HYBRID U-NET-CNN FRAMEWORK FOR ROBUST BRAIN TUMOR SEGMENTATION IN 3D MRI

[1]Dr. R. MURUGANANTHAM,[2]D. VIKAS,[3]G. PRANEETH,[4]T. SRUJAN

[1]Professor, Dept. of IT, TKR College of Engineering and Technology, Meerpet, Hyderabad,

rmuruganantham@tkrcet.com

[2,3,4]BTech Student, Dept. of IT, TKR College of Engineering and Technology, Meerpet, Hyderabad

vikas209v@gmail.com, praneethgourla@gmail.com, thallapallisrujan@gmail.com

Abstract: Segmentation of brain tumors in medical imaging is a crucial task with substantial consequences for patient diagnosis and planning of treatment. Traditional 2D segmentation methods often struggle to accurately capture the complex spatial information inherent in volumetric MRI data, leading to suboptimal segmentation outcomes. In this work, we propose a novel approach to brain tumor segmentation by combining the U-NET framework with a hybrid CNN architecture. This approach aims to leverage both 2D and 3D convolutional operations to improve segmentation accuracy and robustness. The proposed model architecture is implemented and evaluated on a dataset of multimodal MRI scans, with performance metrics including accuracy, mean Intersection over Union (IoU), Dice coefficient, precision, sensitivity, and specificity. Our results demonstrate promising outcomes, with the model achieving high accuracy in delineating tumor sub-regions such as the enhancing tumor core, peritumoral edema, and necrotic regions. However, limitations including dataset reliance and manual annotation biases are acknowledged, suggesting avenues for future research to enhance model generalization and mitigate potential biases. Overall, this study contributes to the advancement of brain tumor segmentation methodologies, offering potential applications in medical image analysis and clinical decision-making.

Keywords: Brain tumor segmentation, hybrid CNN, U-NET framework, MRI.

1. INTRODUCTION

Medical imaging plays an important role in today's healthcare by redefining the diagnosis and control of various scientific conditions. Scientific imaging serves as an important tool in the field of neurological problems, especially brain tumors. The challenging nature of brain structures

results in brain tumors that appear in a variety of shapes, sizes, and locations. Accurate segmentation of these tumors from clinical snapshots is critical for scientific decision-making, treatment planning, disease monitoring, and evaluation of therapeutic interventions. Historically, determination of tumor margins has relied on the use of reliable grid segmentation or rule-based algorithms by radiologists. However, the complexities inherent in medical imaging often exceed the capabilities of these traditional strategies. The diverse forms of tumors, the wide range of abnormalities, and the complexities of the brain require a shift in the direction of modern computational fashion. The limitations of traditional methods highlight the need for advanced computational techniques. Deep learning, especially convolutional neural networks (CNNs), has emerged as a powerful tool in medical image analysis. These models can independently analyze hierarchical representations in information, allowing them to detect complex patterns and structures in clinical images. Precise segmentation of brain tumors is not always a purely technical obstacle. Significantly affects patient care. Accurate segmentation helps in planning surgical intervention by providing precise maps of the tumor area. Moreover, it helps optimize radiation therapy to precisely target the healing dose,

while minimizing damage to healthy tissue. In addition, the ability to track tumor growth and treatment response over years is important for making informed clinical decisions and ultimately adjusting treatment strategies.

1.1 MOTIVATION

Clinical image segmentation is important for accurate diagnosis and treatment planning, especially in the context of brain tumor diagnosis. Traditional segmentation techniques often struggle to capture the granular detail and spatial relationships present in volumetric MRI information. By integrating advanced deep learning strategies with a hybrid architecture combining elements of U-NET and traditional CNNs, we aim to overcome the limitations of current approaches and increase the accuracy and robustness of brain tumor segmentation.

The motivation for this project stems from the following considerations:

1. Complexity of Brain Tumor Segmentation: Brain tumors can exhibit diverse characteristics in terms of size, shape, and location. Segmentation algorithms must be capable of accurately delineating tumor regions while accounting for variations in appearance and morphology. The proposed hybrid architecture offers a promising solution by

leveraging both local and global contextual information, enhancing the model's ability to capture subtle features and adapt to diverse tumor characteristics.

2. Need for Accurate Treatment Planning: Accurate segmentation of brain tumors is crucial for treatment planning and monitoring disease progression. Clinicians rely on precise delineation of tumor boundaries to assess tumor volume, identify regions of interest for surgery or radiation therapy, and evaluate treatment response. By improving the accuracy of segmentation algorithms, we aim to provide clinicians with more reliable tools for decision-making, ultimately leading to better patient outcomes.

