ORIGINAL RESEARCH

Brain Tumor Segmentation from 3D MRI Scans Using U-Net

Sidratul Montaha[1] · Sami Azam[2] · A. K. M. Rakibul Haque Rafid[1] · Md. Zahid Hasan[1] · Asif Karim[2]

Abstract

A fully automated system based on three-dimensional (3D) magnetic resonance imaging (MRI) scans for brain tumor segmentation could be a diagnostic aid to clinical specialists, as manual segmentation is challenging, arduous, tedious and error prone. Employing 3D convolutions requires large computational cost and memory capacity. This study proposes a fully automated approach using 2D U-net architecture on BraTS2020 dataset to extract tumor regions from healthy tissue. All the MRI sequences are experimented with the model to determine for which sequence optimal performance is achieved. After normalization and rescaling, using optimizer Adam with learning rate 0.001 on T1 MRI sequence, we get an accuracy of 99.41% and dice similarity coefficient (DSC) of 93%, demonstrating the effectiveness of our approach. The model is further trained with different hyper-parameters to assess the robustness and performance consistency.

Keywords Brain tumor segmentation · 3D MRI · U-Net · BraTS dataset

Introduction

Brain tumors are masses that arise as a result of irregular brain cell proliferation and the loss of the brain's regulatory systems. Approximately 700,000 people worldwide are suffering from brain tumors, with 86,000 new cases identified in 2019. Nearly 16,380 people died in 2019 from brain tumors [1]. Gliomas are the most prevalent malignant tumors [2], representing 80% of all malignant brain tumors. Early identification of brain tumors is crucial in the diagnosis of cancer as it aids in the selection of the most appropriate treatment strategy to preserve a patient's life.

However, the detection and segmentation of tumors based on imaging and human interpretation is challenging, as tumors can have a mixture of low-grade and high-grade characteristics. In addition, this is often time-consuming and does not always yield high accuracy [2]. A non-invasive computer-aided (CAD) fully automatic diagnostic method might aid clinical specialists in early diagnosis, analysis and treatment planning and decrease the mortality rate by providing more consistent, faster and more correct tumor identification [3]. For this reason, a number of CAD-based image segmentation methods are currently being investigated for this application [4]. The advance of CNN has led to noteworthy advancements on the field of computer vision and different brain tumor segmentation techniques based on deep learning [5].

In this study, we concentrate on tumor segmentation which is regarded as one of the most difficult challenges in multimodal MRI images. The multimodal Brain Tumor Segmentation Challenge 2020 dataset (BraTS 2020) has been used for this research. Four different MRI modalities are included for each patient in the dataset with corresponding manually segmented region of interest (ROI). The main contributions of this research can be stated as follows:

- A convolutional network (U-net model) is designed for relatively quick and accurate image segmentation. In the ISBI challenge, it outperformed the previous best technique for segmenting neuronal structures. The architecture is configured with skip connections to boost performance.
- Instead of employing fixed hyper-parameters, a novel approach is presented to develop the model architecture.

✉ Sami Azam
sami.azam@cdu.edu.au

✉ Asif Karim
asif.karim@cdu.edu.au

1 Department of Computer Science and Engineering, Daffodil International University, 102/1, Sukrabad Mirpur Rd, Dhaka 1207, Bangladesh

2 Faculty of Science and Technology, Charles Darwin University, Ellengowan Drive, Casuarina, NT 0909, Australia

In this regard, the model is optimized with ablation study where a number of experimentations are conducted by altering different hyper-parameters.

- Rather than training the model with 3D images, we have employed single slice of the 3D MRI as the aim of our approach is to minimize the computational cost while achieving the best possible performance.
- In the pre-processing step, the middle slice of 3D MRI is extracted automatically and a normalization approach is carried out which impacts the overall performance.
- The model is trained and validated on each of the four sequences separately to determine which sequence results in the highest accuracy.
- Across different hyper-parameters, the model is able to yield optimal performance which further validates the robustness and performance consistency of the architecture.

Literature Review

Various innovative approaches for automated segmentation of brain tumor have been presented in recent years. Pereira et al. [6], presented an automated segmentation approach using a CNN model. Their approach was tested on the BraTS 2013 dataset, achieving the place with dice coefficient (DC) measures of 88, 83 and 77% for the entire, core and enhancing areas, respectively. Other authors [7] presented a computerized system that can distinguish between a normal and an irregular brain based on MRI images and classify abnormal brain tumors as high-grade glioma (HGG) or low-grade glioma (LGG). With accuracy, specificity and sensitivity scores of 99%, 98.03% and 100%, respectively, their system effectively identified HGGs and LGGs. Noori et al. [8] developed a low-parameter-based network 2D U-Net that uses two different approaches. Using the BraTS 2018 validation data set, they got dice scores of 89.5, 81.3 and 82.3% for whole tumor (WT), enhancing tumor (ET) and tumor core (TC), respectively. Xu et al. [9] enhanced the performance of the segmentation of tumors with a 3D masked U-net multi-scale which captures features by assembling multi-scale architecture images as input, integrating a 3D atrous spatial pyramid pooling layer (ASPP). They obtained dice scores of 80.94, 90.34 and 83.19% on the BraTS 2018 validation set for ET, WT, and TC, respectively. On the BraTS 2018 test data set, this technique achieved dice scores of 76.90, 87.11 and 77.92% for ET, WT, and TC, respectively. Other authors [10] also used the BraTS dataset to evaluate a U-Net model, undertaking two experiments to compare the network's performance. Comparing the results of the first and third sets, the accuracy in detecting HGG patients was 66.96% and 62.22%, respectively. For LGG, the second and third set's accuracies were 63.15% and

