RESEARCH ARTICLE **SPECIAL ISSUE**

RSP Science Hub

International Research Journal on Advanced Science Hub
2582-4376
Vol. 05, Issue 05S May
www.rspsciencehub.com

Check for updates

http://dx.doi.org/10.47392/irjash.2023.S062

International Conference on intelligent COMPUting TEchnologies and Research (i-COMPUTER) 2023

Automated Brain Tumor Segmentation Using Attention gate Inception UNet with Guided Decoder

Amisha P [1], Adersh V R [2]

[1]Department of Electronics and Communication Engineering, College of Engineering Trivandrum, Kerala, India
[2]Assistant Professor, Department of Electronics and Communication Engineering, College of Engineering Trivandrum, Kerala, India

Email: amishaputhalath@gmail.com

Article History

Received: 27 February 2023
Accepted: 21 March 2023

***Keywords**:*

- *BraTS;*
- *Brain tumor segmentation;*
- *Deep learning;*
- *Attention gates;*
- *Inception module;*
- *Guided decoder;*
- *MRI*

Abstract

Brain tumor segmentation technology is a crucial step for the detection and treatment of MRI brain tumors. Tumors can occur in various locations and can be of any size or form. The use of skip connections in MRI brain tumor segmentation approach based on U-Net architecture helps to incorporate low-level and high-level feature information and has recently gained popularity. By introducing an attention mechanism into the UNet architecture, the performance of local feature expression and medical image segmentation can be enhanced. In this paper, we present an innovative deep learning architecture called Attention gate Inception UNet with Guided Decoder for brain tumor segmentation. The backbone of the model is a popular segmentation method called U-Net architecture. While dealing with small-scale tumors, the U-Net network has low segmentation accuracy. Therefore several modifications are made, which results in the integration of attention gates and inception block together with a guided decoder. A sequence of attention gate modules are introduced to the skip connection, that focus on a selected part of an image while ignoring the others. The inception module used will help us to extract further characteristics at each layer. The proposed architecture has the ability of explicitly guiding each decoder layer's learning process and it is supervised by using individual loss function, allowing them to produce efficient feature maps. The proposed model achieved a dice score of 0.9190, 0.9331, 0.8990 for whole tumor, tumor core and enhancing tumor respectively on Brain Tumor Segmentation Challenge (BraTS) 2019 dataset of High Grade Gliomas (HGG).

1. Introduction

A brain tumor is an abnormal growth of cells in the brain. The prevalence of malignant brain tumour is currently high, which has a substantial impact on society and people (Işın, Direkoglu, and Şah). Glioma is the most prevalent malignant brain tumor and is further classified as high-grade glioma (HGG) and low-grade glioma (LGG). In comparison to the LGG, the HGG grows quickly and has a huge level of cell infiltration. Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and Positron Emission Tomography (PET) are some of the non-invasive technologies (Verduin et al.) used to diagnose, monitor, and evaluate cancers. MRI, on the other hand, is a common brain imaging technology since it is an efficient technique of diagnosis approach for the detection of soft tissue cancers. The quantitative analysis of brain tumors can be performed by using multimodal brain images to determine the maximal diameter, amount, and the volume of brain lesions, providing the most accurate diagnosis and treatment approach. The tumor segmentation is the vital step for the detection and treatment of brain cancers (Bjoern et al.). For the diagnosis and assessment of brain tumors, radiologists prefer to use magnetic resonance imaging (MRI). T1-weighted, post-contrast T1- weighted (T1ce), T2-weighted, and FLAIR are some of the complementary 3D MRI modalities available. T1 images are used commonly to differentiate healthy tissues. T2 images are used to provide an underlying assessment to recognise several kinds of tumor and to separate tumors from non-tumor tissues. T1-ce images aid in distinguishing tumor boundaries from adjacent normal tissues. The water molecule signal is decreased in FLAIR images, which aids in identifying the edema zone from the cerebral fluid.

The BraTS 2019 aims to provide the users with a large dataset of 3D MRI scans of both LGG and HGG along with its related ground truth to examine the progress of deep learning models for tumor segmentation. They are employed to train and test the networks that have been developed for specific segmentation tasks (Bjoern et al.). Segmentation techniques are primarily divided into manual, semi-automatic, and fully automatic. The manual segmentation of tumor using MRI data is an intensive. For semi-automatic and fully automatic approaches, the results of manual segmentation are used as Ground Truth (GT). In the semi-automatic technique, human interaction is necessary in the form of manual correction, initialization or modification of the outcomes, resulting in an efficient segmentation. Whereas, in the fully automatic technique, segmentation is completed without the need for human interaction, but it does require anatomical knowledge of the tumor's size, shape, appearance, and location in order to construct a model and complete the work. Automated segmentation will help the physicians to facilitate the diagnosis and surgical planning while also providing a precise, reliable solution for future tumor analysis (Menze et al.). Convolutional neural networks (CNNs) have become more widely used in segmentation process, image classification, etc. For the BraTS challenge dataset, patch wise models account for the majority of CNN-based solutions. These approaches simply feed a small region of image into the network, disregarding its content and correlations between labels. Furthermore, training for these procedures takes a lengthy period. CNN is modified in a number of ways in fully convolutional networks (FCN) (Naceur et al.). Rather than predicting probability distributions patch-by-patch like CNN, the FCN models predict pixel-by-pixel.