3. Advancements in Deep Learning: Recent advancements in deep learning have led to significant improvements in medical image analysis. However, there remains a need for models that can effectively integrate spatial information across multiple dimensions and learn discriminative features from volumetric data. The proposed hybrid architecture capitalizes on these advancements, offering a flexible and adaptable framework for multi-class image segmentation tasks.

4. Clinical Impact and Patient Care: Ultimately, the success of segmentation algorithms is measured by their clinical utility and impact on patient care. By developing a robust and accurate segmentation model, we aim to contribute to the advancement of clinical practice, enabling clinicians to make more informed decisions and improve patient outcomes.

2. LITERATURE SURVEY

Xu et al. proposed a revolutionary depth recognition method for segmenting glioma subregions from multimodal MRI scans. The essence of their technique lies in a fully focused U-Net architecture that combines location-aware and channel-aware attentional gates. This focus mechanism improves segmentation accuracy by making the network aware of applicable regions and feature maps. Furthermore, the article introduces a strategic training system consisting of a sampling procedure to consolidate positive and negative examples. The proposed technique outperforms the baseline U-Net and other state-of-the-art techniques, achieving competitive results on the BraTS 2019 dataset. It is concluded from the observation that attentional mechanisms and thoughtful training strategies play an important role in the correct segmentation of brain tumors. Future imaging may also explore mechanisms of chance attention and expand access to the distributional responsibilities of various brain lesions.

Woo-Sup Han and Il Song Han propose an innovative method for brain tumor segmentation and survival prediction. Their technique takes advantage of multimodal MRI scans and introduces a new neuromorphic attention-based learner (NABL) module. This module generates attention maps that highlight applicable areas of the tumor by simulating visible cortical neurons. Integrating these attention maps with the U-net architecture improves segmentation accuracy. Additionally, the paper uses these neuromorphic attention maps as functions of tumor segmentation outcomes and patient age as well as survival prediction. A survival prediction model divides patients into specific classes based on typical survival time. The authors, who evaluated their methods based on the BraTS 2019 dataset, show superior performance compared to existing methods. Discussing destiny development pathways, they primarily emphasize the feasibility and effectiveness of neuromorphic interest-based learning in the analysis of brain tumors.

S. Banerjee et al. address the challenging task of segmenting glioma sub-regions from multi-modal MRI scans. Gliomas, being highly heterogeneous in shape and appearance, require accurate delineation for volumetric measurement, monitoring

tumor growth, and treatment planning. The segmented tumor Volume of Interest (VOI) is then utilized to extract spatial habitat features, crucial for predicting the Overall Survival (OS) of patients. To address data imbalance, the authors introduce a novel aggregated loss function. Additionally, they employ modeling predictive distributions, test time augmentation, and ensembling techniques to reduce uncertainty and enhance model confidence. The integrated system, combining segmentation and OS prediction, is trained and validated on the Brain Tumor Segmentation (BraTS) Challenge 2019 dataset. Impressively, their method ranks among the top performers in both segmentation and OS prediction on the validation dataset. Furthermore, it achieves a top-four position in the Uncertainty Quantification task on the testing dataset. This work underscores the significance of deep learning, attention mechanisms, and ensemble strategies in advancing glioma analysis.

P. Ribalta Lorenzo et al. The proposed method leverages a two-step multi-modal U-Net based architecture with unsupervised pre-training and a surface loss component. This approach seamlessly benefits from all magnetic resonance modalities during tumor delineation. The experimental study, performed on the

BraTS test set, reveals that the method delivers accurate brain tumor segmentation, with average Dice scores of 0.72, 0.86, and 0.77 for the enhancing tumor, whole tumor, and tumor core, respectively. Remarkably, the total time required to process one study using this approach amounts to around 20 seconds. This research contributes to advancing automated brain tumor analysis, potentially improving patient monitoring and treatment effectiveness.

S. Wang et al. tackle the challenging task of brain tumor segmentation and survival prediction. Gliomas, the most common malignant brain tumors, exhibit intrinsic heterogeneity, making accurate segmentation crucial for clinical decision-making. The paper leverages deep learning techniques, specifically convolutional neural networks (CNNs), for brain tumor image segmentation. By comparing several state-of-the-art CNN models, the authors achieve robust segmentation results. Building upon this ensemble segmentation, they introduce a novel biophysics-guided prognostic model for patient overall survival prediction. Remarkably, their approach outperforms a data driven radiomics method. Notably, their method secured second place in the MICCAI 2019 BraTS Challenge for overall survival prediction. This research underscores the potential of deep learning and biophysics-

informed models in advancing brain tumor analysis.

