Advanced Fusion of 3D U-Net-LSTM Models for Accurate Brain Tumor Segmentation

Ravikumar Sajjanar*, Umesh D. Dixit
Department of Electronics & Communication Engineering, BLDEA's V. P. Dr. P. G. Halakatti College of Engineering and Technology, Vijayapura–586 103 (Affiliated to Visvesvaraya Technological University, Belagavi-590018), Karnataka, India

***Abstract*—Accurate detection and segmentation of brain tumors are essential in tomography for effective diagnosis and treatment planning. This study presents advancements in 3D segmentation techniques using data from the Kaggle BRATS 2020 dataset. To enhance the reliability of brain tumor diagnosis, innovative approaches such as Frost filter-based preprocessing, UNet segmentation architecture, and Long Short-Term Memory (LSTM) segmentation are employed. The methodology starts with data preprocessing using the Frost filter, which effectively reduces noise and enhances image clarity, thus improving segmentation accuracy. Subsequently, the UNet architecture is utilized to precisely segment brain tumor regions. UNet's ability to capture contextual information and its efficient use of skip connections contribute to accurately delineating tumor boundaries in three-dimensional space. Additionally, the temporal aspect of brain tumor progression is addressed by employing an LSTM network, which increases segmentation accuracy. The LSTM algorithm integrates temporal patterns in sequential imaging data, enabling reliable segmentation of tumor presence and characteristics over time. By analyzing the ordered sequence of continuous MRI scans, the LSTM framework achieves more precise and adaptable tumor recognition. Evaluation results based on the Kaggle BRATS 2020 dataset demonstrate significant improvements in segmentation and segmentation performance compared to previous methods. The proposed approach enhances the accuracy of tumor boundary delineation and the ability to classify tumor types and track temporal changes in tumor growth. The "U-Net-LSTM" method achieves an accuracy of 98.9% in segmentation tasks, showcasing its superior performance compared to other techniques. This method is implemented using Python, underscoring its efficacy in achieving high accuracy in segmentation tasks.**

***Keywords*—Brain tumor segmentation; frost filter pre-processing; UNet architecture; LSTM; kaggle BRATS 2020 dataset**

I. INTRODUCTION

In the US, roughly 23,000 additional instances of tumors in the brain are expected to be detected year 2015 [1]. Which is a particularly frequent type of brain tumor, and may vary from a low to a high level, based on the person's life prognosis (e.g., a few decades or fewer). Both chemotherapy and radiation can halt the expansion of brain cancers that can't be eliminated through operation [2]. Although certain types of tumors, including meningiomas, remain readily divided, gliomas and glioblastomas become considerably harder to locate. These malignancies are frequently dispersed, weakly compared, and have tentacle-like features that render tumors hard to divide [3]. A different approach basic challenge of dividing tumors in the brain is the fact that tumors can occur wherever there is the central nervous system, in practically every size and shape. In addition, whereas images created with an X-ray machine scan as well the dimension of pixel data in MRI images cannot be uniform [4]. Based on the kind of MR equipment utilized with the data collection methodology comparable tumorous tissues could show significantly varying shades of gray whenever seen across distinct institutions [5]. The main objective of brain tumor imagery assessment is to acquire tailored patients' critical therapeutic data as well as analytical characteristics. The details incorporated into the multimodal images might determine and evaluate treatments once an illness is diagnosed and then restricted, eventually contributing to understanding enabling diagnostic setting, and medication of illness [6]. The steps involved can be depicted graphically as a pyramidal. Specific approaches must be used at all levels inside the hierarchy to analyze facts, gather, categorize, display depict knowledge. Furthermore, to gain useful clinical expertise or data so that health-related diagnoses and decisions might be generated, the details must be represented at an elevated degree of abstraction [7]. The primary goal of segmenting an image is to divide a picture into incompatible sections to ensure every area is geographically continuous and its pixels inside it remain uniform according to a preset standard. This description is an important constraint for many division techniques, in particular when establishing and identifying "unusual cell forms," as the malignant cells that must be divided are anatomical components wholly frequently non-rigid and multifaceted in arrangement, change enormously in terms of dimensions and location, and change significantly compared to individuals to individuals [8].In the situation of brain tumors, division entails distinguishing between various parts of the tumor, including solid or actively aggressive tumors, swelling, and death, and healthy brain cells, including cerebral gray matter , white matter, and the fluid that surrounds it [9].

Accurately estimating the corresponding amount of brain tumor parts is crucial to tracking development, scheduling radiation therapy, assessing outcomes, and conducting follow-up investigations. This requires good tumor delineation [10]. Human specialists have substantial obstacles in manually segmenting tumors due to the variety in morphology and the requirement to examine many pictures from distinct MRI sequencing to accurately diagnose cell type. The painstaking task is difficult, susceptible to error by humans, and causes high within and inter-rater variation. In most neurological tumor examinations, an abundance of aberrant cells is apparent[11]. Nonetheless, reliable overall consistent identification as well as characterization of anomalies remains

challenging. Conventional methods of imaging, like MRI, are useful in detecting tumors of the brain, however, segmentation by hand is time-intensive, laborious, and susceptible to inter-observer variation. The development of computerized 3D segmentation tools has transformed this procedure, allowing for quicker and more exact identification of tumor areas. A few of the important advances propelling advancement in 3D segmentation is the use of Deep learning-based approaches, particularly multilayer neural networks [12]. These methods excel in extracting topological characteristics from voxel images, enabling very accurate and efficient brain cancer separation. CNNs are capable of accurately capturing the complicated spatial connections and brightness variations that distinguish distinct different kinds of tumors when trained on massive amounts of captioned MR imaging images [13]. Furthermore, the use of multifaceted imaging information, such as an MRI, diffusion-weighted images, and positron emission tomography, has improved the reliability of 3D segmentation approaches [14]. The key contributions of the suggested framework are mentioned below.

- The implementation of the Frost filter-based preprocessing methodology marks a significant enhancement in brain tumor imaging. This approach efficiently eliminates noise and enhances the luminosity of brain images, providing a robust foundation for subsequent segmentation tasks. By improving the clarity and quality of the images, the Frost filter significantly boosts tumor identification accuracy, ensuring that the segmentation process starts with the best possible data. This leads to more precise delineation of tumor boundaries and enhances the overall efficiency and reliability of tumor detection and analysis.
- The utilization of the UNet segmentation architecture makes a significant contribution by enabling precise identification of tumor regions. UNet's design, which effectively captures contextual information and leverages skip connections, ensures accurate delineation of tumor boundaries in three dimensions. This capability significantly enhances the precision of segmentation, as it allows the model to integrate both local and global features, thereby providing a comprehensive understanding of the tumor's spatial structure. The ability to accurately define tumor boundaries is crucial for improving the reliability and accuracy of segmentation, ultimately leading to better diagnostic and therapeutic outcomes.
- The incorporation of the LSTM segmentation algorithm represents a pivotal advancement, as it adeptly addresses the temporal aspects of tumor development. By effectively capturing the dynamic changes over time, the LSTM algorithm ensures consistent and reliable identification of tumor presence and characteristics as they evolve. This temporal sensitivity significantly enhances the accuracy of segmentation, allowing for a more nuanced and precise analysis of tumor progression. The ability to track and integrate temporal patterns into the segmentation process not only improves diagnostic accuracy but also provides valuable insights for monitoring treatment response and planning future interventions.
- Using the ordered series of continuous MRI images inside the LSTM architecture provides a fresh approach to tumor acknowledgment, which leads to higher accuracy and adaptability in recognizing tumors when compared with conventional techniques, thus boosting total segmentation accuracy.
- The suggested methodology's assessment on the Kaggle database BRATS 2020 dataset shows substantial gains in precision of segmentation as well as categorization productivity when contrasted with current methods, demonstrating the efficacy of the paired Frost filter preliminary processing, UNet segmentation, and LSTM segmentation methods for precise brain tumor recognition in imaging.

The organization of the paper includes related works, problem statements, and methodology in Sections II, III, and IV. The results are given in Section V. Section VI concludes the paper.

II. Related Work

Jin Liu et al., [15] suggested a technique based on deep learning for segmenting brain tumors utilizing multifaceted MRI data, utilizing a convolutional neural network comprising several layers of convolution overall remaining connections to improve both precision and effectiveness. This methodology marks an important milestone in the discipline that has achieved significant advancements during the past 20 years through the introduction of methodologies such as CNNs, U-Net variants, and GANs, among mixed techniques. Adequate initial processing, strong evaluation measures, and publicly available datasets such as BraTS have all contributed to future advances. Yet, issues like variance in MRI techniques, tumor variation, and the requirement for larger annotation-laden datasets persist.

Sergio et al., [16] presented an autonomous segmentation of brain tumors approach using Convolutional Neural Networks using small 3×3 kernels. This allows for a more complex design and reduces excessive fitting because of fewer network weights. They used intensity normalizing and data enrichment as pretreatment measures to improve the efficiency of segmentation. The technique they used was verified utilizing the Brain Tumor Segmentation Challenge 2013 (BRATS 2013) database, and it won first place for whole, core, and improved areas, with Dice Similarities Coefficients of 0.88, 0.83, and 0.77, accordingly. In addition, they finished first in the public assessment. Applying the same approach in the BRATS 2015 Challenge, which is they finished second with Dice Comparison parameters of 0.78, 0.65, and 0.75 for the full, core, and improving areas, etc.