62.28%, respectively. Pravitasari et al. [11] proposed a cross between a U-Net and a VGG16 model. It was found that the model was able to detect the tumor region, with an average CCR score of 95.69%. SegNet3, U-Net, SegNet5, Seg-UNet, Res-SegNet and U-SegNet have been used on BraTS datasets in this study [12]. The average accuracy of Res-SegNet, U-SegNet, and Seg-UNet was recorded as 93.3%, 91.6% and 93.1%, respectively. According to Pei et al. [13], feature-based fusion methods can predict a better tumor lesion segmentation for the total tumor (DSCWT = 0.31, $p = 0.15$), tumor core (DSCTC = 0.33, $p = 0.0002$), and the enhanced tumor (DSCET = 0.44, $p = 0.0002$) areas. However, in the real segmentation of WT and ET (DSCWT = 0.85 ± 0.055, DSCET = 0.837 ± 0.074, the approach provided a statistically significant improvement. Lin et al. [14] proposed a 3D deep learning architecture named context deep-supervised U-Net to segment brain tumor from 3D MRI scans. The suggested approach acquired a DSC of 0.92, 0.89, and 0.846 on WT, TC and ET, respectively. Punn et al. [15] proposed a 3D U-net model for brain tumor segmentation using 3D MRI across WT, TC and ET lesions, respectively. The proposed deep learning model performed best achieving the highest DSC of 0.92 on WT, 0.91 on CT and 0.84 on ET. In both of these studies, data normalization was conducted as a pre-processing step. Ullah et al. [16] proposed a 3D U-Net model for the automatic segmentation of brain tumor using 3D MRI scans. A number of image pre-processing techniques were conducted to remove noise and enhance the quality of 3D scans. The suggested method achieved mean DSC of 0.91, 0.86, and 0.70 for the WT, TC and ET, respectively. Table 1 represents the main methodology and limitations of some of these papers.

It is observed from Table 1 that in most of the papers, a lack of image processing and model hyper-parameter tuning exists. A few studies used 3D CNN model which is computationally extensive due to 3D convolutional layer. As 3D brain MRI comprises multiple modalities, these modalities should be explored in experimenting with deep learning models for automated segmentation which is another limitation found in some papers. In this study, we attempt to address these drawbacks in order to develop an optimal U-net model with the highest possible performance.

Dataset

Our system is trained and analyzed with the (BraTS2020 dataset, collected from Kaggle [17]. A total of 473 subjects of 3D images are available in the dataset. For every patient, four MRI sequences: fluid attenuated inversion recovery (FLAIR), T1-contrast-enhanced (T1ce), T1-weighted (T1), T2-weighted (T2) as well as the corresponding ROI (seg) are obtained. The provided ground truths were labeled by the

Table 1 Methodology and limitations of the previous literatures

Paper	Dataset	Methodology	Limitations
Pereira et al. [6]	BraTS 2013	CNN model for tumor segmentation	(i) Use of a single brain MRI modality
Noori et al. [8]	BraTS 2017	Shallow 2D U-Net model for tumor segmentation	(i) Absence of hyper-parameter tuning
	BraTS 2018		(ii) Absence of data-pre-processing
Xu et al. [9]	BraTS 2018	3D U-Net model for tumor segmentation	(i) High computational complexity due to 3D model architecture
Choi et al. [10]	BraTS 2015	U-Net model for tumor segmentation	(i) Use of a single brain MRI modality
Pravitasari et al. [11]	BraTS	UNet-VGG16 hybrid model for tumor segmentation	(i) Use of a single brain MRI modality (ii) absence of data-pre-processing
Lin et al. [14]	BraTS 2020	3D context deep-supervised U-Net model for tumor segmentation	(i) High computational complexity due to 3D model architecture
			(ii) Absence of hyper-parameter tuning
Punn et al. [15]	BraTS 2017	3D U-Net model for tumor segmentation	(i) High computational complexity due to 3D model architecture
	BraTS 2018		(ii) Absence of hyper-parameter tuning
Ullah et al. [16]	BraTS 2018	3D U-Net model for the automatic segmentation of brain tumor	(i) High computational complexity due to 3D model architecture
			(ii) Absence of hyper-parameter tuning

Table 2 Dataset description

Task	Brain tumor segmentation
Dataset	BraTS 2020
Dataset source	Kaggle
Image type	3D brain MRI
Image format	NIfTI
Image size	$224 \times 224 \times 150$
Total number of subjects	473
Number of images in each subject	5
Name of the four sequences and corresponding ROI of each subject	FLAIR, T1, T1ce, T2 and Seg

Fig. 1 Visualization of 3D brain MRI

experts. Each 3D volumes, includes 155 2D slices/images of brain MRIs collected at various locations across the brain. Every slice is 240×240 pixels in size in NIfTI format and is made up of single-channel grayscale pixels. The dataset is summarized in Table 2.