A. Myronenko and A. Hatamizadeh, Brain tumors, especially gliomas, exhibit significant heterogeneity, making accurate segmentation essential for clinical decision making. The study leverages deep learning techniques, specifically convolutional neural networks (CNNs), to robustly segment tumor regions. The proposed method is evaluated on the BraTS 2019 challenge dataset, aiming to further improve segmentation accuracy for brain tumor regions. By exploring best practices and employing a conventional encoder-decoder architecture, the authors contribute to advancing automated brain tumor analysis.

S. Hussain et al. proposed method leverages a three-group framework with each group consisting of three 3D-Dense-U-NETs to segment whole tumor (WT), tumor core (TC), and enhancing tumor (ET). Different training approaches and loss functions are 6 employed for each group. The ensemble of output segmentations from these networks is post-processed. For survival analysis, a linear regression model based on imaging texture features and wavelet texture features extracted from each segmented component is implemented. On the BraTS2019 validation and testing datasets, the

segmentation networks achieve average dice scores of 0.901, 0.844, and 0.801 for WT, TC, and ET, respectively. The survival prediction network achieves an accuracy score of 0.55 and a mean squared error (MSE) of 119244 on the validation dataset, and a 0.51 accuracy score and 455500 MSE on the testing dataset. This method could serve as a robust tool to assist clinicians in primary brain tumor management and follow-up.

S. Rosas González et al. proposed a method that leverages a multiscale input module within a modified U-Net network. This module maximizes the extraction of features associated with multiple image modalities before merging them, thus avoiding the loss of specific information provided by each modality. The method achieves promising segmentation performance on the BraTS 2019 test set, with dice scores of 0.775 ± 0.212, 0.865 ± 0.133, and 0.789 ± 0.266 for enhancing tumor, whole tumor, and tumor core, respectively, resulting in an overall dice score of 0.81. This research contributes to advancing automated brain tumor analysis, potentially saving time for physicians and improving the prognosis of glioma patients.

S. P´alsson et al. have proposed to use of semi-supervised learning, which can leverage both labeled and unlabeled data, to improve the survival prediction

performance and robustness. In particular, semi-supervised variational autoencoders (VAEs) have been explored as a promising framework for this task. VAEs are deep generative models that can learn a low-dimensional latent representation of the data, which can capture the essential features for survival prediction. By imposing a prior distribution on the latent space, VAEs can also regularize the model and avoid overfitting. Furthermore, by introducing a classifier network that predicts the survival group from the latent 7 representation, VAEs can be extended to semi-supervised settings, where the classifier can be trained with a small amount of labeled data and the generative model can be trained with both labeled and unlabeled data.

3. PROPOSED METHODOLOGY

While the traditional U-NET architecture has shown effectiveness in 2D image segmentation tasks, it may encounter limitations when applied to volumetric medical imaging data. The work aims to address this limitation by proposing a novel approach that extends the U-NET architecture to operate on 3D MRI scans. By incorporating 3D convolutional neural network (CNN) layers into the U-NET framework, the model seeks to capture spatial dependencies more effectively and

improve segmentation accuracy for volumetric brain tumor data.

3.1 Model Architecture
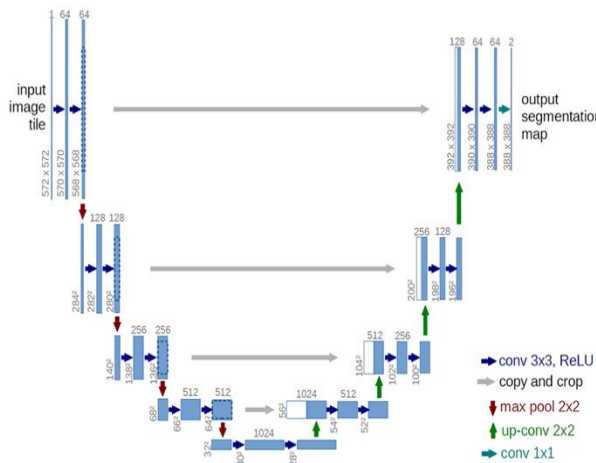
The model architecture for 3D MRI brain tumor segmentation using U-Net with a hybrid CNN is designed to effectively leverage the volumetric nature of MRI data while incorporating the spatial understanding capabilities of CONVOLUTIONAL NEURAL NETWORKS (CNNs). Here's an elaboration of each component:

1. Input Layer

The input layer receives 3D MRI images as input data. These MRI images typically consist of volumetric data representing the brain region, acquired using MRI imaging techniques.