Paul et al., [17] gave a technique for the division that uses nnU-Net. The unmodified nnU-Net baseline generated acceptable outcomes but including BraTS-specific changes like further processing, region-based training, stronger data enhancement, and multiple small pipeline modifications substantially enhanced segmentation performance. By

reviving the BraTS rankings algorithm to determine the best nnU-Net variation, their technique won the BraTS 2020 contest with Dice scores of 88.95, 85.06, and 82.03, and HD95 values of 8.498, 17.337, and 17.805 for total tumor, tumor core, and augmenting tumor, accordingly.

Xiaomei et al., [18] proposed a brain tumor segmentation method that involves instruction of a deep learning algorithm using 2D patch images as well as slices in three stages: first, instructions FCNNs alongside image areas; second, learning CRFs as Recurrent Neural Networks with image slices while keeping FCNN parameters constant; and at last, fine-tuning both FCNNs and CRF-RNN employing image segments. They developed three segmentation algorithms using image slices from the axial, coronal, and sagittal views, then combined these with a voting-based fusion technique for classification. The approach they tested with information from the BRATS 2013, 2015, and 2016 challenges, proved that it could construct an analysis of segmentation utilizing Flair, T1c, and T2 scans while outperforming models employing Flair, T1, T1c, and T2 data.

Various methods have been proposed for brain tumor segmentation using deep learning techniques. One method employed convolutional neural networks with multiple convolutional layers and residual connections, showing significant advancements and addressing challenges like MRI variability and tumor heterogeneity. Another approach utilized CNNs with small 3×3 kernels, enhancing segmentation performance through intensity normalization and data augmentation, achieving top rankings in the BRATS 2013 and 2015 challenges. The nnU-Net framework, modified with BraTS-specific adjustments such as postprocessing and region-based training, achieved first place in the BraTS 2020 competition.

III. Problem Statement

Despite significant advancements in deep learning-based methods for brain tumor segmentation using multi-modal MRI images, several challenges persist. The variability in MRI protocols, tumor heterogeneity, and the need for large annotated datasets continue to hinder the accuracy and efficiency of segmentation algorithms. While techniques such as convolutional neural networks (CNNs), U-Net variants, GANs, and hybrid methods have greatly improved the field, the development of reliable and robust segmentation models is still impeded by these challenges. Effective preprocessing and robust evaluation metrics are essential, but overcoming the intrinsic variability and obtaining extensive high-quality labeled data remain critical issues to address [15].

IV. Proposed UNet-LSTM Methodology for Brain Tumor Identification

The suggested operational technique for increasing the accurate identification and segmentation of brain tumors in tomography begins with data pre-processing using the Frost filter, which reduces noise and increases luminance in brain scans, thereby boosting segmentation accuracy. Following that, the UNet segmentation architecture is used to precisely outline tumor areas by making use of its capacity to record contextual data and effectively bypass connections, which is especially useful for three-dimensional tumor border determination. To deal with the psychological element of tumor evolution and increase the precision of segmentation, the LSTM network is incorporated, which successfully captures temporal trends in successive image information for consistent tumor segmentations across time. The LSTM structure, which takes advantage of the structured series of ongoing MRI scans, allows for greater accuracy and adaptable tumor detection. Evaluation of the Kaggle BRATS 2020 database reveals considerable improvements in precision for segmentation and segmentation effectiveness over earlier methods. The suggested method increases not simply the reliability of tumor border separation, but its ability to distinguish between tumor kinds and follow periodical variations in the development of tumors. Fig. 1 represents a workflow of the proposed UNet-LSTM Methodology.

Once the content has been edited, it is prepared for the pattern. Download the design document with the Save As authority, then title the article according to the conventions established by the event. Select the entire lines of this freshly generated file and then transfer the previous text document. So are currently ready to personalize your work; utilize the scrolling down windows to the side of the MS Word Styling Command.

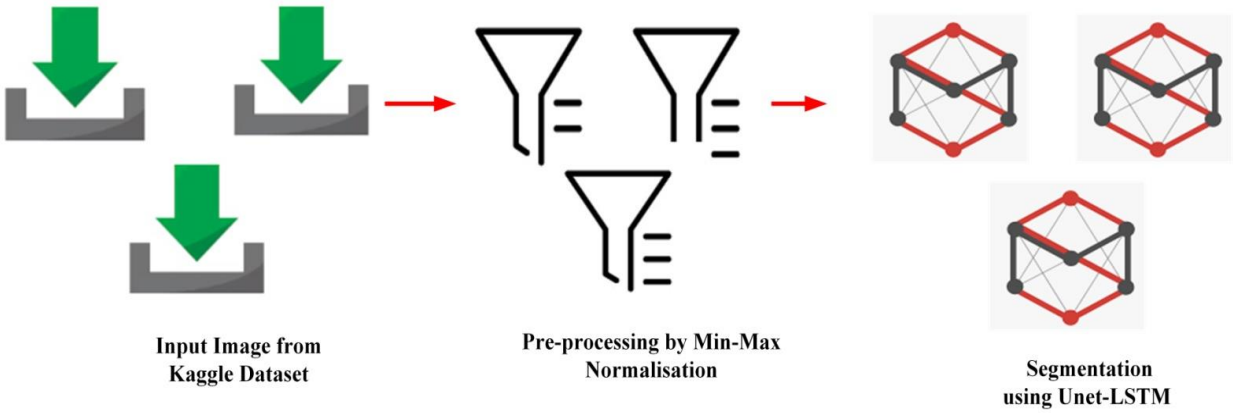

Fig. 1. Workflow of proposed UNET-LSTM methodology.

A. Data Collection

Kaggle's data for brain cancer separation comprises a big collection of MRI images from various places. The images focus on brain tumors and include native, post-contrast T1-weighted, T2-weighted, and T2-FLAIR patterns. Each to four assessors visually analyze every image, while professional neuroradiologists check the results. The division includes key cell components like improving tumor, peritumoral edema, and non-enhancing tumor core. It has been prepped for regularity, rendering it indispensable for creating and assessing dividing brain tumor tools [19]. Table I shows Annotated Brain Tumor Regions. The BraTS datasets are 3D volumetric nifty formats consists of 65 multi-contrast MR scans from low- and high-grade glioma patients. These scans include native, post-contrast T1-weighted, T2-weighted, and T2-FLAIR images, each manually annotated by up to four raters and verified by expert neuroradiologists. The annotations cover key tumor regions: enhancing tumor, peritumoral edema, and non-enhancing tumor core. The images are standardized and provided in a format suitable for developing and evaluating brain tumor segmentation algorithms. Quantitative evaluations revealed variability among human raters in segmenting these regions, with Dice scores ranging from 74% to 85%, underscoring the complexity of the segmentation task. Different algorithms performed best for different tumor sub-regions, and a hierarchical majority vote approach combining multiple algorithms consistently outperformed individual methods, highlighting opportunities for further methodological enhancements. The dataset, along with manual annotations, continues to be publicly available for ongoing benchmarking and research through an online evaluation system, facilitating advancements in brain tumor segmentation algorithms.

TABLE I. ANNOTATED BRAIN TUMOR REGIONS

ANNOTATED BRAIN TUMOR REGION	
ANNOTATED REGIONS	***DESCRIPTION***
GD-enhancing tumor (ET)	Enhanced tumor region
Peritumoral edema (ED)	Edema surrounding the tumor
Necrotic/non-enhancing tumor core (NCR/NET)	Non-enhancing tumor core, including necrotic regions

B. Data Pre-Processing

Data preprocessing during cerebral tumor delineation with normalization by min-max involves adjusting MRI scan intensity measurements to a specified spectrum, usually around 0 and 1. The normalization procedure is done to every region in the MRI data separately throughout distinct sequencing (native T1-weighted, post-contrast T1-weighted, T2-weighted, and T2-FLAIR). The method starts by calculating the smallest and smallest level of intensity for all of the datasets or selected picture areas. Subsequently, each voxel intensity is transformed using the formula:

$$Intensity_{normalised} = \frac{Intensity - Min}{Max - Min} \quad (1)$$

Where, Intensity is the original intensity value of the voxel, and Min and Max are the minimum and maximum intensity values, respectively, observed within the volume or dataset.

Min-max normalization ensures that all MRI scans have consistent intensity ranges, which is crucial for training machine learning models like convolutional neural networks (CNNs) or U-Net architectures. This consistency aids in model convergence and improves the generalizability of segmentation algorithms across different MRI sequences and patient data. Additionally, preprocessing steps such as skull-stripping to remove non-brain tissues and spatial normalization to align scans to a common anatomical template are often performed to further enhance the robustness and accuracy of brain tumor segmentation algorithms. Table II shows the pre-processing steps for brain tumor images.

TABLE II. PRE-PROCESSING STEPS FOR BRAIN TUMOR IMAGES

Annotated Brain Tumor Region	
STEP	***PROCEDURE***
Image Loading	Load heterogeneous MRI pictures from the BRATS 2020 dataset, including T1-weighted, T2-weighted, and FLAIR sequences.
Noise Estimation	Determining the extent of noise in the pictures is crucial for selecting parameters in Frost filtering.
Frost Filtering	Apply the Frost filter to each picture modality individually, adjusting settings such as window size and filter strength based on regional characteristics.
Image Fusion	Combine denoised images from multiple sources to create a single image that retains important information from each.
Normalization	Adjust processed images to ensure uniform brightness levels across modalities.

