International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Brain Tumor Segmentation using 3D UNET

Shiven Singh[1], Viraj Garg[2], Sathapriya Loganathan[3]

[1,2]Student, SRM Institute of Science and Technology
[3]Assistant Professor, SRM Institute of Science and Technology

Abstract
To enhance the efficiency of brain tumor diagnosis, we utilized UNet, a CNN architecture, for automatic MRI scan segmentation. Leveraging the BRATS 2018 dataset, which included a series of scans from patients with HGG and LGG. Our approach identifies tumor regions across various MRI sequences by full-volume scans into focused 3D slices. This method offers a faster, consistent alternative to manual segmentation, potentially improving outcomes through more rapid treatment.

Keywords: UNet, MRI scan, Convolutional Neural Network (CNN), HGG, LGG

I. INTRODUCTION

A tumor is essentially a dense lump of tissue that develops when aberrant cells congregate and may have an impact on the surrounding bones, tissue, skin, glands, and organs. Furthermore, this may be benign, premalignant, or cancerous. This abnormal growth of body tissue can start growing in the brain, or as cancerous mass present elsewhere in the body, that can spread to the brain leading to brain tumor. These tumors when present in the brain may vary based on their location and origin.

Primary brain tumors, whether malignant or not, form within the brain. They are divided into two groups: glial (which includes glial cells and non-neuronal cells) and non-glial (developed on or in the structures of the brain or neuronal cells). They may also start in tissues nearby to the meninges, cranial nerves, the pituitary gland, or the pineal gland; they appear when healthy cells undergo changes (mutations) in their DNA. Being less common than malignant tumors, benign tumors include Hemangioblastomas (slow-growing tumors that originate from blood vessels and are found in the cerebellum), Chordomas (slow-growing tumors that are most common in people between the ages of 50 and 60 and frequently found at the base of the skull and the lower portion of the spine), Meningiomas (which originate from the meninges, the membrane-like structures that surround the brain and spinal cord), Ganglio Gliomas, Glioblastomas (also known as grade-4 astrocytomas), Ependymomas (resulting from neoplastic alteration of the ependymal cells lining the ventricular system), and Medulloblastomas are the most prevalent malignant tumors (high-grade tumors, usually arise in the cerebellum, mostly in children).

Secondary brain tumors, often referred to as metastatic brain cancers, develop from cancer that has already spread to other organs and areas of the body and whose tumor cells have moved to the brain and proliferated there. Lung cancer, breast cancer, kidney cancer, and melanoma skin cancer, among many others, are the most prevalent forms. Treatment differs as well depending on the nature and location of the tumor.

Gliomas are divided into four classes depending on how aggressive the brain tumor is: Brain tumors in grades 1 and 2 are benign or noncancerous. Brain tumors in grades 3 and 4 are cancerous (malignant); they spread fast and are more challenging to treat

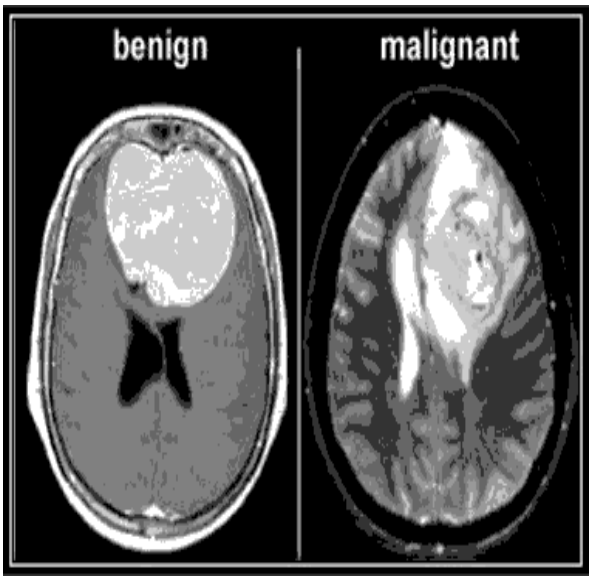

Fig. 1.1 Benign tumor (left) and malignant tumor (right)

