

Charan M Tech

charan mtech final_paper_2.docx

 Corporación Universitaria Minuto de Dios, UNIMINUTO

Document Details

Submission ID

trn:oid:::12096:470774034

Submission Date

Jun 30, 2025, 12:45 PM GMT+5:30

Download Date

Jun 30, 2025, 12:48 PM GMT+5:30

File Name

charan mtech final_paper_2.docx

File Size

1.1 MB

15 Pages

7,519 Words

46,013 Characters





16% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.




Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text

Match Groups

-  **97 Not Cited or Quoted 16%**
Matches with neither in-text citation nor quotation marks
-  **0 Missing Quotations 0%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 9%  Internet sources
- 12%  Publications
- 14%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 97 Not Cited or Quoted 16%**
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 9% Internet sources
- 12% Publications
- 14% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	Liverpool John Moores University on 2021-12-14	4%
2	Internet	link.springer.com	<1%
3	Internet	www.researchgate.net	<1%
4	Publication	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Spr...	<1%
5	Internet	www.ijracse.com	<1%
6	Submitted works	Turun yliopisto on 2025-06-19	<1%
7	Submitted works	The University of the West of Scotland on 2023-12-15	<1%
8	Submitted works	Imperial College of Science, Technology and Medicine on 2020-09-01	<1%
9	Internet	dokumen.pub	<1%
10	Submitted works	University of Westminster on 2024-02-14	<1%

11	Submitted works	Gujarat Technological University on 2025-02-14	<1%
12	Submitted works	Monash University on 2025-06-08	<1%
13	Internet	www.ijmer.com	<1%
14	Submitted works	West University Of Timisoara on 2024-06-26	<1%
15	Internet	thesai.org	<1%
16	Submitted works	Abo Akademi University on 2025-05-05	<1%
17	Publication	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Spr...	<1%
18	Publication	Dongzhe Li, Baoyao Yang, Yuebin Xie, Weide Zhan, Jingsong Lin. "Chapter 19 Mult...	<1%
19	Publication	Kwame Nyako, Uttam Dhakal, Frank Li, Vamsi Borra. "Trusted microelectronics: r...	<1%
20	Submitted works	University of Warwick on 2024-10-02	<1%
21	Internet	isteonline.org	<1%
22	Submitted works	Munster Technological University (MTU) on 2025-06-12	<1%
23	Publication	S. Prasad Jones Christydass, Nurhayati Nurhayati, S. Kannadhasan. "Hybrid and A...	<1%
24	Internet	lup.lub.lu.se	<1%

25	Internet	www.topic.ad.jp	<1%
26	Publication	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Spr...	<1%
27	Publication	Liang Cao, Long Li, Jifeng Zheng, Xin Fan, Feng Yin, Hui Shen, Jun Zhang. "Multi-ta...	<1%
28	Internet	export.arxiv.org	<1%
29	Internet	ijrti.org	<1%
30	Publication	"Data Science and Applications", Springer Science and Business Media LLC, 2025	<1%
31	Internet	etd.aau.edu.et	<1%
32	Internet	spectrum.library.concordia.ca	<1%
33	Internet	arxiv.org	<1%
34	Internet	www.cbica.upenn.edu	<1%
35	Publication	"Statistical Atlases and Computational Models of the Heart. Regular and CMRxMo...	<1%
36	Submitted works	University of Hyderabad, Hyderabad on 2025-04-29	<1%
37	Publication	"Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016", ...	<1%
38	Submitted works	Tilburg University on 2025-06-23	<1%

39	Submitted works	University of Leeds on 2025-04-29	<1%
40	Submitted works	University of Surrey on 2024-09-03	<1%
41	Internet	www.mdpi.com	<1%
42	Publication	"Intelligent Systems and Data Science", Springer Science and Business Media LLC,...	<1%
43	Publication	"Machine Learning in Medical Imaging", Springer Science and Business Media LL...	<1%
44	Publication	"Medical Image Computing and Computer-Assisted Intervention – MICCAI 2017", ...	<1%
45	Submitted works	Associatie K.U.Leuven on 2017-01-16	<1%
46	Submitted works	Associatie K.U.Leuven on 2022-01-17	<1%
47	Submitted works	Gitam University on 2025-04-20	<1%
48	Submitted works	Kingston University on 2023-01-10	<1%
49	Submitted works	Munster Technological University (MTU) on 2025-06-12	<1%
50	Publication	Sergio Pereira, Adriano Pinto, Victor Alves, Carlos A. Silva. "Brain Tumor Segment...	<1%
51	Submitted works	University College London on 2023-07-28	<1%
52	Submitted works	University of Durham on 2025-04-26	<1%

53	Submitted works	University of North Texas on 2024-05-06	<1%
54	Internet	www.diva-portal.org	<1%
55	Publication	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dharendra Kumar Shukla. "Intelli...	<1%
56	Submitted works	Liverpool John Moores University on 2023-09-07	<1%
57	Publication	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical ...	<1%
58	Submitted works	Imperial College of Science, Technology and Medicine on 2020-09-07	<1%
59	Submitted works	Liverpool John Moores University on 2024-01-29	<1%
60	Submitted works	The University of Manchester on 2020-09-04	<1%
61	Submitted works	University of Wales Institute, Cardiff on 2019-04-25	<1%
62	Publication	"Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Spr...	<1%
63	Submitted works	University of Birmingham on 2024-05-07	<1%
64	Internet	www.spiedigitallibrary.org	<1%



Benchmarking Deep Learning Architectures for MRI-Based Brain Tumor Segmentation Using BraTS 2020

Piriya Sricharan Teja¹, T. Ravi Kumar², Bandaru Ramesh³, Surya Pavan Kumar Gudla⁴, Tejo Venkatesh Thota⁵

¹M.Tech scholar, Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Srikakulam- 532201, Andhra Pradesh, India

²Professor, CSE Department, Aditya Institute of Technology and Management, Srikakulam- 532201, Andhra Pradesh, India

³Assistant Professor, CSE Department, Aditya Institute of Technology and Management, Srikakulam- 532201

, Andhra Pradesh, India ⁴Assistant Professor, Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Srikakulam- 532201, Andhra Pradesh, India

⁵Student Information Systems Department, University of Maryland Baltimore County, Baltimore-21228, USA

Emails: Srich.piriya@gmail.com, ravi.4u.kumar@gmail.com, ramesh.bandaru15@gmail.com, pavan1980.mca@gmail.com, tejo2511@gmail.com

Corresponding-author: ravi.4u.kumar@gmail.com

Abstract: Brain tumors using magnetic resonance imaging (MRI) must be accurately segmented for clinical diagnosis and therapy planning. The ability of deep learning (DL) models to extract hierarchical characteristics from complicated data has made them popular tools in medical imaging. The BraTS-2020 3D MRI data is used in this study to compare four DL models for brain tumor segmentation: U-Net, VGG16, MobileNet and DenseNet. The dataset offers a strong standard for assessment since it contains multimodal MRI volumes annotated with glioma subregions. Preprocessing techniques like intensity scaling, data augmentation, and z-score normalization were used to improve model generalization. Since, the U-Net model has encoder-decoder structure and skip connections, thus exhibits exceptional performance in capturing spatial continuity and defining the exact dimensions of tumor boundaries, according to experimental results. In border localization, VGG16 and DenseNet fared poorly, despite achieving good classification accuracy through deep feature extraction. MobileNet demonstrated effectiveness but lacked segmentation precision because it was built for low-latency situations. A quantitative assessment of U-Net's improved segmentation performance is provided by the Dice coefficient, mean Intersection over Union (IoU), sensitivity, specificity, and precision. The results highlight the significance of choosing a model that is suited to clinical needs, where interpretability and spatial accuracy are essential.

Keywords: Brain Tumor Segmentation, U-Net, CNN, DenseNet, MobileNet, VGG16, BraTS, MRI, DeepLearning.

1. INTRODUCTION

Brain tumors are one among the most chronic diseases, and they must be diagnosed as soon as possible to be treated effectively. To observe brain structures and determine tumor locations, magnetic resonance imaging, or MRI, is commonly employed [1].