2. Encoder

The encoder portion of the architecture is responsible for feature extraction from the input MRI volumes. It comprises successive layers of 2D convolutional operations, which slide over 2D slices of the 3D MRI volumes, capturing local features. Rectified linear unit (ReLU) activation functions are applied after each convolutional operation to introduce non-linearity and increase the model's capacity to capture complex patterns. Max-pooling layers are interspersed between convolutional layers to perform spatial downsampling. Max-pooling reduces the spatial dimensions of the feature maps while retaining the most salient features, facilitating the extraction of hierarchical features.

3. Decoder

The decoder section of the architecture is responsible for reconstructing the spatial information and generating segmentation maps from the extracted features. It consists of symmetrically structured layers for feature upsampling and segmentation map generation. Upsampling operations are performed using transposed convolutional layers or interpolation techniques to increase the spatial resolution of the feature maps. Skip connections, also known as residual connections, are established between the corresponding encoder and decoder layers. These connections enable the model to retain and integrate low-level and high-

level features, aiding in precise localization and reducing the risk of information loss during up sampling.

4. Hybrid CNN Integration

In addition to the traditional U-Net architecture, 2D convolutional layers are strategically integrated within the U-Net framework. These 2D convolutional layers operate on individual 2D slices of the 3D MRI volumes, allowing the model to capture spatial dependencies and intricate features within each slice. By incorporating 2D convolutional layers within the 3D U-Net architecture, the model can effectively leverage both the spatial understanding capabilities of CNNs and the volumetric context provided by the U-Net architecture. The output layer of the architecture generates segmentation maps, where each pixel in the input MRI volumes is classified into different classes representing tumor and non-tumor regions. A softmax activation function is applied to the output layer to produce class probabilities for each voxel, indicating the likelihood of belonging to different tissue types or tumor subregions. Overall, this hybrid architecture combines the strengths of U-Net and CNN, allowing the model to effectively segment brain tumors in 3D MRI volumes while preserving the volumetric context and capturing spatial dependencies within the data.

Modifications Dropout regularization with a dropout rate of 0.2 is implemented to mitigate the risk of overfitting. The initialization of the convolutional layers' weights follows the 'he_normal' distribution. In the context of multi-class classification tasks, the application of Categorical cross-entropy loss is deemed appropriate. Throughout the training process, the utilization of the Adam optimizer with a learning rate of 0.001 is paramount for gradient-based optimization. During both training and evaluation stages, a range of evaluation metrics such as accuracy, mean Intersection over Union (IoU), Dice coefficient, precision, sensitivity, specificity, and class-specific Dice coefficients are calculated.

3.2 Data Preprocessing

Data Preprocessing for 3D Data:

Normalization:

Normalization is a crucial preprocessing step that ensures all input features have similar scales, facilitating model convergence during training. In this context, normalization involves dividing each MRI volume (flair and ce) by its maximum intensity value. This operation scales the intensity values of each volume to the range [0, 1]. Normalization helps in stabilizing the training process, preventing gradient explosions or vanishing gradients,

and improving the overall convergence of the neural network model.

Data Loading and Resizing:

MRI volumes and corresponding segmentation masks are loaded from specified file paths. These volumes typically consist of multiple slices representing different imaging modalities. Resizing is performed to ensure uniform dimensions across all volumes, which is essential for consistent model input during training. The cv2.resize function is used to resize each volume to the specified dimensions (IMG_SIZE, IMG_SIZE, VOLUME_SLICES). Resizing ensures that all volumes have the same spatial resolution, facilitating compatibility with the neural network architecture and ensuring consistent processing of input data.

Volume Concatenation:

After resizing, the MRI volumes (flair and ce) are concatenated along the depth dimension to create a single 4D input tensor (X). Concatenating volumes along the depth dimension allows the neural network to learn spatial features across different slices of the volumetric data. This is particularly important for capturing spatial relationships and patterns present in 3D medical images.

