MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL, INDIA, 462003

BRAIN TUMOR SEGMENTATION USING DEEP LEARNING APPROACH

Minor Project Report
VIthSemester

Submitted by:

Archit Badola	201112278
Shubham Singh	201112251
Saurav Pawar	201112253
Prabhudeba Nath	201112203

Under the Guidance of
Dr. Vaibhav Soni

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Session: 2022-2023

MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL, INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the project entitled "**BRAIN TUMOR SEGMENTATION USING DEEP LEARNING APPROACH"** submitted by:

Archit Badola	201112278
Shubham Singh	201112251
Saurav Pawar	201112253
Prabhudeba Nath	201112203

is the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering is authentic work carried out by them under my supervision and guidance.

Dr. Vaibhav Soni
(Minor Project Supervisor)

DECLARATION

We, hereby declare that the following report which is being presented in the Minor Project entitled as "**BRAIN TUMOR SEGMENTATION USING DEEP LEARNING APPROACH**" is an authentic documentation of our own original work to the best of our knowledge. The following project and its report, in part or whole, has not been presented or submitted by us for any purpose in any other institute or organization. Any contribution made to the research by others, with whom we have worked at Maulana Azad National Institute of Technology, Bhopal or elsewhere, is explicitly acknowledged in the report.

Archit Badola	201112278
Shubham Singh	201112251
Saurav Pawar	201112253
Prabhudeba Nath	201112203

ACKNOWLEDGEMENT

With due respect, we express our deep sense of gratitude to our respected guide and coordinator Dr. Vaibhav Soni, for his valuable help and guidance. We are thankful for the encouragement that he has given us in completing this project successfully.

It is imperative for us to mention the fact that the report of a minor project could not have been accomplished without the periodic suggestions and advice of our project guide Dr. Vaibhav Soni and project coordinators Dr. Sanyam Shukla and Dr. Namita Tiwari.

We are also grateful to our respected HOD, Dr. Deepak Singh Tomar and HOD, AI, Dr. Nilay Khare and Chairperson, Central Computing Facility, Dr. Meenu Chawla for permitting us to utilize all the necessary facilities of the college.

We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not the least; we would like to express our deep appreciation towards our family members and batch mates for providing support and encouragement.

ABSTRACT

A brain tumor, known as an intracranial tumor, is an abnormal mass of tissue in which cells grow and multiply uncontrollably, seemingly unchecked by the mechanisms that control normal cells. Imaging tests are generally prescribed by doctors for detecting tumors. Imaging tests show pictures of the inside of the body. Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs - and take actions or make recommendations based on that information. CT scan, MRI, PET are commonly used imaging methods among which MRI is considered the reliable source. Magnetic resonance imaging (MRI) is used to measure the tumor's size as it detects very minute objects. The aim of our project is to segment brain tumors using machine learning.

Preprocessing is performed on an MRI image in order to remove the impulsive noises and resize the image for improving the image quality. We divided our dataset into two parts: training and testing. The outcome of the project is examined using various performance metrics: accuracy and dice score. It shows the efficiency of deep learning techniques for the detection of brain cancer from the MRI images of the brain.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

1.1 BRAIN TUMOR AND DIAGNOSIS METHODS

The diagnosis and treatment of brain tumors have always been critical challenges in the field of medical imaging. Brain tumors are abnormal growths of cells in the brain that can be either benign or malignant. They can arise from different types of brain tissue, such as the brain parenchyma, meninges, or pituitary gland. Brain tumors can cause various symptoms depending on their location and size, such as headaches, seizures, changes in vision, speech difficulties, or motor deficits.

The diagnosis of brain tumors typically involves medical imaging techniques, such as magnetic resonance imaging (MRI) and computed tomography (CT) scans, which provide detailed images of the brain structures. These imaging modalities allow for the visualization of tumor size, location, shape, and relationship with surrounding brain tissue.However, the interpretation of these images can be complex and time-consuming, requiring the expertise of trained radiologists. This has led to an increasing interest in developing automated methods for brain tumor detection and segmentation using machine learning techniques.

The process of brain tumor diagnosis using medical imaging typically involves several stages of image preprocessing. The first stage involves the acquisition of raw image data, which may be in the form of MRI or CT scans. The raw images are then subjected to various preprocessing steps to enhance image quality, remove noise, and correct artifacts. The preprocessing steps are crucial to ensure that the images used for tumor segmentation are of high quality and free from artifacts that could affect the accuracy of the segmentation results.

After image preprocessing, the next stage involves the identification and segmentation of brain tumors from the preprocessed images. This is typically done using image analysis techniques, including image thresholding, region growing, edge detection, or machine learning-based approaches. Image thresholding involves setting a threshold value to separate tumor regions from healthy brain tissue based on intensity values. Region growing involves starting from a seed point and growing regions based on intensity or other image features. Edge detection involves identifying boundaries between tumor and healthy tissue based on intensity gradients. Machine learning-based approaches involve training algorithms on labeled data to learn patterns and features that can distinguish tumor regions from healthy brain tissue. These approaches can provide automated and objective tumor segmentation results, which can aid in the diagnosis and treatment planning for brain tumors.