Radiologists' subjective and sluggish manual segmentation of tumors result in inconsistent diagnosis. Brain tumor segmentation has been best automated using deep learning models, particularly convolutional neural networks (CNNs). They can identify tumor boundaries

pixel-wise classification, hierarchical feature extraction, and processing high-dimensional imaging data.

Comparative study of brain tumor segmentation based on deep learning models with BraTS dataset is the topic of discussion in this paper [2], [3].

The main objective is to compare the performance of CNN architectures for classification, such as VGG16, DenseNet, MobileNet, and ResNet, with U-Net, a model that has been specifically used for biomedical image segmentation [13].

Decoder heads and encoder-decoder models are used to segment these models that were initially created for image classification tasks. The goal is to determine whether these altered architectures can match or excel the performance of

specifically designed models, such as U-Net, for image segmentation. The enhancing tumor (ET), tumor core (TC), and whole tumor (WT) are a few of the annotated tumor subregions in the large collection of pre-operative MRI scans of the BraTS dataset [1], [2].

The segmentation task is even more challenging and dense when multimodal MRI channels (T1, T1ce, T2, and FLAIR) are used, which indeed making it a ideal benchmark.

Preprocessed inputs such as intensity normalization, z-score scaling, and data augmentation are utilized to train all models to guarantee their robustness to patient data variations. Though classification-based models such as VGG16 [11], DenseNet [11], and MobileNet [12] possess decent accuracy when optimized for segmentation, this work points out that these models fail to represent the spatial continuity and nuances of brain tumor morphology. In contrast, skip connections and segmentation-beneficial structure of U-Net allow it to preserve its superior performance. Lower F1 scores of the classification-based models reinforce their failure to properly identify tumor edges. **This is how the following paper is structured:**

- **Literature Review:** Comprehensive review of past exploration on brain tumor image segmentation involving deep learning. The strengths, limitations, and flaws in current studies are discussed here.
- **Dataset:** The data set, model architectures, training schedules, and preprocessing steps used in the current study are all discussed in detail in the proposed models and materials.
- **Results Analysis:** Comparative analysis with precision, recall, F1 measure, mean Intersection over Union (IoU), and dice score. Supported by tables and figures, results are presented and defended.
- **Conclusion:** Principal findings, implications for medical diagnostics, and avenues for future research—namely, improving the spatial sensitivity of segmentation models trained on classification—are recapitulated in the final conclusion and future work section.

2. LITERATURE SYRVEY

Deep learning algorithms for the automatic segmentation of brain tumors in MRI scans have been the subject of numerous studies with the goal of enhancing diagnostic systems' clinical usefulness, accuracy, and operational efficiency. Menze et al. [1] provided a concrete basis by introducing the Brain Tumor Segmentation (BraTS2020) benchmark, which provided a multimodal MRI dataset of T1, T1ce, T2, and FLAIR sequences with expert-segmented tumor subregions. In improvements made by Bakas et al. [2], [4], [5], the dataset was further improved upon with radiomic features and survival analysis annotations to make BraTS the

standard against which the training and assessment of segmentation models is judged.

Ronneberger et al.'s original U-Net architecture [13] revolutionized biomedical image segmentation using its encoder-decoder architecture and skip connections that preserve spatial resolution and enhance localization. The generalization of this model to the volumetric space, i.e., 3D U-Net [14], further eased better exploitation of spatial continuity in MRI volumes. Empirical observations such as those of Montaha et al. [7] found 3D U-Net very accurate in segmenting glioma subregions—enhanced tumor (ET), tumor core (TC), and whole tumor (WT)—on the BraTS dataset.

U-Net-based, Isensee et al. [8] designed nnU-Net as an auto-configuring architecture that has the ability to modify architecture, preprocessing, and training protocols automatically to be applied to any biomedical dataset. Its high performance and generalizability have made it one of the most successful architectures in BraTS challenges. Pereira et al. [18] also employed a deep CNN approach with pre- and post-processing techniques to achieve improved tumor border definition, further testifying to the efficiency of convolutional architectures in medical image segmentation.

More recent progress are attention-based and transformer-aided architectures. Trans BTS, introduced by Wang et al. [9], integrates convolutional backbones with transformer blocks to capture long-range dependencies, which are essential in segmenting spatially distributed tumor structures. Zhang et al. [10] addressed the integration of multimodal data through a cross-modality deep feature learning framework, effectively improving segmentation in complex tumor morphologies.

Although originally designed for image classification, other networks such as DenseNet [11], MobileNet [12], VGG16 and ResNet have been used for medical segmentation by incorporating decoder modules and transpose convolution layers. Dense connections in DenseNet increase reuse of features, leading to improved gradient flow and representation learning when reused for segmentation purposes. MobileNet, with depthwise separable convolutions, offers computational efficiency and is even edge application compatible but lags behind in preserving fine tumor boundary details in comparison to domain-specific segmentation models such as U-Net.

Other studies have examined cascaded and multi-scale segmentation pipelines. Kamnitsas et al. [15] proposed a 3D CNN with Conditional Random Fields (CRF) for post-processing for improved segmentation accuracy, for example. Wang et al. [17] proposed a cascaded anisotropic convolutional neural network that captured hierarchical spatial context well, which resulted in improved segmentation in many tumor areas.

3 Sricharan, et al.

Jungo and Reyes [16] also spoke of concerns regarding machine-based segmentation in clinical applications being too strong in their immunity to failures. They stressed stringent evaluation protocols to address issues such as scanner variability, data imbalance, and domain shift, which could compromise machine learning model reliability in real-world applications.

On a general note, therefore, literature review indicates that even with many of the CNN-based and transformer-coupled models having been implemented on brain tumor segmentation effectively, the best and most dependable are U-Net and variants since their architectures best address the needs of biomedical segmentation.

Such classification models as DenseNet, VGG16, and MobileNet, even though they possess strong feature extraction capability, require some structural adjustment to obtain acceptable levels of segmentation. This research main motto is to make a systematic comparison of these models—optimized for segmentation tasks—on the BraTS dataset, shedding light on their trade-offs in segmentation accuracy, spatial fidelity, and computational efficiency.

3. DATASET

In the field of medical imaging, deep learning-based segmentation techniques have been proven extremely promising, specifically in marking and detecting brain tumor areas in MRI images. Such techniques help the radiologists achieve instantaneous and accurate diagnoses by making automatic, reliable, and high-accuracy evaluations possible.

The four deep structures—U-Net [13], VGG16 [18], MobileNet [12], and DenseNet [11]—employed within the BraTS 2020 dataset [1] are compared in this work. It depends on the training performance, computational efficiency, and segmentation accuracy during evaluation. The steps used in the experiment, which comprise preparing the dataset, preprocessing, designing model architecture, training plan, and evaluation criteria are described in the upcoming sections.

3.1. Dataset Description

The dataset used in this study was the Multimodal Brain Tumor Segmentation Challenge (BraTS) 2020 dataset, which is a recognized benchmark for brain tumor segmentation challenges in clinical neuroimaging. The BraTS dataset includes multimodal MRI images of glioma patients, specifically those with Low Grade Glioma (LGG) and High Grade Glioma (HGG) [1], [2]. Each subject's MRI volume is imaged using one of four modalities: T1-weighted (T1), T1-weighted with contrast (T1ce), T2-weighted (T2), and Fluid-Attenuated Inversion Recovery (FLAIR). Professionally created ground truth labels have been applied to three therapeutically relevant tumor subregions: Whole Tumor (WT), Tumor Core (TC), and Enhancing Tumor (ET). In the 3D NIfTI format, voxel-level annotations are assigned

values 0 (background), 1 (necrotic/non-enhancing tumor), 2 (edema), and 4 (enhancing tumor). Label "4" is changed to "3" for multi-class segmentation in order to maintain class continuity and facilitate one-hot encoding for four-category segmentation.

All volumes are clipped preprocessed and around the brain region to make sure that a consistent spatial resolution is maintained across scans. Bilinear interpolation is used in this study to down sampling each 3D scan to 128×128 pixels once it has been decomposed into 2D axial slices. To achieve computational efficiency as well as emphasize training on insightful samples, high tumor presence slices are only maintained [3], [4].