Visualization of 3D MRI Scans

A 2D image consists of single or multi-channel pixels whereas 3D images comprise 3D cubes or voxels. While reading a NIfTI file, all the information in the file is encoded, where each detail is known as an attribute. When visualizing a 3D image (Fig. 1), a list is initialized in which, whenever a volume is read, it iterates over all the 155 slices of the 3D volume to append each slice sequentially in the list. The number of voxels in a 3D image is calculated with Eq. (1).
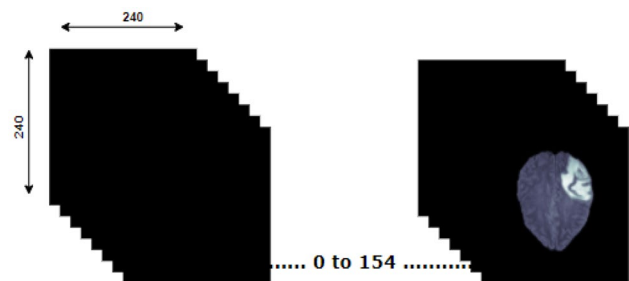
$$V_t = S_t \times H_s \times W_s, \tag{1}$$

where V_t is the total voxel number of an image, S_t is the number of 2D slices of a 3D image, H_s is the height of each slice and W_s is the width of each slice. Figure 1 illustrates the visualization of a 3D MRI.

Description of MRI Sequences and Corresponding ROI

The FLAIR sequence is generated with a very long time to echo (TE) and repetition time (TR) and abnormalities, such as edematous tissues are found bright, while the normal cerebrospinal fluid (CSF) fluid remains comparatively dark. In T1-weighted MRI sequences, TR and TE are both kept shorter with the result that both tumor regions and CSF appear dark. T2-weighted sequences are produced using long TE and TR times, which make both tumor and CSF bright. In the T1ce sequence, tumor and CSF both appear
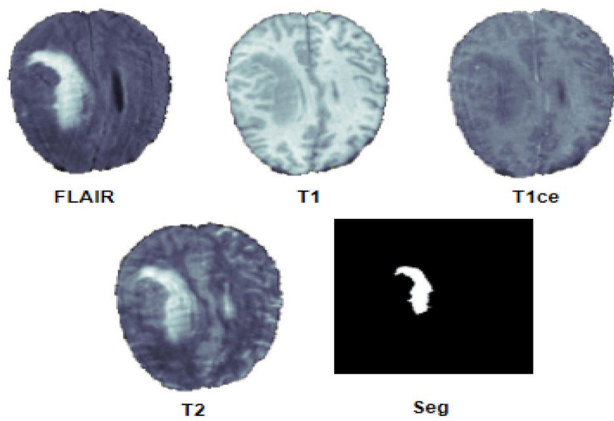
Fig. 2 2D representation of all four sequences and segmented ROI of a 3D MRI

Fig. 3 Proposed methodology

dark. All the sequences therefore contain different characteristics, and it is useful that all of these sequences are used for training to determine which sequence results in the best segmentation performance in terms of accuracy. 2D slices of all the sequences are shown in Fig. 2 for a single subject.

Proposed Methodology

In the data pre-processing step, normalization and rescaling are implemented on each of the four MRI sequences. Afterwards, every sequence is passed as the training dataset into the model. The corresponding manually segmented ROI (Seg) is employed as ground truth for each of the sequences separately. Adam optimizer with a learning rate of 0.001 is used for training the model. For four sequences, four results are produced namely res-1, res-2, res-3 and res-4 (Fig. 3). Finally, based on the accuracy for each particular sequence, the optimal resulting model is obtained.

Data Preprocessing

Brain tumor classification using 3D MRI scans is usually difficult and computationally complex, requiring pre-processing techniques to improve the model's performance [18, 19]. In this research, two pre-processing steps are performed for every MRI sequence before training the model: intensity normalization and rescaling.

Normalization

As MRI intensities differ based on manufacturers, procurement parameters, and different sequences are taken over different periods, the 3D images need to be normalized before feeding them to a model. As scanning of patients is likely
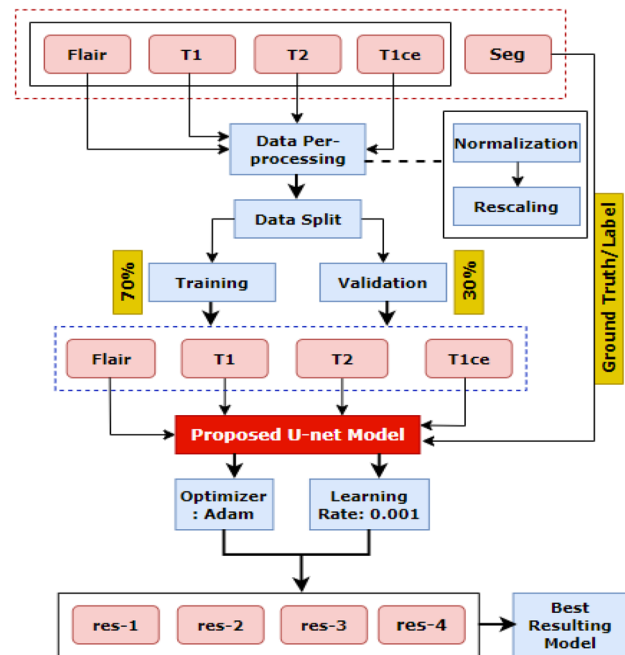
to be performed in different lighting environments, intensity normalization plays an imperative role in brain tumor segmentation.