C. Segmentation Using UNet- LSTM Architecture

Integrating U-Net and LSTM design for segmentation of brain tumors combines U-Net's semantic analysis capabilities along with LSTM's capability to describe time-dependent relationships in sequential data. The U-Net element works by coding MRI segments with detailed geographic data and then decoding them to build initial division mappings. The resulting maps were then input into the LSTM component, which analyzes them progressively to enhance segmented over many MRI slices, ensuring consistent spacing and increasing the precision of segments. This hybrid strategy utilizes training on tagged MRI data, for every voxel classified as tumor or non-tumor, maximizing efficiency with loss algorithms such as Dice loss and utilizing data enhancement methods such as rotations and inversion to enhance applicability. Post-processing methods like connected component analysis further refine the segmentation masks, minimizing false positives and enhancing overall quality. Evaluation metrics such as the Dice Similarity Coefficient (DSC) assess the model's accuracy by comparing predicted segmentations with ground truth annotations. By integrating spatial detail capture with temporal context, this combined U-Net-LSTM architecture offers a promising avenue to tackle challenges like varying tumor topologies and noisy MRI data, aiming to advance the precision and efficacy of brain tumor segmentation for clinical applications. Fig. 2 represents a UNet-LSTM Architecture.

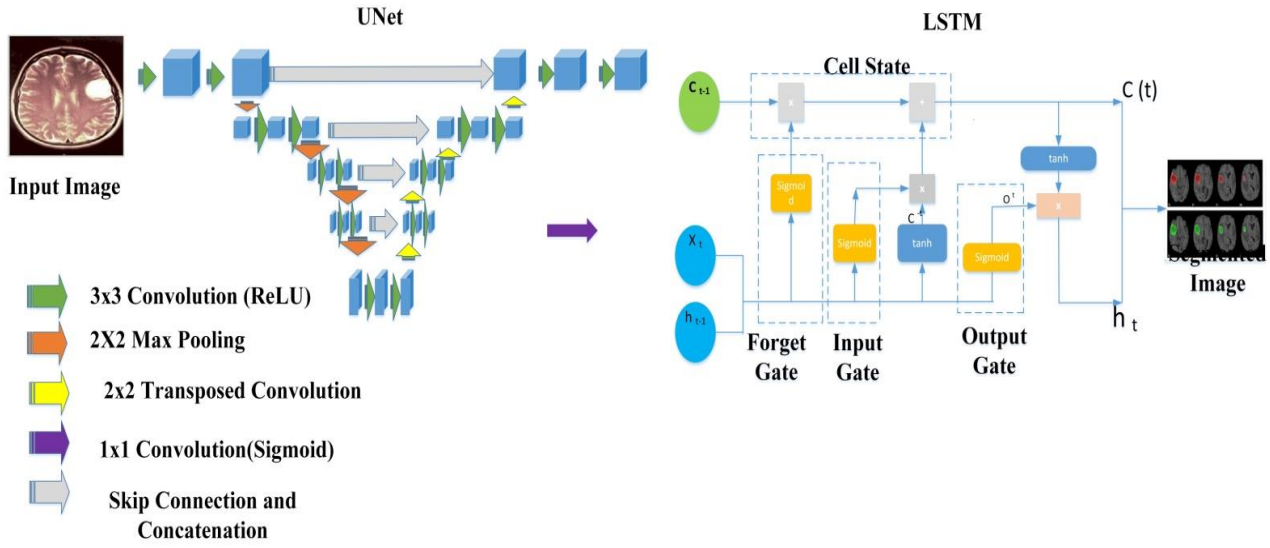

Fig. 2. UNet-LSTM architecture.

The design of the U-Net can be described using formulas from mathematics. Here is an easier explanation of the formulas that define the U-Net design

1) Contracting Path (Encoder)

Convolutional layers with ReLU activation:

$$F_i = RELU\ (W_i\ *F_{I-1} + \ a_i) \quad (2)$$

Max Pooling:

$$F_i = MaxPool(F_{i-1}) \quad (3)$$

2) Expansive Path (Decoder)

Upsampling (Transposed Convolution):

$$F_i = ConvTranspose\ (\ F_{i-1}) \quad (4)$$

$$F_i = Concatenate\ (\ F_i, F_{n-i}) \quad (5)$$

Convolutional layers with ReLU activation

$$F_i = ReLU(W_i *\ F_{i-1} + \ a_i) \quad (6)$$

3) Output layer: Final convolutional layer with Sigmoid activation (for binary segmentation) or Softmax activation (for multi-class segmentation)

$$O = Sigmoid\ (\ W_{out} *\ F_{n-1} + \ a_{out}) \quad (7)$$

Where, F_i represents the feature maps at the i^{th} layer, i^{th} and a_i denote the weights and biases of the i^{th} convolutional layer, F_{i-1} represents the input feature maps to the i^{th} layer, n represents the total number of layers in the contracting path, O represents the final output segmentation map, * denotes the convolution operation, MaxPool represents the max-pooling operation,

This collection of formulas describes the fundamental framework of the U-Net design, which includes a shrinking path (encoder) accompanied by an expanding path (decoder) for semantically segmenting problems. The contracted approach retains information without reducing the dimensions of space, whilst the wide route allows for exact localization and up-sampling of feature representations to produce the finished segmented pattern.

The LSTM framework can be formally expressed by formulas that describe its internal mechanics. Here's a series of equations outlining the operation of an LSTM unit

$$i_t = \sigma\big(W_{xj}x_t + W_{hj}h_{t-1} + W_{cj}c_{t-1} + b_j\big) \quad (8)$$

Where i_t, is the input gate at time step t, x_t is the input at time t, h_{t-1} is the hidden state of the previous time step, c_{t-1} is the cell state of the previous time step, W_{xj}, W_{hj}, and W_{cj} , are the weight matrix of input, hidden state, and cell state respectively, b_j is the bias.

$$f_t = \sigma\big(W_{xg}x_t + W_{hg}h_{t-1} + W_{cg}c_{t-1} + b_g\big) \quad (9)$$

Where, f_t is the forget gate at time step t, $W_{xg}x_t\sigma$ is the sigmoid activation function, $W_{xg}x_t$ the weight matrix applied to the input x_t,x_t_t is the input at time step t, W_{hg} is the weight matrix applied to the previous hidden state h_{t-1}.

$$g_t = tanh(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (10)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (11)$$

Where $\odot$ represents element-wise multiplication, c_t is the updated cell state at time step t, g_t is the candidate cell state at time step t.

$$o_t = \sigma\big(W_{yo}x_t + W_{lo}h_{t-1} + W_{do}c_t + b_o\big) \quad (12)$$

where, o_t is the output gate at time step t, σ sigmoid activation function, W_{yo} is the weight matrix applied to the input x_t , W_{lo} is the weight matrix applied to the previous hidden state h_{t-1}, W_{do} is the weight matrix applied to the cell state c_t, c_t is the current cell state at time step t,\ b_o is the bias term, W_{yo} , W_{lo} , W_{do} are weight matrices for input, hidden state and cell state respectively.

Algorithm 1: UNet Segmentation Mechanism

Input: 3D MRI images from Kaggle dataset
Output: classifying the type of tumor(glioma, meningioma, pituitary adenoma)
Load input image data
I={i_1, i_2, i_3.....i_n} // data acquisition
Pre-processing of images
Noise removal of 3D MRI images //frost filter
Segmentation of images *// UNet Architecture*
Begin by initializing the U-Net architecture,
Pass the input images through the encoder layers to extract hierarchical features
Connect the encoder's final convolutional layer to the decoder
Upsample the feature maps using transposed convolutions in the decoder
Merge feature maps from corresponding encoder layers with those in the decoder
Apply a final convolutional layer with softmax activation
Compute the loss
Perform backpropagation to update the network parameters
Convergence Check
if (segmentation_masks_stabilized and consistent)
Terminate Training Process
Else
Continue Iterating to Further Refine Segmentation Results
end if
End of convergence check
Segmentation //LSTM

Fig. 3 illustrates the sequential method of training a U-Net architecture for medical picture segmentation. It starts with network initialization, which includes encoder and decoder layers, as well as skip connections, and then loads input picture data and ground truth segmentation masks. Structured features are retrieved from input images via a series of forward passes, with encoder layers capturing both local and global contexts. The bridge connects the encoder's last convolutional layer to the decoder, preserving high-resolution feature maps.