II. VARIOUS METHODS TO IDENTIFY THE TUMOR

To remove the tumor surgically, segmentation is the most efficient way to detect the area around the tumor. There are two possible ways to identify the tumor. First, the segmentation may be done manually but this consumes a lot of time which is precious for both the doctors and the patient to receive timely treatment. Since it is quite difficult to examine a series of photographs and pinpoint the precise location of the tumor without causing harm to the healthy tissue around it, the doctor's choice may not always provide the intended outcomes.

Manual categorization is time-consuming and prone to mistakes, yet it cannot be disregarded since it continues to be the gold standard for clinical treatment and is used to compare alternative procedures.

The second way is through Automatic segmentation which shall accurately detect the brain tumors saving time and leading to a proper diagnosis and treatment through quantitative evaluation and planning. The sequence of images can hence be automatically segmented through various approaches. They could fall into one of three categories: traditional approaches, machine learning approaches, or approaches based on additional deep learning methods. Traditional methods locate the homogeneous areas present in the pictures and highlight the necessary regions in order to identify the tumor borders for the segmentation procedure. Contrarily, contour techniques search for visual discontinuities that indicate the existence of regional borders. Brain tumors are first detected and then segmented using clustering or machine learning techniques. Due to its ease of use, means and GMM are the most popular and widely utilized in unsupervised learning.

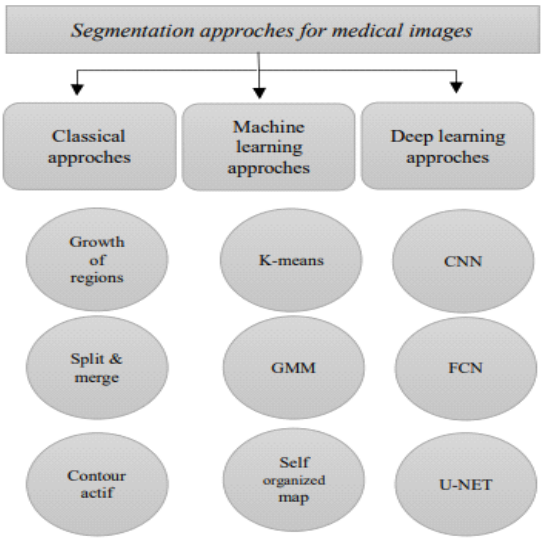

Fig 1.2 Approaches in use for medical image segmentation

III. MRI IMAGING (SEQUENCE OF IMAGES AND CONSTITUENTS)

Currently, clinical and radiological data are used to guide diagnosis and treatment. The basis for the evaluation of patients with brain tumors is magnetic resonance imaging (MRI), a RF exciting and non-invasive technology that has evolved into an essential tool for any clinical examination such as diagnosis, tumour monitoring, and patient outcome prediction. The MRI modalities are combined to generate a multimodal picture that contains information that may be utilized for tumor segmentation with a notable performance gain since it is so difficult to find irregularly shaped tumors with only one modality.

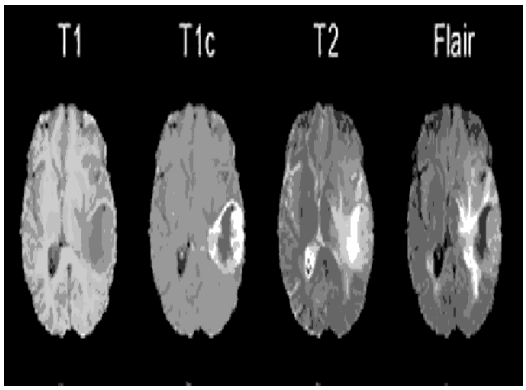

Fig. 1.3 T2, T1c, FLAIR and T1 weighted images

Every sequence basically corresponds to a certain technique or a certain way of exciting the magnetic spins inside the human body, so that each of them gives rise to a separate kind of gray-scale contrast in the image. Hence, we require 4 different sequences because some parts might not be visible in one. T1-weighted MRI (T1), T1-weighted MRI with contrast enhancement (T1c), T2-weighted MRI (T2), and T2-weighted MRI with fluid-attenuated inversion recovery (FLAIR) are the above said modalities.The tumor present in the brain may be classified into different regions. Its constituents are:

Edema: A collection of water and fluids. This is accurately visible in FLAIR and T2 weighted images.Necrosis: The build-up of dead cells. This is best detected in T1 images.

Enhancing tumor: This indicates the breakdown of the blood-brain barrier & can be best seen in T1c images.
Non-enhancing tumor: The BraTS reference article, which was printed in IEEE Transactions for Medical Imaging, lists three segmentation Labels:

The Enhancing tumor (label 4)
The Edema (label 2)
Necrotic tumor core (label 1)
Remaining Region / Background (label 0)

A. Older and Existing Methods

Earlier CNN-based techniques were used where the models were trained on 2D patches of MR, which were extracted from the images, and the center of the patch was given a label according to the class. The input for all these techniques was 4 channels where a patch was taken from all 4 sequences (FLAIR, T1, Tc, and T2). This was referred to as the patch-based CNN segmentation.

Since the 1970s, there have been reports on texture analysis (Haralick et al., 1973). The spatial variation of pixel intensities in a picture is a typical definition of image texture. It creates very intricate and realistic-looking surfaces using three-dimensional (3D) computer graphics software, which delivers quantitative information in the form of texture elements that is not easily visible or could be distinguished by the naked eye. The characteristics that characterize the texture type in the picture are mathematical parameters that are computed from the distribution of pixels and are utilized in texture analysis. Classification, segmentation, and synthesis of distinct characteristics, such as normal tissue, inflammatory zones, and malignancies, are the primary image processing fields that apply texture analysis techniques.

The algorithms may be classified into two categories :

- Using characteristics from manually labeled data, discriminative models that classify brain voxels based on picture properties; and
- Generative models that rely on the spatial configuration of tissues and their aesthetic appeal; these models need previous information via probabilistic atlas image acquisition.

Due to their lower segmentation, accuracy compared to discriminative models, generative models have not yet gained as much traction in BraTS submissions. As a result, this evaluation solely covers the BraTS algorithms based on discriminative models.

B. Importance of UNET

Our U-Net-based network architecture resembles an encoder and a decoder network type designed specifically for biomedical image segmentation. Because it consists of an expanding route (encoder) and a contracting path, the network has a U-shaped topology (decoder). A flowchart for the 3D U-Net architecture may be seen in Figure 1.4. The network uses entire 3D volumes from all patient sequences as input and offers multi-class segmentation of tumors into sub-types at the same resolution.