Based on the design of FCN, a symmetric fully convolutional network termed U-Net is suggested, which comprises a contracting path which helps us to gather contextual information and an expanding path to ensure precise location. In the field of segmenting brain tumors, U-Net has a great deal of success. U-Net will consistently reduce the image's dimension during downsampling, resulting in the reduction of accuracy for small-scale tumors. To overcome this problem, the attention mechanism is used which can improve the local feature expression (Noori, Bahri, and Mohammadi). The attention mechanism, which is crucial to human perception, may efficiently use the needed information while suppressing the irrelevant information. This Attention Network can have a residual block added to it, which efficiently adjusts in response to the deeper layers that improves the accuracy in classification.

The proposed work presents a novel deep learning network called the Attention gate Inception UNet with Guided Decoder, which incorporates guided decoding and the properties of Inception module with the attention gates. The network design includes a guided decoder that supervises decoder's learning process and facilitates the production of improved features. The prediction capabilities of each decoder layer is enhanced by using weighted guided loss and thus improves the final output layer's prediction accuracy (Maji, Sigedar, and Singh). Attention Gates are incorporated into a hybrid network design with a backbone of Inception module with UNet architecture, that emphasis on important regions of the images. The proposed model is evaluated using BraTS 2019 dataset of HGG. The important contributions of this paper includes :

- The encoder path of the network is made of inception modules which in turn improves the segmentation accuracy.

- Use of less number of filters in each layer of encoder path and decoder path reduces the computational requirements.
- Each decoder layer is trained individually with its loss function which results in guided decoder.
- Evaluation made on BraTS 2019 dataset (HGG) reveals that the network gives better results as compared to other existing segmentation networks.

2. Related works

In the field of medical imaging, the segmentation of brain tumor has now become an essential element . Because of its excellent capacity to extract high discriminant characteristics automatically, deep neural networks have become popular in the field of image processing techniques (Esteva et al.). Meanwhile, deep learning-based computer-aided MRI brain tumor diagnosis has gained a lot of interest. Pereira et al. (Pereira et al.) developed an automated segmentation network using a 3 X 3 convolutional kernel design based on VGGNet. To create a dual-path 2D CNN network that integrates local and global routes, Havaei et al. (Havaei et al.) employed various size convolutional kernels to capture distinct contextual local features. FCN based tumor segmentation architecture categorise and estimate each pixel of entire image of brain to enhance the efficiency of segmentation process. Badrinarayanan et al. (Badrinarayanan, Kendall, and Cipolla) proposed a SegNet architecture based on FCNs that performs non-linear upsampling using a unique approach of unpooling. For brain tumor segmentations, Alqazzaz et al. constructed the SegNet (Alqazzaz et al.) which uses four units of SegNet models, each of which was evaluated on the four distinct modalities separately. Ronneberger et al. (Ronneberger, Fischer, and Brox) introduced a FCN known as U-Net, that employed in various segmentation areas. In the area of brain tumor segmentation, researchers introduce the UNet based design. Dong et al. (Dong et al.) used a 2D UNet based architecture with data augmentation to improve brain tumor segmentation performance. Yang et al. (Yang et al.) improved the U-Net by including residual module that helps in the extraction of more features at each layer. ResUNet is extensively utilised as the basis model for several Deep Learning networks because of its superior performance and efficiency in feature extraction. Kermi et al. (Kermi, Mahmoudi, and Khadir) improved segmentation accuracy by using a slightly modified U-Net architecture with a mixed loss function comprising a weighted cross-entropy loss and generalised dice loss. These approaches on the other hand send all the features that are extracted to the decoder stage via skip connection.

The attention process, which is crucial for human perception, may efficiently capture the relevant information while suppressing redundant information. The attention gate was first employed for machine translation in the area of natural language processing (Bahdanau, Cho, and Bengio). This approach decreases the likelihood of false positives that assists in directing the model's focus to a particular task, hence improving the network performance. Wang et al. (Wang et al.) employed a Residual Attention module to produce various attention features from multiple networks, which adjust when layers go deeper and boost the classification performance. Hu et al. (Hu, Shen, and Sun) proposed a Squeeze-and-Excitation (SE) block attention module that concentrates on channel relation and performs feature recalibration on a channel-by-channel basis to increase feature expression. For semantic segmentation, Fu et al. (Fu et al.) used two attention modules that are composed of spatial and channels attention, where the attention blocks are identical to non-local operation. Zhou et al. (Zhou et al.) investigated a guided attention module for brain tumor segmentation that adaptively recalibrates channel-wise feature outputs. Zhang et al. (T. Zhang et al.) combined numerous attention modules on Res-UNet resulting in excellent ventricular segmentation performance.

