



Multimodal MRI Brain Tumor Segmentation Using a 3D LinkNet–ResNet152 Architecture

Felicia Rajathi P¹., Dr.D.Gladis²

¹ Department of Computer Science, Madras Christian College (Autonomous), Chennai, Tamil Nadu, India

² Bharathi Women's College (Autonomous), Chennai – 600108, Tamil Nadu, India

Abstract

Accurate segmentation of brain tumor regions from multimodal magnetic resonance imaging (MRI) is an important step in computer-assisted brain tumor analysis. However, the heterogeneous appearance and complex spatial structure of tumor regions make automated three-dimensional segmentation challenging. This study presents a multimodal three-dimensional LinkNet architecture with a ResNet152 encoder for semantic segmentation of brain tumor MRI volumes. The proposed framework utilizes multimodal MRI data from the BraTS 2020 dataset, comprising FLAIR, post-contrast T1-weighted, and T2-weighted modalities. The input volumes are intensity-normalized, spatially cropped, and combined to form a three-channel representation with a standardized volume size of $128 \times 128 \times 128$ voxels. A LinkNet encoder-decoder architecture is employed, with ResNet152 serving as the feature extraction backbone and ImageNet-pretrained encoder weights used for initialization. The network performs four-class voxel-level segmentation corresponding to background, necrotic/non-enhancing tumor, peritumoral edema, and enhancing tumor. The trained model is evaluated using accuracy, precision, recall, F1-score, confusion matrix, and receiver operating characteristic analysis. The implementation also identifies tumor-containing slices and extracts region-based quantitative characteristics from the segmentation output. In addition, a Django-based interface is developed to facilitate MRI input, segmentation visualization, slice selection, and generation of a PDF-based analysis report. The results demonstrate the applicability of the proposed LinkNet–ResNet152 framework for automated multimodal brain tumor segmentation and subsequent tumor-region analysis, providing an integrated workflow from MRI preprocessing to segmentation visualization and reporting.

Keywords: Brain tumor segmentation; Magnetic resonance imaging; Multimodal MRI; LinkNet; ResNet152; BraTS 2020; Semantic segmentation; Deep learning; 3D segmentation.

1. Introduction

Brain tumors are among the challenging neurological disorders that require accurate and timely assessment for effective clinical management. Magnetic Resonance Imaging (MRI) is widely used for brain tumor evaluation because it provides detailed anatomical and pathological information. However, manual delineation of tumor regions from MRI is time-consuming and requires considerable clinical expertise. Recent advances in artificial intelligence and deep learning have enabled automated approaches for brain tumor segmentation, quantification, and classification, making deep learning an important research direction in MRI-based brain tumor analysis [1], [2].

Multimodal MRI provides complementary information about brain tumor regions through different imaging sequences. FLAIR, post-contrast T1-weighted (T1ce), and T2-weighted images provide complementary representations of tumor tissue, edema, and surrounding anatomical structures. The combined use of multiple MRI modalities has consequently become an important approach for automated brain tumor segmentation. Recent surveys have reported extensive use of convolutional neural networks, vision transformers, and hybrid architectures for multimodal brain tumor segmentation [2], [3].

The Brain Tumor Segmentation (BraTS) benchmark has played an important role in the development and evaluation of automated brain tumor segmentation methods. The benchmark provides multimodal magnetic resonance images together with expert annotations and has been widely used for assessing

algorithms for different tumor subregions [3]. In this study, the BraTS 2020 dataset is used as the source of multimodal MRI volumes for developing the proposed segmentation framework.

Deep convolutional neural networks (CNNs) can learn hierarchical representations from medical images and have been widely adopted for semantic segmentation. Encoder-decoder architectures are particularly suitable for segmentation because the encoder progressively extracts high-level features, whereas the decoder reconstructs spatial information to produce a dense segmentation map. Recent reviews indicate that CNN-based encoder-decoder architectures remain important for multimodal brain tumor segmentation, while newer transformer-based and hybrid approaches are also being investigated [1], [3]. LinkNet is an encoder-decoder architecture developed for efficient semantic segmentation. It incorporates connections between corresponding encoder and decoder stages, allowing feature representations from the encoder to be transferred to the decoder during spatial reconstruction [4]. Such connections are useful for preserving spatial information while performing semantic segmentation. In this study, the LinkNet architecture is adapted to a three-dimensional multimodal MRI segmentation task.

The selection of the encoder is an important component of an encoder-decoder segmentation framework. Residual networks introduced residual learning to facilitate the training of substantially deeper neural networks. The original ResNet study demonstrated networks with depths of up to 152 layers, establishing ResNet152 as a deep feature-extraction architecture [5]. In the proposed framework, ResNet152 is employed as the encoder backbone of LinkNet to extract hierarchical representations from multimodal brain MRI volumes.

In this work, a three-dimensional LinkNet architecture with a ResNet152 encoder is developed for semantic segmentation of brain tumor MRI volumes. The framework utilizes multimodal MRI data from the BraTS 2020 dataset and represents the input as a three-channel volume containing FLAIR, post-contrast T1-weighted, and T2-weighted images. The input volumes are intensity-normalized and spatially cropped to a standardized size of $128 \times 128 \times 128$ voxels. The segmentation network produces four classes corresponding to background, necrotic/non-enhancing tumor, peritumoral edema, and enhancing tumor.

The proposed workflow is designed not only to generate voxel-level segmentation maps but also to support subsequent tumor-region analysis. Following segmentation, tumor-containing slices are identified and quantitative characteristics are extracted from the predicted regions. The complete workflow is further integrated into a Django-based interface for MRI input, segmentation visualization, slice selection, and generation of a PDF-based analysis report.

Despite the progress achieved by recent deep learning methods, multimodal brain tumor segmentation continues to involve challenges related to the heterogeneous appearance of tumor regions, differences between MRI modalities, class imbalance, and reliable segmentation of individual tumor subregions [1],[3]. These challenges motivate the investigation of three-dimensional encoder-decoder frameworks capable of utilizing complementary information from multimodal MRI volumes while preserving spatial details.

Therefore, this study investigates a three-dimensional LinkNet-ResNet152 architecture for multimodal MRI brain tumor segmentation. The proposed workflow combines MRI preprocessing, four-class voxel-level segmentation, performance evaluation, tumor-containing slice identification, region-based analysis, visualization, and automated reporting within an integrated framework. The study specifically focuses on the application of LinkNet with a deep residual encoder for three-dimensional multimodal brain tumor segmentation.

2. Related Work

Deep learning-based brain tumor segmentation has developed rapidly with the increasing availability of multimodal MRI datasets and advances in convolutional neural network architectures. Recent reviews indicate that U-Net-based encoder-decoder networks, transformer-based architectures, hybrid models, and emerging state-space models have been extensively investigated for multimodal brain tumor segmentation [6], [7]. The main objective of these approaches is to exploit complementary information from different MRI modalities while preserving the spatial characteristics of tumor boundaries and subregions.