Data normalization means transforming floating-point feature values from their regular range into a new arbitrary standard range that is usually between 0 and 1 [21]. Min–max normalization is a widely used technique for normalization that is performed in our research on each MRI sequence independently. In this process, for each feature, the minimum value is converted to 0 and the maximum value is converted to 1. The remaining values are transformed into a range of values between 0 and 1. Normalization is performed using Eq. (2).

$$zi = \frac{pi - \min(p)}{\max(p) - \min(p)}, \tag{2}$$

where p refers to the pixel values means $p = (p1,\ldots, pn)$ and zi refers to i th resultant normalized data.

Rescaling

Due to the GPU memory limitation, after normalization, the dataset is resampled to 128×128×1 voxels [9]. Only the middle single slices are utilized instead of all 155 slices of the brain where the original dimensions of 3D brain MRI image are 240 × 240 × 155 pixels. As experimenting with 3D data requires high computational

resources, we use Intel Core i5-8400 Processor, NVidia GeForce GTX 1660 GPU, 8 GB of Memory, and 256 GB DDR4 SSD for storage.

Data Split

Each of the preprocessed sequences and corresponding ROI images is split into two sets (training and validation) maintaining a ratio of training to validation of 70:30. The segmented ROI images (seg) are used as training and validation labels while training and validating the model. The diagram in Fig. 4 is shown to provide better insight.
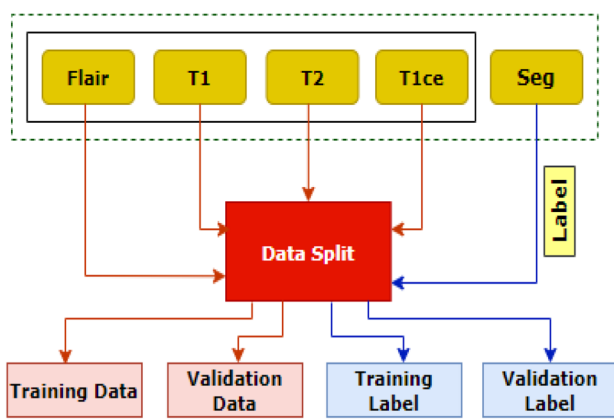

Fig. 4 Dataset split

Proposed Architecture

U-Net Architecture

For the task of segmenting tumor regions from MRI images, the U-Net segmentation model is used in this research which is a widely accepted architecture [20], developed by Olaf Ronneberger et al. [21] for the purpose of segmenting Biomedical images. U-Net utilizes a compact encoder–decoder structure that concatenates the features from many encoder levels with the decoder layers twice [22]. Figure 5 illustrates the U-Net architecture.

U-Net consists of a total of 23 convolutional layers and is made up of two paths: a contraction path (the encoder) and a symmetric expanding path (the decoder) (Fig. 5). The encoder or contraction path captures image contexts and it is made of repeated stacks of convolution and maxpooling layers alike CNN. In the encoder, the dimensions of the image gradually decrease. The encoder has a pair of 3×3 unpadded convolutional layers, each of them followed by a rectifier linear unit (ReLU) and a maxpooling layer of 2×2 sized kernel. Maxpooling layers are used for down sampling the image. The first set of convolutional layers has 64 filters and the filter number is doubled in each next set of convolutional layers (e.g., second set of convolution layers with filters $64 \times 2 = 128$). The decoder of the symmetric expanding path uses transposed convolutions (deconvolution layer) that enable precise localization of the ROI in images. The size of images increases gradually while the depth of the image decreases because of the
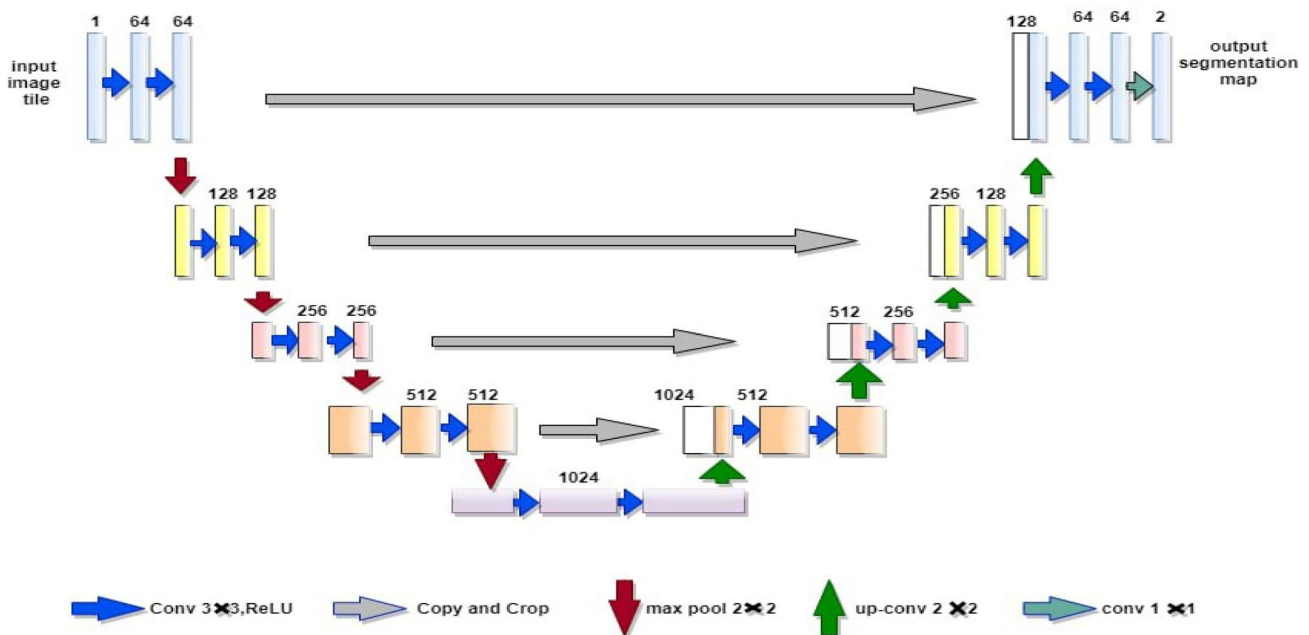