The attention gate mechanisms can improve the local feature expression to overcome the issue of poor segmentation accuracy in UNet model for small scale tumors. The attention gate is combined with the Residual U-Net model and implemented in an AGResU-Net architecture (J. Zhang et al.). The Attention gate with ResUNet and guided decoder (Maji, Sigedar, and Singh) are incorporated to outperform existing state of art models in the segmentation process.

2.1. Network

The proposed architecture of Attention gate Inception UNet with Guided Decoder model is shown in

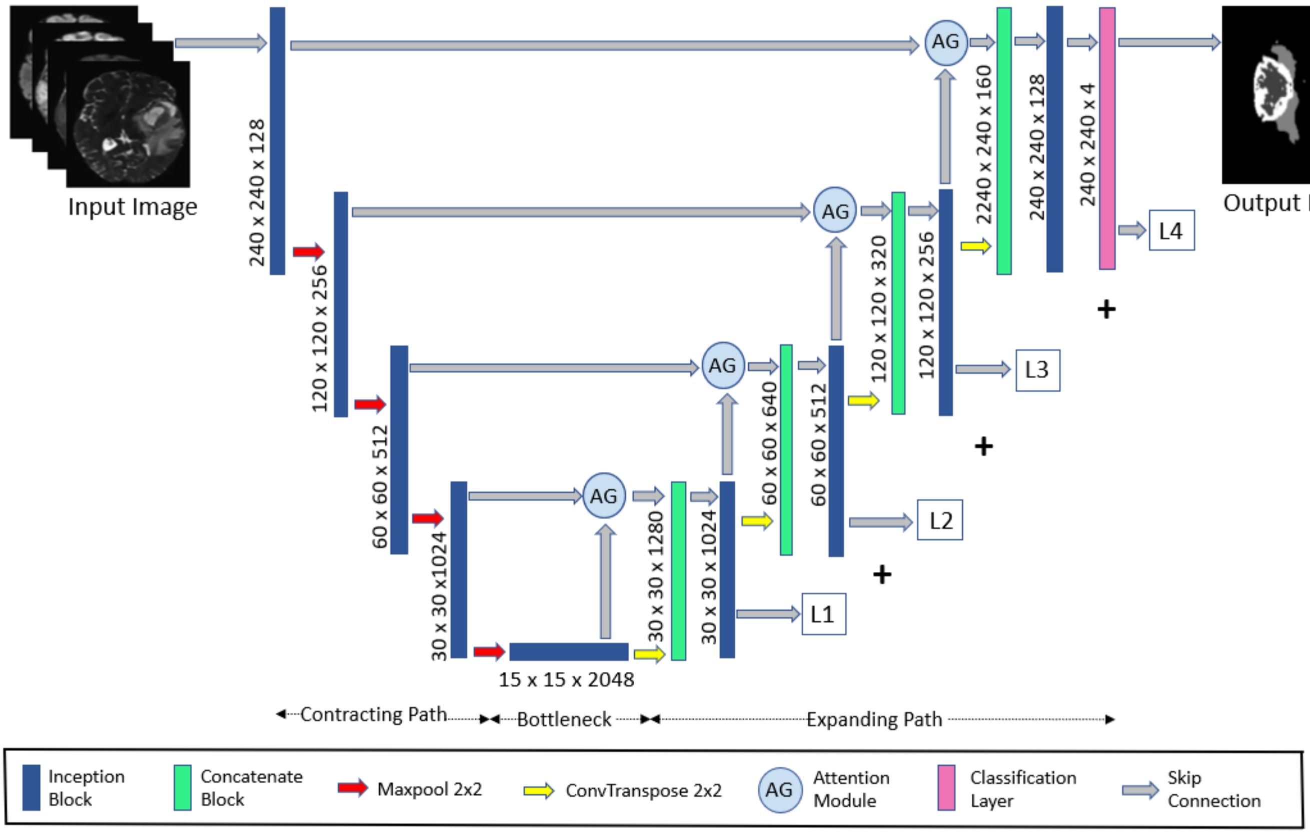

FIGURE 1. **Attention gate Inception UNet with Guided Decoder**

FIGURE 1. Popular segmentation model, UNet act as the backbone of this architecture. The network has an expanding path and a contracting path, one of which recovers the original resolution and the other of which encodes the complete input image. The encoder path helps to capture contextual and spatial features of the images and decoder path enable the precise localization. To access the low level features produced by the encoder, skip connections are used and all the information is passed through this skip connection. In our model, inception block is used together with an attention gate and a guided decoder. Inception blocks are mainly added to capture more characteristics from each layer. The depiction of an inception module is depicted in FIGURE 2, that includes filters of various sizes (1x1, 3x3, 5x5) which are used to extract better features. The attention mechanism concentrates on particular part of the image and ignores the rest. By inhibiting feature activations in irrelevant areas of an image, the attention gates added to the model can lower the false positives. The main highlight of the model is Guided decoder together with the inception block which have the ability to explicitly control each decoder layer's learning process. Every decoder layer's individual loss function aids in monitoring the learning process and producing a combined loss function. The model predicts the output at each layer by training the various layers of decoder with its own loss function and passing to the final layer. The model consists of four encoder layers at contracting path and four decoder layers at expanding path as shown in FIGURE 1. The result of the first 3 decoder layers are rescaled to input size prior to getting compared to GT. The network's performance of segmentation is enhanced by the predictions made at each decoder layer as well as the transmission of weighted loss from every intermediate layer towards the final layer. The output image which is produced at the final decoder layer is regarded as a segmented result of the network.