Multimodal MRI segmentation is particularly challenging because each imaging sequence provides different tissue contrasts. T1-weighted, T1ce, T2-weighted, and FLAIR images provide complementary information for identifying tumor components and surrounding edema. Recent studies have therefore investigated modality-fusion strategies to improve feature representation. Jiang *et al.* reviewed multimodal brain tumor segmentation approaches and highlighted modality fusion, missing-modality handling, and generalization across datasets as important research challenges [6]. Similarly, recent feature-fusion approaches have attempted to learn modality-specific and spatial-channel representations to improve segmentation of heterogeneous tumor regions [8].

CNN-based encoder-decoder architectures remain widely used because they can learn hierarchical spatial features while reconstructing dense segmentation maps. U-Net and its variants have provided the foundation for many brain tumor segmentation methods, with subsequent studies introducing residual blocks, attention mechanisms, multi-scale feature extraction, and other architectural modifications [6], [7].

Recent research has also investigated transformer-based models and hybrid CNN-transformer frameworks to capture long-range contextual information. However, these approaches can introduce additional architectural and computational complexity.

The use of pretrained CNN backbones provides another approach for improving feature extraction within encoder-decoder networks. Residual architectures are particularly suitable for deep feature extraction because residual connections facilitate the training of deeper networks. ResNet152 therefore provides a high-capacity encoder for extracting hierarchical representations from multimodal MRI volumes. In the present study, ResNet152 is incorporated into LinkNet as the encoder backbone rather than developing a new feature-extraction network from the beginning.

LinkNet was originally proposed as an efficient semantic segmentation architecture based on encoder representations and encoder-decoder shortcut connections [4]. The architecture transfers information from encoder stages to corresponding decoder stages, allowing the decoder to recover spatial information while maintaining computational efficiency. Although LinkNet was originally developed for general semantic segmentation, its encoder-decoder structure provides a suitable basis for adapting the architecture to medical image segmentation.

Recent brain tumor segmentation research has increasingly focused on sophisticated attention, multi-scale fusion, transformer, and hybrid architectures. For example, recent work on multimodal MRI segmentation has introduced multi-scale contextual aggregation and attention-based fusion to improve the representation of heterogeneous tumor regions on the BraTS 2020 dataset [10]. Other recent approaches have investigated multimodal feature-fusion mechanisms and knowledge-transfer strategies to address the difficulty of combining MRI modalities and the practical problem of incomplete modalities [8]. These studies demonstrate the continuing importance of effective multimodal feature representation in automated brain tumor segmentation.

Despite these advances, there remains a need to investigate relatively structured encoder-decoder architectures that can process three-dimensional multimodal MRI volumes while maintaining a practical segmentation workflow. In particular, this study investigates the application of LinkNet with a deep ResNet152 encoder to three-dimensional multimodal brain tumor MRI. Therefore, this study focuses on a LinkNet-ResNet152 framework for four-class segmentation of multimodal MRI volumes, followed by tumor-containing slice identification, region-based analysis, visualization, and automated reporting.

The present work investigates a LinkNet encoder-decoder framework with ResNet152 for three-dimensional multimodal MRI segmentation and extends the segmentation output to tumor-containing slice identification, region-based quantitative analysis, visualization, and automated reporting.

3. Materials and Methods

3.1 Dataset

The proposed framework was developed using the BraTS 2020 (Brain Tumor Segmentation 2020) dataset, which provides multimodal magnetic resonance imaging (MRI) volumes and corresponding expert-annotated tumor segmentation masks. The dataset contains multiple MRI sequences that provide complementary information about brain tumor regions and surrounding tissues.

For the present implementation, three MRI modalities were used as input channels: Fluid-Attenuated Inversion Recovery (FLAIR), post-contrast T1-weighted (T1ce), and T2-weighted MRI. These modalities were spatially combined to construct a three-dimensional multimodal representation for segmentation.

The corresponding segmentation labels were converted into four classes representing the background and different tumor regions. The original label value 4 was remapped to label 3 during preprocessing. Consequently, the final label representation consisted of four classes: background (0), necrotic/non-enhancing tumor (1), peritumoral edema (2), and enhancing tumor (3).

3.2 Multimodal MRI Input Representation

Each subject was represented as a three-dimensional multimodal MRI volume. The resulting input volume had a size of $128 \times 128 \times 128 \times 3$ voxels.

$$X \in \mathbb{R}^{128 \times 128 \times 128 \times 3}$$

where the first three dimensions represent the spatial volume and the fourth dimension represents the three MRI modalities: FLAIR, T1ce, and T2.

The use of multimodal MRI enables the network to simultaneously learn complementary intensity and structural information from the different imaging sequences. Thus, the final input to the proposed LinkNet-ResNet152 model was a three-dimensional volume of size $128 \times 128 \times 128$ with three MRI modalities.

3.3 MRI Preprocessing

Preprocessing was performed to standardize the MRI volumes before model training. The implemented preprocessing procedure consisted of intensity normalization and spatial cropping.

First, the MRI intensities were normalized using Min-Max scaling so that the voxel intensity values were mapped to the range [0, 1]. This normalization provides a consistent intensity range across the input MRI volumes.

The volumes were subsequently spatially cropped to obtain a standardized spatial dimension of $128 \times 128 \times 128$ voxels.

The preprocessed FLAIR, T1ce, and T2 volumes were then stacked to form the final three-channel input volume.

Therefore, the final input volume supplied to the segmentation model had a size of $128 \times 128 \times 128 \times 3$.

3.4 Label Preprocessing

The segmentation masks provided with the dataset were converted into a four-class representation. The original label value 4 was remapped to label 3 to obtain consecutive class indices.

The resulting segmentation classes were defined as follows:

- Class 0: Background
- Class 1: Necrotic/non-enhancing tumor
- Class 2: Peritumoral edema
- Class 3: Enhancing tumor

This four-class representation was used as the target output for voxel-level semantic segmentation.

3.5 Training and Validation Data Preparation

The preprocessed dataset was divided into training and validation subsets using a 75:25 ratio. A fixed random seed of 42 was used during the splitting process to provide reproducibility.

The training and validation samples consisted of the preprocessed three-dimensional MRI volumes and their corresponding segmentation masks. Each input sample was represented as a three-dimensional volume of size $128 \times 128 \times 128$ with three MRI modalities.

3.6 Three-Dimensional LinkNet Architecture

The proposed segmentation framework is based on the LinkNet encoder-decoder architecture. LinkNet was adapted in this study for three-dimensional multimodal brain MRI segmentation.

The architecture consists of an encoder for hierarchical feature extraction and a decoder for spatial reconstruction. Feature information from corresponding encoder stages is transferred to the decoder through link connections. These connections enable the decoder to utilize information from earlier feature representations during reconstruction of the segmentation output.

The implemented model was configured with a ResNet152 encoder, four segmentation classes, and an input shape of $128 \times 128 \times 128 \times 3$. The network therefore receives a three-channel volumetric MRI input and generates four-class voxel-level segmentation predictions.

3.7 ResNet152 Encoder

A ResNet152 network was employed as the encoder backbone of the LinkNet architecture. ResNet152 is a deep residual network that uses residual connections to facilitate hierarchical feature learning in deep neural networks.