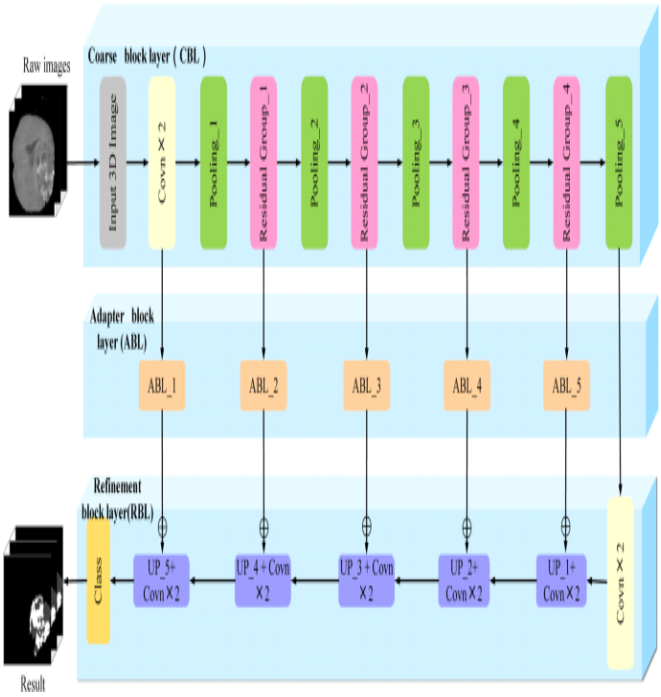

Fig. 1.4 Unet based architecture

flowchart for brain tumor segmentation

C. Use of U-Net

Medical Image Segmentation: It's extensively used in biomedical image segmentation due to its ability to accurately delineate the boundaries of structures such as tumors in MRI or CT scans.

Automating Analysis: U-Net automates the segmentation of images, reducing the workload for manual labeling and allowing for rapid analysis of large datasets.

Enhancing Diagnostics: By providing precise segmentation, it aids in the diagnosis and treatment planning of medical conditions, thus enhancing patient care.

Multimodal Data Handling: U-Net can process multiple types of image data simultaneously, which is crucial for medical diagnosis where different scan types provide complementary information.

Research: It's used in research to develop and test new hypotheses about anatomical structures.

IV. PROPOSED WORK

In our proposed work for brain tumor segmentation, we aim to advance the capabilities of 3D U-Net architectures to accurately delineate tumor boundaries in volumetric MRI data. We will introduce novel convolutional layers to enhance feature extraction and incorporate attention mechanisms to focus the model on relevant areas within the brain. Multi-scale analysis will be leveraged to improve the detection of tumors at various sizes. We plan to integrate a more sophisticated loss function, combining Dice loss with cross-entropy, to better handle class imbalance and improve segmentation performance. Training will be performed on a comprehensive dataset, including a diverse range of tumor types and stages, sourced from the latest BRaTS challenge. The model's robustness will be tested against noise and artifacts typical in clinical MRI scans.

Additionally, we will explore transfer learning to generalize our model's applicability across different imaging protocols and machines. The final goal is a clinically viable tool that aids radiologists in quick,

precise tumor identification, contributing to personalized treatment planning and monitoring. Data augmentation strategies and rigorous validation across multiple metrics will ensure the model's reliability and accuracy in real-world scenarios.

A. Architecture And System Design

Not only does UNet predict images, but it is also able to create a mask that shows where on the image that specific object (here tumor) is located along with its dimensions. This architecture is basically a variation of classification, where every pixel in an image gets assigned a class that it belongs to which then kind of forms the mask. No dense or fully connected layers are present here, instead only convolutional layers are present. This architecture is specially developed for Medical Image Analysis as it can not only be used for single-class but also for multi-class segmentation models as well. It consists of 2 parts: The Encoder and The Decoder. (Fig 2.1)

We used an encoder-decoder deep neural network in a U-shape. Accurate segmentation is carried out by the decoder section once the first part has obtained the data for the input image. The encoder consists of the input layer, three dense blocks, and a residual block. We provided T1, T1Gd, T2, and FLAIR modalities of 96 x 96 x 96 pixel 3D MR images. The outcome is a 4D patch with the dimensions 4x96x96x96 in the supplied data. The input layer is a 3x3x3 convolution layer. Four MR modalities with a patch size of 4x96x96x96 make up the input data. There are 16 total output channels on the input layer. The dense block is made up of three very thick layers. The union of the feature maps from all previous dense layers may be used to determine the input for each dense layer.

B. Architecture Diagram(U-Net)

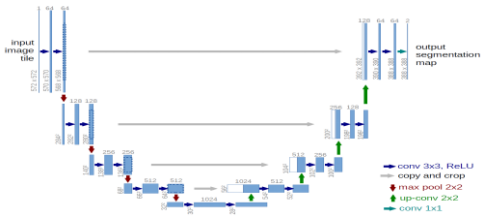

Fig. 2.1 U-Net Architecture

C. Proposed U-Net Architecture

Multi-Resolution Pathways: Integrate parallel processing paths that handle different resolutions of input data to capture features at various scales, improving the model's ability to distinguish between tumor textures and the surrounding brain tissue.