Each layers of encoder path consist of an inception block and it is followed by Maxpool 2 x 2 down sampling layer. The size of feature map is decreased by half and feature map count is doubled as we move towards the bottleneck layer. The model takes a 240 x 240 x 4 input and has four channels that correspond to four modalities. The first inception block of proposed model contains 128 feature maps with size of 240 x 240, then it is passed to 2 x 2 down sampling layer. The inception module of second layer consist of 256 feature maps with 120 x 120 size. The output of third inception module is of size 60 x 60 x 512. The inception block of fourth layer comprises of 1024 feature maps with a size of 30 x 30.

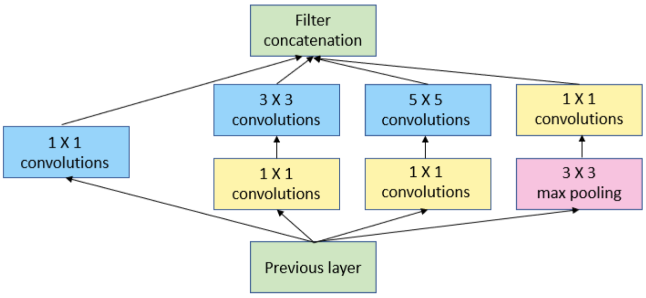

FIGURE 2. Depiction of Inception module (Szegedy et al.).

The encoder and decoder layers of the model are connected by the bottleneck layer. The output from last encoder layer is passed to bottleneck layer after down-sampling operation. The reduction in the number of filters helps to decrease the computational complexity of the model. Bottleneck layer comprises of an inception module with 2048 feature maps with size of 15 x 15. The output from that layer flows in two directions. One goes into the decoder path through the convolutional transpose and the second one is used as gating signal of attention gate. In skip connections, the attention gate helps in the connection of the encoder and its corresponding decoder layers. Dual inputs are used by the attention gate, one from the appropriate encoder, which includes every contextual and spatial features in that specific layer, and other from the decoder layer beneath it. After concatenation, the output produced at the attention gate is transmitted to the decoder layer.

The decoder path comprises four layers, all of which include an inception block which is preceded by a 2 x 2 convolutional transpose layer used for upsampling. Attention gate output are concatenated with the preceding decoder layer's output. The results are passed to the inception blocks after concatenation. The feature map size is doubled and feature map count is decreased by half as we move away from the bottleneck layer. The output of each decoder layer undergoes upsampling to obtain a size 240 x 240, which is finally passed to the classification layer. The feature maps are converted into probabilities using a softmax activation. The output from last decoder layer is passed to the classification layer, from where we obtain the segmented prediction result of the network. The analysis of the architecture is provided in the TABLE I, including different layers and their output shape.

TABLE 1. Analysis of Architecture

Sl.No.	Layer	Output shape
1	Input layer	240 x 240 x 4
2	Encoder layer 1	240 x 240 x 128
3	Encoder layer 2	120 x 120 x 256
4	Encoder layer 3	60 x 60 x 512
5	Encoder layer 4	30 x 30 x 1024
6	Bottleneck layer	15 x 15 x 2048
7	Decoder layer 1	30 x 30 x 1024
8	Decoder layer 2	60 x 60 x 512
9	Decoder layer 3	120 x 120 x 256
10	Decoder layer 4	240 x 240 x 128
11	Classification layer	240 x 240 x 4

2.2. Combined Loss function

The proposed study uses weighted cross entropy (WCE) loss and weighted dice loss (WDL) to segment the MRI data. WDL helps to reduce the gap existing between evaluation metric and training set while also being immune to data imbalance. Furthermore, the WCE loss (Adel, Mahmoudi, and Khadir) has been shown to be useful for multi-task learning and the problem of class imbalance. As a result, we use a blend of WDL and WCE loss functions to give improved model training supervision. This combined loss function (CL) is expressed as follows (Maji, Sigedar, and Singh) :

$$CL = WDL + WCE \tag{1}$$

Where WDL and WCE depict the Weighted dice loss and the Weighted cross entropy loss. The Weighted dice loss is given in Equation 2 (Maji, Sigedar, and Singh) as follows:

$$WDL = \ 1 - DS_{Weighted} \tag{2}$$

Where weighted Dice Score (Maji, Sigedar, and Singh) is expressed as follows:

$$DS_{Weighted} = \frac{\sum_n W_n DS_n}{\sum_n W_n} \tag{3}$$

Here, W_n is the class n weight factor and DS_n is the class n dice score. Each class's Dice Scores are determined separately and are calculated as follows (Maji, Sigedar, and Singh):

$$DS_n = \frac{2\sum p_{n g_n}}{\sum p_n + \sum g_n} \tag{4}$$

Here, p_n is the predicted output and g_n is the ground truth (GT) of class n. The term class refers to the various tumor regions. WCE loss (Adel, Mahmoudi, and Khadir) is used together with WDL and is given as:

$$WCE = -\textstyle\sum_{n \in N} w_n g_n log(p_n) \tag{5}$$

Where w_n, g_n and p_n represents the weight, GT and predicted output for class n respectively. Thus CL funtion of both WDL and WCE is used in the training of our proposed model.

2.3. Dataset

The proposed architecture is evaluated by using the BraTS 2019 training dataset. There are 259 HGG scans and 76 LGG scans. HGG scans are used in the proposed model. Each patient's MRI scan includes four sequences: T1, T1ce, T2, and FLAIR as shown in FIGURE 3. Segmented images of GT are labelled as four classes, namely background and non tumor region (label 0), Necrosis (label 1), Edema (label 2) and Enhancing tumor region (label 4). The dimensions of the images are (240, 240, 155). The proposed model uses 150 scans of HGG and 25 slices are selected from 255 slices to decrease the memory requirements. Thus a total of 3750 images are collected from each modality.

2.4. Implementation details

Each MRI modality of the brain tumor is with size (240, 240, 155). The study chooses 25 slices from each volume and identifies slices with more unique characteristics. Then alternate slices are choosen from that are concentrated at the center, because the slices farther from center contain less region of interest. As we employ the four MRI modalities together to detect various tumor locations, the slices from the four sequences are concatenated. The input data is normalized using Z-score to convert all the data with mean 0 and standard deviation as 1. To avoid the problem of overfitting, the data augmentations such as horizontal and vertical flips, height and width shifts, shear, zoom and rotations are applied to our training set. Then they are randomly divided for training, testing and for validation. The different regions of brain tumor are detected by using four modalities of MRI, all the slices of four modalities are combined to form channels. The model is trained by using batch size 4 and 100 epochs with Adam optimizer of 0.0001 learning rate.

3. Results and Discussion

The proposed model is examined on 150 patients BraTS 2019 dataset. We randomly split 2500 images for training, 750 images for testing and 500 for validation. This segmentation of brain tumor experiment is conducted by using 4 modalities of brain tumor. Firstly we have done the preprocessing phase for resizing the actual size of the image and for changing the resolutions. The augmentation process is performed using image data generator to improve the performance of network. After the training process it produces a mask of given brain tumor which indicates the range of infection. The prediction result of randomly selected dataset obtained after training and the corresponding ground truth is depicted in FIGURE 4.

In the segmented output image, white color represents Enhancing Tumor (ET), dark grey represents necrosis and light grey represents edema. The tumors are divided into three categories as : (a) Whole Tumor (WT), that comprises necrosis, edema, and ET (labels 1, 2, and 4), (b) Tumor Core (TC), which contains necrosis and ET (labels 1 and 4), and (c) ET (label 4). This section comprises of experiments conducted for evaluating the dice score of proposed model and also includes the comparative study with existing architectures.

3.1. Evaluation metrics

For the assessment of the proposed model, Dice Score (DS) coefficient and Intersection over Union (IoU) metric are used. The DS was employed as a statistical validation parameter to assess the effectiveness of automated segmentation of MR images. DS can be stated as the ratio between twice the overlap area of ground truth and predicted output with the total number of pixels. The range of the DS value is 0 to 1, with 0 representing no spatial overlap and 1 representing perfect overlap. (Zou et al.). The

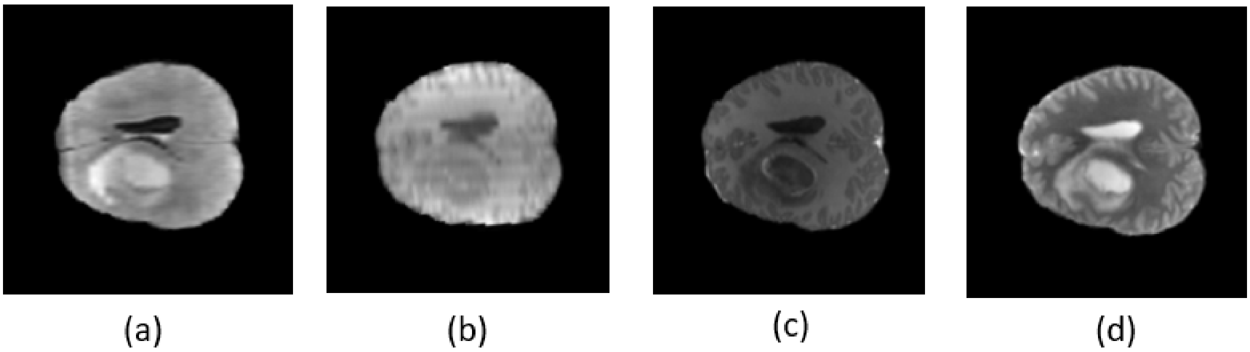

FIGURE 3. **Different modalities of MRI used : (a) Flair, (b) T1, (c) Contrast Enhanced T1 and (d) T2.**

