

Digital Object Identifier

Magnetic Resonance Imaging Brain Segmentation using Bi-Directional Convolutional Long Short-Term Memory U-Net with Densely Connected Convolutions

MESHARI D. ALANAZI¹, AMNA MARAOUI², IMEN WERDA³, (Member, IEEE), AHMED BEN ATITALLAH¹, TURKI M. ALANAZI⁴, MOHAMMED ALBEKAIRI¹, ANIS SAHBANI⁵, and AMR YOUSEF^{6,7}, (Member, IEEE)

¹Department of Electrical Engineering, College of Engineering, Jouf University, Sakakah 72388, Saudi Arabia

²Networked Objects, Control, and Communication Systems Laboratory (NOCCS), National engineering school of Sousse, Sousse, Tunisia

³Electronic and Information Technology Laboratory, University of Sfax, Sfax, Tunisia

⁴Department of Electrical Engineering, College of Engineering, University of Hafr Al Batin, Hafr Al Batin 39524, Saudi Arabia.

⁵Institute for Intelligent Systems and Robotics (ISIR), Centre National de la Recherche Scientifique, Sorbonne University, 75006 Paris, France.

⁶Engineering Department, University of Business and Technology, Ar Rawdah, Jeddah 23435, Saudi Arabia

⁷Engineering Mathematics Department, Alexandria University, Lotfy El-Sied St. off Gamal Abd El-Naser, Alexandria 11432, Egypt

Corresponding author: Ahmed Ben Atitallah (abenatitallah@ju.edu.sa).

This work was funded by the Deanship of Graduate Studies and Scientific Research at Jouf University under grant No. (DGSSR-2023-02-02291)

ABSTRACT Despite advancements in imaging technologies, traditional diagnostic methods relying on manual interpretation by clinicians remain prone to errors. This underscores the need for robust automated segmentation techniques. Accurate three-dimensional (3D) Magnetic Resonance Imaging (MRI) brain tumor segmentation is critical for effective diagnosis in neuroimaging. Given the complexity of brain tumors, deep learning approaches have emerged as more suitable than traditional methods for efficient and precise segmentation. The dual difficulties of high computing needs in 3D convolutional networks and the necessity of more precise brain tumor diagnosis are addressed by this work. We provide an artificial intelligence (AI) framework combining bi-directional convolutional long short-term memory (ConvLSTM) layers with densely connected convolutions (DCC). Our method enhances spatial feature learning employing dense connections, and catches complex temporal links across MRI slices. Particularly for the identification of the Enhancing Tumor (ET) region, this novel combination solves the complexity limitation of convolutional techniques, hence improving segmentation robustness. Our approach is validated by extensive testing on the Brain Tumor Segmentation (BraTS) 2020 and 2021 datasets. On the BraTS 2021 dataset, the proposed M-BDCU-Net model beats many state-of-the-art approaches for particular tumor sub-regions with Dice Similarity Coefficient (DSC) scores of 0.81 for Tumor Core (TC), 0.85 for Enhancing Tumor (ET), and 0.82 for Whole Tumor (WT). With 8.81 million parameters and 60.34 GFLOPs, our model also greatly lowers complexity by up to 15% and balances accuracy with efficiency. It is thus much lighter than current models. These results highlight how rapid, more dependable, and exact brain tumor detection our system could allow. This work fills important voids in present research by combining methodological innovation with exhaustive evaluation and provides a useful answer for therapeutic uses.

INDEX TERMS Bi-Directional, Brain Segmentation, Convolutional Long Short-Term Memory, Densely Connected Convolutions, Magnetic Resonance Imaging, U-Net

I. INTRODUCTION

The human body operates through a finely tuned system of cell growth, death, and regeneration. However, cells mutations lead to the development of tumors. Due to the brain's vital role in regulating the neurological system, brain tumors are among the most destructive.

In their early stages, brain tumors may not contain cancerous cells. However, as they progress, they can become malignant, characterized by uncontrolled cell division and the potential to spread to other parts of the body. Early detection and accurate segmentation of these tumors are crucial for effective treatment. Unfortunately, brain cancer remains one of the deadliest forms of cancer, highlighting the urgent need for advanced diagnostic tools.

Magnetic Resonance Imaging (MRI) has become a cornerstone in the diagnosis of brain tumors. By providing detailed, high-resolution images of the brain's internal structures, MRI enables clinicians to assess the shape, location, type, and size of tumors. Accurate diagnosis, treatment planning, and tracking disease progression all depend on this information. However, manually segmenting brain tumors from 3D MRI scans is a time-consuming and error-prone task that requires significant expertise. Consequently, automated segmentation techniques are essential to enhance the precision and efficiency of clinical practice.