In the proposed framework, the ResNet152 encoder extracts hierarchical representations from the multimodal MRI volume. The extracted features are subsequently transferred to the decoder through the LinkNet feature connections.

The implementation used ImageNet-pretrained encoder weights for initialization. This provides pretrained feature representations at the beginning of the model training process. The layer-wise configuration of the ResNet152 encoder used in the proposed architecture is presented in Table 1.

Table 1. Description of the ResNet152 Encoder

Layer Name	Input Shape	Output Shape	Stride
Input Layer	(128, 128, 128, 3)	(128, 128, 128, 3)	–
Conv1	(128, 128, 128, 3)	(64, 64, 64, 64)	2
BatchNorm1	(64, 64, 64, 64)	(64, 64, 64, 64)	–
Activation1	(64, 64, 64, 64)	(64, 64, 64, 64)	–
Conv2_x	(64, 64, 64, 64)	(64, 64, 64, 64)	1
Conv3_x	(64, 64, 64, 64)	(32, 32, 32, 128)	2
BatchNorm3	(32, 32, 32, 128)	(32, 32, 32, 128)	–
Activation3	(32, 32, 32, 128)	(32, 32, 32, 128)	–

Layer Name	Input Shape	Output Shape	Stride
Conv4_x	(32, 32, 32, 128)	(16, 16, 16, 256)	2
BatchNorm4	(16, 16, 16, 256)	(16, 16, 16, 256)	–
Activation4	(16, 16, 16, 256)	(16, 16, 16, 256)	–
Conv5_x	(16, 16, 16, 256)	(8, 8, 8, 512)	2
BatchNorm5	(8, 8, 8, 512)	(8, 8, 8, 512)	–
Activation5	(8, 8, 8, 512)	(8, 8, 8, 512)	–

3.8 LinkNet Decoder and Feature Reconstruction

The decoder reconstructs the spatial representation from the encoded features to generate a dense voxel-level segmentation map. LinkNet feature connections transfer information from corresponding encoder stages to the decoder.

These connections assist in recovering spatial information that may be reduced during the encoding process. Consequently, the encoder-decoder structure combines hierarchical feature representations with spatial information required for voxel-level segmentation.

For the proposed application, the decoder generates predictions corresponding to the four segmentation classes defined during label preprocessing. The layer-wise configuration of the LinkNet decoder used for spatial reconstruction is presented in Table 2.

Table 2. Description of the LinkNet Decoder

Layer Name	Input Shape	Output Shape	Stride
UpConv8_x	(32, 32, 32, 128)	(64, 64, 64, 64)	2
BatchNorm8	(64, 64, 64, 64)	(64, 64, 64, 64)	–
Activation8	(64, 64, 64, 64)	(64, 64, 64, 64)	–
UpConv9_x	(64, 64, 64, 64)	(128, 128, 128, 32)	2
BatchNorm9	(128, 128, 128, 32)	(128, 128, 128, 32)	–
Activation9	(128, 128, 128, 32)	(128, 128, 128, 32)	–
OutputConv	(128, 128, 128, 32)	(128, 128, 128, 4)	1

3.9 Model Training

The LinkNet-ResNet152 model was trained using the prepared three-dimensional multimodal MRI volumes. The implemented training configuration used a batch size of 2 and a total of 70 training epochs. The model used a combined loss consisting of Dice Loss and Categorical Focal Loss. The model was trained using the training subset and evaluated using the validation subset during the training process. The main model attributes and training configuration are summarized in Table 3.

Table 3. Model Attributes and Training Configuration

Parameter	Configuration
Architecture	3D LinkNet
Backbone	ResNet152
Encoder Weights	ImageNet
Input Shape	$128 \times 128 \times 128 \times 3$

Parameter	Configuration
MRI Modalities	FLAIR, T1ce, T2
Number of Classes	4
Activation Function	Softmax
Batch Size	2
Training Epochs	70
Loss Function	Dice Loss + Categorical Focal Loss

3.10 Voxel-Level Segmentation

After training, the model generated a class prediction for each voxel of the input MRI volume. The predicted output represented the four segmentation categories corresponding to background, necrotic/non-enhancing tumor, peritumoral edema, and enhancing tumor.

The predicted segmentation maps were compared with the corresponding reference labels at the voxel level. This enabled quantitative evaluation of the model's classification performance.

3.11 Performance Evaluation

The implemented evaluation procedure considered multiple performance measures, including accuracy, precision, recall, F1-score, confusion matrix, and receiver operating characteristic (ROC) analysis.

The classification metrics were calculated from the predicted and reference voxel labels. The voxel predictions were flattened before generating the classification report, allowing the performance of the four classes to be examined at the voxel level.

The confusion matrix was used to examine the distribution of correctly and incorrectly classified voxels across the four classes. ROC analysis was additionally performed to investigate the class-wise discrimination capability of the model.

The reported accuracy represents voxel-level classification accuracy and is therefore not interpreted as a Dice coefficient or as a direct tumor-region overlap measure.

3.12 Tumor-Region Analysis

Following segmentation, the framework performed additional analysis of tumor-containing slices. The segmentation output was examined to identify slices containing tumor regions, allowing the relevant portions of the three-dimensional volume to be visualized and analyzed.

Region-based characteristics were subsequently extracted from selected tumor-containing slices. The implemented analysis included intensity- and mask-derived measurements based on the segmented region. These measurements provided additional quantitative information for the subsequent tumor analysis and reporting workflow.

3.13 Visualization and System Integration

The complete workflow was integrated into a Django-based application to provide an interface for MRI input, segmentation visualization, tumor-containing slice selection, and report generation.

The application supports visualization of the processed MRI data and segmentation results and provides a mechanism for generating a PDF-based analysis report. Thus, the overall system integrates preprocessing, model inference, visualization, quantitative region analysis, and reporting into a single application. The overall workflow of the proposed multimodal MRI brain tumor segmentation and analysis framework is illustrated in Figure 1.