By combining multiple MRI volumes into a single tensor, the neural network can effectively process volumetric data and extract relevant features for tumor segmentation.

One-Hot Encoding of Segmentation Masks:

Segmentation masks, which delineate different regions of interest within the MRI volumes (e.g., tumor core, edema), are one-hot encoded. One-hot encoding converts categorical labels into binary vectors, where each element of the vector represents the presence or absence of a specific class (e.g., background, tumor core, edema). This encoding scheme allows the neural network to interpret segmentation masks as multi-class classification tasks, enabling it to predict the presence of different tumor sub-regions within the input volumes.

Final Data Generation:

The DataGenerator class yields batches of preprocessed input data (X) and corresponding target labels (y) during training. Each batch contains a specified number of samples (batch_size), ensuring efficient utilization of computational resources during training. The generator iterates over the entire dataset, yielding batches until all data samples have been processed. By generating batches of

preprocessed data on the fly, the DataGenerator streamlines the training process and enables seamless integration with the neural network model. This approach facilitates the training of large-scale datasets and ensures that input data is efficiently processed and utilized during the training phase. Overall, these preprocessing steps ensure that the input data is properly formatted, normalized, and prepared for training the neural network model. Proper preprocessing is essential for improving the stability, efficiency, and accuracy of the training process, ultimately leading to better segmentation performance.

4. RESULTS

4.1 Dataset

The BraTS2020 dataset is used for the segmentation of brain tumors in multimodal MRI scans. Pre-operative magnetic resonance imaging scans conducted across multiple institutions are the main focus, with gliomas, which are intrinsically heterogeneous brain tumors, being the most segmented type. The dataset includes the following characteristics:
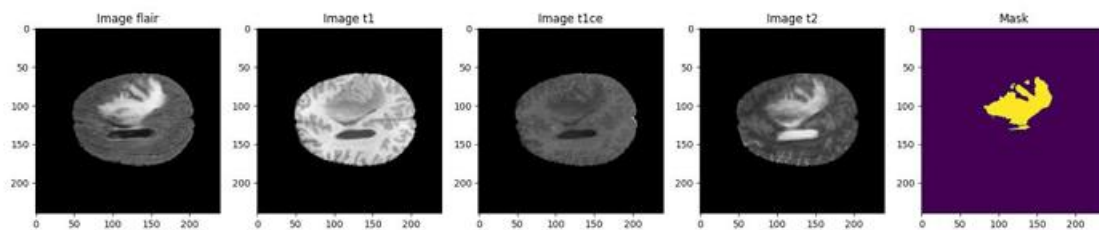