In conclusion, brain tumor diagnosis relies heavily on medical imaging techniques, such as MRI and CT scans, which provide detailed images for tumor visualization. The process of brain tumor diagnosis involves several stages of image preprocessing, including image acquisition, preprocessing, and tumor segmentation. Preprocessing steps are critical to ensure high-quality images for accurate tumor segmentation. Image analysis techniques, including image thresholding, region growing, edge detection, and machine learning-based approaches, are used for tumor segmentation from preprocessed images. These methods contribute to the accurate identification and delineation of brain tumor regions from healthy brain tissue, supporting clinical decision-making in brain tumor diagnosis and treatment planning.

1.2 MOTIVATION FOR THE WORK

The primary motivation for brain tumor segmentation is the development of a computer-based approach capable of efficiently and accurately identifying brain tumors from MRI scans. The study intends to employ two distinct methods to determine the optimal approach. The objective is to create a dependable and precise method that can aid doctors in diagnosing and treating brain tumors with greater effectiveness. The project aims to contribute to the development of more efficient and accurate methods for automated brain tumor segmentation in medical imaging.

1.3 PROBLEM STATEMENT

The problem addressed in this project is the accurate segmentation of brain tumors from magnetic resonance imaging (MRI) scans using deep learning techniques. Brain tumor segmentation is a critical task in medical image analysis, as it helps in the precise diagnosis and treatment planning of brain tumors. However, the manual segmentation of brain tumors is a time-consuming and error-prone process, which necessitates the need for automated segmentation methods. Deep learning-based methods, such as the UNet and ResUNet architectures, have shown promising results in accurately segmenting brain tumors from MRI scans. In this project, we aim to compare the performance of two deep learning architectures, UNet and ResUNet, in brain tumor segmentation tasks using two different datasets, BraTS 2020 and LGG-MRI.

1.4 SCOPE

The scope of this project is to develop and compare two different deep learning models for brain tumor segmentation from MRI scans. The first model will use the UNet architecture and the BraTS 2020 dataset, while the second model will use the ResUNet architecture and the LGG-MRI dataset. The project involves pre-processing the data, training the models, and evaluating their performance. The project is limited by the availability of datasets and computational resources. The BraTS 2020 and LGG-MRI datasets were selected for this study, and the performance of the models is limited to the performance of the datasets used. Additionally, the project is limited by the availability of computational resources, which may impact the training and evaluation of the models.

2. LITERATURE SURVEY

The research proposes a deep learning architecture Attention Res-Unet with Guided Decoder (ARU-GD) for the semantic segmentation of brain tumors. The architecture had a new guided decoder, and inculcated attention gates. These changes in the base network Res-UNet improved the learning process by generating superior feature maps at the decoder and allowed only activations from relevant regions at the encoder side, and they worked together to increase the performance of segmentation. The proposed ARU-GD network is trained and tested on the BRATS 2019 dataset. [1] [8]

Different from recent fully convolutional network (FCN)-based methods, which directly perform segmentation on whole input volume, the research introduces a cascaded network including two U-Nets. The 2D U-Net aims to roughly determine the location (i.e., ROI) of brain tumor and the 3D U-Net aims to perform fine segmentation on the ROI. [2] [6]

The major aim of the paper was to establish a comparative study on the performance of the base or the original U-Net model by tweaking various components such as activation function, filter size, pooling strategies, use of batch normalization and dropout and addition of extra layer to up sampling and downsampling path. The experimentation results show that models with ReLU activation function have performed better considering results of both the dataset. Again, a model with average pooling with ReLU activation function has shown the best result when compared with the max pooling. When the kernel size was increased to 5×5 and 7×7, there was an increase in execution time. When an extra block was added to the up sampling and downsampling path with 32 as the least number of feature channels, a slight increase in performance was observed. Finally, we found that batch normalization and dropout implemented individually or in combination do not enhance the performance of the U-Net model in this case. [3] [9]

The research paper introduces an improved U-Net architecture for brain tumor segmentation. It performs a k-fold cross validation approach on the TCGA-LGG dataset using an improved U-Net architecture. The paper proposes modification (a) utilizes dense-convolutional blocks, instead of convolutional blocks during down-sampling, which enhances feature re-usability, (b) VGG16 pretrained layers for encoder path, and (c) Batch Normalization (BN) layers in dense blocks to improving model stability and performance. [4] [7]

In this study, two-step pre-processing was used to discard the black slices of the MRI image. Then, the proposed models were used to segment the brain tumor image. These models were obtained by adding two-pathway-residual (TPR) blocks to the UNet structure. Two-pathway-residual blocks exploit both local features as well as more global features, simultaneously. By the presence of TPR blocks in the UNet structure, not only the evaluation criteria such as DSC and sensitivity have been improved, but also the number of parameters of the proposed models has been reduced. [5] [10]

3. RESEARCH GAPS IDENTIFIED

Limited dataset: The use of only two datasets: BraTS 2020 and LGG-MRI may result in limitations in terms of diversity and representativeness of the brain tumor cases, potentially affecting the generalizability of the models.