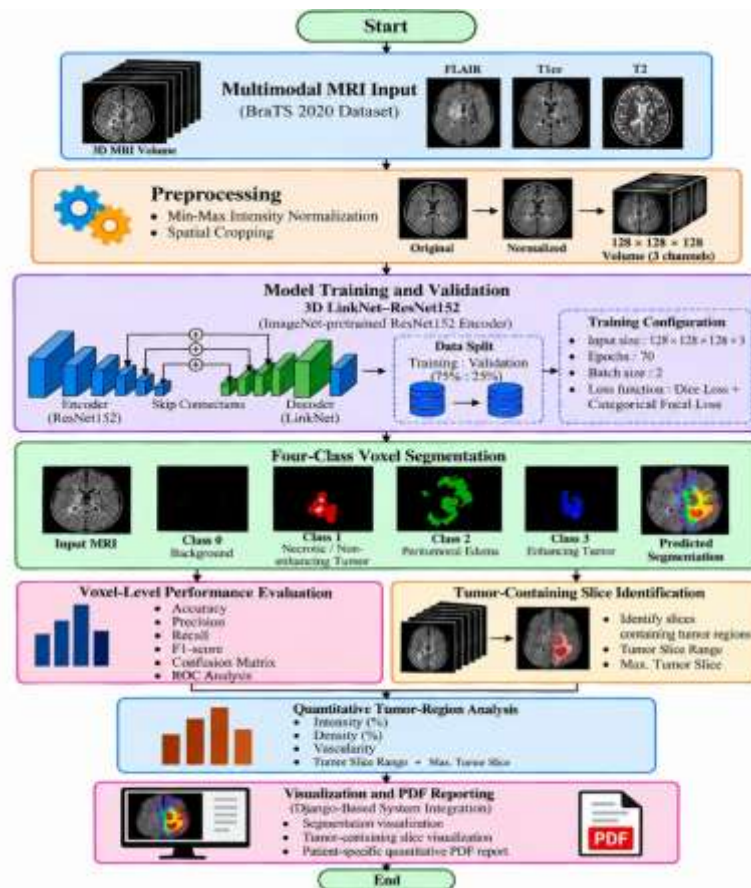


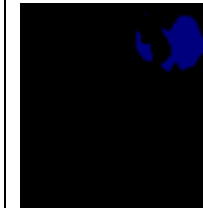
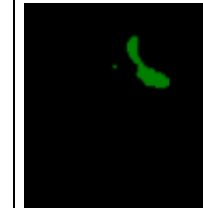
Figure 1. Overall workflow of the proposed multimodal MRI brain tumor segmentation and analysis framework.

4. Results and Discussion

4.1 Segmentation Results

The developed three-dimensional LinkNet architecture with a ResNet152 encoder was evaluated for multimodal brain tumor segmentation using the BraTS 2020 dataset. The network generated a four-class segmentation mask corresponding to background, non-enhancing tumor, peritumoral edema, and enhancing tumor. The predicted mask provided voxel-level class predictions that were subsequently used for visualization and quantitative tumor-region analysis. Representative segmentation results for six patient cases are summarized in Table 4.

Table 4. Representative Four-Class Segmentation Results and Whole-Tumor Masks for Six Selected Patients

Patient Number	Max Tumor Slice	Non-Tumor (Class 0)	Non-Enhancing Tumor (Class 1)	Peritumoral Edema (Class 2)	Enhancing Tumor (Class 3)	Whole Tumor
9	74					

12	70					
20	76					
32	87					
63	52					
72	65					

The segmentation results demonstrate that the LinkNet-ResNet152 architecture generated voxel-level tumor-region predictions from multimodal MRI volumes. The predicted masks were further processed to identify slices containing tumor regions, enabling patient-level analysis of the segmented regions.

4.2 Voxel-Level Classification Performance

The quantitative performance of the segmentation model was evaluated using a voxel-level classification report. The evaluation included accuracy, precision, recall, and F1-score for the four segmentation classes. The model achieved an overall voxel-level classification accuracy of **97.56% (0.9756)**. The classification report indicates comparatively strong performance for the background class, whereas the non-enhancing tumor and enhancing tumor classes showed lower class-wise performance. The macro-averaged F1-score was **0.49**, while the weighted-average F1-score was **0.97**.

The difference between the macro and weighted F1-scores is important because the voxel distribution is highly imbalanced, with background voxels constituting a large proportion of the complete MRI volume. Consequently, the high overall accuracy and weighted F1-score should be interpreted together with the lower macro F1-score rather than being considered evidence of uniformly high performance across all tumor classes. The voxel-level classification report is shown in Figure 2. The overall voxel-level classification performance is summarized in Table 5.

```

                precision    recall  f1-score   support

   class0       0.98         1.00         0.99    32399030
   class1       0.60         0.17         0.27     154771
   class2       0.62         0.48         0.54     790678
   class3       0.34         0.11         0.17     209953

 accuracy              0.98    33554432
 macro avg           0.63    0.44    0.49    33554432
 weighted avg        0.97    0.98    0.97    33554432

Accuracy score: 0.9756289720535278
    
```

Figure 2. Voxel-level classification report for the four segmentation classes.

Table 5. Overall voxel-level classification performance

Metric	Value
Accuracy	0.9756
Accuracy (%)	97.56
Macro F1-score	0.49
Weighted F1-score	0.97

It should be noted that the reported **97.56% value represents voxel-level classification accuracy and is not a Dice coefficient**. Therefore, the result is reported as classification accuracy rather than segmentation Dice.

4.3 Confusion Matrix Analysis

A confusion matrix was generated to examine the distribution of correct and incorrect predictions among the four segmentation classes. The matrix provides class-wise information regarding the correspondence between the ground-truth labels and predicted voxel classes.

The classification results indicate that the background class contributes substantially to the high overall accuracy, whereas the tumor-specific classes are more challenging to classify. In particular, the relatively lower performance associated with classes 1 and 3 indicates lower classification performance for these tumor subregions than for the background class.

This observation is relevant to multimodal brain tumor segmentation because tumor subregions can exhibit heterogeneous appearance and occupy substantially fewer voxels than background tissue. Thus, class-specific evaluation is necessary in addition to overall accuracy. The corresponding confusion matrix is shown in Figure 3.

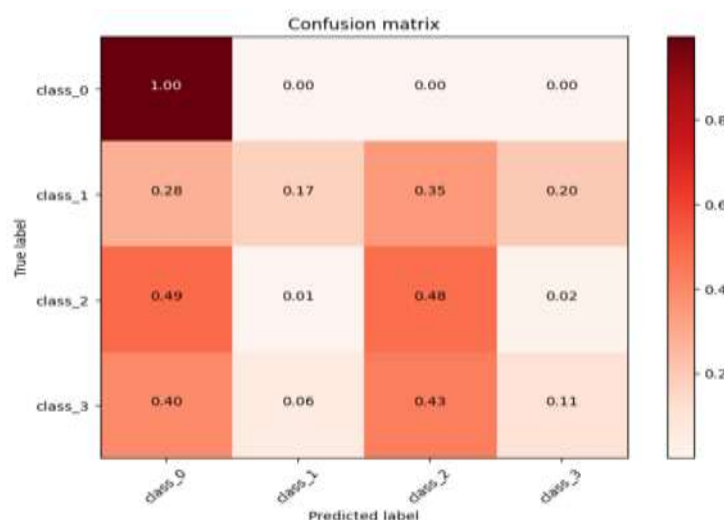


Figure 3. Confusion matrix for voxel-level classification of the segmented tumor regions.

4.4 ROC-AUC Analysis

Receiver operating characteristic (ROC) analysis was additionally performed to assess the discriminative behaviour of the four predicted classes. The corresponding ROC curves are shown in Figure 4. The obtained class-wise ROC-AUC values are presented in Table 6.