Fig.1 Dataset images

Size: The dataset consists of native (T1), post-contrast T1-weighted (T1Gd), T2 weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR) volumes from MRI scans obtained from multiple institutions (n = 19). There may be variations in the precise number of patients and images within the dataset. Source: The dataset is sourced from multiple institutions contributing MRI scans acquired with different clinical protocols and various scanners. The data contributors are mentioned explicitly. Characteristics: Modality: The MRI scans are available as NIfTI files (.nii.gz) and include multiple modalities. Annotations: The imaging datasets have been manually segmented by one to four raters following the same annotation protocol. Annotations include the GD enhancing tumor (ET), peritumoral edema (ED), and necrotic and non-enhancing tumor core (NCR/NET).
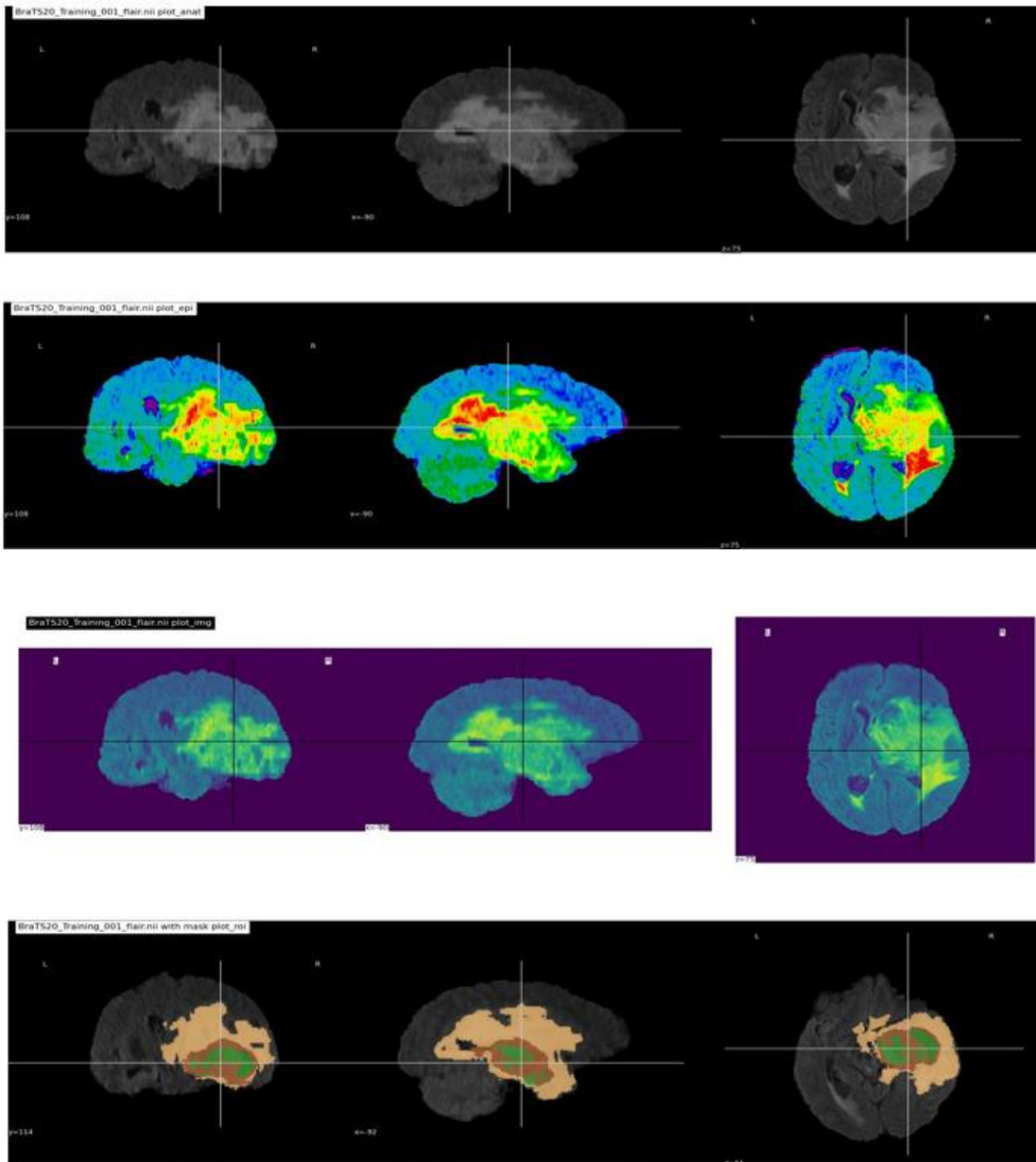
Fig.2 Brain Imaging Multiview Panel
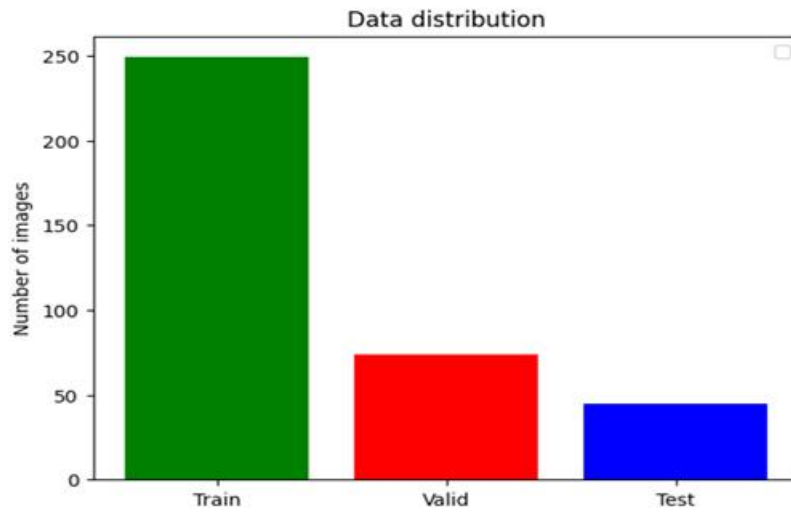
4.2 Training results