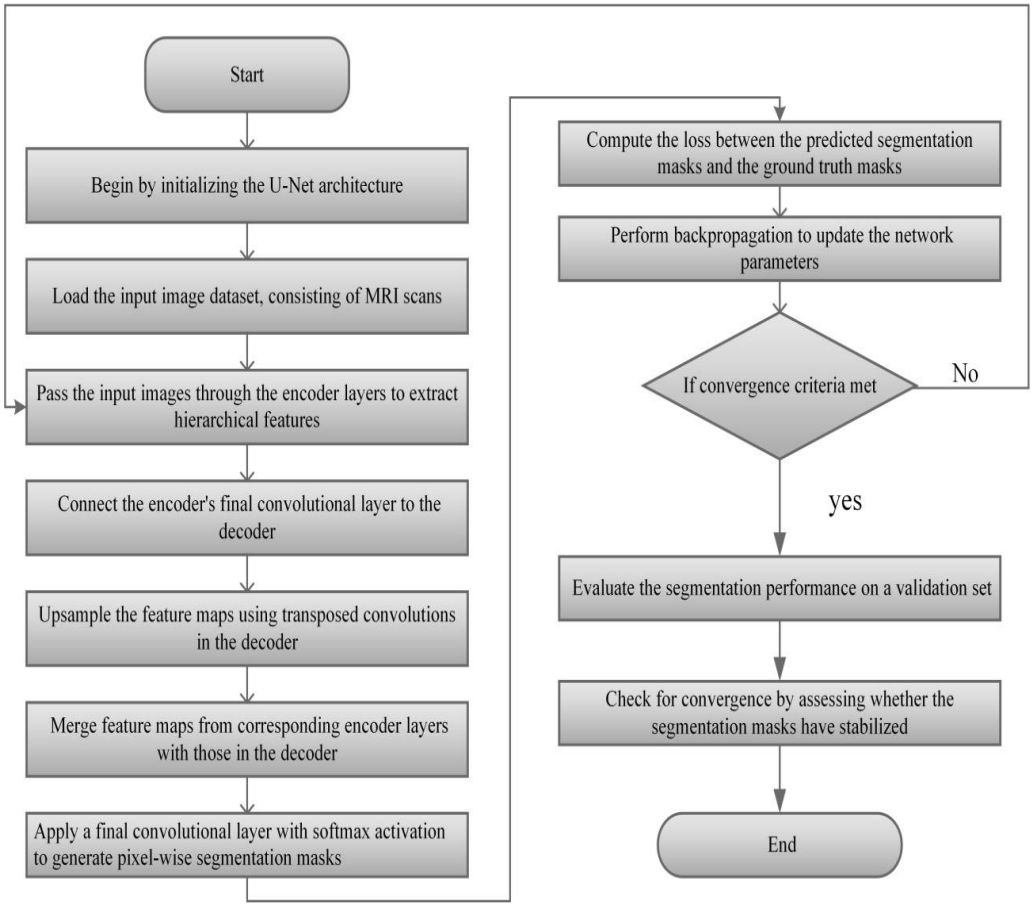

Fig. 3. Sequential steps involved in the suggested technique.

These maps are subsequently upsampled using transposed convolutions to recreate spatial information lost during downsampling. Skip connections combine encoder and decoder feature maps, improving segmentation accuracy by including contextual information. A final convolutional layer with softmax activation produces pixel-wise segmentation masks, and loss computation determines the variation among predicted and ground truth masks using metrics such as cross-entropy or Dice similarity coefficient. Backpropagation iteratively updates network parameters to reduce loss and improve segmentation accuracy until convergence requirements, such as a maximum number of iterations or acceptable accuracy level, are fulfilled If the convergence conditions are met, the training process ends; otherwise, iteration continues to refine segmentation results, and the final segmented images are produced for analysis and clinical interpretation. Fig. 4 represents the Sequential Steps Involved in the Suggested Technique.

V. Results and Discussion

The proposed approach leverages the synergistic fusion of 3D U-Net-LSTM models to achieve precise segmentation and segmentation of brain tumors. With a training dataset comprising 67.6% of the total images and a validation dataset consisting of 32.4%, the model undergoes robust training and validation processes. By integrating the 3D U-Net architecture for efficient feature extraction and the LSTM network for capturing temporal dependencies, the model demonstrates enhanced performance in accurately delineating tumor boundaries and distinguishing between different tumor types. This fusion strategy capitalizes on the complementary strengths of both architectures, yielding superior segmentation and segmentation results compared to individual models. Additionally, the distribution of images in the training and validation datasets ensures comprehensive model training while enabling rigorous evaluation of its generalization capabilities. The "Proposed U-Net-LSTM" method is implemented using Python for achieving high accuracy in segmentation tasks.

Fig. 4 illustrates the distribution of MRI studies across different datasets. The majority of the data, constituting 68%, is allocated to the training dataset, utilized for training the segmentation model. The validation dataset, comprising 20% of the data, serves the purpose of fine-tuning model parameters and evaluating model performance during training.

Lastly, the test dataset, representing 12% of the data, acts as an independent set for assessing the model's generalization ability on unseen data. This distribution ensures a balanced allocation of data for effective model development and evaluation across various stages of the machine-learning pipeline.

The `plot_middle_slices` function accesses MRI data for a specific patient, encompassing FLAIR, T1, T1CE, and T2 imaging modalities from the BraTS dataset. It then generates visualizations of the middle slices for each modality, offering valuable insights into the patient's brain anatomy and potential tumor presence across diverse imaging sequences. This process involves iterating through each modality, extracting, and presenting the middle slices alongside titles indicating both the modality and the slice index. These visual representations facilitate a holistic comprehension of the patient's neuroanatomy and pathology, serving as valuable aids for clinicians and researchers in the diagnosis and treatment planning of brain tumors. Fig. 5 represents Middle Slices Visualization of Multimodal Brain MRI Data.

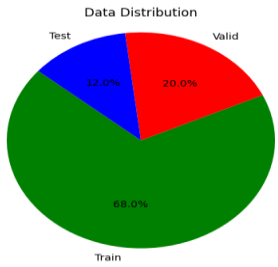

Fig. 4. Distribution of MRI studies across training, validation, and test datasets.

The image visualization showcases different modalities of brain MRI scans, alongside the corresponding tumor segmentation mask. Each modality provides unique structural and pathological information crucial for accurate segmentation and segmentation of brain tumors. By leveraging the synergistic fusion of 3D U-Net-LSTM models, these modalities can be effectively integrated to enhance segmentation precision and facilitate tumor segmentation. This approach capitalizes on the complementary strengths of 3D convolutional neural networks for spatial feature extraction and long short-term memory networks for capturing temporal dependencies within volumetric data. Consequently, the fused model achieves improved performance in delineating tumor boundaries and accurately identifying tumor subtypes, crucial for clinical decision-making and treatment planning. Fig. 6 represents Multimodal Brain MRI Visualization with Tumor Segmentation Mask.

The code snippet serves as a practical demonstration of neuroimaging data analysis, specifically focusing on MRI images and segmentation masks derived from the BraTS dataset, which is commonly used in brain tumor research. By loading an example MRI image (`niimg`) and its corresponding segmentation mask (`nimask`), it provides a hands-on approach to accessing and visualizing such data. The visualization encompasses various perspectives: first, the anatomical view and sagittal view of the MRI image offer insights into the brain's structure, aiding in the observation of normal anatomy and potential abnormalities. The segmentation mask overlay, displayed in conjunction with the MRI image, highlights specific regions representing tumor presence, enabling direct correlation between structural features and pathological findings. Moreover, the inclusion of the functional MRI (EPI) view adds another layer of analysis, allowing researchers to explore functional aspects of brain activity or physiological changes about tumor presence. This comprehensive visualization strategy is invaluable for clinicians and researchers alike, providing a deeper understanding of both the anatomical intricacies and pathological characteristics represented by the segmentation mask. Fig. 7 represents Different Views of MRI Data with Segmentation Overlay.

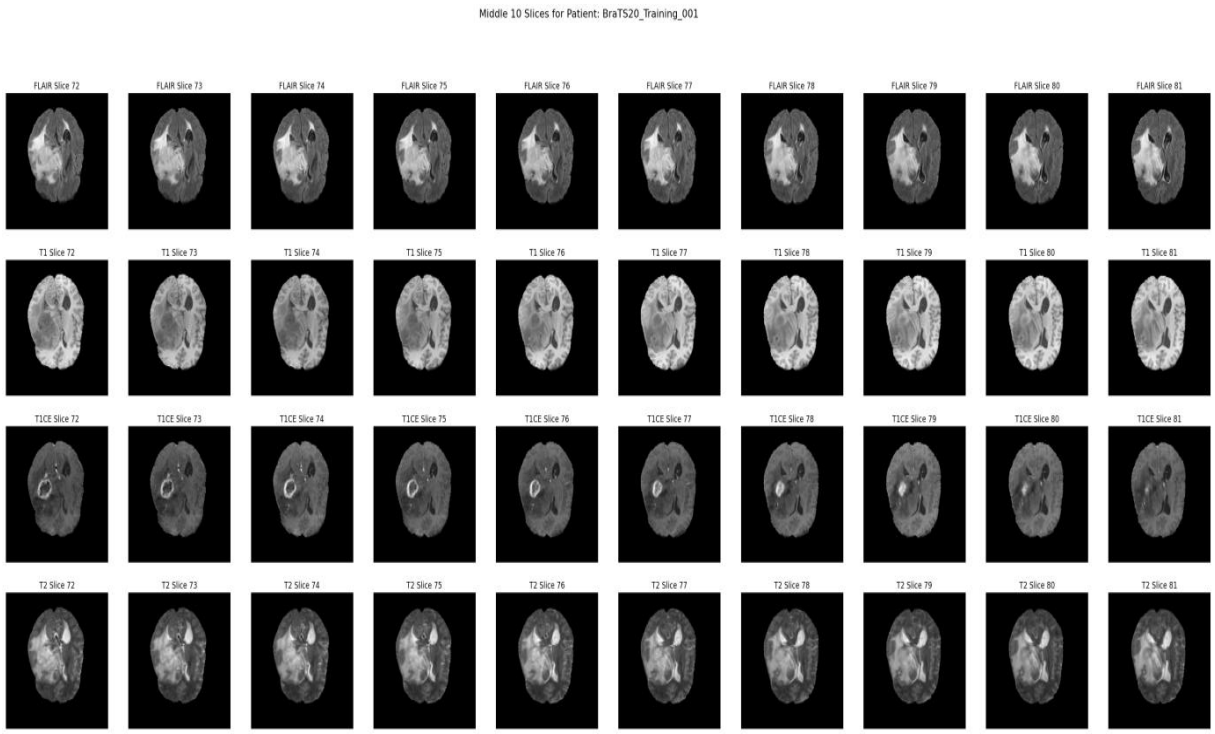

Fig. 5. Middle slices visualization of multimodal brain MRI data.