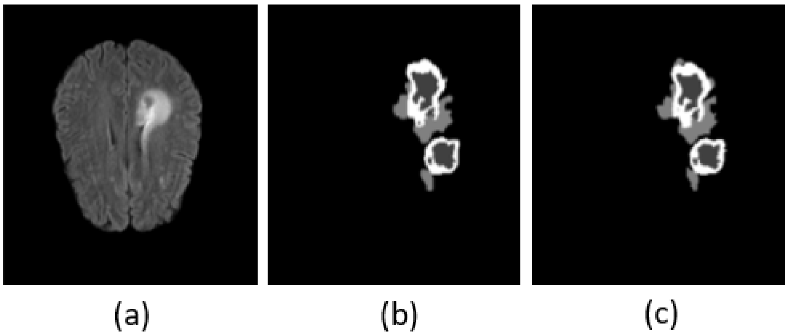

FIGURE 4. **Prediction result of proposed model (a) Input Flair Sequence, (b) Segmentation output of model, (c) Ground Truth.**

DS is calculated as follows (Taha and Hanbury):

$$DS = \frac{2TP}{FN + FP + 2TP} \tag{6}$$

Where TP, FP, FN represent true positive, false positive and false negative prediction respectively. IoU is stated as the ratio between the overlap area of predicted segmentation and the ground truth to the area of union in between them. The IoU is also called as Jaccard index. The IoU is calculated as follows (Zou et al.):

$$IoU = \frac{TP}{TP + FP + FN} \tag{7}$$

Both these evaluation metrics are inter-related as shown below and both has a narrow range of [0,1] and is frequently near to the value of 1.

$$IoU = \frac{DICE}{2 - DICE} \tag{8}$$

In all the experiments we use DS and mean IoU as evaluation metric for WT, TC and ET regions.

3.2. Ablation Study

Attention gate Inception UNet with Guided Decoder model is the modification of UNet architecture with Attention gate, inception block and guided decoder. UNet includes skip connections, which serve as communication channels between encoder and decoder part. The backbone of the architecture is the UNet design, which lacks distinct highlights in the network, resulting in a segmented image that is less accurate. This limitation is overcome in the proposed model by using attention gate that focus on features which are relevant to the task of segmentation and inception block that extract characteristics at each layer (Latif et al.). The losses in the guided decoder will helps to produce better features in the model.

This model is compared with the Attention ResUNet with Guided Decoder (ARU-GD) with attention gate together with residual module and guided decoder (Maji, Sigedar, and Singh). Adding the inception blocks to UNet architecture with attention gate and guided decoder, results in improved segmentation performance than ARU-GD model. The count of filters in every layer at encoder and decoder path is less in the proposed model in order to reduce the computational complexity. The count of feature maps are increasing in the order of 128, 256, 512, 1024, etc at encoder path with inception module followed by Maxpool 2 x 2 down sampling layer and decreases vice versa at the decoder path. In the residual block, only one type of filter size can be used, while in the inception block, numerous sets of 3 × 3 convolutions, 1 × 1 convolutions, 3 × 3 max pooling, and cascaded 3 × 3 convolutions are used which helps to extract additional features of each layer. (Cahall et al.). The segmented prediction results of the model is shown in Fig. 5, which is compared to ARU-GD and ground truth of randomly selected test images.

The tumor segmentation is affected by the problem of class imbalance. Since the Dice Score (DS) coefficient tackles this issue, it has been used as an assessment metric in all the conducted experiments together with the Intersection over Union (IoU). The model has attained DS of 0.9190, 0.9331, 0.8990 and mean IoU of 0.8502, 0.8745, 0.8165 on WT, TC and ET as shown in TABLE 2 and TABLE 3 respectively.

TABLE 2. Comparison of dice score of proposed model .

Method	WT	TC	ET
ARU-GD Model	0.9189	0.9305	0.8997
Attention gate Inception U-Net with Guided decoder	0.9190	0.9331	0.8990

TABLE 3. Comparison of Mean IoU of proposed model

Method	WT	TC	ET
ARU-GD Model	0.8500	0.8700	0.8177
Attention gate Inception U-Net with Guided decoder	0.8502	0.8745	0.8165

3.3. Comparative study

The proposed model is compared with various existing models to analyse the segmentation performance. The comparative study of the model with ARU-GD in terms DS and mean IoU is provided in the Table II and Table III respectively. In this model with inception module, the number of filters used are less as compared to that of ARU-GD model and also has achieved better segmentation results as well. The model consists of less number of trainable parameters, which aids to reduce the computational complexity. Both the DS and Mean IoU of WT and TC is more for the proposed model, only in case of ET region the DS and Mean IoU of proposed model is inferior to ARU-GD model.