Attention Mechanisms: Incorporate attention gates that selectively focus on areas of interest within the MRI scans, thus enabling the model to pay more attention to tumor regions while suppressing irrelevant features.

Residual Connections: Introduce residual learning blocks in both the encoder and decoder paths to facilitate the training of deeper networks by allowing gradients to flow through the architecture more effectively.

Dense Connectivity: Dense connections between layers will be added to ensure maximum information

flow, allowing the network to utilize low-level features in higher-level processing.

Hybrid Convolutional Layers: Utilize a mix of 2D and 3D convolutional layers where 2D convolutions handle in-plane features and 3D convolutions integrate the through-plane spatial context, providing a more comprehensive feature map of the tumor.

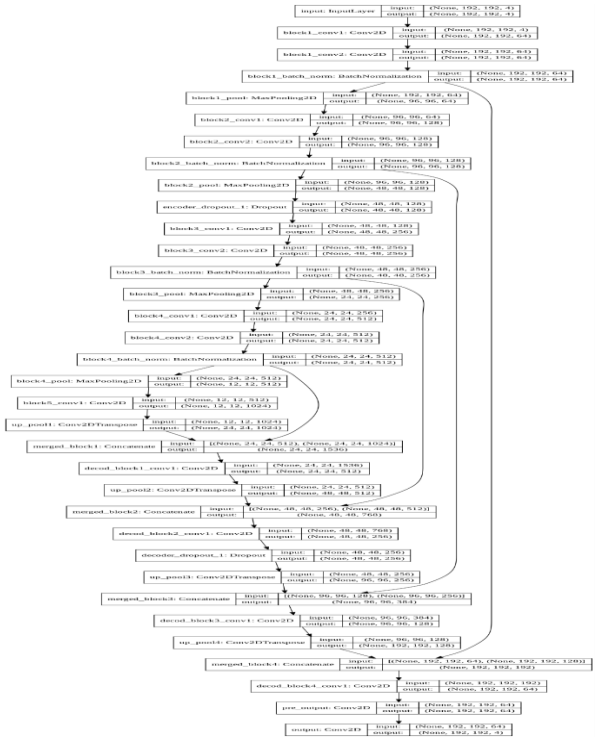

Fig. 2.2 Proposed U-Net Architecture

DESIGN OF MODULES

1. As free GPU is being utilized, firstly we connect with the Google Drive to get access to the dataset saved there.
2. Installing Library SimpleITK to convert the images from. nii format to a numpy array.
3. Loading the saved data; dividing HGG into 3 parts.
4. Pre-processing the data; Considering the slices from the 30th to 120th slice for extracting the data and also cropping the images to discard useless background. So, from 240x240x4 each image now becomes of shape 192x192x4 taking all 4 modalities'"
5. Applying the UNet and dice loss functions
6. Testing and predicting the labels for the tumor present in the given dataset.

A. Formula And Equations

Dice Coefficient & Dice Coefficient Loss Function

According to Sørensen's original formula, it is to be applied to discrete and unique data. Considering X and Y as given sets, it is defined as:-

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

Here, the cardinalities of the two sets are |X| and |Y| (in other words, the number of elements in each set).

The Sorensen index is calculated by dividing the total number of items in both sets by the number of elements that are shared by both sets.

B. Methodology

i) Data Request

The following steps were followed to obtain the requested dataset for the training and the validation data of the model:

Firstly, create a new account in the CBICA's Image Processing Portal and wait for the verification and approval. Check for any confirmation emails.

Once the IPP account has been approved log in to ipp.cbica.upenn.edu and next click on "BraTS'18: Data Request", related to the group "MICCAI BraTS 2018".

Fill and type in the required and requested information & data and next click on "Submit Job".