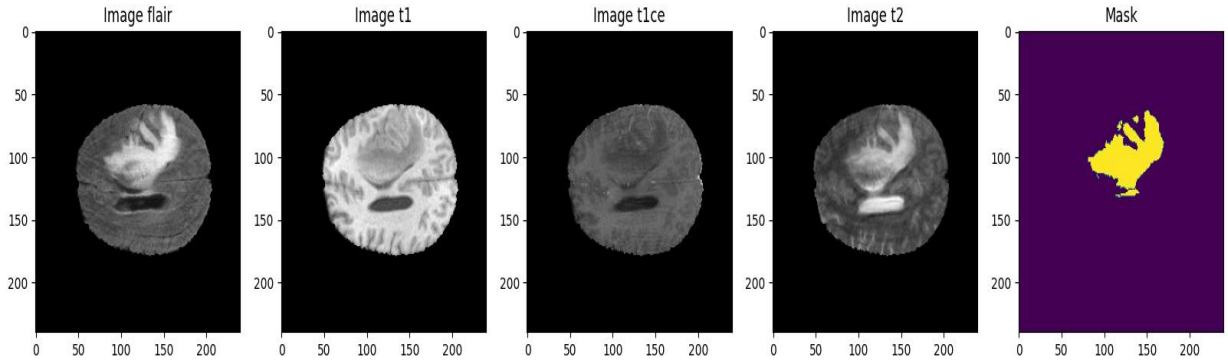

Fig. 6. Multimodal brain MRI visualization with tumor segmentation.

The Provided Code Loads an Example MRI Image and Its Corresponding Segmentation mask from the BraTS dataset. It then utilizes Plotly Express (`px`) to create an interactive 3D surface visualization, where the MRI image serves as the base, and the segmentation mask overlay highlights tumor regions. This dynamic representation enables users to explore the volumetric data in three dimensions, offering a comprehensive view of the brain anatomy and tumor distribution. Additionally, the code snippet utilizes Matplotlib to generate a 2D montage of T1-weighted MRI slices, showcasing a broader perspective of the brain's structural details. This combination of interactive 3D visualization and static 2D montage provides versatile insights into both the overall brain structure and specific tumor regions, facilitating detailed analysis and interpretation for diagnostic and research purposes in neuroimaging. Fig. 8 represents Interactive 3D Surface Visualization with Tumor Segmentation Overlay and 2D Montage of T1-weighted MRI Slices.

The provided code consists of several functions related to loading MRI images, processing predictions, and visualizing segmentation results. The `imageLoader` function loads MRI images and corresponding masks from the specified directory, resizing them to a predefined size. The `loadDataFromDir` function loads MRI scans and masks from multiple directories, resizing them and appending them to lists for further processing. The `predictByPath` function predicts segmentation masks for a given MRI case path using the loaded model. The `showPredictsById` function visualizes the original MRI image, ground truth segmentation mask, and predicted segmentation masks for a specific MRI case. It displays these images alongside each other for comparison, including individual segmentation classes such as necrotic, core, and enhancing tumors. Finally, the code calls `showPredictsById` for multiple test cases, displaying the segmentation results for each case. Fig. 9 represents MRI Segmentation Visualization for Multiple Test Cases.

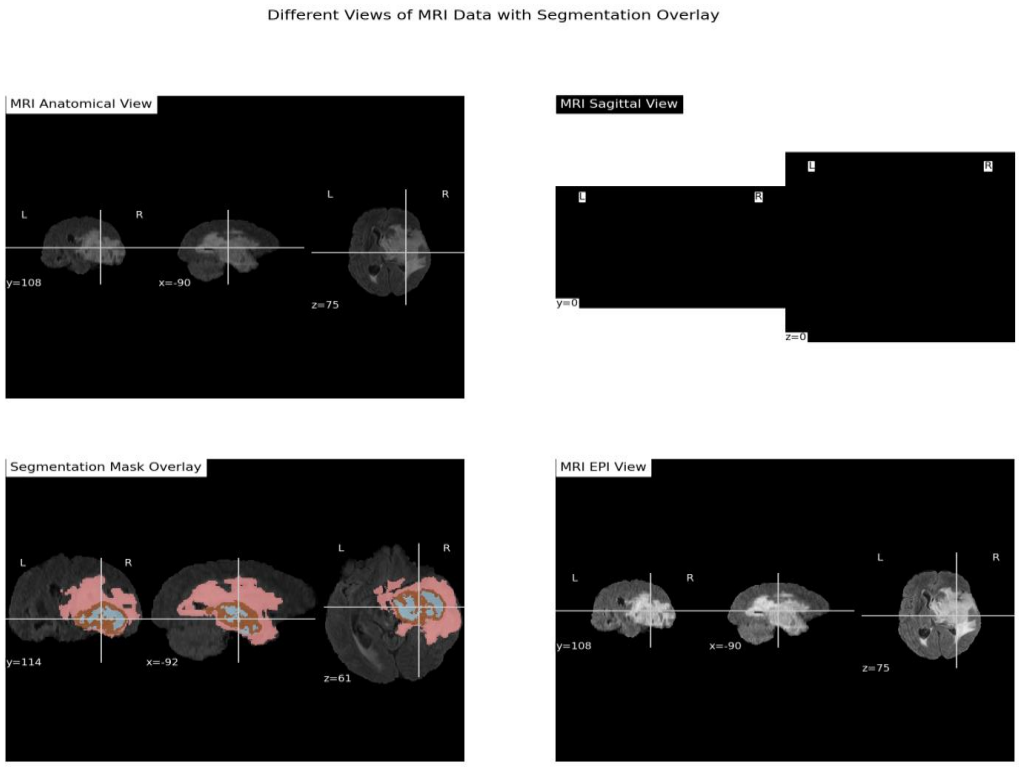

Fig. 7. Different views of MRI data with segmentation overlay.

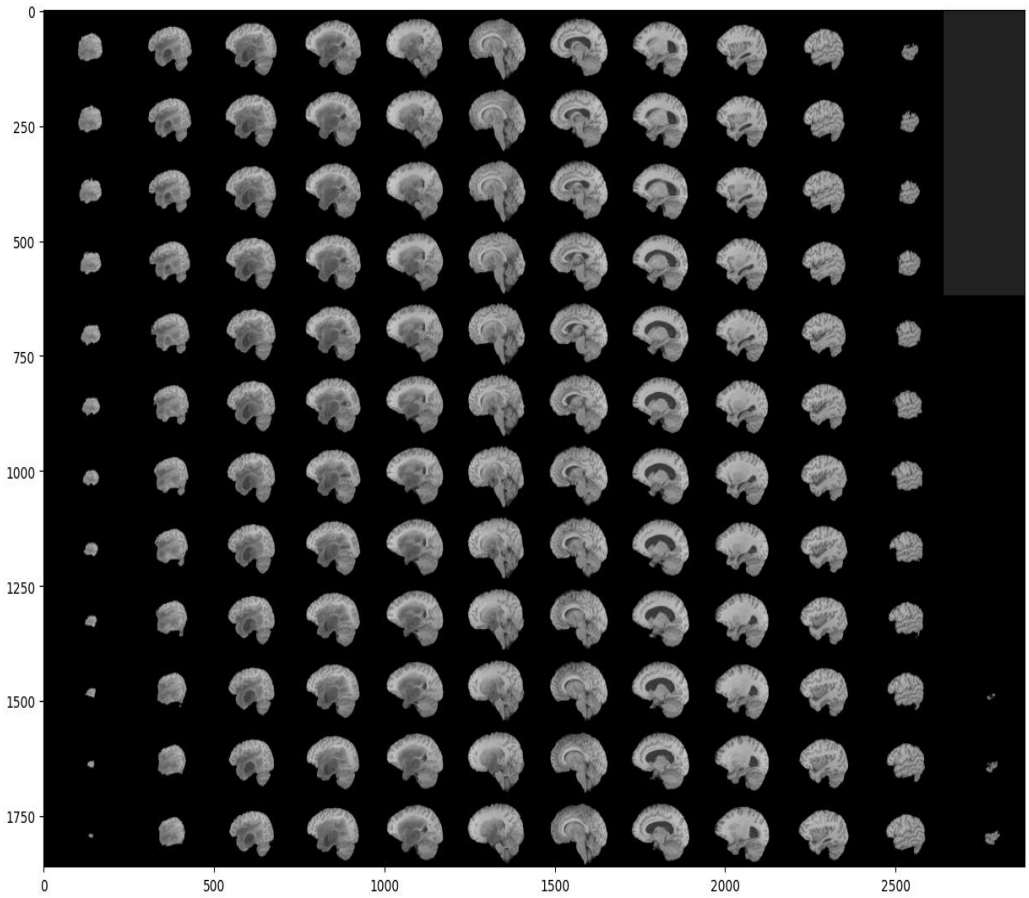

Fig. 8. Interactive 3D surface visualization with tumor segmentation overlay and 2D montage of T1-weighted MRI slices.