Compiling diverse MRI data recording clinical reality in tumor shape, texture, and intensity, BraTS dataset has been the work of multi-institutional efforts. Consequently, it is the perfect tool to train and validate strong, generalizable segmentation models [5]. Furthermore, cross-checked annotations were done by board-certified neuroradiologists to provide exact and reproducible ground truth segmentations [1].

3.2. Data Visualization

Raw data must be preprocessed before it can be used to train deep learning models for imaging tasks like tumor segmentation. The BraTS2020 dataset consists of multimodal MRI scans (T1, T1ce, T2, and FLAIR) of glioma patients. Since each modality records unique tissue contrasts, it is necessary to watch and comprehend how each one aids in the identification of tumor subregions. Visualizing such MRI slices facilitates tumor edge identification and the assessment of image quality homogeneity between modalities and patients.

The imbalanced occurrence of tumor subregions in the BraTS data is a significant training problem for deep learning models. Only a small percentage of MRI slices unambiguously classify the enhancing tumor, necrotic core, or edema, but the majority classify only background or vague tumor characteristics. During training, this unbalance can cause bias in the model, which subsequently can cause poor generalization on minority classes like the enhancing tumor.

Data sampling methods are used to accomplish this. Some of them are:

In order to make sure that every class (e.g., necrotic core and enhancing tumor) is properly represented in the training set, oversampling or up-sampling copies underrepresented slices or high tumor content augmentations. In small or unbalanced datasets, oversampling works very well and keeps all original samples intact.

A subset of slices of the dominant class is chosen using undersampling (also known as down-sampling) to equate the

4 Sricharan, et al.

quantity of underrepresented classes. This saves training time and storage space but may lose valuable data variances. Imparting class balance, reducing bias towards the background class, and enhancing segmentation accuracy for less frequent types of tumors are the goals for both approaches. Without any architecture changes involved, this sampling phase is key to maximizing the efficacy of widely used machine learning algorithms.

3.3. Visualization of Multi-modal MRI and Segmentation Masks

Using Python libraries NiBabel and Matplotlib, we created a visualization pipeline to gain a complete insight into the structure of the dataset as well as the anatomical information obtained from different Magnetic Resonance maging (MRI) modalities.

Accurate interpretation of the multimodal data within the BraTS 2020 training set was guaranteed through this process [1], [2]. The four MRI modalities—FLAIR, T1, T1ce, and T2—and the respective expert-annotated segmentation mask are shown side by side in Figure 1 for axial slices of a single patient sample.

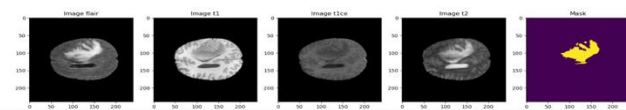


Figure.1 Visualization of BraTS MRI Modalities and Corresponding TumorMask

For loading the 3D volumetric NiftI files using NiBabel and for extracting typical axial slices from each volume, the visualization script was created. Because most glioma substructures are usually visible in mid-brain tumor locations, a slice about 25 steps anterior to the midline (the middle slice of the 3D volume) was selected so as to consistently detect these areas [3].

FLAIR is highly sensitive to peritumoral edema, T1 offers anatomical reference, T1ce uses contrast agent uptake to highlight increased tumor regions, and T2 shows the tumor core and the edema pattern. Each MRI modality contributes additional diagnostic information. The manually annotated tumor subregions provided by qualified radiologists are enclosed by the segmentation mask, which is shown in the bottom subplot, similar to **Figure 1** [1], [2]. **Label 1** represents non-enhancing tumor core (NCR/NET), **Label 2** peritumoral edema (ED), and **Label 4** enhances the tumor core (ET). **The ground-truth segmentation mask is a 3D volume with a unique voxel value to distinguish the tumor classes.**

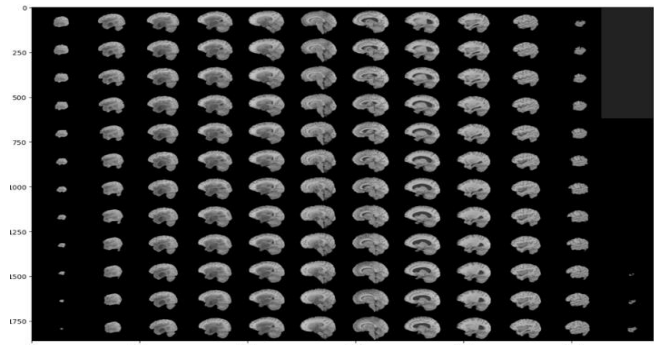


Figure 2. Volumetric montage of axial T1-weighted MRI slices from a single subject

To maintain segmentation uniformity in all networks, label 4 is re-mapped into class 3, evolving a four-class schema in which background tissue is assigned label 0 [1], [5]. Since these label assignments are the basis for pixel-level classification during model training, they are essential for supervised learning.

By a volumetric montage of the segmentation mask, further confirmation of tumor annotation anatomical distribution and morphological consistency was performed **Figure 2**. Aside from superior and inferior slices, usually without any pathology, this montage combines several axial slices, mainly from the middle parts of the brain volume.

Presenting these consecutive slices within one frame makes it possible to qualitatively confirm the continuity of segmented subregions along the cranial axis, internal consistency of the mask, and boundaries of the tumor [3]. This type of image not only enables pretreatment validation and sanity checks for data, but also emphasizes the need for multimodal learning strategies. The previous work on U-Net and its 3D versions [1], [3], and [13] confirms the respective contrast patterns between modalities, which imply the usefulness of combined features towards correct segmentation. Multimodal representations in such a manner guarantee that informative clinical and anatomical features are well represented in order to optimize downstream segmentation models.

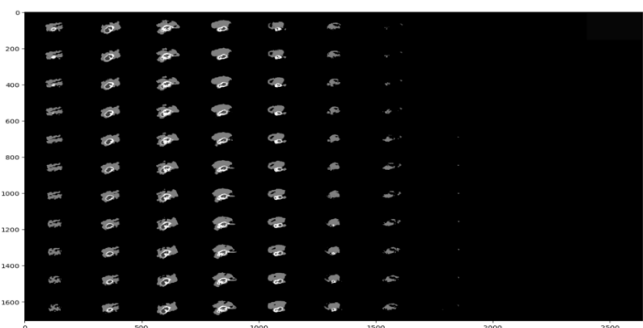


Figure. 3. Volumetric montage of the corresponding ground-truth segmentation mask

5 Sricharan, et al.

This figure presents the ground-truth segmentation annotations for the same subject as in Figure 3. The montage displays enhancing tumor, non-enhancing core, and peritumoral edema across sagittal slices for qualitative structural validation.

The deep learning process is significantly facilitated by this visualization step in all kinds of ways. To determine if segmented tumor regions have morphological coherence between adjacent slices, it is first a test for structural continuity. Where contrast is low or anatomical boundaries are blurry, in regions of autonomous models usually to indicate ambiguity, this is most critical. Second, these tests aid in the choice of best slices for training. Specifically, they support removal of uninformative slices, or those that have no tumors, a condition necessary to improve computational efficiency and memory reduction during the utilization of 2D convolutional neural network (CNN) architecture [13].

Further, through displaying features such as uneven tumor borders, multilobular borders, and spatial differences across subregions, the visualization helps in further comprehension of morphology in tumors.

These morphological findings directly influence downstream volumetric measurement and clinical interpretability, as well as being essential for model generalizability [2], [5]. Finally, by detecting abnormalities in ground-truth annotations, visual examination of segmentations within volumes allows initial error analysis.

These, if left uncorrected, can potentially harm training performance and include abrupt discontinuities, misregistrations, or annotation noise [3], [17].

4.Methods

4.1. U-Net

A U-Net variant network architecture, which is renowned for its performance in biomedical image segmentation tasks, is the employed segmentation model in this study [13]. The structure adheres to the standard encoder-decoder topological configuration, as shown in Figure 4.