Comparison with other methods: The lack of a comprehensive comparison with other existing methods for brain tumor segmentation using deep learning may limit the understanding of the relative performance and advantages of your proposed UNet and ResUNet architectures.

Evaluation metrics: The choice of evaluation metrics used to assess the performance of the models may be limited or not fully justified, potentially affecting the reliability and interpretability of the results obtained.

Clinical relevance: The potential clinical relevance and real-world implications of the proposed models, including their potential benefits and challenges in a clinical setting, may not be fully addressed, limiting the overall practical significance of the project.

By addressing these gaps in the project report, a more comprehensive and robust analysis of the brain tumor segmentation using deep learning models can be achieved, providing a more well-rounded understanding of the strengths and limitations of the proposed approach.

4. PROPOSED WORK AND METHODOLOGY

4.1 DATASET

Model 1: UNet

The BraTS 2020 dataset is a medical imaging dataset consisting of brain tumor magnetic resonance imaging (MRI) scans. The BraTS 2020 dataset contains MRI scans from patients with brain tumors. The dataset includes four MRI modalities: T1-weighted, T1-weighted contrast-enhanced, T2-weighted, and Fluid Attenuated Inversion Recovery (FLAIR). The images are provided in three-dimensional NIfTI format.

The dataset consists of two parts: a training set and a validation set. The training set contains MRI scans from 369 patients, while the validation set contains MRI scans from 125 patients. Each MRI scan in the dataset has a corresponding manual segmentation of the tumor regions performed by medical experts.

The dataset is used to train and machine learning models for brain tumor segmentation and classification, and to advance the state-of-the-art in the field of medical image analysis.

Model 2: ResUNet

The LGG-MRI Segmentation dataset is a publicly available dataset of magnetic resonance imaging (MRI) scans of brain tumors, specifically lower grade gliomas (LGGs). The dataset contains preoperative MRI scans of 110 patients with LGGs, along with segmentation masks for the tumors. The images were acquired using three different MRI protocols: T1-weighted, T1-weighted with contrast enhancement, and T2-weighted.

4.2 IMAGE PREPROCESSING

Images come in different shapes and sizes. They also come through different sources. Taking all these variations into consideration, we need to perform some pre-processing on any image data. RGB is the most popular encoding format, and most "natural images" we encounter are in RGB. Also, among the steps of data pre-processing is to make the images of the same size. We used cropping techniques to achieve this. Preprocessing is required to remove impulsive noises and resize the image for improving the image quality. We need to convert the MRI image into a grayscale image. So, after removal of noises and scaling we can increase the classification accuracy rate. Thresholding is used where we take all the pixels whose intensities are above a certain threshold and convert them to ones; the pixels having value less than the threshold are converted to zero. This results in a binary image. Normalization is the most crucial step in the pre-processing part. This refers to rescaling the pixel values so that they lie within a confined range. So preprocessing helps in overall enhancement of image and increases the classification accuracy rate.

Under this step, the input data was preprocessed to ensure it is ready for training the 3D UNet model. This included standardizing the image sizes and cropping the images. As the format of the images in the dataset is .NIfTI and there is no image loader for this format, we converted the images to .npy format.

4.3 DISPLAYING THE SAMPLE DATASET

The image that is acquired is completely unprocessed. Here, we process the image using the file path from the local device. Finally, all the images of the dataset are displayed.