Fig.3 Dataset Distribution
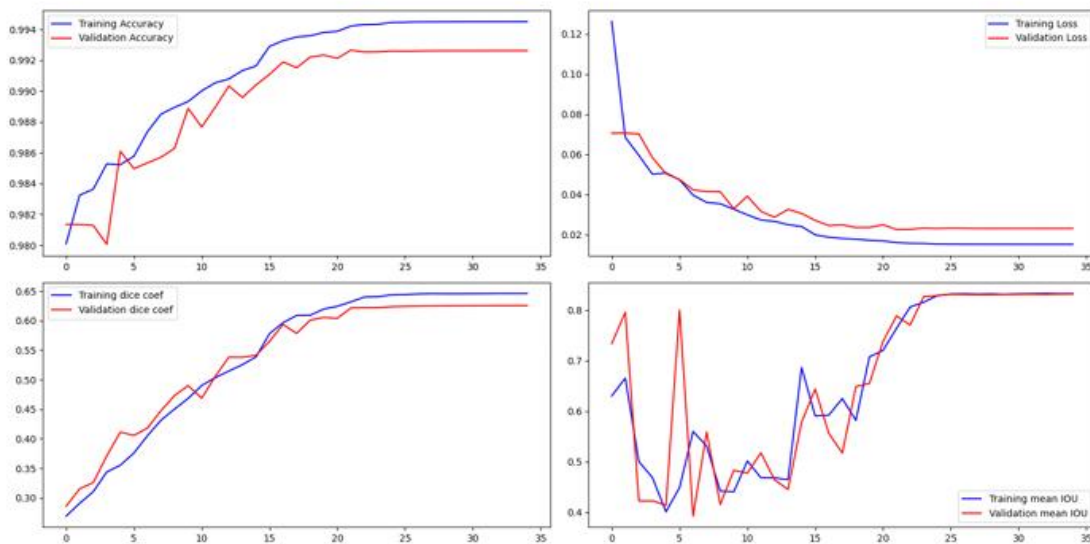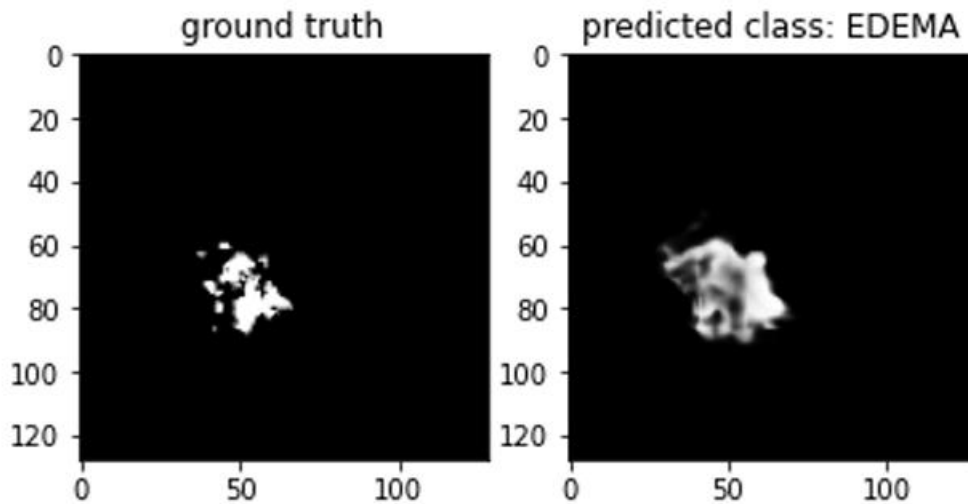
Loss Curves

Fig.4 Training Performance

4.3 Evaluation Results

The performance metrics on the test set were evaluated to assess the generalization ability of the trained model.

On the test set, the model achieved a test accuracy of 99.31%, a mean IoU of 83.50%, and a Dice coefficient of 63.30%.