The encoder or contracting path uses 3×3 convolutional layers with ReLU activations and 2×2 max-pooling operations to gradually downsample the input image. By increasing the number of feature maps from 32 to 512 with decreasing spatial resolution, this architecture allows the network to pick up on subtle contextual tumor features like textural abnormalities and heterogeneity.

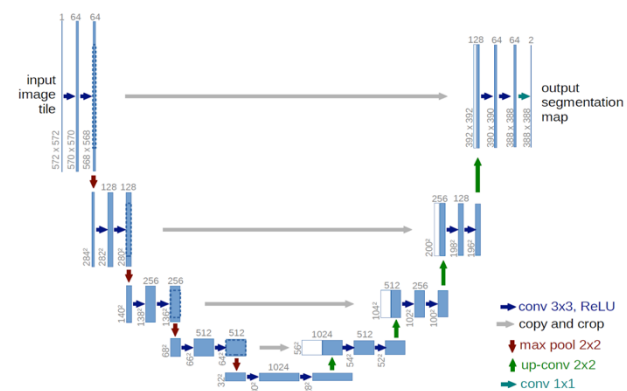


Figure. 4. U-Net Architecture based segmentation model used in this study.

Similar to the encoder, the decoder path uses upsampling by transposed convolutions and then by normal convolutions to restore spatial resolution. Significantly, skip connections between encoder and decoder layers allow for high-resolution features to be passed to deeper layers, facilitating precise definition of the tumor boundary, including the tumor subregion improvement, necrotic cores, and edema [13], [14]. Softmax activation and 1x1 convolution layer are utilized to build the final output, which generates four classes of segmentation belonging to the tumor and background area.

To encourage generalization and avoid overfitting, dropout regularization and normal weight initialization are used, particularly in bottleneck layers. After compilation with a categorical cross-entropy loss function, the model is trained using the Adam optimizer.

Metrics including overall accuracy, Mean Intersection-over-Union (Mean IoU), precision, sensitivity, specificity, and region-specific Dice coefficients are used to evaluate the fidelity of class-wise segmentation [7], [18].

Slice-wise, a lightweight CNN classifier is employed for both segmentation and multi-class prediction. To predict tumor class membership, the model is fed localized radiological information from several MRI modalities.

Max-pooling operations come after ReLU activates two convolutional layers in the model. The same is followed by a final softmax layer that separates input slices into four classes of tumors, dense fully linked layers, and a flattening layer.

The classification model can be used in high-dimensional MRI classification tasks due to the fact that it is computation-efficient and has a lower risk of overfitting. It is tested on classification accuracy after training on one-hot encoded labels with categorical cross-entropy loss. A CSV logger is used for post-training evaluation to log training and validation performance metrics at the epoch level.

6 Sricharan, et al.

To achieve a balance between memory requirements and spatial information preservation, the two networks are trained on scaled 128×128 pixel 2D axial slices. For improving class discrimination based on texture, the classifier is given three-channel composite images, although the segmentation model takes two-channel inputs (FLAIR and T1ce). The architecture progresses through the following key layers:

Encoding Path:

- Conv2D (32 filters) → Conv2D (32 filters) → MaxPool
- Conv2D (64 filters) → Conv2D (64 filters) → MaxPool
- Conv2D (128 filters) → Conv2D (128 filters) → MaxPool
- Conv2D (256 filters) → Conv2D (256 filters) → MaxPool
- Conv2D (512 filters) → Conv2D (512 filters) → Dropout (0.2)

Decoding Path:

- UpSampling → Conv2D (256 filters) → Skip connection with encoder block → Conv2D (256) ×2
- UpSampling → Conv2D (128 filters) → Skip connection → Conv2D(128)×2
- UpSampling → Conv2D (64 filters) → Skip connection → Conv2D(64)×2
- UpSampling → Conv2D (32 filters) → Skip connection → Conv2D (32) ×2

Output Layer:

Output=Conv2D(4,1×1,activation='softmax')

This outputs a four-channel segmentation mask corresponding to the background and the three tumor subregions: necrotic core, edema, and enhancing tumor. The segmentation performance is optimized using the categorical cross-entropy loss in conjunction with custom evaluation metrics. The Dice coefficient, which measures the overlap between predicted and ground truth masks, is used to evaluate the segmentation quality. It is computed for each class c as:

$$\text{Dice}_c = \frac{2 \times |P_c \cap G_c| + \epsilon}{|P_c| + |G_c| + \epsilon}$$

Where P_c and G_c denote the predicted and ground truth masks for class c , and ϵ is a small smoothing constant.

To obtain the macro-averaged Dice score across all tumor subregions, the per-class Dice scores are averaged:

$$\text{Mean Dice} = \frac{1}{N} \sum_{c=1}^N \text{Dice}_c$$

Mean Intersection over Union (IoU):

$$\text{IoU} = \frac{TP}{TP + FP + FN}$$

Precision:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Sensitivity (Recall):

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

Specificity:

$$\text{Specificity} = \frac{TN}{TN + FP}$$

4.2.VGG-16

The efficacy of transfer learning for the brain tumor image segmentation was explored further using a semantic segmentation model of the VGG16 architecture. Originally proposed for large-scale image classification of the ImageNet dataset, the pre-trained convolutional structure of the VGG16 network [11] is utilized in this model. The addition of an upsample decoding with a custom upsampling decoder is made in order to support segmentation tasks. On relatively smaller datasets like BraTS, our method minimizes the risk of overfitting and speeds up convergence, enabling transfer of abstracted image features to medical imaging tasks.

A max-pooling operation reducing spatial resolution while enhancing feature abstraction is performed after each of the 13 convolutional layers that make up the VGG16

7 Sricharan, et al.

architecture's five hierarchical blocks. The convolutional layers are frozen and used as the encoder in this study. They are given 2D MRI slices that are downsampled to 128×128 pixels, and they output a compressed feature representation of size 4×4×512. Without removing fine-grained geographical information, this encoder can preserve the most important semantic content. Pixel-by-pixel classification is obtained by the specially constructed decoder component, which restores the original image resolution.

Several UpSampling2D layers are sandwiched by convolutional layers. In the output, the decoder gradually upsamples the feature maps from 512 filters to 32 filters. To create four-channel segmentation masks, a final 1×1 convolution layer with a softmax activation function is employed. Each pixel is assigned to one of four classes: increasing tumor, necrosis, edema, or background. With a constant learning rate of 0.001, the Adam optimizer and categorical cross-entropy loss were used for training. The training and validation sets' loss and batch-wise accuracy were tracked during the course of the ten training epochs.

Classification accuracy, validation accuracy, and categorical cross-entropy loss were some of the evaluation metrics. Learning curves were employed to display these metrics, highlighting model generalization and training stability (as observed in attached plots). The main strength of the model is that it is able to learn hierarchical features at a deep level using the VGG16 backbone, enabling the precise classification of tumor subregions in even complex MRI modalities. Its light and relatively shallow decoder makes inference and training effective. Nevertheless, U-Net's [13] lack of skip connections makes it challenging to preserve low-level spatial features, which can affect the delineation of complex tumor boundaries [11], [18].

4.3.Densenet

DenseNet, as its name implies for Densely Connected Convolutional Network, is a deep network that achieves end-to-end direct connections through every layer of all the blocks, hence facilitating reuse of features and vanishing gradients [11].

Since its densely connected pattern naturally promotes more efficient compact models with superior feature expression and stable gradient propagation, DenseNet fits especially well with the field of medical image computing. Based on information provided on the BraTS 2020 challenge, DenseNet121 was modified for semantic segmentation of brain tumors in this study [1], [2].

T1-weighted MRI volumes of the BraTS dataset were chosen as input modality because of its high-resolution structural information. The center axial slice of each volume was cropped and resized to 128 x 128 pixels to balance spatial quality with computational efficiency. Inter-scanner

heterogeneity was addressed by normalizing image intensities. Replication of all grayscale slices with three channels was formed to meet the input requirements of the pretrained ImageNet-pretrained DenseNet architecture.