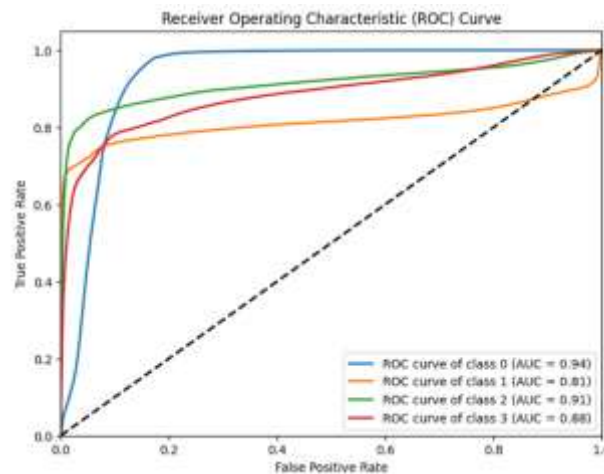


Figure 4. Receiver Operating Characteristic (ROC) curves for the four segmentation classes with corresponding AUC values.

Table 6. Class-wise ROC-AUC performance

Class	ROC-AUC
Class 0	0.9367
Class 1	0.8143
Class 2	0.9131
Class 3	0.8838

The ROC-AUC values indicate class-dependent discriminative performance. Class 0 obtained an AUC of **0.9367**, while class 1 produced the lowest value of **0.8143**. Classes 2 and 3 achieved AUC values of **0.9131** and **0.8838**, respectively.

The variation in AUC is consistent with the classification results, where tumor-specific classes are more challenging than the background class. Therefore, ROC analysis provides additional information about class discrimination beyond overall voxel-level accuracy.

4.5 Tumor-Containing Slice Identification

Following voxel-level segmentation, the predicted masks were processed to identify MRI slices containing tumor regions. This step enabled the three-dimensional segmentation output to be examined at the individual-slice level and provided a basis for subsequent patient-specific quantitative analysis.

For each selected patient, the system identified a range of slices containing predicted tumor regions and determined the slice with the maximum tumor-region representation. The maximum tumor slice was subsequently used as a representative slice for region-based analysis.

The identified tumor-containing slice range is illustrated in Figure 5, while representative tumor-containing MRI slices are shown in Figure 6.

```
Tumor present in slices: [36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
60 61 62 63 64]
Range of slices with tumor: 36 - 64
```

Figure 5. Tumor-containing slice range identified from the segmented 3D MRI volume.

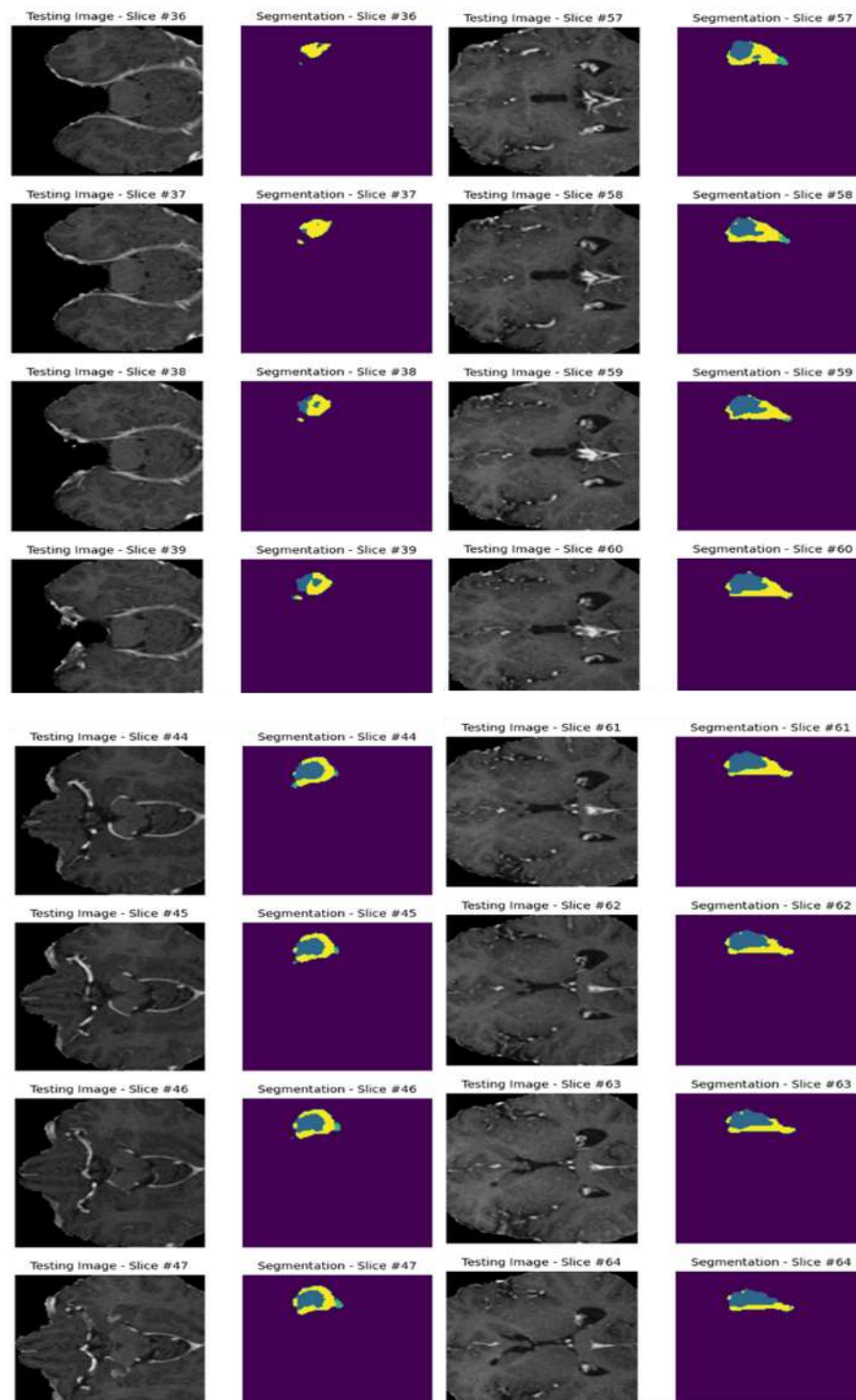


Figure 6. Representative tumor-containing MRI slices identified from the segmented 3D volume.

4.6 Quantitative Tumor-Region Analysis

Following segmentation, the predicted tumor regions were further analyzed to obtain patient-level quantitative characteristics. The implemented system generated intensity, density, vascularity-related values, tumor slice range, and maximum tumor slice information for the selected MRI volumes.

For the present analysis, results from 20 distinct patient cases were combined to provide a broader patient-level summary of the extracted quantitative characteristics. The extracted quantitative characteristics for the 20 patient cases are summarized in Table 7.

Table 7. Quantitative tumor-region analysis for 20 representative patients.