While these approaches are simple, they typically fail to capture the complex and varied patterns seen in brain tumors. Traditional segmentation techniques, including statistical-based techniques, depend on parameters like mean, median, variance, and standard deviation [1]. Support Vector Machines (SVM) and Random Forests (RF), among other machine learning techniques, have shown promise enhancement in segmentation accuracy and efficiency [2], [3]. These techniques find difficulty with the variety and complexity of medical imaging databases.

Deep learning provides unheard-of precision and resilience, hence transforming medical picture segmentation. Particular convolutional neural networks (CNNs) have shown amazing success since they can automatically extract high-level characteristics from images [4]–[7]. Many deep learning models, particularly 3D CNNs and transformer-based architectures, are computationally expensive and complicated even with their outstanding performance. Their great computing needs render them unworkable for real-world therapeutic uses, in which case computational resources are sometimes constrained.

The gap between research and actual application emphasizes the necessity of a solution that balances computing performance with efficiency. Although state-of-the-art models have great segmentation accuracy, their complexity limits their scalability and implementation in environments with limited resources. Rapid diagnosis and treatment planning depend on accurate tools that must also be easily available and efficient for clinicians. Motivated by these challenges, we propose a Modified Bi-Directional ConvLSTM U-Net (M-BDCU-Net) architecture combining dense connectivity and

temporal information to improve segmentation robustness and accuracy while greatly lowering computing costs. Our approach addresses the trade-off between efficiency and performance, especially for the identification of the Enhancing Tumor (ET) region, therefore offering a reasonable strategy for brain tumor segmentation that can be easily applied in clinical situations. This work aims to enhance patient outcomes and progress the field of medical imaging by bridging the gap between innovative research methods and the pragmatic needs of medical practitioners.

This paper is arranged as follows: Section II looks at related medical image analysis and brain tumor segmentation related work. Section III clarifies the architecture of the Modified Bi-Directional ConvLSTM U-Net (M-BDCU-Net). Section IV details the experimental design, together with evaluation measures and training settings. Section V ends the work by compiling our contributions and their possible influence on brain tumor identification and treatment.

II. RELATED WORKS

Brain tumor segmentation in medical imaging has been a subject of sustained and profound research interest for several decades. Over this period, a multitude of techniques and methodologies have emerged, all sharing the common objective of enhancing the precision and effectiveness of this critical task [8], [9].

In recent years, significant advancements have been made in cancer diagnosis through segmentation, leading to improved brain tumor detection. In particular, machine learning has had a transformative impact on the medical field. One notable approach was proposed in [6], where the authors employed SVM, K-Nearest Neighbors (KNN), and RF algorithms to detect brain tumors, testing their method on the BraTS2013 dataset. Similarly, in [7], researchers introduced a hybrid model that integrated Convolutional Neural Networks (CNN) with SVM on the BraTS2015 dataset. Their model effectively classified tumors as benign or malignant using a threshold-based algorithm for brain tumor detection. Recently, deep learning has revolutionized brain tumor segmentation. One of the most popular architectures is U-Net, which uses an encoder-decoder structure for precise segmentation. For instance, [10] applied a Fully Convolutional Network (FCN) with U-Net to segment brain tumors in the BraTS 2017 and BraTS 2019 datasets, achieving high accuracy. Several research studies have enhanced U-Net by including advanced features. [11] proposed the Self-Calibrated Attention U-Net (SCAU-Net), which uses external attention and self-calibrated convolutions. The model substitutes self-calibrated convolution modules to enhance feature extraction for conventional self-attention in skip connections with external attention. [12] introduced Edge U-Net, which incorporates MRI data connected to boundaries during decoding. Combining border data with contextual information at several levels improves tumor localization. [13] designed the Attention Gate Residual U-Net (AGResU-Net), which includes residual modules and attention gates to skip con-

nections to improve segmentation performance. This design reduces noise and highlights significant features better on the BraTS 2017, 2018, and 2019 datasets than conventional U-Net and Residual U-Net. A model combining multi-view attention and multi-scale feature interaction into U-Net was proposed in [14]. Interesting results were obtained on the BraTS 2019, 2020, and 2021 datasets.

Recently, researchers have applied transformers, originally developed for natural language processing, to brain tumor segmentation. With a model using a U-shaped network architecture was proposed in [15] a Swin transformer as the encoder and a CNN-based decoder, the efficacy of this methodology has been tested on the BraTS 2021. Other research [16] introduced Swin Transformer Brain Tumor Segmentation (SwinBTS), which uses a 3D Swin Transformer for encoding and decoding combined with convolutional operations. Another notable model is the Nested Modality-Aware Transformer (NestedFormer) [17]. It uses spatial and cross-modality attention transformers to improve relationships between and within modes in multi-modal MRIs. Three-dimensional models have become increasingly popular for brain tumor segmentation as they leverage spatial information from volumetric medical images. [18] proposed an end-to-end framework combining deep residual networks with U-Net for 3D segmentation. This approach improves

efficiency and accuracy by using residual connections to enhance feature learning.