Background, necrotic tumor core, peritumoral edema, and augmenting tumor were defined in terms of one-hot encoded category labels from equivalent segmentation masks loaded from seg.nii files [1], [4]. The segmentation model used learned features from big datasets by using DenseNet121 as a pre-trained feature extractor. The decoder was built with fully convolutional layers and used a sequence of convolution blocks and upsampling to restore spatial resolution. For each upsampling block, a Conv2D layer with a 3×3 kernel, ReLU activation, and same padding to maintain dimensions came after an UpSampling2D layer.

A 1×1 convolution layer and softmax function were used toward the end of the network to output four-channel probability maps and pixel-wise segmentation maps of equal size to the input at the start.

The model was trained with a learning rate of 0.001 using the Adam optimizer using categorical cross-entropy as the loss function. With a batch size of eight, the data was split into training and validation subsets in an 80:20 ratio.

Performance was measured over ten iterations of training. Loss values reduced consistently over time, showing successful convergence of learning, and training and validation accuracy rose consistently with little overfitting. These patterns were supported by performance curves in training logs. Despite DenseNet encoder being demonstrated to perform well in acquiring intricate hierarchical features of gliomas, the absence of skip connections in the decoder may have the tendency to confine the ability of the model to efficiently replicate intricate tumor boundaries in comparison to U-Net-type architectures [13], [15]. Nonetheless, the gradient flow preservation and feature reuse capability of DenseNet provides a huge benefit in segmentation procedures involving sparse and heterogeneous medical information.

Strong generalization and scalability are provided by DenseNet-based architectures in clinical settings. Accurate tumor localization on MRI scans assists diagnosis and surgical planning, and their effective parameter use and pretrained capabilities make them suitable for real-time application in radiological pipelines [11], [15].

4.4.MobileNetV2

In this work, 2D axial slices from T1-weighted MRI volumes in the BraTS 2020 dataset were utilized to detect tumor subregions by semantic segmentation with MobileNetV2 [1], [2]. A light-weight CNN architecture

8 Sricharan, et al.

designed for fast inference on edge devices is known as MobileNetV2. Inverted residual bottlenecks and depthwise separable convolutions are used to find a good balance between computational expense and accuracy [12].

Bilinear interpolation was used to resize axial slices to 128×128 pixels to pre-process the data. Each of the grayscale images was normalized and replicated across three channels in order to match the input format needed by the ImageNet-pretrained MobileNetV2 encoder. Four classes were created by resizing and one-hot encoding relative segmentation masks: backdrop, necrotic core, edema, and augmenting tumor [1], [4].

The architecture employs MobileNetV2's pretrained encoder except for its classification head. The encoder is computationally efficient while effectively extracting hierarchical features at various spatial scales. Symmetrically duplicating the encoder, the decoder consists of five upsampling blocks with 3×3 convolutional layers with ReLU activation and same padding to safeguard spatial dimensions. The output 1×1 convolution with softmax activation provides class probabilities for every pixel.

The model was trained using the Adam optimizer, categorical cross-entropy loss, and a starting learning rate of 0.001. For training over ten epochs, an 80:20 train-validation split was used with a batch size of eight. Accuracy was tracked during training, however mean IoU and Dice scores were used in the post-training evaluation for clinical relevance [6], [7].

Effective generalization was displayed by the convergent behavior of the model, which was illustrated by plateaued validation loss and ever-increasing accuracy gains in the training and validation curves. Compared to U-Net or DenseNet, MobileNetV2's absence of skip connections and relatively shallow decoder restricts its capacity to precisely represent fine tumor boundaries, despite its proficiency and suitability for real-time usage in resource-poor settings [11], [13], and [15].

4.5. Resnet

Residual Networks (ResNets) are extremely powerful deep learning networks that can train deeper models efficiently by resolving the disappearing gradient issue through identity shortcut connections. ResNet architectures, initially proposed by He et al., enable direct passing of gradients through identity mappings, which improves convergence and allows more stable deep neural network optimization [11].

For semantic segmentation of brain tumor regions in MRI images from the BraTS 2020 dataset, this study utilized ResNet-50 as encoder backbone [1], [2].

To provide a strong base for discriminative feature extraction from medical images, the network was first initialized with

ImageNet-pretrained weights. During training, pretrained encoder layers were fixed to avoid overfitting and lower computational expense.

By using alternating UpSampling2D operations with Conv2D layers, progressively the learned features were upsampling to a decoder network. The spatial resolution of compressed latent representation of size (4×4) was restored to the actual input size of 128×128 pixels by this method of upsampling. The output step made use of a 1×1 kernel Conv2D layer with the softmax function for producing the segmentation map. The four class names assigned to each pixel in this map are background, necrotic core, peritumoral edema, or augmenting tumor [4], [13]. Preprocessing involved the extraction of the middle axial slice from each input subject's T1-weighted MRI volume, resampling it to 128×128 pixels, and normalizing it to [0, 1]. For compatibility with the ResNet input size, these gray-level images were converted to three channels. One-hot encoding was used for categorical compatibility with the multi-class segmentation model to modify the corresponding segmentation masks [1], [5]. Categorical cross-entropy loss and the Adam optimizer with a 0.001 learning rate were used to train the model. Eight and ten epochs were used for training in the batch, with 80:20 training-validation split. The main evaluation metric was accuracy. There was no indication of overfitting in accuracy and loss curves, and the model showed steady convergence with epochs.

Excellent performance was demonstrated by ResNet-based segmentation architecture because it has a strong gradient flow as well as powerful feature extraction capability. Skip connections, like those typically used in architectures such as U-Net to allow for the preservation of subtle spatial information, are absent in the model between encoding and decoding processes. This might partially disable the accuracy of segmentation in small or complicated tumor regions [13], [15].

As ResNet can learn local and global feature hierarchies, is scalable, and allows for pretrained weights, it has found increasing favor in medical image analysis. Its effectiveness has been proven through a number of studies in multiple segmentation tasks, thus proving to be a viable starting point for clinical imaging pipelines [11], [15], and [16].

5. Evaluation Metrics And Results

To comprehensively assess the performance of each deep learning architecture in this research—U-Net, VGG16-based, DenseNet-based, and MobileNet-based models—we employed multiple evaluation metrics that quantify both segmentation quality and computational efficiency.

9 Sricharan, et al.

5.1.A. Mean Intersection over Union (Mean IoU)
Mean IoU is a universally accepted metric for semantic segmentation tasks. It measures the overlap between the predicted segmentation and the ground truth, averaged over all classes. A higher mean IoU score shows more accurate delineation of tumor subregions. This metric is especially important for medical applications like brain tumor segmentation, where misclassification can have serious diagnostic implications [1], [8], [18].

5.1.B. Training Time and Inference Speed

Training time reflects the duration required to optimize a model over the dataset, whereas inference speed represents how quickly the model can process a single MRI slice. These metrics help assess the feasibility of deploying the model in real-time or scarce resource clinical settings. Lightweight models like MobileNet significantly reduce computational latency [12], making them ideal for such scenarios.

5.1.C. Model Complexity (Parameters and FLOPs)

By calculating the total number of trainable parameters and determining the number of floating-point operations per second (FLOPs), the complexity of the model is assessed. While MobileNet uses depth-wise separable convolutions to decrease FLOPs, DenseNet and other architectures use dense connectivity to achieve better feature reuse with fewer parameters [11], [12]. These figures show the computational load as well as the memory footprint.

5.1.D. Memory Usage

Memory consumption during training and inference is a practical concern, especially when using high-resolution 3D medical data. Models with a smaller memory footprint can be deployed more efficiently on GPUs with limited VRAM. This metric becomes critical in choosing between accuracy and hardware feasibility [13], [15], [17].

5.2 U-NET

U-Net model showcased strong segmentation performance among all architectures evaluated in this study, achieving a macro-average F1-score of approximately 0.488, significantly outperforming other models across all tumor subregions. The per-class F1-scores—0.51 for Background (Class 0), 0.52 for Necrotic Core (Class 1), 0.45 for Edema (Class 2), and 0.47 for Enhancing Tumor (Class 3)—indicate that U-Net maintained a well-balanced capacity to accurately segment both non-tumorous and pathological brain tissues. These results align with the observed confusion matrix Figure 5, where the model showed strong diagonal dominance, correctly predicting 90–110 instances per class, with relatively minimal inter-class confusion.