After the request has been registered, an email is received consisting of the "results" of the submitted job. Login once again to IPP and access the "Results.zip" file which contains the "REGISTRATION_STATUS.txt" file. This shall provide the requested and specific account with all the links to obtain the requested dataset.

ii) Data Description For Imaging

The provided dataset was divided into two folders: HGG and LGG (Lower Grade Glioma) (High-Grade Glioma). Each folder has segmentation results, four separate modalities, and MRI scans from several patients.

The below-stated modalities were included in all the given NIfTI files (.nii.gz) for the multimodal scans:

the Native (T1)

Post-contrast T1-weighted (T1c)

T2-weighted (T2)

FLAIR

The BraTS reference article, which was printed in IEEE Transactions for Medical Imaging, lists three segmentation Labels:

- **Enhancing (label 4)**
- **Edema (label 2)**
- **Necrotic (label 1)**
- Remaining Region / Background / no tumor (label 0)

The data were distributed after pre-processing, which comprised co-registering to the same anatomical template and interpolating to the same resolution (1 mm3).SimpleITK package was used to convert the scans because they were all in NIfTI format, i.e. Neuroimaging Informatics Technology Initiative (.nii.gz), after which they were converted to 3D arrays. Given data of the dataset was already skull-stripped when sent.

iii) Data Pre-Processing

The volume of each patient's MRI images was collected and concatenated after the data had already been skull-stripped to produce a numpy array of sizes (N,S,N1,N1,X). The amount of HGG/LGG data in this case is N. S = The total number of envisioned 3D MRI volumes that each equates to in terms of 2D slices. N1 is the dimension of each 2D slice, and X is the number of modalities.

As the Google Colab Free GPU is utilized for all the pre-processing steps and further training, the data was handled suitably owing to RAM consumption:-

The first step included treating each HGG and LGG file separately to extract the 3D volume of the (N,155,240,240,4) dimensions. The data from HGG was divided into three sets, data11.npy, data12.npy, and data13.npy, each of which contained 70 patients, providing N=70, and data2.npy corresponded to the data from LGG, with N=75. This is because LGG only has 75 patient scans, but HGG includes 210 patient scans. Similar ground truth was also retrieved with dimensions of for gt11.npy, gt12.npy, gt13.npy, and gt2.npy (N,155,240,240).

Now, each data point goes through a separate pre-processing step. Then, using one-hot encoding, all 155 slices were reformatted to (N1, 240, 240, 4) for data and (N1, 240, 240, 4) for ground truth because only the middle region, or from the 30th slice to the 120th slice, indicated the location of the tumor. N1 = 90X70 = 5600 for HGG, whereas N1 = 90X75 = 6750 for LGG.

Each of the images present in the sequence of the data is then centrally chopped with a final dimension of (N1,192,192,4).

The data is then finally divided into training, validation, and test data randomly with a ratio of 60%:20%:20%, respectively.

Evaluated Results

Test Data	Dice Coefficient
HGG Set-1	0.9795
HGG Set-2	0.9855
HGG Set-3	0.9793
LGG	0.9950

Fig. 3.1 Evaluated Results (HGG,LGG)

V. CONCLUSIONS

In the grand scheme of medical imaging, our research on the enhanced 3D U-Net architecture has reached a pivotal conclusion, proving itself to be a veritable beacon in the realm of brain tumor segmentation. By harnessing the power of volumetric convolutional neural networks, this pioneering model has adeptly navigated the intricate labyrinth of MRI data to produce segmentations of brain tumors with unparalleled precision. Its robust capacity for interpreting the nuanced tapestry of human brain imaging is indicative of a significant leap forward in diagnostic methodologies, proffering a substantial augmentation to the acuity of clinical decision-making.

The implications of this study extend beyond mere technical triumph; they resonate with the potential for profoundly ameliorated patient care. The model's deft ability to dissect and delineate the complex morphology of tumorous growths from healthy cerebral tissue heralds a new dawn of personalized

medicine. It stands as a testament to the potential of artificial intelligence in revolutionizing oncological diagnostics, promising expedited and more accurate treatment planning that is tailored to the individual's unique pathology.