TABLE 4. Comparison of dice score after changing parameters.

Method	**WT**	**TC**	ET
ARU-GD Model	0.9156	0.9312	0.8951
Attention gate Inception U-Net with Guided decoder	0.9190	0.9331	0.8990

TABLE 5. Comparison of Mean IoU after changing parameters.

Method	**WT**	**TC**	**ET**
ARU-GD Model	0.8443	0.8712	0.8120
Attention gate Inception U-Net with Guided decoder	0.8502	0.8745	0.8165

Then the comparison is made after reducing the filter counts at each layer of ARU-GD model and made them same as that of our model. Both models are trained using a batch size of 4 over 100 iterations. The number of filters at encoding path of both models are made to 32, 64, 128, 256, etc, where it decreases vice versa at each layer of decoder path.

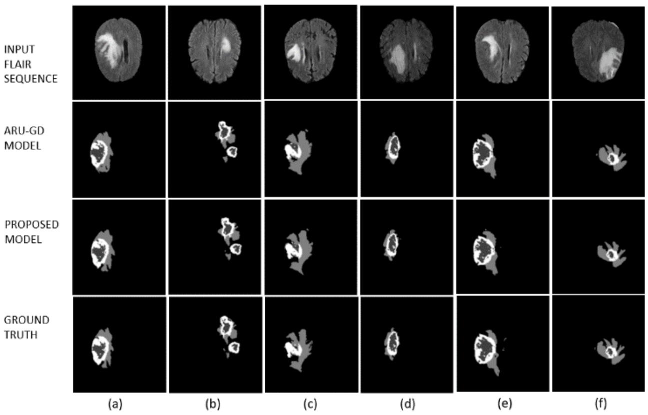

FIGURE 5. **Comparison of prediction results onrandomly selected test images with the ground truth. Whitecolor, dark grey,light grey in segmented image corresponds to enhancing tumor, necrosis and-edema respectively.**

The comparison on DS and mean IoU of both model after making the parameters same is

shown in the TABLE 4 and TABLE 5 respectively. It is evident from the results that the DS and mean IOU of WT, TC and ET region in proposed model is superior as compared to the existing model. The comparative prediction result on randomly selected test images (a-f) is depicted in FIGURE 5. The model is capable of segmenting individual tumor region as closely as possible to the ground truth is evident in FIGURE 5. The model segment all the three regions using different labels.

The dice score of model is compared with existing models of Ronneberger et al. (Ronneberger, Fischer, and Brox) employed with U-Net architecture, Pereira et al. (Pereira et al.), Yang et al. (Yang et al.), Zhou et al. (Zhou et al.), Maji et al. (Maji, Sigedar, and Singh). From the Table VI, it is evident that, dice score of proposed model is high while compared to existing models of tumor segmentation process including U-Net, modified CNN, etc. Among all these 6 models, our model outperforms in all the segmentation of three regions. The segmentation of ET and TC is more difficult, even though all segmentation models are difficult to distinguish between enhancing and tumor core and have less dice score, the suggested Attention gate Inception

UNet with Guided Decoder model successfully segmented these regions with a better segmentation performance.

TABLE 6. **Comparison of proposed model with existing models.**

Method	WT	TC	ET
Ronneberger et al. (Ronneberger, Fischer, and Brox)	0.85	0.82	0.70
Pereira et al. (Pereira et al.)	0.78	0.65	0.75
Yang et al. (Yang et al.)	0.87	0.77	0.75
Zhou et al. (Zhou et al.)	0.88	0.79	0.77
Maji et al. (Maji, Sigedar, and Singh)	0.91	0.87	0.80
Proposed model	0.91	0.93	0.89

4. Conclusion

The ability to segment brain tumors has a significant impact on diagnosis, growth rate prediction and treatment planning. The manual segmentation is often ineffective since it is time consuming and it differs from observer to observer. In this work, a novel deep learning network called Attention gate Inception UNet with Guided Decoder for brain tumor segmentation is proposed. A guided decoder and attention gates have been incorporated into the architecture. These modifications to the baseline network of inception module together with

UNet architecture helps to enhance the learning process through superior feature maps in decoder path and allow only the relevant information at the encoder side. The model is trained and tested using BraTS 2019 dataset, and has achieved better dice score as compared to existing segmentation models. The model with inception block, Attention gate and Guided Decoder will improve the performance and help for the reduction of loss function. The proposed model has obtained dice score of 0.9190, 0.9331, 0.8990 and mean IoU of 0.8502, 0.8745, 0.8165 on WT, TC and ET respectively.

5. Future Works

In future, the tumor segmentation performance can be extended by including other image modalities and through further modification in the architecture, thereby resulting in the improvement of clinically approved automatic segmentation methods for efficient diagnosis.