Fig. 5 U-Net model architecture

deconvolutional layers. In the decoder, upsampling of feature maps commences. This task is carried out by deconvolution layers of 2×2 which halve the number of feature channels. Subsequently, concatenation with the correspondingly cropped feature map of the encoder is done, followed by two 3×3 deconvolution layers and ReLU Basic working formula of ReLU is [23]:

$$\text{ReLU}(q) = \begin{cases} 0, & \text{if } q \leq 0 \\ q, & \text{otherwise.} \end{cases} \tag{3}$$

The output layer of this architecture consists of a 1×1 deconvolutional layer. This maps all the feature vectors into two classes (ROI and background of image) outputting an image consisting of segmented ROI (tumor) in one color and the rest of the image in another color. A sigmoid activation function Eq. (4) is used to improve accuracy. The provided pre segmented ROI images are used as validation data [23].

$$\text{Sigmoid}(q) = \frac{1}{1 + \exp(-q)}. \tag{4}$$

As this architecture does not contain any dense layers, it is a fully convolutional network (FCN) capable of taking any size of image as input. Moreover, the model uses skip connectors to reduce the loss of information. In developing a deep learning architecture, skip connections are introduced to address different issues. Particularly, skip connections allow feature reusability, while alleviating model training and convergence. In this study, skip connection is employed to improve the overall performance and competence of the model. As skip connection has a continual gradient flow from the first layer to the last, it addresses the vanishing gradient problem of the proposed U-net model effectively.

Training Strategy

The model proposed in this paper is separately trained and validated four times with a total of 473 images, containing only a single grayscale channel, for four MRI sequences. The corresponding manually segmented ROIs are used as true labels. While training, the optimizer Adam is used with a learning rate of 0.001 [24]. Equation (5) is the mathematical formula of the Adam optimizer [25].

$$\omega_t = \omega_{t-1} - \eta \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t} + \varepsilon}, \tag{5}$$

where ω is the weight of the model, η is the size of each step, and β, ε are the required hyper-parameters. Here $m_t = \beta_1 m_{t-1} + (1 - \beta_1)g_t$ and squared gradient $v_t = \beta_2 v_{t-1} + (1 - \beta_{21})g_t{}^2$, estimation bias corrected of first moments $\widehat{m}_t$, and similarly, second moments $\widehat{v}_{t=} v_t / (1 - \beta_{t1}{}^2)$. The batch size is set to 32 and

binary_crossentropy is employed as loss function. Binary_crossentropy [22] is defined as:

$$L_c = -\frac{1}{n} \sum_{i=1}^{n} \left\{ y_i \log(f_i) + (1 - y_i)\log(1 - f_i) \right\}, \tag{6}$$

where n is the number of samples, y_i is the true label of a specific sample and f_i is its predicted label. The model is trained for 50 epochs [26]. All the training parameters for training the network are described in Table 3.

Results and Discussion

DSC score, sensitivity [27] and specificity [28] are used to evaluate the model besides accuracy and validation accuracy [29] based on the provided segmented ground truth of the tumor part in the MRI.

Evaluation Metrics

DSC performance metric computes the similarity percentage between the ground truth and the output of a model. Suppose, P and Q are two sets, the dice similarity of these two sets are then calculated with Eq. (7) [23].

$$DSC = \frac{2 \times |P \cap Q|}{|P| + |Q|}, \tag{7}$$

$$\text{Sensitivity}(P, T) = \frac{|P_1 \wedge T_1|}{|T_1|}. \tag{8}$$

Sensitivity is calculated with Eq. (8) where, cardinalities of sets P and Q are denoted with |P| and |Q|, respectively, T_1 representing the proportion of tumor regions of ground truth images and P_1 represents tumor regions that were predicted by the model [5].

$$\text{Specificity}(P, T) = \frac{|P_0 \wedge T_0|}{|T_0|}. \tag{9}$$

Specificity is calculated with Eq. (9) where T_0 represents non-tumor tissue regions of the ground truth and P_0

Table 3 Training hyper-parameters

Loss function	Binary_crossentropy
Optimizer	Adam
Learning rate	0.001
Batch size	32

represents the non-tumor tissue regions predicted by the model [5].