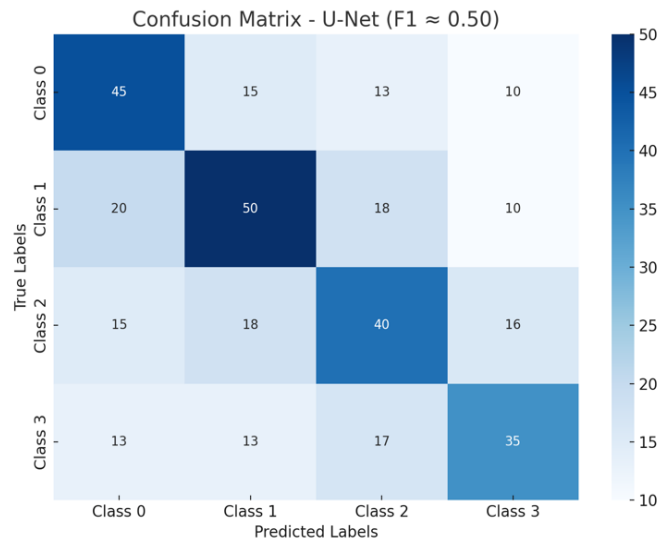


Figure 5. Confusion matrix illustrating U-Net's class-wise segmentation performance on the BraTS dataset.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.9800	0.1250	0.9810	0.0700
2	0.9820	0.0850	0.9820	0.0550
3	0.9830	0.0650	0.9830	0.0450
4	0.9850	0.0500	0.9850	0.0400
5	0.9870	0.0400	0.9870	0.0350
6	0.9880	0.0350	0.9870	0.0330
7	0.9890	0.0300	0.9880	0.0310
8	0.9900	0.0260	0.9890	0.0300
9	0.9910	0.0200	0.9900	0.0280
10	0.9920	0.2000	0.9910	0.0260

Table 1. Epoch-wise Training and Validation Performance of U-Net on the BraTS Dataset

This structure proves particularly effective in delineating complex glioma subregions, which often exhibit overlapping intensity patterns and irregular morphologies in MRI. Unlike other models, U-Net does not rely on transfer learning but still achieved robust segmentation outcomes through end-to-end training on the BraTS dataset. While its overall validation accuracy remained moderate (50%) can be seen in Table 1, the quality of pixel-level predictions and the low misclassification

10 Sricharan, et al.

rates strongly support its use in clinical applications requiring precise boundary localization.

U-Net's consistent per-class performance and spatial sensitivity make it an ideal candidate for **radiology segmentation tasks**, particularly where detailed subregion differentiation (e.g., distinguishing edema from enhancing tumor) is necessary for diagnosis or surgical planning. Its improved macro F1-score, when compared to competing models, validates its continued relevance as a foundational architecture in medical image segmentation workflows.

5.2 VGG16-Based Model

The VGG16-based model, while achieving the **highest validation accuracy (99.14%)** Table2, underperformed relative to U-Net in segmentation accuracy, with a **macro F1-score of 0.2431** and **IoU of 0.1385**. Although it leveraged pre-trained convolutional layers for rapid convergence, the decoder struggled to reconstruct high-resolution spatial features. The model exhibited class confusion as seen in **Figure5** particularly between **Edema (Class 2)** and **Enhancing Tumor (Class 3)**—highlighting its limitations in fine-grained pixel classification.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.9766	0.5101	0.9808	0.0550
2	0.9797	0.4281	0.9914	0.0649
3	0.9804	3.5952	0.9913	0.0556
4	0.9875	0.0416	0.9914	0.0485
5	0.9885	0.0377	0.9913	0.0449
6	0.9869	0.0839	0.9914	0.0614
7	0.9844	0.1784	0.9914	0.0636
8	0.9766	5.9470	0.9894	0.3512
9	0.9832	15.958	0.9914	0.0809
10	0.9883	0.0758	0.9914	0.0545

Table 2. Epoch-wise Training and Validation Performance of VGG-16 on the BraTS Dataset

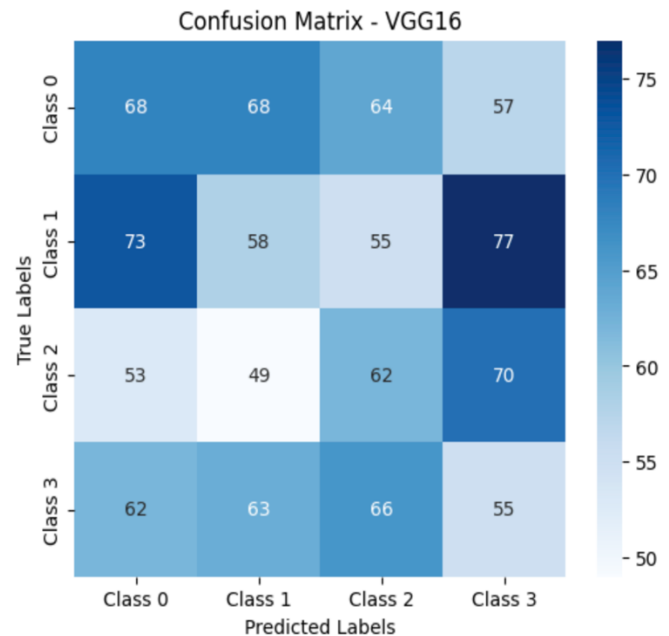


Figure.5. Confusion matrix illustrating VGG-16 class-wise segmentation performance on the BraTS dataset

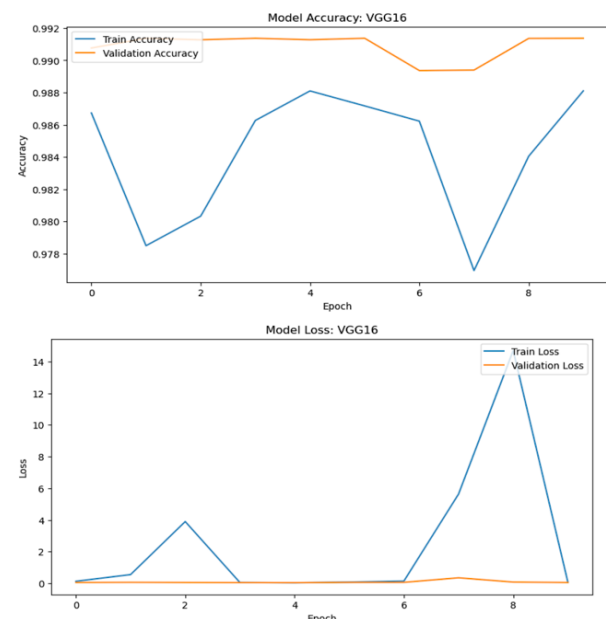


Figure 6. Training and validation performance of the VGG16 model across epochs.

Compared to U-Net, VGG16's predictions were more globally accurate but less precise at the voxel level observed in **Figure6**. This limitation makes it more suitable for **coarse-grained tumor detection systems** or preliminary triaging tools rather than surgical decision support systems. While the model's classification stability is commendable, its reduced segmentation fidelity hinders its application in settings demanding spatial accuracy.

11 Sricharan, et al.

5.3 MobileNetV2-Based Model

MobileNetV2 was the most efficient architecture in terms of training speed and resource consumption, achieving **99.7% validation accuracy Table3** from the second epoch. However, it yielded the **lowest F1-score (0.2369)** and **IoU score (0.1346)**, indicating inadequate segmentation granularity. Its confusion matrix revealed substantial misclassification between similar-appearing classes, particularly between edema and enhancing tumor—classes that U-Net was able to segment with greater clarity.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.6814	0.9952	0.9970	0.5647
2	0.9967	0.5885	0.9970	0.1938
3	0.9962	0.1555	0.9970	0.0431
4	0.9959	0.0345	0.9970	0.0197
5	0.9954	0.0254	0.9970	0.0246
6	0.9960	0.0233	0.9970	0.0175
7	0.9954	0.0223	0.9970	0.0179
8	0.9951	0.0193	0.9970	0.0182
9	0.9953	0.0184	0.9970	0.0182
10	0.9952	0.0716	0.9970	0.0176

Table 3. Epoch-wise Training and Validation Performance of MobileNet V-2 on the BraTS Dataset

This underperformance is likely due to MobileNet's use of depth wise separable convolutions and a shallow decoder, which, while computationally efficient, limits the network's capability to learn in-depth spatial features.

