] Raut, G., Raut, A., Bhagade, J., Bhagade, J. and Gavhane, S., 2020, February. Deep learning approach for brain tumor detection and segmentation. In 2020 International Conference on Convergence to Digital World-Quo Vadis (ICCDW) (pp. 1-5). IEEE.

[14] Yogananda, C.G.B., Shah, B.R., Vejdani-Jahromi, M., Nalawade, S.S., Murugesan, G.K., Yu, F.F., Pinho, M.C., Wagner, B.C., Emblem, K.E., Bjørnerud, A. and Fei, B., 2020. A fully automated deep learning network for brain tumor segmentation. Tomography, 6(2), pp.186-193.

[15] Naser, M.A. and 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, p.103758.

[16] Akil, M., Saouli, R. and Kachouri, R., 2020. Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy. Medical image analysis, 63, p.101692.

[17] Magadza, T. and Viriri, S., 2021. Deep learning for brain tumor segmentation: a survey of state-of-the-art. Journal of Imaging, 7(2), p.19.

[18] Aboussaleh, I., Riffi, J., Mahraz, A.M. and Tairi, H., 2021. Brain tumor segmentation based on deep learning's feature representation. Journal of Imaging, 7(12), p.269.

[19] Liu, Z., Tong, L., Chen, L., Jiang, Z., Zhou, F., Zhang, Q., Zhang, X., Jin, Y. and Zhou, H., 2023. Deep learning based brain tumor segmentation: a survey. Complex & intelligent systems, 9(1), pp.1001-1026.

[20] Huang, P., Li, D., Jiao, Z., Wei, D., Cao, B., Mo, Z., Wang, Q., Zhang, H. and Shen, D., 2022. Common feature learning for brain tumor MRI synthesis by context-aware generative adversarial network. Medical Image Analysis, 79, p.102472.