Model 1: BraTS 2020 Dataset

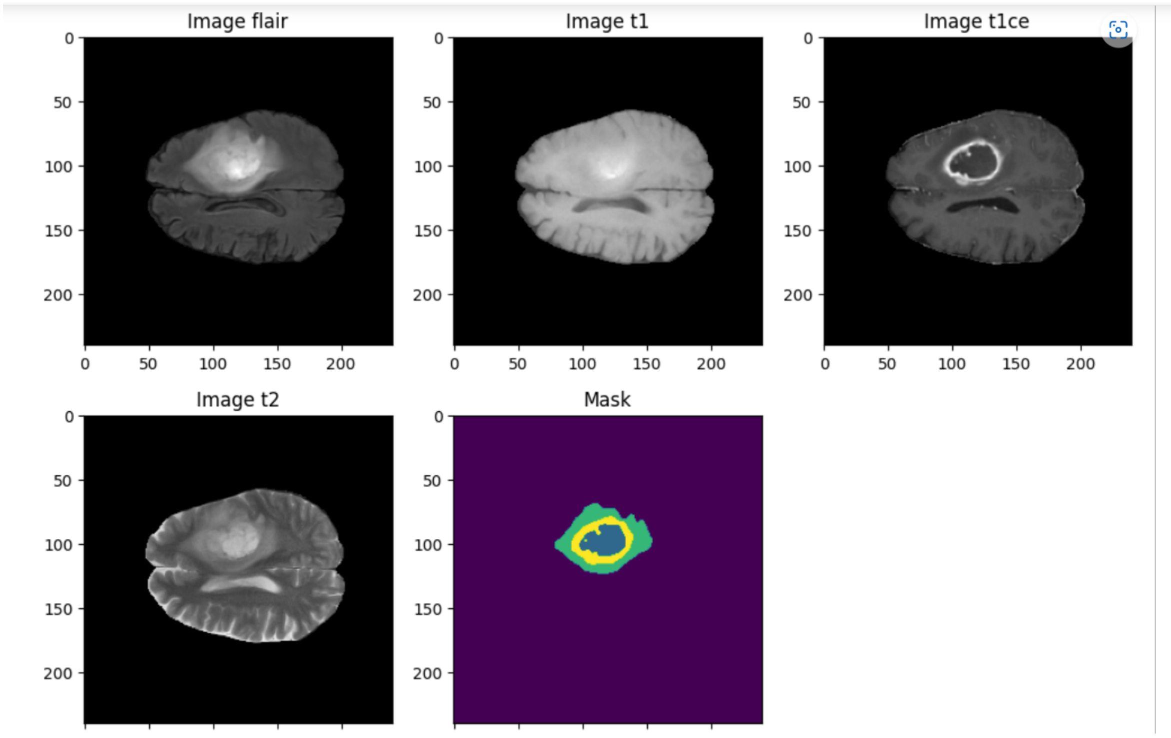

Fig 4.3.1 MRI images from BraTS 2020 dataset

Model 2: LGG-MRI Dataset

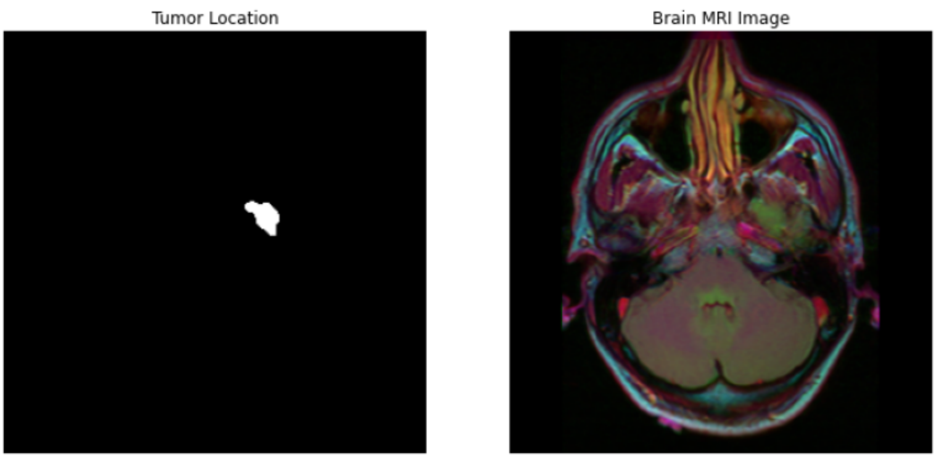

Fig 4.3.2 MRI images from LGG-MRI dataset

4.4 SPLITTING THE DATASET

To train the model the dataset is split into two parts which are training and testing dataset.

Model 1: UNet

The BraTS 2020 dataset training data was splitted into 2 parts namely, training and testing, training contains 258 instances and testing contains 86 instances.

Model 2: ResUNet

The LGG-MRI dataset consisted of 1373 images. They were splitted into 3 parts, training , testing and validation, where training contains 1167 images while testing and validation contains 103 images each.

4.5 BUILDING THE MODEL

Model 1: UNet

The architecture of the 3D UNet model consists of an encoder and a decoder, similar to the 2D UNet model. However, the 3D UNet model operates on 3D volumes instead of 2D images. The encoder is composed of multiple layers of 3D convolutional and max pooling operations that gradually reduce the spatial dimensions of the input volume while increasing the number of feature maps. The decoder is made up of upsampling and 3D convolutional layers that continuously reduce the number of feature maps while increasing the spatial dimensions of the feature maps.