Patient	Intensity (%)	Density (%)	Vascularity	Tumor Slice Range	Max. Tumor Slice
9	31.78	0.42	0.50	61–82	74
11	17.28	1.46	0.28	45–67	53

Patient	Intensity (%)	Density (%)	Vascularity	Tumor Slice Range	Max. Tumor Slice
12	22.26	4.04	0.45	57–85	70
19	37.55	1.02	0.40	77–103	93
20	24.25	3.78	0.45	57–102	76
21	25.95	2.53	0.46	66–100	85
24	20.30	2.55	0.49	23–69	36
32	26.21	4.35	0.31	68–121	87
35	21.55	4.47	0.23	46–71	60
36	18.00	2.20	0.37	44–96	60
37	24.32	2.61	0.31	50–85	71
39	20.79	1.29	0.25	17–55	33
62	28.00	9.76	0.41	15–68	49
63	23.18	0.78	0.39	44–61	52
72	28.45	4.00	0.39	49–90	65
78	26.34	2.18	0.39	40–65	54
86	23.18	1.04	0.40	36–64	47
90	20.43	3.72	0.40	57–110	75
93	28.01	4.12	0.30	60–103	79
99	24.26	1.65	0.24	31–68	48

The patient-wise results demonstrate variation in the extracted quantitative characteristics. The intensity values range from **17.28% to 37.55%**, while density ranges from **0.42% to 9.76%**. The vascularity-related value ranges from **0.23 to 0.50**.

Patient 19 has the highest reported intensity value of **37.55%**, whereas patient 11 has the lowest value of **17.28%**. The highest density value is observed for patient 62 (**9.76%**), while the lowest is observed for patient 9 (**0.42%**). The highest vascularity-related value is **0.50 for patient 9**, whereas the lowest is **0.23 for patient 35**.

The detected tumor-slice ranges also vary across patients. Patient 86 has a range of **36–64**, whereas patients 90 and 32 have relatively extended detected ranges of **57–110** and **68–121**, respectively. The maximum tumor slice ranges from **slice 33 for patient 39** to **slice 93 for patient 19**.

These measurements demonstrate how the segmentation output can be extended beyond voxel-level classification to obtain patient-specific spatial and region-based information. The reported vascularity value is a **mask-derived computational feature generated by the implemented system** and should therefore not be interpreted as a clinical measurement of tumor vascularity.

4.7 Comparison of LinkNet Backbone Architectures

The original project also investigated LinkNet with three different encoder backbones: ResNet152, InceptionV3, and DenseNet121. The comparative experiment was performed to examine the effect of the backbone architecture on segmentation performance. The available results provide a comparative evaluation of these three backbone configurations. Representative prediction outputs for the three LinkNet backbone configurations are shown in Figures 7–9. The corresponding comparative accuracy, precision, recall, and F1-score results are presented graphically in Figures 10–13.

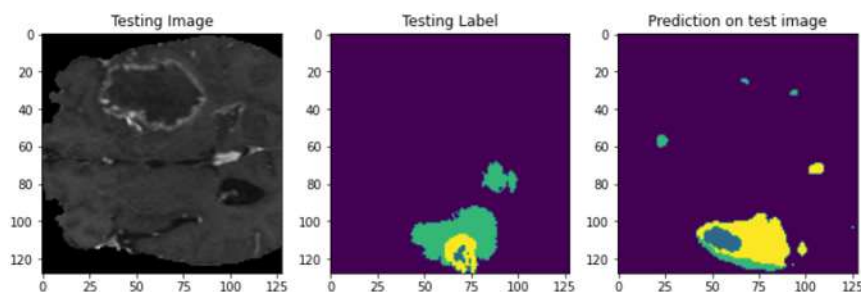


Figure 7. Representative prediction output of LinkNet with ResNet152 backbone.

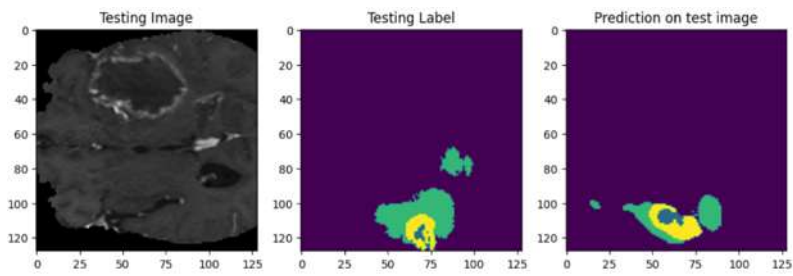


Figure 8. Representative prediction output of LinkNet with InceptionV3 backbone.

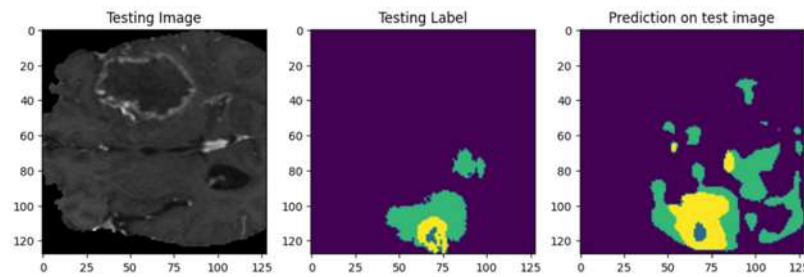


Figure 9. Representative prediction output of LinkNet with DenseNet121 backbone.

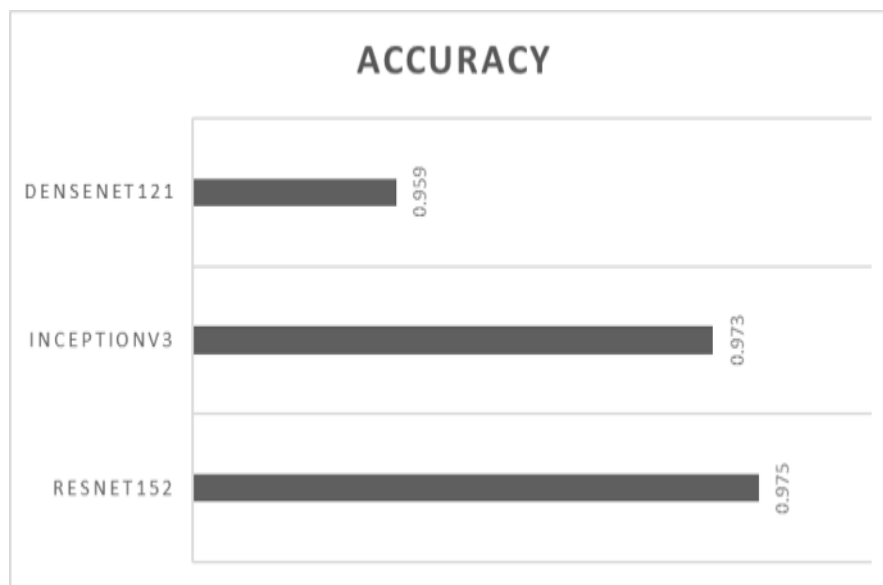


Figure 10. Accuracy comparison of the LinkNet models using different backbone architectures.



Figure 11. Precision comparison of the LinkNet models using different backbone architectures.