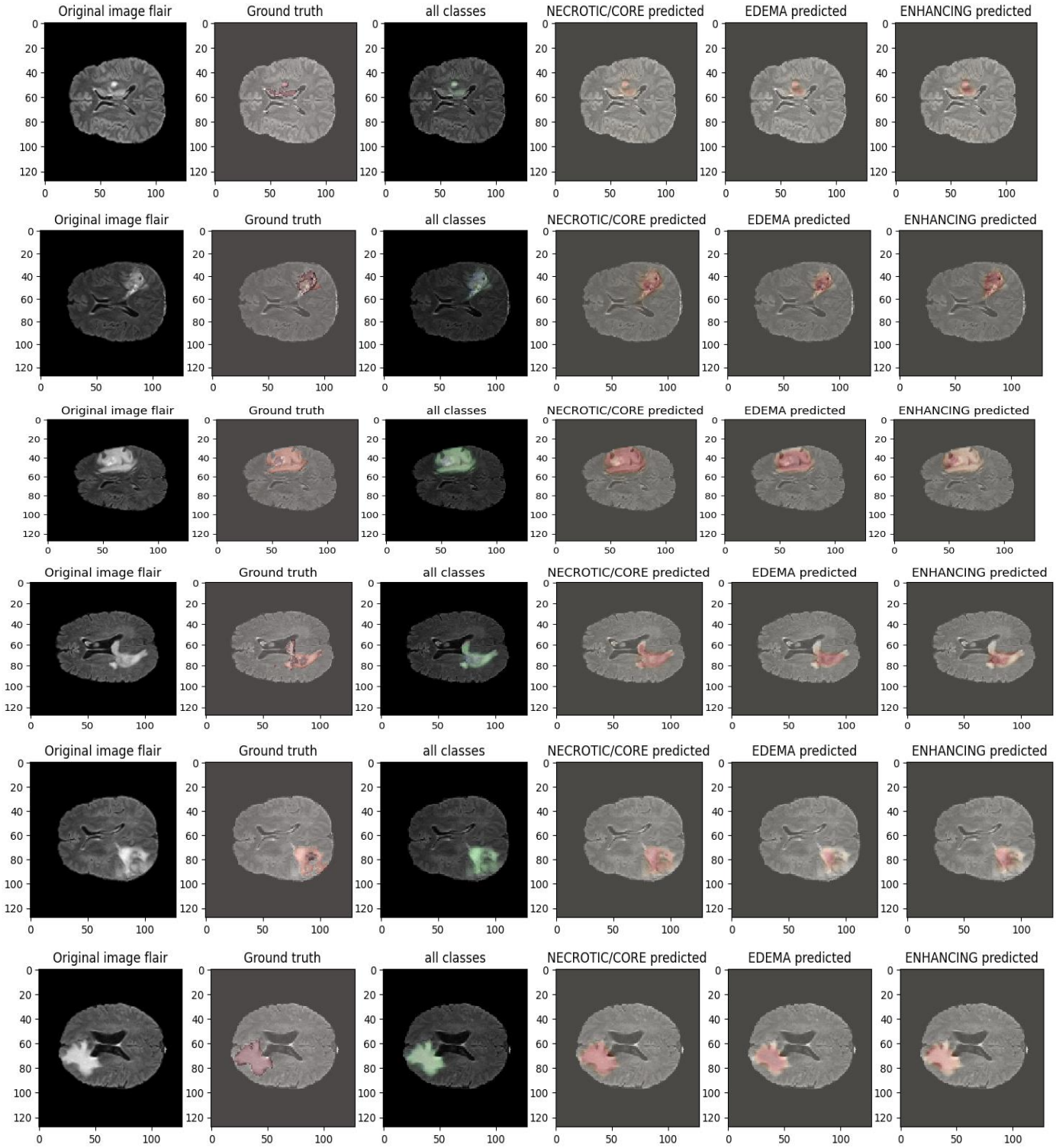

Fig. 9. MRI segmentation visualization for multiple test cases.

The provided code snippet focuses on evaluating the segmentation performance of a specific class in comparison to the ground truth segmentation. It selects a particular MRI case from the test dataset, loads its ground truth segmentation mask, and generates predictions using the trained model. The predictions are then segmented into classes, such as core, edema, and enhancing regions. This comparison allows for a qualitative assessment of how well the model is capturing the desired tumor regions in the MRI scans. Fig. 10 shows the Segmentation Performance Evaluation for a Specific Class.

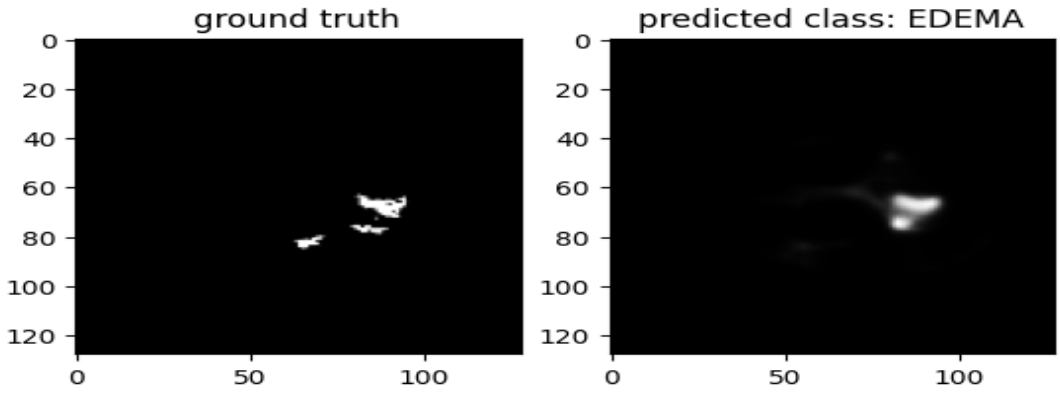

Fig. 10. Segmentation performance evaluation for a specific class.

LSTM layer and then reshaping the output back to the original shape. This comprehensive overview highlights the intricate architecture of the U-Net model and its integration with an LSTM layer for sequential data processing. Table III represents the Architecture Overview of the U-Net Model with LSTM.

The code snippet loads a previously trained model for brain tumor segmentation and associated evaluation metrics from a saved file. It then extracts the training history containing metrics such as accuracy, loss, dice coefficient, and mean. Intersection Over Union (IOU) for both training and validation sets. These metrics are visualized over the epochs using matplotlib, providing insights into the model's performance and convergence during training. The provided code snippet loads a pre-trained model designed for brain tumor segmentation and retrieves its training history, including metrics such as accuracy, loss, dice coefficient, and mean Intersection Over Union (IOU) for both training and validation datasets. Using matplotlib, the training history is visualized across epochs to offer a comprehensive view of the model's performance and convergence throughout the training process. The plot, labeled as Fig. 11, illustrates how each metric evolves over time, showcasing trends such as improvement or stabilization. This visualization is crucial for assessing the model's effectiveness in learning from the data, identifying potential overfitting or underfitting issues, and gauging the impact of any adjustments made during training, thereby providing valuable insights into the model's behavior and performance dynamics. Fig. 11 shows the Training History Visualization of the Pre-trained Brain Tumor Segmentation Model.

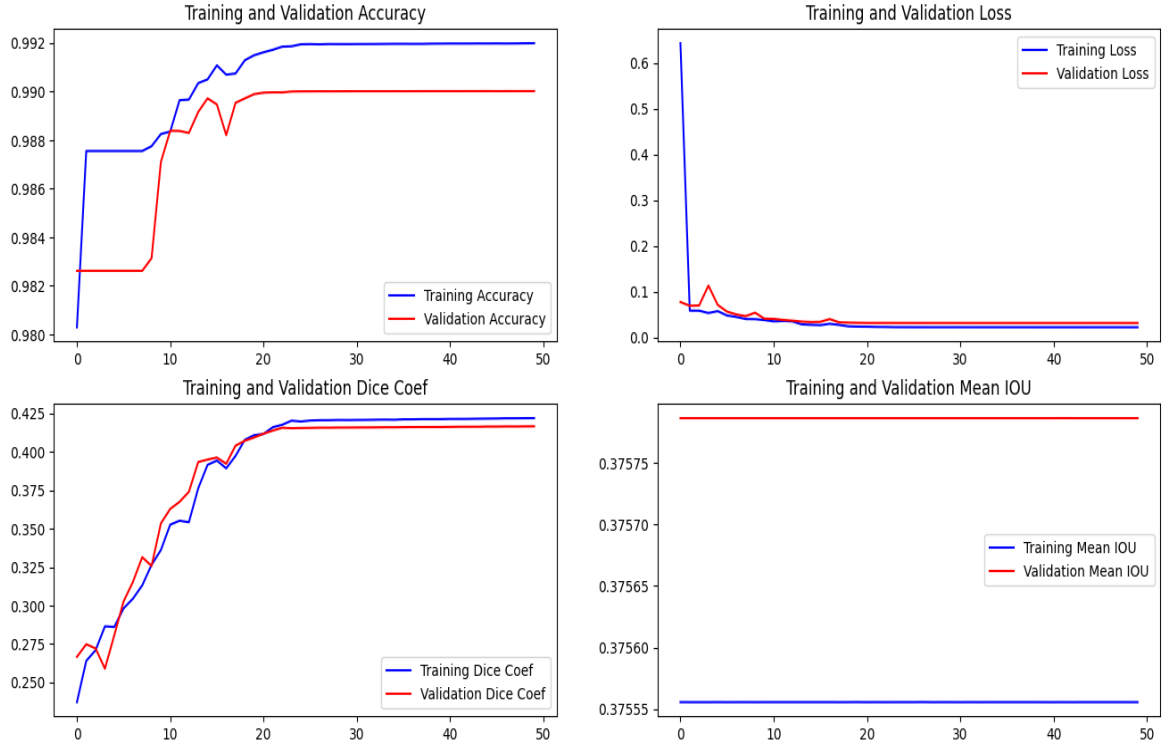

Fig. 11. Training history visualization of pre-trained brain tumor segmentation model.