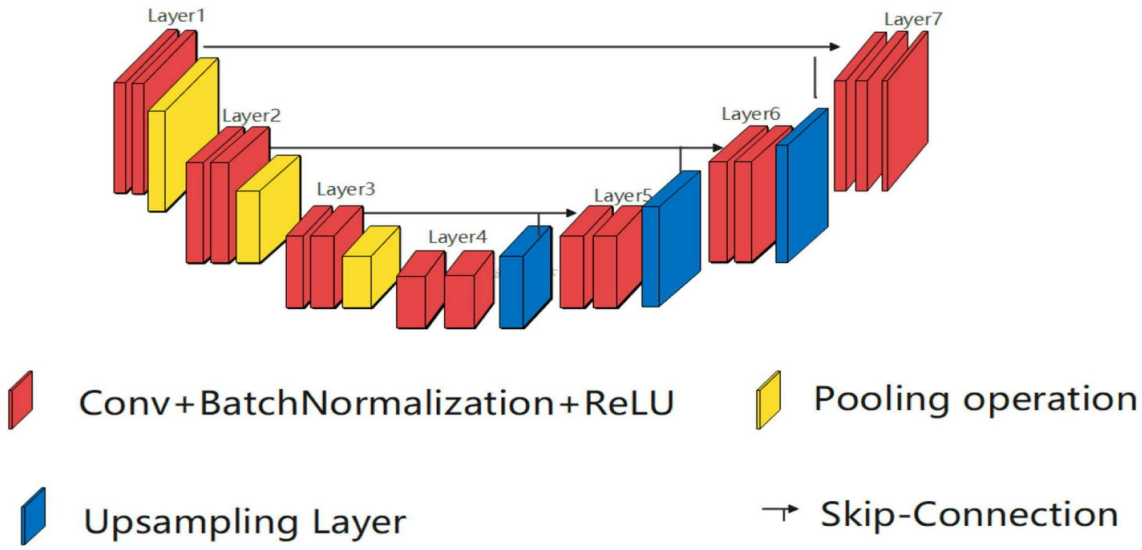

Fig 4.5.1 UNet architecture

Here is a high-level overview of the architecture of the 3D UNet model:

Input layer: The input to the model is a 3D volume representing the medical image to be segmented.

Encoding layers: The encoding layers consist of multiple blocks of 3D convolutional and max pooling operations that reduce the spatial dimensions of the input volume while increasing the number of feature maps. Each block typically consists of two 3D convolutional layers followed by a 3D max pooling layer.

Decoding layers: The decoding layers consist of multiple blocks of upsampling and 3D convolutional layers that gradually increase the spatial dimensions of the feature maps while reducing the number of feature maps. Each block typically consists of an upsampling layer, a concatenation layer that combines the feature maps from the corresponding encoding layer, and two 3D convolutional layers.

Output layer: The output of the model is a 3D volume representing the segmentation mask of the input volume.

Model 2: ResUNet

ResUNet is a modified version of the UNet architecture that incorporates residual connections which are skip connections that allow the model to bypass certain layers and pass the input data directly to deeper layers. This helps to mitigate the vanishing gradient problem and allows the model to capture both low-level and high-level features effectively. The model consists of an encoder branch and a decoder branch connected by skip connections. The skip connections allow for the fusion of features from different resolutions, enhancing the model's ability to capture fine details.

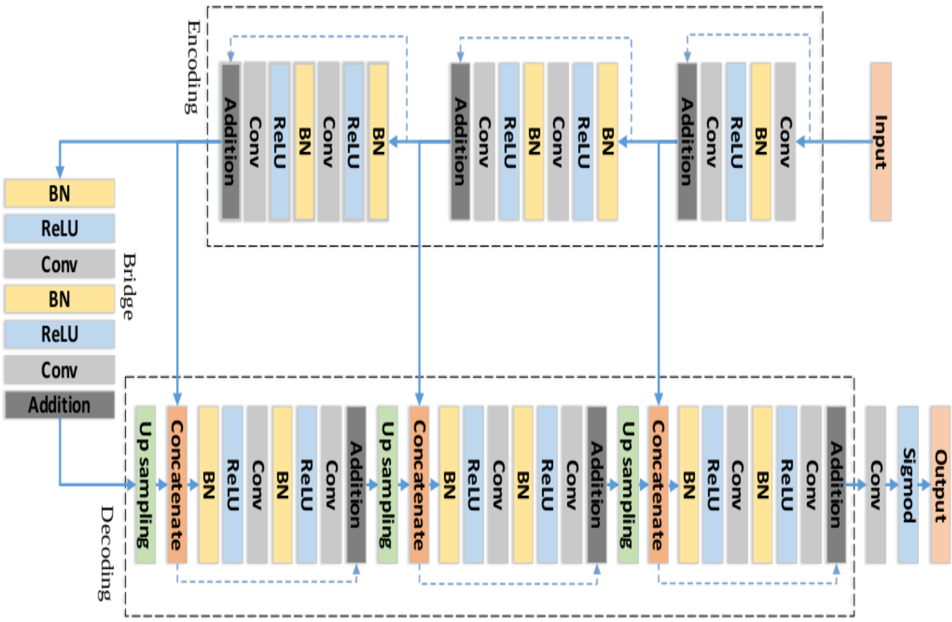

Fig 4.5.2 ResUNet architecture

Here is a high-level overview of the architecture of the ResUNet model:

Encoder Branch:

The encoder branch comprises 4 convolutional layers, each followed by a max pooling layer, to extract features from the input image.The encoder branch is responsible for extracting features from the input image. The activation function used in the convolutional layers is ReLU (Rectified Linear Unit), which introduces non-linearity into the model and helps in capturing complex patterns.