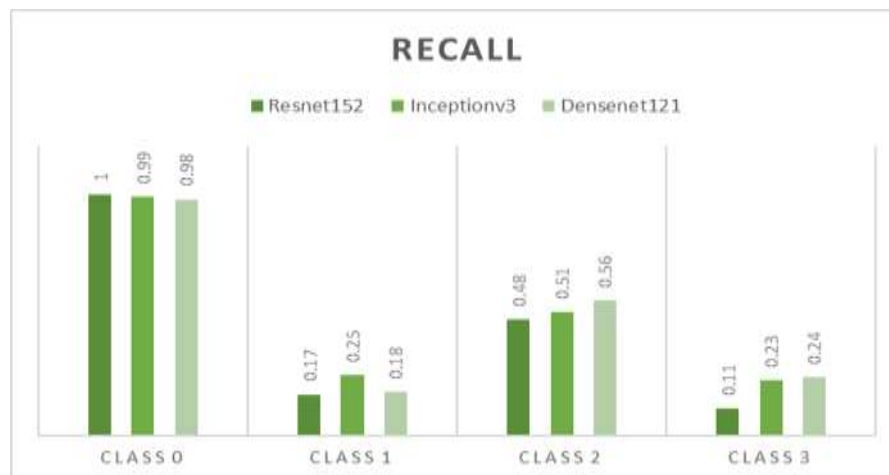


Figure 12. Recall comparison of the LinkNet models using different backbone architectures.

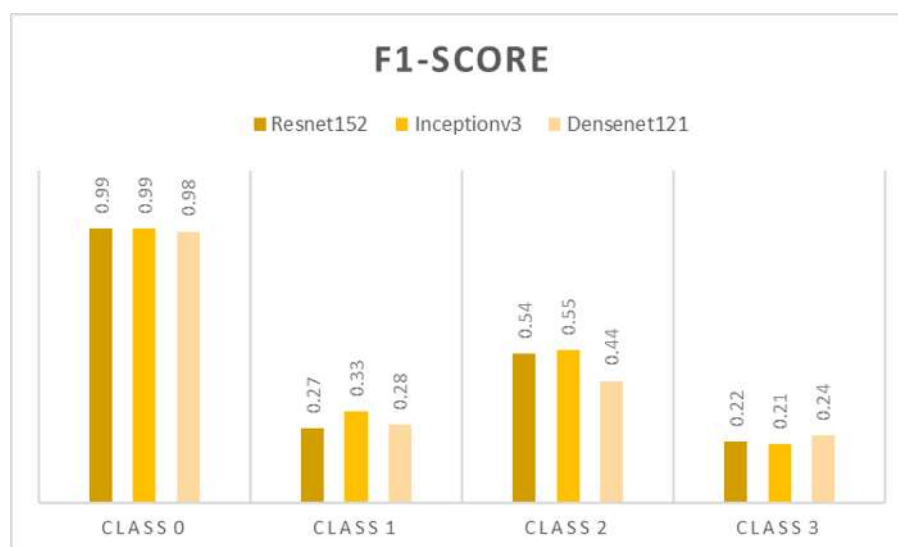


Figure 13. F1-score comparison of the LinkNet models using different backbone architectures.

The present study focuses on the LinkNet–ResNet152 configuration for the investigated multimodal brain MRI segmentation task. The use of a deep residual encoder provides hierarchical feature extraction within the LinkNet encoder-decoder framework, while the pretrained backbone provides an initialized feature representation.

The comparative results for the three LinkNet backbone configurations are presented graphically in Figures 10–13.

4.8 Integrated Visualization and Reporting

The segmentation framework was further integrated into a Django-based interface. The system supports MRI input, segmentation visualization, extraction of tumor-containing slices, and generation of a patient-specific PDF report. The MRI modality image upload interface is shown in Figure 14. The slice selection interface used for viewing the segmentation results is shown in Figure 15. The resulting segmentation output is displayed through the interface, as shown in Figure 16.

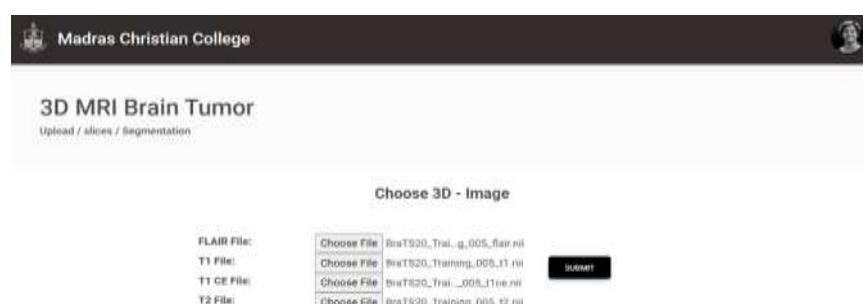


Figure 14. MRI modality image upload interface of the Django-based system.

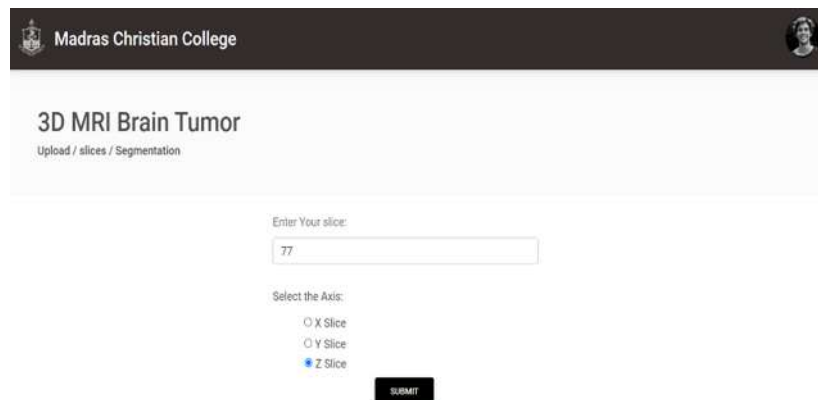


Figure 15. Slice selection interface of the Django-based system.

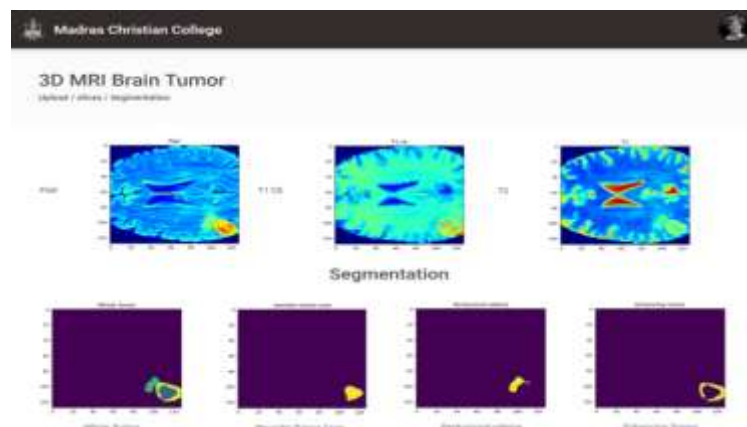


Figure 16. Display of the segmentation output in the Django-based system.

The predicted segmentation mask can be used for visualization and subsequent quantitative analysis. An example of the generated patient-specific PDF report is shown in Figure 17.

Tumor Report

Tumor Slices Range:	(57, '-', 110)
Max Tumor Slice:	75
Min Tumor Slice:	0
Intensity:	26.68284804931935
Density:	1.2001624474158654
Vascularity:	0.39248299417353116

Figure 17. Example of the generated PDF report containing the segmented tumor results and quantitative analysis.