6. Acknowledgement

We would like to acknowledge the BRATS research team for their efforts and for providing the data that allowed us to perform this research.

7. Authors' Note

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

References

Adel, Kermi, Issam Mahmoudi, and Mohamed Tarek Khadir. "Deep Convolutional Neural Networks Using U-Net for Automatic Brain Tumor Segmentation in Multimodal MRI Volumes". *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Springer International Publishing, 2019. 37–48.

Alqazzaz, Salma, et al. "Automated brain tumor segmentation on multi-modal MR image using SegNet". *Computational Visual Media* 5.2 (2019): 209–219.

Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.12 (2017): 2481–2495.

Bjoern, H Menze, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)". *IEEE Transactions on Medical Imaging* 34.10 (2015): 1993–2024.

Cahall, Daniel E, et al. "Inception Modules Enhance Brain Tumor Segmentation". *Frontiers in Computational Neuroscience* 13 (2019): 44–44.

Dong, Hao, et al. "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks". *Communications in Computer and Information Science* (2017): 506–517.

Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks". *Nature* 542.7639 (2017): 115–118.

Fu, Jun, et al. "Dual Attention Network for Scene Segmentation". *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019): 3146–3154.

Havaei, Mohammad, et al. "Brain tumor segmentation with Deep Neural Networks". *Medical Image Analysis* 35 (2017): 18–31.

Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-Excitation Networks". *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018): 7132–7141.

Işın, A, C Direkoglu, and M Şah. "Review of MRI-based brain tumor ˘image segmentation using deep learning methods". *Procedia Comput. Sci* 102 (2016): 317–324.

Kermi, Adel, Issam Mahmoudi, and Mohamed Tarek Khadir. "Deep Convolutional Neural Networks Using U-Net for Automatic Brain Tumor Segmentation in Multimodal MRI Volumes". *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (2019): 37–48.

Latif, Urva, et al. "An ¡scp¿end-to-end¡/scp¿ brain tumor segmentation system using ¡scp¿multi-inception-UNET¡/scp¿". *International Journal of Imaging Systems and Technology* 31.4 (2021): 1803–1816.

Maji, Dhiraj, Prarthana Sigedar, and Munendra Singh. “Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors”. *Biomedical Signal Processing and Control* 71 (2022): 103077–103077.

Menze, Bjoern H, et al. “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”. *IEEE Transactions on Medical Imaging* 34.10 (2015): 1993–2024.

Naceur, Mostefa Ben, et al. “Fully Automatic Brain Tumor Segmentation using End-To-End Incremental Deep Neural Networks in MRI images”. *Computer Methods and Programs in Biomedicine* 166 (2018): 39–49.

Noori, Mehrdad, Ali Bahri, and Karim Mohammadi. “Attention-Guided Version of 2D UNet for Automatic Brain Tumor Segmentation”. *2019 9th International Conference on Computer and Knowledge Engineering (ICCKE)* (2019).

Pereira, Sergio, et al. “Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images”. *IEEE Transactions on Medical Imaging* 35.5 (2016): 1240–1251.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. *Lecture Notes in Computer Science* 9351 (2015): 234–241.

Szegedy, Christian, et al. “Going deeper with convolutions”. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2015. 1–9.

Taha, Abdel Aziz and Allan Hanbury. “Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool”. *BMC Medical Imaging* 15.1 (2015): 29–29.

Verduin, Maikel, et al. “Noninvasive Glioblastoma Testing: Multimodal Approach to Monitoring and Predicting Treatment Response”. *Disease Markers* 2018 (2018): 1–11.

Wang, Fei, et al. “Residual Attention Network for Image Classification”. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017): 3156–3164.

Yang, Chushu, et al. “Automatic Brain Tumor Segmentation Method Based on Modified Convolutional Neural Network”. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2019): 998–1001.

Zhang, Jianxin, et al. “Attention Gate ResU-Net for Automatic MRI Brain Tumor Segmentation”. *IEEE Access* 8 (2020): 58533–58545.

Zhang, Tinghong, et al. “Multiple Attention Fully Convolutional Network For Automated Ventricle Segmentation in Cardiac Magnetic Resonance Imaging”. *Journal of Medical Imaging and Health Informatics* 9.5 (2019): 1037–1045.

Zhou, Chenhong, et al. “Learning Contextual and Attentive Information for Brain Tumor Segmentation”. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (2019): 497–507.

Zou, Kelly H, et al. “Statistical validation of image segmentation quality based on a spatial overlap index1”. *Academic Radiology* 11.2 (2004): 178–189.

Embargo period: The article has no embargo period.

To cite this Article: , Amisha P, and Adersh V R . “Automated Brain Tumor Segmentation Using Attention gate Inception UNet with Guided Decoder.” International Research Journal on Advanced Science Hub 05.05S May (2023): 463–473. http://dx.doi.org/10.47392/irjash.2023.S062