Performance Computation of all Trained Models

In Table 4, validation loss is referred as 'V_Loss', validation accuracy as 'V_Acc', specificity as 'Spe', sensitivity as 'Sen' and lastly dice similarity coefficient score as 'DSC'.

Among all the model configurations trained with different MRI sequences, U-Net trained with the T1 MRI sequence results in the best validation accuracy of 99.41% with a validation loss of 0.037. The T2 MRI sequence results in the lowest validation score of 98.25% with a validation loss of 0.08. The validation accuracies of the FLAIR and T1ce sequences are 98.95% and 98.68%, respectively. In terms of specificity and sensitivity, U-Net trained on T1 MRI sequence has the best performance with scores of 99.68% and 98.97% for specificity and sensitivity, respectively. The T2 MRI sequence has the lowest scores in this regard with a specificity of 98.37% and a sensitivity of 98.49%. The most important measure regarding the performance of a segmentation model is the DSC for this metric the best scores (93.86%) are also achieved with the T1 MRI sequence trained on the U-Net segmentation model. The FLAIR MRI sequence results in a DSC of 91.23% which is slightly less than the DSC for the T1 MRI sequence. The lowest DSC of 79.32% is recorded with the T2 MRI sequence.

Moreover, further examination is conducted to assess the impact of model hyper-parameters on performance [30]. As for T1 modality, the highest performance is achieved, and the experiments concerning hyper-parameters are conducted using the dataset of T1. In this regard, different batch sizes, number of epochs, loss functions, optimizers, learning rate and dropout factors are investigated. Table 5 shows the results regarding changing these hyper-parameters.

It is observed from Table 5 that across different hyper-parameters, the model is able to provide optimal outcomes. For batch size 32 the highest DSC of 93.86% is obtained whereas near highest performance is achieved for batch size of 16 and 64. However, while using batch size 128, the performance drops slightly. Across all three configurations regarding epoch number (50, 100 and 150), the highest

Table 4 Validation accuracy (V_Acc), validation loss (V_Loss), specificity (Spe), sensitivity (Sen) and dice similarity coefficient (DSC) scores of the approach

MRI sequence	V_Loss	V_Acc	Spe	Sen	DSC
FLAIR	0.061	98.95	99.14	98.53	91.23
T1	**0.037**	**99.41**	**99.68**	**98.97**	**93.86**
T1ce	0.072	98.68	98.52	98.72	85.67
T2	0.08	98.25	98.37	98.49	79.32

T1 MRI Sequence Performed Best

DSC of 93.86% is acquired. However, to minimize the training time, epoch number 50 is kept. For both loss functions binary cross-entropy and categorical cross-entropy, the highest DSC of 93.86% is acquired is recorded. Regarding optimizer, for Adam, Nadam and Adamax, DSC > 93% is achieved where for Adam the highest DSC is found. The SGD optimizer records a slightly lower DSC of 92.43%. Across all the learning rates except 0.01 DSC > 93% is achieved. A learning rate of 0.001 is selected since for this configuration the highest DSC is acquired and lower training time is required comparatively. Three dropout factors of 0.2, 0.5 and 0.8 are experimented with where for 0.2 and 0.5 the highest DSC of 93.86% is obtained. However, for dropout factor of 0.8 performance falls to 91.38%. From the results of these experiments, it can be concluded that the model is stable and robust enough to yield optimal performance across different hyper-parameters. No sign of overfitting is perceived while training the model altering hyper-parameters. Furthermore, over different dropout factors, the model can maintain its performance consistency with no occurrence of overfitting.

Comparison with Some Existing Literatures

The proposed segmentation approach of brain MRI is compared with some recent studies conducted on similar datasets on the basis of DC scores on Table 6.

With our proposed approach, it was made possible to achieve a dice score of 93%. Wei et al. [5] experimented with BraTS 2018 dataset and proposed a 3D segmentation model (S3D-UNet) which is a U-Net-based segmentation model. Using this model, they achieved a DSC of 83%. A similar approach was taken by Yanwu et al. [9] who proposed a 3D segmentation model (3D U-Net) based on an existing U-Net segmentation model, achieving a DSC of 89% with the BraTS dataset. A similar approach to our study was taken by [31] where a U-Net segmentation model was utilized for segmenting MRI images from the BraTS 2018 dataset, achieving a DSC of 82% (Table 6). A better result was obtained by [23] who proposed a BU-Net segmentation model using the BraTS 2017 dataset for training and validation. The BraTS 2018 dataset was also used to test their proposed model resulting in a 90% DSC score. Ghosh et al. [20] experimented with two models named baseline U-Net and U-Net with ResNeXt50 for brain tumor segmentation. LGG segmentation dataset (TCGA-LGG) was used in this study. Results suggest that U-Net with ResNeXt50 outperformed baseline U-Net with DSC of 93.2%. The approach proposed in this paper outperforms all studies shown in Table 6 with a DSC of 93.9%. Moreover, as stated 3D data processing and developing a deep learning model accordingly requires high computational power, most of these prior studies employed

Table 5 Performance of the model across different hyper-parameter

Experimentations with various batch sizes

No.	Batch size	DSC (%)	Findings
1	16	93.17	Near highest performance
2	32	93.86	Highest performance
3	64	93.62	Near highest performance
4	128	92.84	Performance fallen