4.9 Discussion

The experimental results demonstrate the applicability of a LinkNet architecture with a ResNet152 encoder for three-dimensional multimodal brain tumor segmentation. The model achieved **97.56% voxel-level classification accuracy** and a **0.97 weighted F1-score**, while the macro F1-score was **0.49**. The difference between these measures highlights the influence of class imbalance and indicates that overall accuracy alone is insufficient for characterizing performance across individual tumor classes.

The ROC-AUC analysis further showed class-dependent discrimination, with AUC values ranging from **0.8143 to 0.9367**. The lower performance observed for some tumor classes indicates that segmentation of individual tumor subregions remains more challenging than background classification.

An additional component of the implemented workflow is the extraction of patient-level information from the predicted segmentation masks. Analysis of 20 patient cases demonstrated variation in intensity, density,

mask-derived vascularity-related values, tumor-containing slice ranges, and maximum tumor slices. These results illustrate how the segmentation output can be extended to downstream quantitative analysis rather than being limited to generation of a multiclass segmentation mask.

The backbone comparison conducted in the project provided a comparative evaluation of the LinkNet configurations using ResNet152, InceptionV3, and DenseNet121. The present study focuses on the LinkNet-ResNet152 configuration.

Overall, the results demonstrate the applicability of a three-dimensional LinkNet-ResNet152 framework for multimodal MRI segmentation, tumor-containing slice identification, region-based analysis, visualization, and automated reporting. At the same time, the relatively lower macro F1-score indicates the need for further investigation into the representation and classification of individual tumor subregions, particularly the less frequent classes.

5. Conclusion

This study presented a three-dimensional LinkNet architecture with a ResNet152 encoder for multimodal brain tumor segmentation from MRI volumes. The framework was developed using the BraTS 2020 dataset and incorporated MRI preprocessing, three-channel multimodal volume construction, four-class voxel-level segmentation, performance evaluation, tumor-containing slice identification, and region-based quantitative analysis.

The proposed LinkNet-ResNet152 framework achieved an overall voxel-level classification accuracy of 97.56%, with a weighted F1-score of 0.97. The ROC-AUC analysis further demonstrated class-dependent discrimination, with AUC values of 0.9367, 0.8143, 0.9131, and 0.8838 for classes 0, 1, 2, and 3, respectively. The results indicate strong overall voxel-level classification performance, with class-dependent performance reflected in the reported F1-scores and ROC-AUC values. The analysis of tumor-containing slices also provided a mechanism for identifying relevant regions within the three-dimensional MRI volume and extracting mask-derived quantitative characteristics.

A further contribution of the work is the integration of the segmentation model with a Django-based user interface. The system supports MRI input, visualization of segmentation results, identification of tumor-containing slices, and generation of a PDF-based patient analysis report. Thus, the developed workflow connects deep learning-based segmentation with subsequent visualization and reporting functions.

The comparative investigation included different LinkNet encoder backbones, namely ResNet152, InceptionV3, and DenseNet121. The present study focuses on the LinkNet-ResNet152 configuration for the investigated multimodal brain MRI segmentation task.

Overall, the study demonstrates the applicability of a LinkNet-ResNet152 architecture for three-dimensional multimodal brain tumor segmentation and integrated tumor-region analysis. The framework provides a structured pipeline from MRI preprocessing and voxel-level segmentation to visualization and automated reporting, providing a basis for further investigation of deep learning approaches for volumetric brain tumor analysis.

6. Limitations and Future Work

The present study has several limitations that should be considered when interpreting the results. First, the evaluation was performed using the available BraTS 2020 data, and the reported performance is primarily based on voxel-level classification measures. Although the overall voxel-level accuracy was high, the macro F1-score was substantially lower than the weighted F1-score, indicating differences in performance among the segmentation classes. This highlights the effect of class imbalance and the difficulty of accurately identifying smaller tumor subregions.

Second, the present implementation uses a three-channel multimodal representation for the trained segmentation model, comprising FLAIR, T1ce, and T2 MRI modalities. Consequently, the findings should not be generalized directly to configurations using all four commonly available BraTS modalities without further experimental evaluation.

Third, the quantitative tumor characteristics extracted after segmentation, including intensity, density, and the reported vascularity-related measure, are computational features derived from segmentation masks. They should therefore be interpreted as image-derived measurements rather than clinical biomarkers. Further clinical validation would be required before such measurements could be considered for clinical decision-making.

Future work will investigate additional three-dimensional segmentation architectures, including **Attention U-Net** and a standard **3D U-Net baseline**, using the same multimodal brain MRI data. Attention U-Net will be explored to examine the potential contribution of attention mechanisms for focusing on relevant tumor regions, while 3D U-Net will serve as a baseline volumetric encoder-decoder architecture. Their performance will be evaluated and compared with the proposed LinkNet-ResNet152 framework using appropriate segmentation metrics.

Further work can also investigate evaluation on independent datasets and the use of additional segmentation metrics to assess the generalizability and robustness of the proposed framework. The integrated Django interface can be extended to provide a more comprehensive visualization and reporting environment, including patient information, imaging results, segmentation outputs, and quantitative measurements. Further validation with independent datasets and appropriate clinical evaluation would be required before practical application.

References

- [1] F. J. Dorfner, J. B. Patel, J. Kalpathy-Cramer, E. R. Gerstner, and C. P. Bridge, "A review of deep learning for brain tumor analysis in MRI," **npj Precision Oncology**, vol. 9, Art. no. 2, 2025, doi: 10.1038/s41698-024-00789-2.
- [2] Z. U. Abidin, R. A. Naqvi, A. Haider, H. S. Kim, D. Jeong, and S. W. Lee, "Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: A prospective survey," **Frontiers in Bioengineering and Biotechnology**, vol. 12, Art. no. 1392807, 2024, doi: 10.3389/fbioe.2024.1392807.
- [3] B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," **IEEE Transactions on Medical Imaging**, vol. 34, no. 10, pp. 1993–2024, 2015, doi: 10.1109/TMI.2014.2377694.
- [4] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," in **Proceedings of the IEEE Visual Communications and Image Processing (VCIP)**, 2017, pp. 1–4, doi: 10.1109/VCIP.2017.8305148.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [6] B. Jiang, M. Liao, Y. Zhao, G. Li, S. Cheng, X. Wang, and Q. Xia, "Deep learning for brain tumor segmentation in multimodal MRI images: A review of methods and advances," **Image and Vision Computing**, vol. 156, Art. no. 105463, 2025, doi: 10.1016/j.imavis.2025.105463.
- [7] T. Zhou, Z. Wang, X. Liu, W. Liu, and S. Zhu, "Learning deep feature representations for multi-modal MR brain tumor segmentation," **Neurocomputing**, vol. 638, Art. no. 130162, 2025, doi: 10.1016/j.neucom.2025.130162.
- [8] W. Liu, J. Hou, X. Zhong, H. Zhan, J. Cheng, Y. Fang, and G. Yue, "Improving multi-modal brain tumor segmentation via pre-training and