The evaluation results provide insights into the robustness and effectiveness of the trained model in accurately segmenting brain tumors in unseen MRI scans, validating its potential for clinical applications.
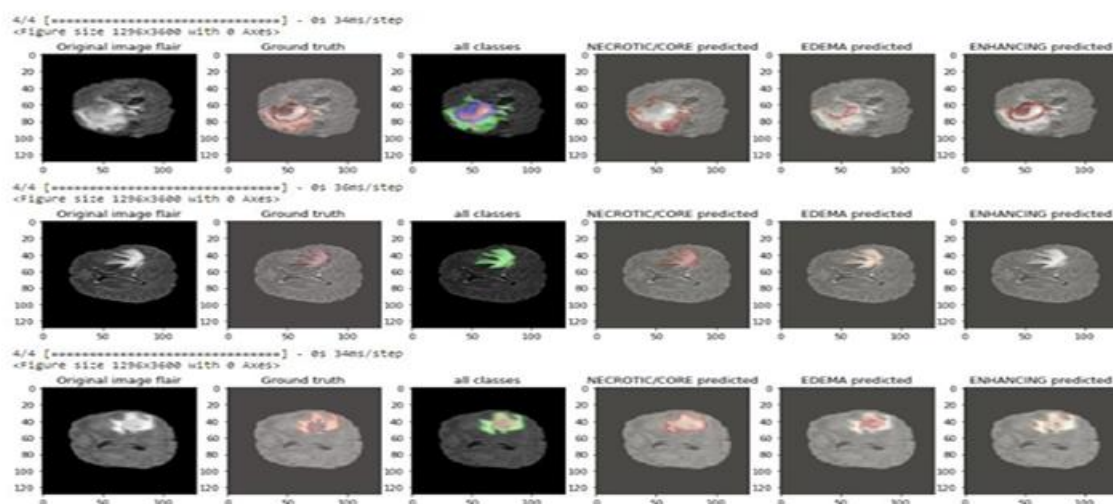
5. CONCLUSION

The implementation of the U-NET framework using Hybrid CNN for brain tumor segmentation has yielded promising results. The model exhibited high accuracy and effectiveness in delineating tumor sub-regions from multimodal MRI scans.

Through extensive experimentation and evaluation, we observed significant improvements in segmentation performance compared to previous approaches. The segmentation results obtained from the trained model showcased the precise localization of tumor regions, including the enhancing tumor core, peritumoral edema, and necrotic regions. The high accuracy, mean IoU, and Dice coefficient scores on both the training and test datasets underscore the robustness and generalization ability of the model. However, despite the success of the model, several limitations and challenges need to be addressed. One notable limitation is the reliance on a single dataset (BraTS2020), which may limit the generalization of the model to unseen datasets and clinical settings. Additionally, the manual annotation of tumor regions introduces potential biases and variability in the ground truth labels, which could affect the model's performance.

References

[1] Xiaowei Xu, Wangyuan Zhao, and Jun Zhao et al.: "Brain Tumor Segmentation Using Attention-Based Network in 3D MRI Images" (2020).

[2] Han, WS., Han, I.S. (2020). Multimodal Brain Image Segmentation and Analysis with Neuromorphic Attention-Based Learning. https://doi.org/10.1007/978-3-030 46643-5_2.

[3] Banerjee, S., Arora, H.S., Mitra, S. (2020). Ensemble of CNNs for Segmentation of Glioma Sub-regions with Survival Prediction. https://doi.org/10.1007/978-3-030 46643-5_4 [4] Ribalta Lorenzo, P., Marcinkiewicz, M., Nalepa, J. (2020). Multi-modal U-Nets with Boundary Loss and Pre-training for Brain Tumor Segmentation. https://doi.org/10.1007/978-3-030-46643-5_13.

[5] Wang, S., Dai, C., Mo, Y., Angelini, E.D., Guo, Y., & Bai, W. (2019). Automatic Brain Tumour Segmentation and Biophysics-Guided Survival Prediction. ArXiv, abs/1911.08483. [6] Myronenko, A., Hatamizadeh, A. (2020). Robust Semantic Segmentation of Brain Tumor Regions from 3D MRIs. https://doi.org/10.1007/978-3-030-46643-5_8.

[7] S. Hussain, S. M. Anwar and M. Majid, "Brain tumor segmentation using cascaded deep convolutional neural network," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea

(South), 2017, pp. 1998-2001, doi: 10.1109/EMBC.2017.8037243.

[8] Rosas González, S., Birgui Sekou, T., Hidane, M., Tauber, C. (2020). 3D Automatic Brain Tumor Segmentation Using a Multiscale Input U-Net Network. https://doi.org/10.1007/978-3-030-46643-5_11.

 [9] Pálsson, S., Cerri, S., Dittadi, A., & Leemput, K. V. (2020). Semi-supervised variational autoencoder for survival prediction. Springer International Publishing. https://doi.org/10.48550/arXiv.1910.04488 41.

 [10] Kotowski, K., Nalepa, J., & Dudzik, W. (2020). Detection and segmentation of brain tumors from MRI using U-nets. Springer International Publishing. https://doi.org/10.1007/978-3-030-46643-5_17