Experimentations with various number of epochs

No.	Epochs	DSC (%)	Findings
1	50	93.86	Highest performance
2	100	93.86	Highest performance
3	150	93.86	Highest performance

Experimentations with various loss functions

No.	Loss Function	DSC (%)	Findings
1	Categorical Cross-entropy	93.86	Highest performance
2	Binary Cross-entropy	93.86	Highest performance

Experimentations with various optimizers

No.	Optimizer	DSC (%)	Findings
1	Adam	93.86	Highest performance
2	Nadam	93.72	Near highest performance
3	Adamax	93.36	Near highest performance
4	SGD	92.43	Performance fallen

Experimentations with various learning rates

No.	Learning rate	DSC (%)	Findings
1	0.001	93.86	Highest performance
2	0.01	92.84	Performance fallen
3	0.0001	93.61	Near highest performance
4	0.007	93.77	Near highest performance
5	0.0007	93.44	Near highest performance

Experimentations with various dropouts

No.	Dropout	DSC (%)	Findings
1	0.2	93.86	Highest performance
2	0.5	93.86	Highest performance
3	0.8	91.38	Performance fallen

efficient hardware configuration including processor, memory and GPU.

Conclusion

In this paper, a CNN-based 2D U-net segmentation model is proposed for segmenting brain tumor from MRI images. This method receives each of the MRI sequences individually as training data and extracts ROIs (tumor regions in Brain MRI). To decrease computational cost and increase efficiency, the input images are normalized and rescaled into single 128 x 128 images. We incorporate 2D layers, to better integrate the information. This model is trained with the brain MRI dataset BraTS 2020 for each of the four MRI sequences (FLAIR, T1, T1ce, T2) to find which sequence gives the best segmentation performance based on the ground truth of the MRI images. The highest DSC score (93.9%) is recorded with the T1 MRI sequence training on U-Net model. Results are compared with previous research demonstrating that this is a promising approach while decreasing computational complexity.

Table 6 Comparison with existing literature

Paper	Dataset	Model	DC
Wei et al. 2019	BraTS, 2018	S3D-U-Net	83%
Yanwu et al. 2019	BraTS	3D U-Net	89%
Tran et al. 2019	BraTS, 2018	2D U-Net	82%
Mobeen et al.	BraTS, 2017	BU-Net	90%
	BraTS, 2018 for validation dataset		
Ghosh et al. 2021	LGG segmentation dataset (TCGA-LGG)	U-Net and U-Net with ResNeXt50	93.2%
This paper	**BraTS, 2020**	**U-Net**	**93.9%**

The proposed Model Performed Best

Limitation and Future Work

The paper presents a novel approach for the automated segmentation of brain tumor from 3D MRI scans using optimized U-net model. In future study, the research can be extended by increasing the number of images. In this regard, some other 3D datasets of brain MRI can be explored. The deep learning architecture can be further optimized with hybrid CNN or attention mechanism-based approach. Moreover, we aim to present a classification of different types of brain tumors after segmenting the tumor region from 3D MRI scans.

Data availability The Dataset has been Collected from: Brain Tumor Segmentation 2020 Dataset. [Online]. https://www.kaggle.com/awsaf49/brats20-dataset-training-validation. It is Publicly available and free to use.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Rehman A, Khan MA, Saba T, Mehmood Z, Tariq U, Ayesha N. Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture. Microsc Res Tech. 2021;84(1):133–49. https://doi.org/10.1002/jemt.23597.
2. Irmak E. Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework. Iran J Sci Technol. 2021;45(3):1015–36. https://doi.org/10.1007/s40998-021-00426-9.
3. Sharif MI, Khan MA, Alhussein M, Aurangzeb K, Raza M. A decision support system for multimodal brain tumor classification using deep learning. Complex Intell Syst. 2022;8(4):3007–20. https://doi.org/10.1007/s40747-021-00321-0.
4. H. N. T. K. Kaldera, S. R. Gunasekara, and M. B. DIssanayake, "Brain tumor Classification and Segmentation using Faster R-CNN," *2019 Adv. Sci. Eng. Technol. Int. Conf. ASET 2019*, pp. 1–6, 2019, doi: https://doi.org/10.1109/ICASET.2019.8714263.
5. Crimi A, et al. Glioma, multiple sclerosis, stroke and traumatic brain injuries, vol. 2. Cham: Springer International Publishing; 2018.
6. Pereira S, Pinto A, Alves V, Silva CA. Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging. 2016;35(5):1240–51. https://doi.org/10.1109/TMI.2016.2538465.
7. Han'guk Chŏngbo Kwahakhoe, IEEE Computer Society, and Institute of Electrical and Electronics Engineers, "The 32nd International Conference on Information Networking (ICOIN 2018) : January 10 (Wed.)-12 (Fri.), 2018, Holiday Inn Chiang Mai, Chiang Mai, Thailand," pp. 473–475, 2018.
8. Noori M, Bahri A, Mohammadi K. Attention-guided version of 2D UNet for automatic brain tumor segmentation. In: 2019 9th international conference on computer and knowledge engineering. Piscataway: IEEE; 2019. p. 269–75. https://doi.org/10.1109/ICCKE48569.2019.8964956.
9. Le XH, Ho HV, Lee G, Jung S. Application of long short-term memory (LSTM) neural network for flood forecasting. Water (Switzerland). 2019;11(7):1387. https://doi.org/10.3390/w11071387.
10. Choi SG, Sohn CB. Detection of HGG and LGG brain tumors using u-net. Medico-Legal Updat. 2019;19(1):560–5. https://doi.org/10.5958/0974-1283.2019.00132.4.
11. Pravitasari AA, et al. "UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation. Telkomnika (Telecommun Comput Electron Control). 2020;18(3):1310–8. https://doi.org/10.12928/TELKOMNIKA.v18i3.14753.
12. Daimary D, Bora MB, Amitab K, Kandar D. Brain tumor segmentation from MRI images using hybrid convolutional neural networks. Proc Comput Sci. 2020;167(2019):2419–28. https://doi.org/10.1016/j.procs.2020.03.295.
13. Pei L, Bakas S, Vossough A, Reza SMS, Davatzikos C, Iftekharuddin KM. Longitudinal brain tumor segmentation prediction in MRI using feature and label fusion. Biomed Signal Process Control. 2020;55:101648. https://doi.org/10.1016/j.bspc.2019.101648.
14. Lin M, et al. Fully automated segmentation of brain tumor from multiparametric MRI using 3D context deep supervised U-Net. Med Phys. 2021;48(8):4365–74. https://doi.org/10.1002/mp.15032.
15. Punn NS, Agarwal S. Multi-modality encoded fusion with 3D inception U-net and decoder model for brain tumor segmentation. Multimed Tools Appl. 2021;80(20):30305–20. https://doi.org/10.1007/s11042-020-09271-0.
16. Ullah F, et al. Brain MR image enhancement for tumor segmentation using 3d u-net. Sensors. 2021;21(22):1–14. https://doi.org/10.3390/s21227528.