Decoder Branch:

The decoder branch consists of 4 upsampling (transposed convolutional) layers and 4 convolutional layers, in reverse order compared to the encoder branch, to upsample the feature maps and reconstruct the segmentation output. The decoder branch is responsible for reconstructing the segmentation output from the features extracted by the encoder branch. The skip connections, which connect corresponding layers of the encoder and decoder branches, allow for the fusion of features from different resolutions, enabling the model to capture fine details and enhance segmentation accuracy.

4.6 TRAINING THE MODEL

Model 1: UNet

The 3D UNet model is trained using:

Optimizer: Adam optimizer

Learning rate: A learning rate of 0.0001 was used

Batch size: A batch size of 16 was used during training. A smaller batch size allows for more frequent updates of the model weights but may increase the training time, while a larger batch size reduces the frequency of updates but may cause less variability in the weight updates. The batch size of 16 was chosen based on experimentation and available computational resources.

Number of epochs: The UNet model was trained for 75 epochs

Throughout the training process, the model's performance was monitored using validation data, which was kept separate from the training data. The model's accuracy, loss, and dice coefficient were tracked to evaluate its performance and make necessary adjustments to the hyperparameters or model architecture, if required.

Model 2: ResUNet

The ResUNet model is trained using:

Optimizer: Adam optimizer

Learning rate: A learning rate of 0.05 was used

Batch size: A batch size of 32 was used during training

Number of epochs: The ResUNet model was trained for 75 epochs

Throughout the training process, the model's performance was monitored using validation data, which was kept separate from the training data. The model's accuracy, loss, and dice coefficient were tracked to evaluate its performance and make necessary adjustments to the hyperparameters or model architecture, if required.

4.7 EVALUATING THE MODEL

Model 1: UNet

The trained UNet model was evaluated on separate images to generate the predicted masks and then compare them with the label of the images. The model's performance on the test images was analyzed in terms of accuracy and dice score.

Accuracy: 0.9522

Dice Score: 0.82

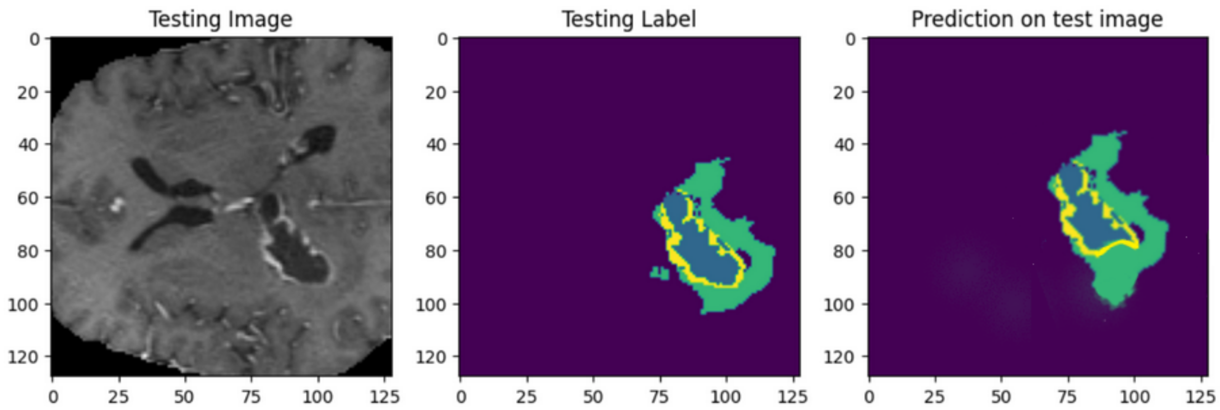

Fig 4.7.1 Mask generation using UNet model

Model 2: ResUNet

The trained ResUNet model was evaluated on separate images to generate the predicted masks and then compare them with the label of the images. The model's performance on the test images was analyzed in terms of accuracy and dice score.

Accuracy: 0.9939

Dice Score: 0.86

Fig 4.7.2 Mask generation using ResUNet model

5. RESULTS AND DISCUSSION

In Image Segmentation, accuracy may not be the most suitable metric to evaluate the performance of the model. This is because accuracy measures the percentage of correctly classified pixels, which may not be an adequate representation of the quality of the segmentation, especially if the foreground and background classes are imbalanced.