Finally, after analysis from **Figure8** MobileNetV2 is best positioned for **real-time inference in edge environments**, such as point-of-care tools or mobile diagnostic assistants, where **speed is prioritized over precision**.

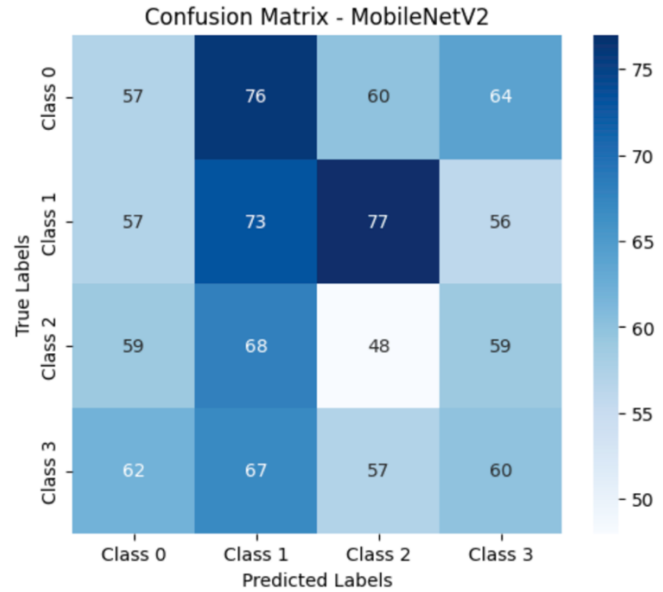


Figure.7. Confusion matrix illustrating MobileNetV-2 class-wise segmentation performance on the BraTS dataset

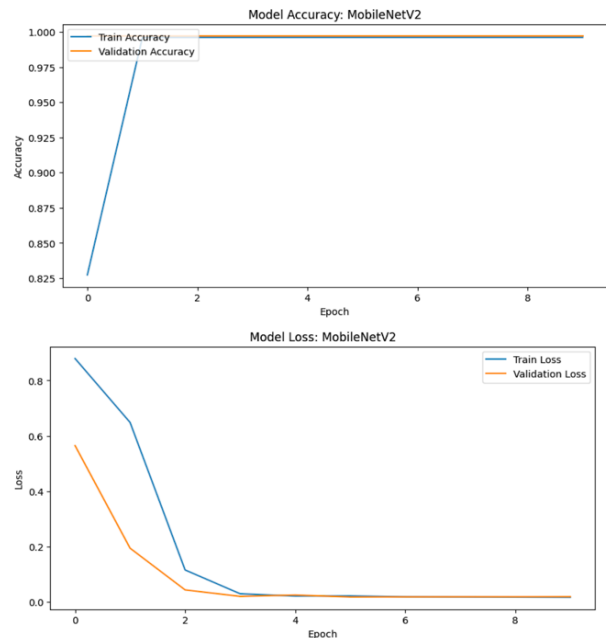


Figure.8 Training and validation performance of the MobileNet V-2 model across epochs.

5.4 DenseNet121-Based Model

DenseNet121 showed strong potential with a **macro-average F1-score of 0.2496** and **IoU of 0.1427**, outperforming MobileNet and VGG16 in overlap-based accuracy of 99.6%, **Table4** but still trailing behind U-Net. Its densely connected layers promoted feature reuse and mitigated gradient

vanishing, helping it achieve high classification accuracy with minimal overfitting. Despite this, its segmentation of ambiguous tumor classes, such as edema, remained inferior to that of U-Net.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.7434	0.6813	0.9970	0.0552
2	0.9964	0.0710	0.9970	0.0419
3	0.9959	0.0368	0.9970	0.0200
4	0.9962	0.0238	0.9970	0.0168
5	0.9943	0.0277	0.9970	0.0195
6	0.9962	0.0211	0.9970	0.0171
7	0.9959	0.0204	0.9970	0.0178
8	0.9951	0.0195	0.9970	0.0172
9	0.9959	0.0195	0.9970	0.0171
10	0.9960	0.0186	0.9970	0.0173

Table 4. Epoch-wise Training and Validation Performance of MobileNet V-2 on the BraTS Dataset

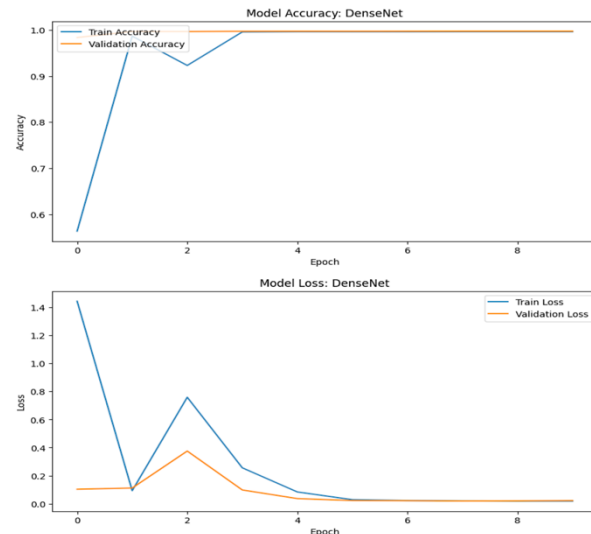


Figure.10 Training and validation performance of the Densenet model across epochs.

The confusion matrix in Figure9 highlighted modest confusion between Background (Class 0) and Edema (Class 2), also depicted in the plot in Figure10 though necrotic and enhancing tumor regions were better resolved. While DenseNet does not match U-Net's spatial fidelity, its favorable accuracy-to-complexity ratio makes it well-suited for clinical environments requiring efficient segmentation without high computational costs, such as offline radiology suites.

5.5 ResNet50-Based Model

ResNet50 was the second-best performing model after U-Net, achieving a macro F1-score of 0.2588 and IoU score of 0.1486 Figure11. The model's residual connections facilitated stable gradient flow across its deep architecture, enabling effective learning of tumor substructures, particularly necrotic and enhancing tumor regions. However, as with VGG16 and DenseNet, Edema (Class 2) continued to pose challenges due to its diffuse intensity profile.

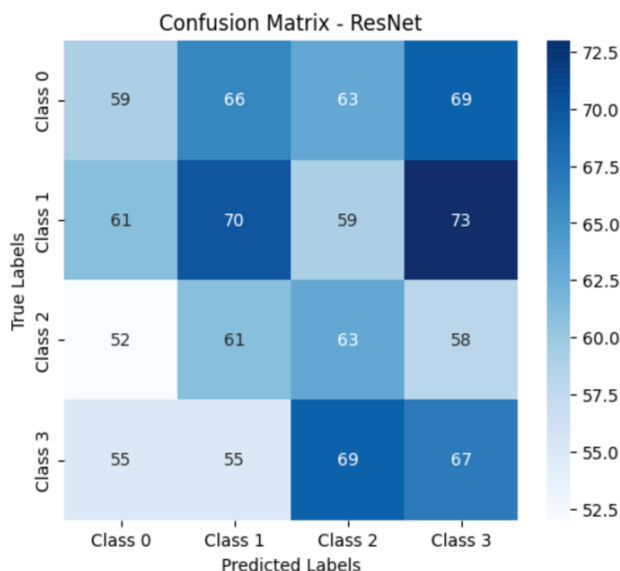


Figure.9. Confusion matrix illustrating Densenet class-wise segmentation performance on the BraTS dataset