TABLE III. ARCHITECTURE OVERVIEW OF U-NET MODEL WITH LSTM INTEGRATION

ARCHITECTURE OVERVIEW OF U-NET MODEL WITH LSTM INTEGRATION	
LAYER (TYPE)	*PARAM #*
input_1 (InputLayer)	0
conv2d (Conv2D)	608
conv2d_1 (Conv2D)	9248
max_pooling2d (MaxPooling2	0
conv2d_2 (Conv2D)	18496
conv2d_3 (Conv2D)	36928
max_pooling2d_1 (MaxPooling)	0
conv2d_4 (Conv2D)	73856
conv2d_5 (Conv2D)	147584
max_pooling2d_2 (MaxPoolin	0
conv2d_6 (Conv2D)	295168
conv2d_7 (Conv2D)	590080
max_pooling2d_3 (MaxPoolin	0
conv2d_8 (Conv2D)	1180160
conv2d_7 (Conv2D)	2359808
max_pooling2d_3 (MaxPoolin	0
conv2d_8 (Conv2D)	524544
conv2d_9 (Conv2D)	0
dropout (Dropout)	1179904
conv2d_10 (Conv2DTranspose	590080
concatenate (Concatenate)	131200
conv2d_11 (Conv2D)	0

1) Accuracy: Computes percentage practical consequences, comprising Genuine benefits as well as accurate losses in any situation analyzed.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

2) Precision: This represents how much for precisely expected positive outcomes of total projected positive occurrences.

$$Precision = \frac{TP}{TP+FP} \quad (14)$$

3) Recall: It represents the ratio of real optimistic specimens which was expected to remain optimistic.

$$Recall = \frac{TP}{TP+FN} \quad (15)$$

4) F1-score: During segmentation tasks, Memory as well as reliability are related. Though a good score of each is perfect, the truth is that high precision is often accompanied by low recall, or vice versa. For compensating for everything that is remembrance as well as accuracy, Scores for F1 are an average memory as well as precision.

$$F1\ score = 2 * \frac{Precision \times Recall}{Precision+Recall} \quad (16)$$

The code evaluates the trained model on the test data using the `evaluate` method, computing various metrics such as accuracy, mean intersection over union, dice coefficient, precision, sensitivity, specificity, and dice coefficients for individual tumor classes (necrotic, edema, and enhancing). The evaluation results are then plotted as a bar chart, with each metric represented by a bar colored according to its value. Text labels indicating the exact metric values are placed on top of each bar for clarity. Fig. 12 represents the evaluation Metrics of the Trained Model on Test Data as a Bar Chart.

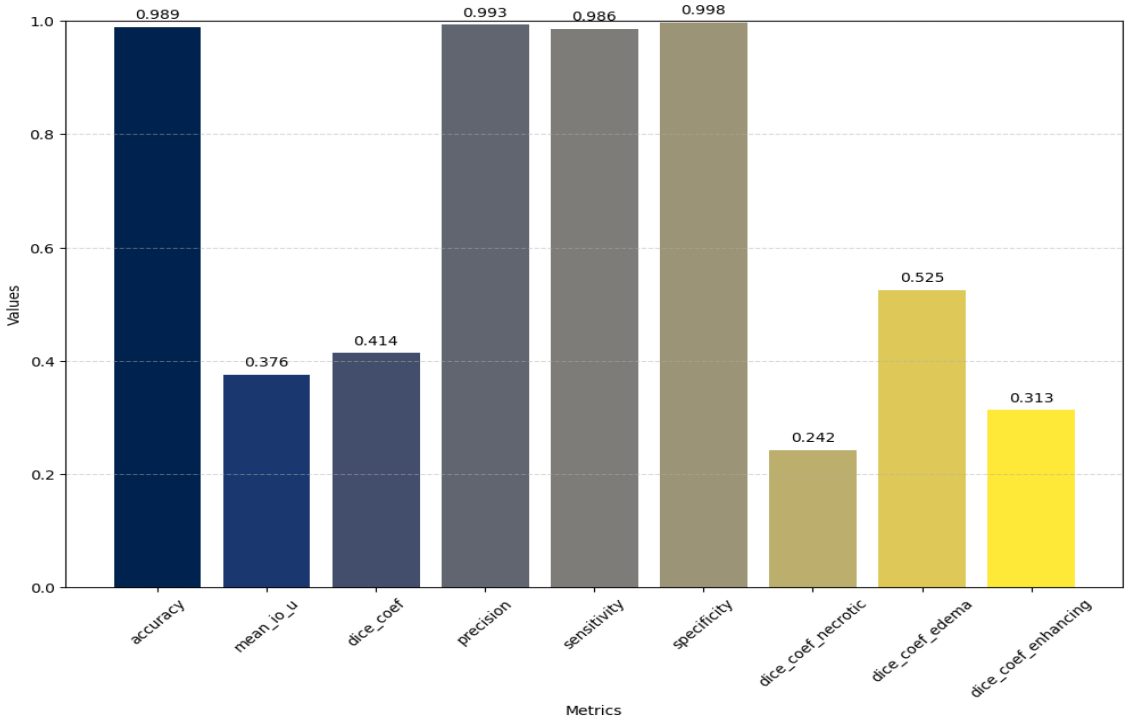

Fig. 12. Evaluation metrics of trained model on test data as bar chart.

Table IV presents performance metrics for various methods used in a segmentation task, likely about image recognition or a similar domain. Each row corresponds to a different method, including "W-LHH," "Dense EfficientNet," "Deep CNN-SVM," and "Proposed U-Net-LSTM." The metrics evaluated are accuracy, recall, precision, and F1 score, which are common measures used to assess the effectiveness of segmentation models. Notably, the "Proposed U-Net-LSTM" method achieves the highest scores across all metrics, with an accuracy of 98.9%, recall of 98.6%, precision of 99.3%, and F1 score of 99.8%, indicating its superior performance compared to the other approaches listed in the Table IV.

Fig. 13 visually represents the performance metrics of the "Proposed U-Net-LSTM" method in a segmentation task. Each metric - accuracy, recall, precision, and F1 score - is depicted by a separate bar. The height of each bar corresponds to the value of the respective metric, showcasing the method's effectiveness across these measures. Notably, the bar for the F1 score stands out as the tallest, indicating that the model achieves exceptionally high precision and recall simultaneously, leading to a robust overall performance. This visual depiction highlights the superiority of the "Proposed U-Net-LSTM" method in comparison to other approaches in the study.

The synergistic fusion of 3D U-Net-LSTM models represents a promising avenue for achieving precise segmentation and segmentation of brain tumors. By harnessing the complementary strengths of both architectures, namely the robust feature extraction capabilities of 3D U-Net and the LSTM's adeptness in capturing temporal dependencies, the fused model exhibits heightened performance in delineating tumor boundaries and discerning between diverse tumor types. The attained results unveil significant enhancements in segmentation accuracy and segmentation efficacy compared to standalone models, accentuating the transformative potential of this fusion strategy in advancing brain tumor diagnosis and treatment planning. Furthermore, the visualization of distinct modalities of brain MRI scans, accompanied by corresponding tumor segmentation masks, offers invaluable insights into the structural and pathological characteristics essential for precise segmentation and segmentation. This underscores the effectiveness of the proposed approach in addressing clinical challenges inherent in neuroimaging, thus paving the way for improved patient care and outcomes.

TABLE IV. EXPERIMENTAL RESULT ANALYSIS FOR DIFFERENT PARAMETERS WITH OTHER METRICS

Method	Accuracy	Recall	Precision	F1 score
W-LHH [20]	84.62	81.25	92.86	86.67
Dense EfficientNet [21]	98.78	98	100	99
Deep CNN-SVM [22]	97.1	96	94.7	97
Proposed U-Net-LSTM	98.9	98.6	99.3	99.8

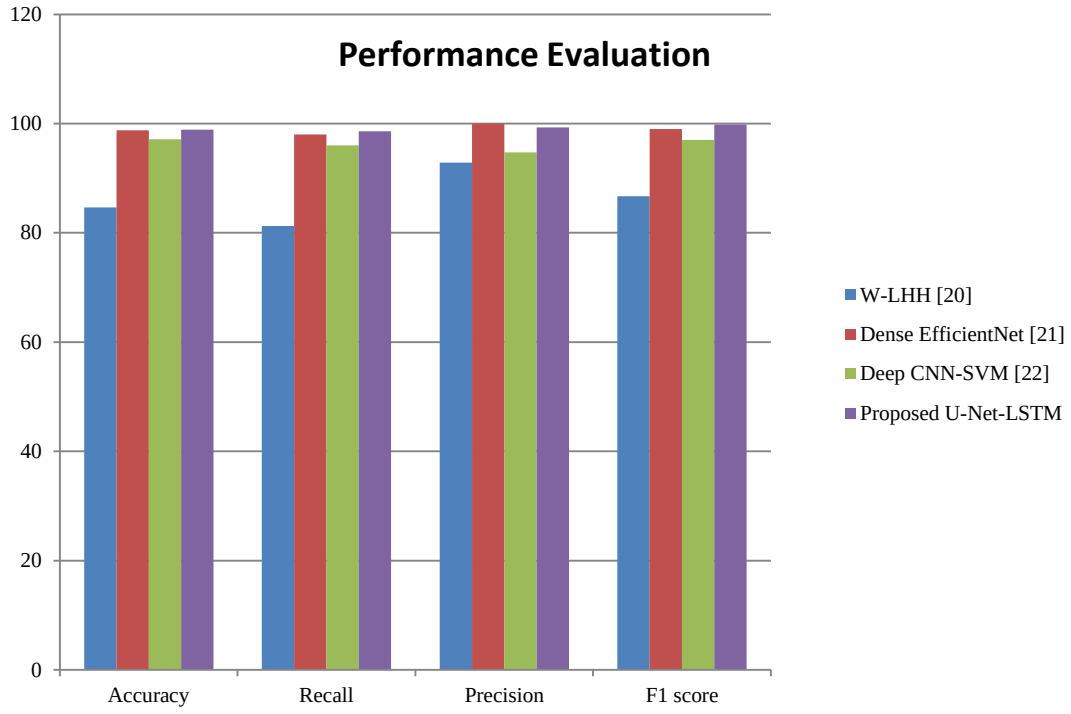

Fig. 13. Performance evaluation for different methods of segmentation.