Dice Score is a common evaluation metric for image segmentation tasks. It is based on the overlap between the predicted segmentation mask and the ground truth mask. The Dice Score measures the similarity between the two masks, with a value of 1 indicating a perfect match between the two masks.

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)}$$

where,

TP stands for True Positive
FP stands for False Positive
FN stands for False Negative

The Dice Score takes into account both true positives and false positives, which is particularly useful in imbalanced datasets where there are many more background pixels than foreground pixels. In such cases, a high accuracy may be achieved by simply predicting all pixels as background, but the Dice score will penalize such a prediction as it fails to capture the foreground objects.

Therefore, Dice Score is a more appropriate evaluation metric for image segmentation as it provides a more accurate measure of the model's performance.

Model 1: UNet

Dice Score: 0.82

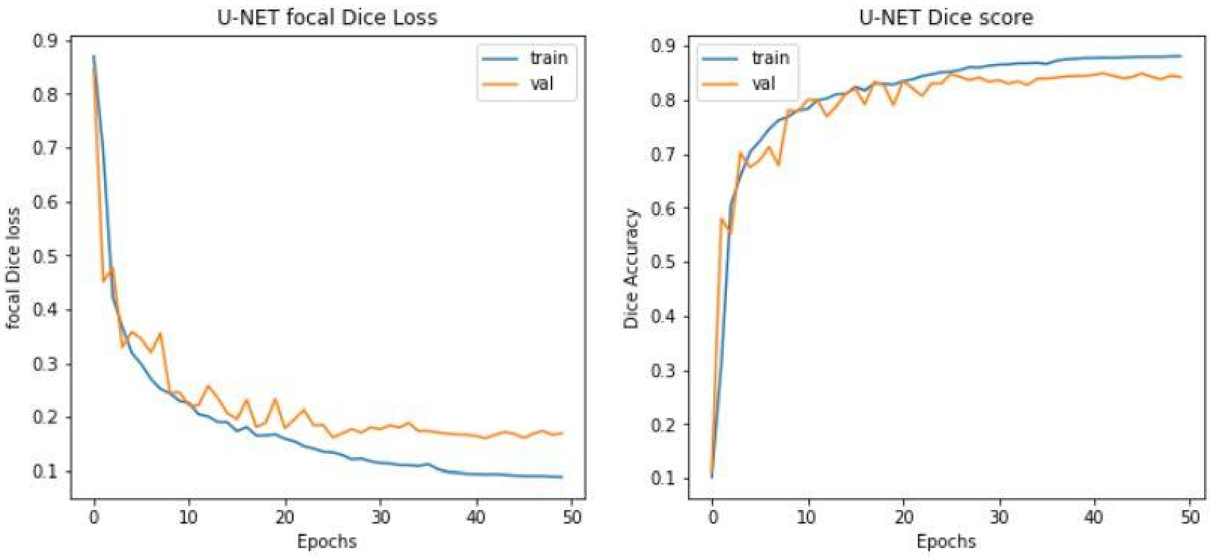

Fig 5.1 Focal Dice Loss and Dice Score of UNet

Model 2: ResUNet

Dice Score: 0.86

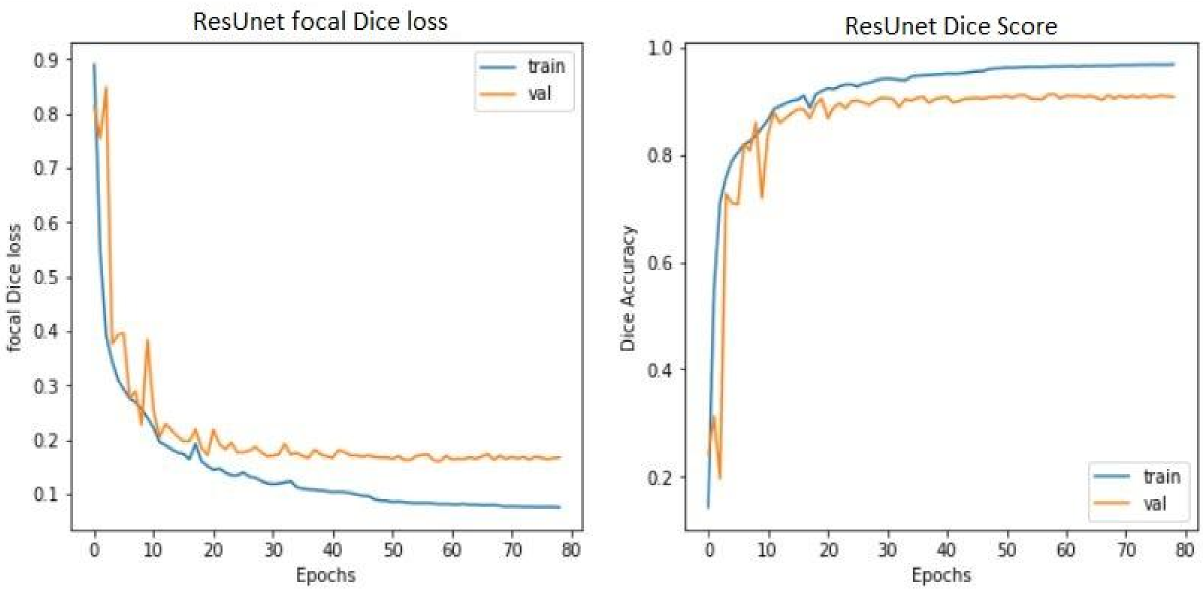

Fig 5.2 Focal Dice Loss and Dice Score of ResUNet

6. CONCLUSION AND FUTURE SCOPE

6.1 CONCLUSION

In this project we developed two models for brain tumor segmentation: one using UNet and another using ResUNet. The BraTS 2020 dataset was used for Model 1, while the LGG-MRI Segmentation dataset was used for Model 2. The models were implemented and achieved promising results with dice scores of 0.82 and 0.86, respectively. The obtained Dice scores of 0.82 and 0.86 respectively, indicate the potential of these models to assist medical professionals in accurate tumor detection and diagnosis.

Moreover, the results of this study can be used as a baseline for future research in this field. The implementation of deep learning models can significantly improve the accuracy of tumor segmentation, thus enabling doctors to provide better treatment and care for patients suffering from brain tumors.

In addition, the development of deep learning models for brain tumor segmentation can help in the creation of computer-assisted diagnosis (CAD) systems. These systems can analyze medical images, and provide suggestions and insights to doctors for making informed decisions.

Overall, this study provides a promising framework for the development of deep learning models for brain tumor segmentation, and we hope that our work can inspire more research in this important field.

6.2 FUTURE SCOPE

Additionally, the interpretability of deep learning models remains a challenge, and future studies should focus on developing models that can provide explanations for their predictions.