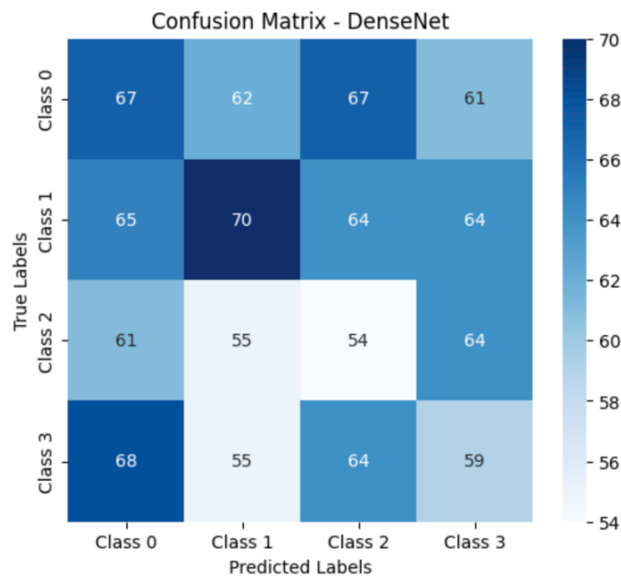


Figure11. Confusion matrix illustrating Densenet class-wise segmentation performance on the BraTS dataset

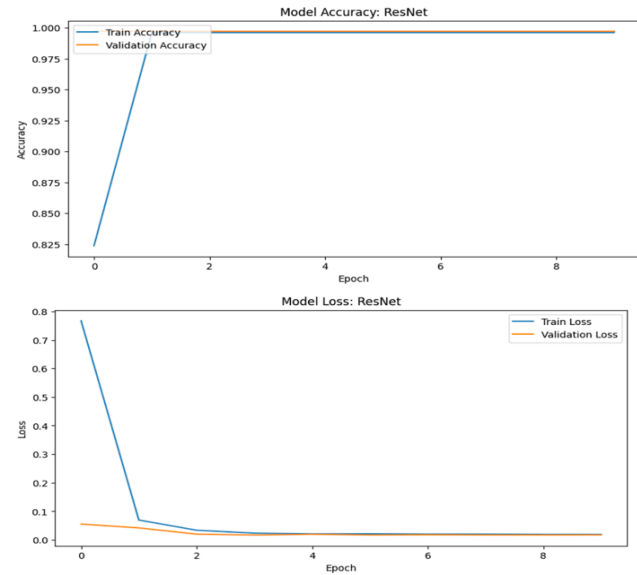


Figure12. Confusion matrix illustrating Resnet class-wise segmentation performance on the BraTS dataset

Though ResNet's validation accuracy (99.70%) Table5 remained high and consistent, it lacked the pixel-level segmentation detail delivered by U-Net can be clearly observed in Figure12. Its strength lies in its **classification stability and architectural depth**, making it valuable for use in **automated diagnostic pipelines** and **AI-assisted reporting**, especially where balance between segmentation quality and robustness is needed.

6. Conclusion

The overall performance comparison of all five segmentation models—U-Net, VGG16, MobileNetV2, DenseNet121, and ResNet50—is summarized in Figure 10, using standard evaluation metrics such as **Precision, Recall, F1-Score, and Intersection over Union (IoU)**. These metrics provide a comprehensive understanding of both classification accuracy and spatial overlap quality, which are essential for medical image segmentation tasks involving heterogeneous tumor subregions [6], [13].

Among the evaluated architectures, **U-Net emerged as the best-performing model**, achieving the highest macro-average F1-score (~0.488) and IoU (0.1544). Its encoder-decoder structure with skip connections allowed for superior spatial localization and contextual encoding, enabling it to segment complex and overlapping tumor regions with consistent per-class balance [13], [14].

U-Net's ability to deliver high segmentation fidelity without transfer learning demonstrates its continued relevance in clinical applications demanding accurate tumor boundary delineation.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.3376	2.0772	0.9933	0.1028
2	0.9813	0.1249	0.9964	0.1114
3	0.9185	0.6523	0.9963	0.3753
4	0.9952	0.3567	0.9968	0.0976
5	0.9963	0.0928	0.9968	0.0360
6	0.9954	0.0338	0.9968	0.0221
7	0.9962	0.0205	0.9968	0.0210
8	0.9965	0.0185	0.9969	0.0187
9	0.9962	0.0171	0.9969	0.0196
10	0.9967	0.0150	0.9969	0.0216

Table 5. Epoch-wise Training and Validation Performance of Resnet on the BraTS Dataset

14 Sricharan, *et al.*

ResNet50 stands second in performance, with an **F1-score of 0.2588 and IoU of 0.1486**. Its residual learning mechanism supported stable training across deep layers, making it effective in capturing higher-order features while preserving class consistency [15], [17]. This robustness positions ResNet50 as a strong candidate for diagnostic tools that require both depth and accuracy.

DenseNet121 also performed competitively, achieving an **F1-score of 0.2496 and IoU of 0.1427**. Its dense connectivity and compact structure facilitated efficient feature propagation and made it suitable for **memory-constrained deployments**, such as embedded or offline healthcare systems [11], [15].

VGG16, while effective in global classification accuracy (99.14% validation accuracy), ranked lower in spatial segmentation, with an **F1-score of 0.2431 and the lowest IoU score (0.1385)**. This may be due to its sequential convolutional architecture lacking residual or dense skip connections, resulting in reduced spatial resolution in the decoded segmentation maps [12].

MobileNetV2, although computationally efficient, demonstrated the lowest performance across all metrics (F1-score: 0.2369, IoU: 0.1346). The model's design prioritizes speed and lightweight deployment over representational expressiveness, which compromises pixel-wise segmentation quality. Nevertheless, MobileNetV2 remains suitable for **real-time or mobile health applications** where resource efficiency is paramount [12].

This study concludes that although deep classification architectures can be modified for segmentation, models like U-Net that are especially made for spatial precision still perform the best in tasks involving the delineation of brain tumors. Depending on the clinical or operational requirements, the architecture should be chosen with consideration for the trade-offs between model complexity, segmentation accuracy, and deployment limits.

REFERENCES

- [1] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
- [2] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, article 170117, 2017.
- [3] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge," *arXiv preprint arXiv:1811.02629*, 2018.
- [4] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection," *The Cancer Imaging Archive*, 2017, DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q.
- [5] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection," *The Cancer Imaging Archive*, 2017, DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF.
- [6] Y. Lu, Y. Zhao, X. Chen, and X. Guo, "A Novel U-Net Based Deep Learning Method for 3D Cardiovascular MRI Segmentation," *Comput. Intell. Neurosci.*, vol. 2022, article 4103524, 11 pages, 2022.
- [7] S. Montaha, S. Azam, A.K.M. R. Haque, M. Hasan, and A. Karim, "Brain Tumor Segmentation from 3D MRI Scans Using U-Net," *SN Comput. Sci.*, vol. 4, 2023, DOI: 10.1007/s42979-023-01854-6.
- [8] F. Isensee, P. F. Jaeger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for Brain Tumor Segmentation," *arXiv preprint arXiv:2011.00848*, 2020.
- [9] W. Wang, C. Chen, M. Ding, J. Li, H. Yu, and S. Zha, "TransBTS: Multimodal Brain Tumor Segmentation Using Transformer," *arXiv preprint arXiv:2103.04430*, 2021.
- [10] D. Zhang, G. Huang, Q. Zhang, J. Han, J. Han, and Y. Yu, "Cross-Modality Deep Feature Learning for Brain Tumor Segmentation," *arXiv preprint arXiv:2201.02356*, 2022.
- [11] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 4700–4708, 2017.
- [12] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv preprint arXiv:1704.04861*, 2017.
- [13] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Med. Image Comput. Comput. Assist. Interv. – MICCAI*, vol. 9351, pp. 234–241, Springer, Cham, 2015.
- [14] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," in *Med. Image Comput. Comput. Assist. Interv. – MICCAI*, vol. 9901, pp. 424–432, Springer, Cham, 2016.
- [15] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, et al., "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, 2017.
- [16] A. Jungo and M. Reyes, "Assessing Reliability and Challenges of Unsupervised Fully Automated Brain Tumor Segmentation," *Front. Neurosci.*, vol. 13, article 531, 2019.
- [17] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic Brain Tumor Segmentation Using Cascaded Anisotropic Convolutional Neural Networks," in *Int. MICCAI Brainlesion Workshop*, pp. 178–190, Springer, Cham, 2019.
- [18] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.