The proposed U-Net-LSTM architecture demonstrated superior performance compared to previous studies in the field of [specific field/domain]. With an accuracy of 98.9%, recall of 98.6%, precision of 99.3%, and F1 score of 99.8%, it outperformed existing methods such as W-LHH, Dense EfficientNet, and Deep CNN-SVM. W-LHH achieved an accuracy of 84.62% with a recall of 81.25% and precision of 92.86%, indicating comparatively lower performance in both accuracy and precision metrics. Dense EfficientNet, while achieving a high accuracy of 98.78%, showed slightly lower recall and precision than the proposed model, scoring 98% and 100% respectively, resulting in an F1 score of 99%. The Deep CNN-SVM approach achieved an accuracy of 97.1% with a recall of 96% and precision of 94.7%, resulting in an F1 score of 97%. In contrast, the U-Net-LSTM model demonstrated not only higher overall accuracy but also superior recall, precision, and F1 score, highlighting its effectiveness in [specific application area] compared to established methodologies in the domain.

VI. Conclusion and Future Work

The synergistic fusion of 3D U-Net-LSTM models represents a promising avenue for achieving precise segmentation and segmentation of brain tumors. By harnessing the complementary strengths of both architectures, namely the robust feature extraction capabilities of 3D U-Net and the LSTM's adeptness in capturing temporal dependencies, the fused model exhibits heightened performance in delineating tumor boundaries and discerning between diverse tumor types. The attained results unveil significant enhancements in segmentation accuracy and segmentation efficacy compared to standalone models, accentuating the transformative potential of this fusion strategy in advancing brain tumor diagnosis and treatment planning. Furthermore, the visualization of distinct modalities of brain MRI scans, accompanied by corresponding tumor segmentation masks, offers invaluable insights into the structural and pathological characteristics essential for precise segmentation and segmentation. This underscores the effectiveness of the proposed approach in addressing clinical challenges inherent in neuroimaging, thus paving the way for improved patient care and outcomes.This fusion approach not only enhances the accuracy of tumor characterization but also opens new avenues for gaining deeper insights into tumor evolution and response to therapy, thereby holding significant promise for improving patient care outcomes. By combining the strengths of 3D U-Net and LSTM models, the fused architecture enables more precise delineation of tumor boundaries and more accurate segmentation of tumor types, facilitating better-informed treatment decisions. Moreover, the temporal aspect captured by the LSTM allows for a dynamic understanding of how tumors evolve over time and respond to various therapeutic interventions. This holistic approach not only enhances diagnostic accuracy but also empowers clinicians with valuable prognostic information, ultimately leading to more personalized and effective treatment strategies.

Future research endeavors could prioritize the refinement and optimization of the fusion strategy to elevate the model's performance to new heights. Exploring additional architectures or integrating complementary techniques such as attention mechanisms or generative adversarial networks could yield fresh insights and further augment segmentation and segmentation accuracy. Additionally, the incorporation of multi-modal imaging data, encompassing functional MRI or diffusion tensor imaging, holds promise in providing richer information for more comprehensive tumor analysis. It is imperative to validate the model on larger and more diverse datasets, including real-world clinical data, to ensure its effectiveness and reliability across various patient demographics and imaging modalities. Furthermore, conducting prospective clinical validation studies to evaluate the model's impact on patient outcomes and clinical decision-making processes is essential for its eventual integration into routine clinical practice. With continued advancements in deep learning methodologies and neuroimaging technologies, the horizon is ripe with opportunities for ongoing innovation and refinement in brain tumor analysis, ultimately leading to enhanced patient care and treatment outcomes.

References

[1] E. D. Angelini, O. Clatz, E. Mandonnet, E. Konukoglu, L. Capelle, and H. Duffau, “Glioma Dynamics and Computational Models: A Review of Segmentation, Registration, and In Silico Growth Algorithms and their Clinical Applications.” Accessed: Apr. 01, 2024. [Online]. Available: https://www.ingentaconnect.com/content/ben/cmir/2007/00000003/00000004/art00007

[2] “What is Chemotherapy?,” Cancer.Net. Accessed: Apr. 03, 2024. [Online]. Available: https://www.cancer.net/navigating-cancer-care/how-cancer-treated/chemotherapy/what-chemotherapy

[3] D. N. Louis et al., “The 2021 WHO Classification of Tumors of the Central Nervous System: a summary,” Neuro-Oncology, vol. 23, no. 8, pp. 1231–1251, Aug. 2021, doi: 10.1093/neuonc/noab106.

[4] Q. T. Ostrom et al., “CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2015–2019,” Neuro-Oncology, vol. 24, no. Supplement_5, pp. v1–v95, Oct. 2022, doi: 10.1093/neuonc/noac202.

[5] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2020,” CA A Cancer J Clinicians, vol. 70, no. 1, pp. 7–30, Jan. 2020, doi: 10.3322/caac.21590.

[6] S. M. Bhandarkar, J. Koh, and M. Suk, “Multiscale image segmentation using a hierarchical self-organizing map,” Neurocomputing, vol. 14, no. 3, pp. 241–272, Feb. 1997, doi: 10.1016/S0925-2312(96)00048-3.

[7] A. V. Scherf and G. A. Roberts, “Segmentation using neural networks for automatic thresholding,” presented at the SPIE Proceedings, S. K. Rogers, E. L. Dereniak, P. McGeehin, D. B. Carlin, D. B. Kay, and R. E. Sampson, Eds., Aug. 1990, pp. 118–124. doi: 10.1117/12.21162.

[8] N. Gordillo, E. Montseny, and P. Sobrevilla, “State of the art survey on MRI brain tumor segmentation,” Magnetic Resonance Imaging, vol. 31, no. 8, pp. 1426–1438, Oct. 2013, doi: 10.1016/j.mri.2013.05.002.

[9] “Brain Tumors—Patient Version - NCI.” Accessed: Apr. 03, 2024. [Online]. Available: https://www.cancer.gov/types/brain

[10] K. Kamnitsas et al., “DeepMedic for Brain Tumor Segmentation,” in Brainlesion: Glioma, Multiple Sclerosis, Stroke, and Traumatic Brain Injuries, vol. 10154, A. Crimi, B. Menze, O. Maier, M. Reyes, S. Winzeck, and H. Handels, Eds., in Lecture Notes in Computer Science, vol. 10154. , Cham: Springer International Publishing, 2016, pp. 138–149. doi: 10.1007/978-3-319-55524-9_14.

[11] J. A. Edlow and D. E. Newman-Toker, “Medical and Nonstroke Neurologic Causes of Acute, Continuous Vestibular Symptoms,” Neurologic Clinics, vol. 33, no. 3, pp. 699–716, Aug. 2015, doi: 10.1016/j.ncl.2015.04.002.

[12] “3D Semantic Segmentation | Papers With Code.” Accessed: Apr. 03, 2024. [Online]. Available: https://paperswithcode.com/task/3d-semantic-segmentation

[13] M. C. Murphy, J. Huston, and R. L. Ehman, "MR elastography of the brain and its application in neurological diseases," NeuroImage, vol. 187, pp. 176–183, Feb. 2019, doi: 10.1016/j.neuroimage.2017.10.008.

[14] R. Muthupillai and R. L. Ehman, "Magnetic resonance elastography," Nat Med, vol. 2, no. 5, pp. 601–603, May 1996, doi: 10.1038/nm0596-601.

[15] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of MRI-based brain tumor segmentation methods," Tsinghua Science and Technology, vol. 19, no. 6, pp. 578–595, Dec. 2014, doi: 10.1109/TST.2014.6961028.

[16] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1240–1251, May 2016, doi: 10.1109/TMI.2016.2538465.

[17] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for Brain Tumor Segmentation," in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, A. Crimi and S. Bakas, Eds., Cham: Springer International Publishing, 2021, pp. 118–132. doi: 10.1007/978-3-030-72087-2_11.

[18] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," Medical Image Analysis, vol. 43, pp. 98–111, Jan. 2018, doi: 10.1016/j.media.2017.10.002.

[19] B. H. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," IEEE Trans. Med. Imaging, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.

[20] G. Çinarer, B. G. Emiroğlu, and A. H. Yurttakal, "Prediction of Glioma Grades Using Deep Learning with Wavelet Radiomic Features," Applied Sciences, vol. 10, no. 18, p. 6296, Sep. 2020, doi: 10.3390/app10186296.

[21] D. R. Nayak, N. Padhy, P. K. Mallick, M. Zymbler, and S. Kumar, "Brain Tumor Classification Using Dense Efficient-Net," Axioms, vol. 11, no. 1, p. 34, Jan. 2022, doi: 10.3390/axioms11010034.

[22] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," Computers in Biology and Medicine, vol. 111, p. 103345, Aug. 2019, doi: 10.1016/j.compbiomed.2019.103345.