There are several areas for improvement and future work that can be explored in this project, including:

- ➢ Exploring the potential of incorporating attention mechanisms to improve the segmentation accuracy.
- ➢ Evaluating the models on a larger dataset with a greater variety of tumor types to ensure the generalization of the models.
- ➢ Investigating the potential of using multi-modal imaging data such as MRI, CT and PET scans to improve the accuracy of segmentation.

7. REFERENCES

[1] Maji, D., Sigedar, P., & Singh, M. (2022). Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors. *Biomedical Signal Processing and Control*, *71*, 103077.

[2] Li, S., Liu, J., & Song, Z. (2022). Brain tumor segmentation based on region of interest-aided localization and segmentation U-Net. *International Journal of Machine Learning and Cybernetics*, *13*(9), 2435-2445.

[3] Das, S., Swain, M. K., Nayak, G. K., Saxena, S., & Satpathy, S. C. (2022). Effect of learning parameters on the performance of U-Net Model in segmentation of Brain tumor. *Multimedia Tools and Applications*, *81*(24), 34717-34735.

[4] Ghosh, S., Chaki, A., & Santosh, K. C. (2021). Improved U-Net architecture with VGG-16 for brain tumor segmentation. *Physical and Engineering Sciences in Medicine*, *44*(3), 703-712.

[5] Aghalari, M., Aghagolzadeh, A., & Ezoji, M. (2021). Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks. *Biomedical Signal Processing and Control*, *69*, 102841.

[6] Al Nasim, M.A., Al Munem, A., Islam, M., Palash, M.A.H., Haque, M.M.A. and Shah, F.M., 2022, December. Brain Tumor Segmentation using Enhanced U-Net Model with Empirical Analysis. In *2022 25th International Conference on Computer and Information Technology (ICCIT)* (pp. 1027-1032). IEEE.

[7] Fidon, L., Ourselin, S. and Vercauteren, T., 2021. Generalized wasserstein dice score, distributionally robust deep learning, and ranger for brain tumor segmentation: BraTS 2020 challenge. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6* (pp. 200-214). Springer International Publishing.

[8] Kumar, P.S., Sakthivel, V.P., Raju, M. and Satya, P.D., 2023. Brain tumor segmentation of the FLAIR MRI images using novel ResUnet. *Biomedical Signal Processing and Control*, *82*, p.104586.

[9] Yang, T., Song, J., Li, L. and Tang, Q., 2020. Improving brain tumor segmentation on MRI based on the deep U-net and residual units. *Journal of X-ray Science and Technology*, *28*(1), pp.95-110.

[10] Arora, A., Jayal, A., Gupta, M., Mittal, P. and Satapathy, S.C., 2021. Brain tumor segmentation of mri images using processed image driven u-net architecture. *Computers*, *10*(11), p.139.