17. Brain Tumor Segmentation 2020 Dataset. [Online]. https://www.kaggle.com/awsaf49/brats20-dataset-training-validation.

18. Montaha S, Azam S, Rafid AKMRH, Hasan MZ, Karim A, Islam A. TimeDistributed-CNN-LSTM: a hybrid approach combining CNN and LSTM to classify brain tumor on 3D MRI scans performing ablation study. IEEE Access. 2022;10:60039–59. https://doi.org/10.1109/ACCESS.2022.3179577.

19. Khalil HA, Darwish S, Ibrahim YM, Hassan OF. 3D-MRI brain tumor detection model using modified version of level set segmentation based on dragonfly algorithm. Symmetry (Basel). 2020;12(8):1256. https://doi.org/10.3390/SYM12081256.

20. Gadosey PK, et al. SD-UNET: stripping down U-net for segmentation of biomedical images on platforms with low computational budgets. Diagnostics. 2020;10(2):1–18. https://doi.org/10.3390/diagnostics10020110.

21. Weng W, Zhu X. INet: convolutional networks for biomedical image segmentation. IEEE Access. 2021;9:16591–603. https://doi.org/10.1109/ACCESS.2021.3053408.

22. S. Banerjee, S. Mitra, F. Masulli, and S. Rovetta, "Deep Radiomics for Brain Tumor Detection and Classification from Multi-Sequence MRI," pp. 1–15, 2019, [Online]. http://arxiv.org/abs/1903.09240.

23. Rehman MU, Cho S, Kim JH, Chong KT. Bu-net: brain tumor segmentation using modified u-net architecture. Electron. 2020;9(12):1–12. https://doi.org/10.3390/electronics9122203.

24. Asim M, Rashid A, Ahmad T. Scour modeling using deep neural networks based on hyperparameter optimization. ICT Express. 2021;8:357–62. https://doi.org/10.1016/j.icte.2021.09.012.

25. Ding P, Li J, Wang L, Wen M, Guan Y. HYBRID-CNN: an efficient scheme for abnormal flow detection in the SDN-based smart grid. Secur Commun Netw. 2020;2020:1–20. https://doi.org/10.1155/2020/8850550.

26. Kandel I, Castelli M. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. ICT Express. 2020;6(4):312–5. https://doi.org/10.1016/j.icte.2020.04.010.

27. Khan MA, Sharif MI, Raza M, Anjum A, Saba T, Shad SA. Skin lesion segmentation and classification: a unified framework of deep neural network features fusion and selection. Expert Syst. 2022;39(7):1–21. https://doi.org/10.1111/exsy.12497.

28. Khan MA, et al. Lungs cancer classification from CT images: an integrated design of contrast based classical features fusion and selection. Pattern Recognit Lett. 2020;129:77–85. https://doi.org/10.1016/j.patrec.2019.11.014.

29. Ghosh S, Bandyopadhyay A, Sahay S, Ghosh R, Kundu I, Santosh KC. Colorectal histology tumor detection using ensemble deep neural network. Eng Appl Artif Intell. 2021;100:104202. https://doi.org/10.1016/j.engappai.2021.104202.

30. Montaha S, et al. MNet-10: a robust shallow convolutional neural network model performing ablation study on medical images assessing the effectiveness of applying optimal data augmentation technique. Front Med. 2022. https://doi.org/10.3389/fmed.2022.924979.

31. Naser MA, Deen MJ. Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. Comput Biol Med. 2020;121:103758. https://doi.org/10.1016/j.compbiomed.2020.103758.