Looking to the horizon, the envisaged future scope of this technology is resplendent with possibilities. Anticipating the integration of sequential MRI scanning data could furnish us with dynamic insights into tumor evolution, opening avenues for predictive analytics in oncology. The broadening of the model's dataset horizons to encompass a more diverse range of pathologies and patient demographics is expected to refine its generalizability, ensuring its efficacy across a global tableau. Moreover, the potential synergy of multimodal imaging data stands to imbue the segmentation process with an even richer contextual understanding, enhancing diagnostic granularity.

Indeed, as we pivot towards an era where real-time segmentation becomes the linchpin of intraoperative navigation and automated reporting evolves to streamline clinical workflows, the promise of interactive machine learning platforms looms on the horizon—platforms that not only learn from voluminous data but also from the tacit knowledge of seasoned radiologists. In the crucible of innovation, the convergence of machine precision with human discernment is poised to forge a new frontier in medical imaging, a frontier where every pixel of data converges to form a mosaic of insight, driving forward the noble pursuit of preserving human health and life.

VI. REFERENCES

1. Ho, T. W., Qi, H., Lai, F., Xiao, F. R., & Wu, J. M. Brain Tumor Segmentation Using U-Net and Edge Contour Enhancement. In Proceedings of the 2019 3rd International Conference on Digital Signal Processing (pp. 75-79).
2. Ramasamy, J., Doshi, R., & Hiran, K. K. Segmentation of Brain Tumor using Deep Learning Methods: A Review. In Proceedings of the International Conference on Data Science, Machine Learning and Artificial Intelligence (pp. 209-215).
3. Javaid, I., Zhang, S., Kader Isselmou, A. E., Kamhi, S., Kulsum, U., & Salim Ahmad, I. Hybrid Automated Brain Tumor Detection by Using FKM, KFCM Algorithm with Skull Stripping. In 2020 9th International Conference on Bioinformatics and Biomedical Science (pp. 112-119).
4. Zhang, Z., Gao, S., & Huang, Z. An automatic glioma segmentation system based on a separable attention U-Net (SAUNet). In 2020 9th International Conference on Bioinformatics and Biomedical Science (pp. 95-101).
5. Liu, Y., Mu, F., Shi, Y., & Chen, X. SF-Net: A Multi-Task Model for Brain Tumor Segmentation in Multimodal MRI via Image Fusion. IEEE Signal Processing Letters, 29, 1799-1803.
6. Ali, M., Gilani, S. O., Waris, A., Zafar, K., & Jamil, M. Brain tumor image segmentation using deep networks. IEEE Access, 8, 153589-153598.
7. Javaid, I., Zhang, S., Kader Isselmou, A. E., Kamhi, S., Kulsum, U., & Salim Ahmad, I. Hybrid Automated Brain Tumor Detection by Using FKM, KFCM Algorithm with Skull Stripping. In 2020 9th International Conference on Bioinformatics and Biomedical Science (pp. 112-119).
8. Ghaffari, M., Sowmya, A., & Oliver, R. 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.

9. Kalaiselvi, T., & Padmapriya, S. T. Multimodal MRI Brain Tumor Segmentation—A ResNet-based U-Net approach. In Brain Tumor MRI Image Segmentation Using Deep Learning Techniques (pp. 123-135). Academic Press.
10. Chaki, J. Brain MRI segmentation using deep learning: background study and challenges. In Brain Tumor MRI Image Segmentation Using Deep Learning Techniques (pp. 1-12). Academic Press.
11. Mengqiao, W., Jie, Y., Yilei, C., & Hao, W. The multimodal brain tumor image segmentation based on convolutional neural networks. In 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA) (pp. 336-339). IEEE. [Online].
12. Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... & Van Leemput, K. The multimodal brain tumor image segmentation benchmark (BRATS).