Despite these advancements, several challenges remain. Many models struggle with tumors of varying sizes, shapes, and textures. Additionally, models often fail to generalize well across different datasets and accomplish a poor performance when tested on data they were not trained on. Brain tumors also have multiple sub-regions, such as the necrotic core, active tumor, and edema, each with unique properties. While advanced techniques like attention mechanisms and transformers have improved segmentation, a comprehensive solution to these challenges is still needed.

Another critical issue is the trade-off between model performance and computational complexity. Many state-of-the-art models, such as transformer-based architectures, achieve high accuracy but require significant computational resources. This makes them less practical for real-world clinical applications, where computational efficiency is crucial. For example, while models like NestedFormer and SwinBTS deliver impressive results, their high complexity requirements limit their scalability and deployment in resource-constrained environments. To address these gaps, we propose a new model that integrates BConvLSTM and DCC networks. Our method aims to make segmentation more accurate and reliable across a wide range of datasets

TABLE 1: Proposed architecture hyperparameters

Phase	Layer Type	Output Shape	Kernel Size / Strides	Activation / Other
Encoding	Input Layer	(None, 240, 240, 4)	-	-
	Convolutional Layer	(None, 240, 240, 64)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 240, 240, 64)	3×3, strides=1	ReLU
	Pooling Layer	(None, 120, 120, 64)	2×2, strides=2	MaxPooling
	Convolutional Layer	(None, 120, 120, 128)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 120, 120, 128)	3×3, strides=1	ReLU
	Pooling Layer	(None, 60, 60, 128)	2×2, strides=2	MaxPooling
	Convolutional Layer	(None, 60, 60, 256)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 60, 60, 256)	3×3, strides=1	ReLU
	Dropout Layer		-	Dropout(0.5)
	Pooling Layer	(None, 30, 30, 256)	2×2, strides=2	MaxPooling
	Convolutional Layer	(None, 30, 30, 512)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 30, 30, 512)	3×3, strides=1	ReLU
	Dropout Layer	(None, 30, 30, 512)	-	Dropout(0.5)
	Convolutional Layer	(None, 30, 30, 512)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 30, 30, 512)	3×3, strides=1	ReLU
	Dropout Layer	(None, 30, 30, 512)	-	Dropout(0.5)
	Concatenate Layer	(None, 30, 30, 1024)	-	Concatenate
	Convolutional Layer	(None, 30, 30, 512)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 30, 30, 512)	3×3, strides=1	ReLU
	Dropout Layer	(None, 30, 30, 512)	-	Dropout(0.5)
Decoding	Transposed Conv Layer	(None, 60, 60, 256)	2×2, strides=2	ReLU
	Batch Norm Layer	(None, 60, 60, 256)	-	BatchNormalization
	Activation Layer	(None, 60, 60, 256)	-	Activation(ReLU)
	Reshape Layer	(None, 1, 60, 60, 256)	-	Reshape
	Reshape Layer	(None, 1, 60, 60, 256)	-	Reshape
	Concatenate Layer	(None, 2, 60, 60, 256)	-	Concatenate
	ConvLSTM2D Layer	(None, 60, 60, 128)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 60, 60, 256)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 60, 60, 256)	3×3, strides=1	ReLU
	Transposed Conv Layer	(None, 120, 120, 128)	2×2, strides=2	ReLU
	Batch Norm Layer	(None, 120, 120, 128)	-	BatchNormalization
	Activation Layer	(None, 120, 120, 128)	-	Activation(ReLU)
	Reshape Layer	(None, 1, 120, 120, 128)	-	Reshape
	Reshape Layer	(None, 1, 120, 120, 128)	-	Reshape
	Concatenate Layer	(None, 2, 120, 120, 128)	-	Concatenate
	ConvLSTM2D Layer	(None, 120, 120, 64)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 120, 120, 128)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 120, 120, 128)	3×3, strides=1	ReLU
	Transposed Conv Layer	(None, 240, 240, 64)	2×2, strides=2	ReLU
	Batch Norm Layer	(None, 240, 240, 64)	-	BatchNormalization
	Activation Layer	(None, 240, 240, 64)	-	Activation(ReLU)
	Reshape Layer	(None, 1, 240, 240, 64)	-	Reshape
	Reshape Layer	(None, 1, 240, 240, 64)	-	Reshape
	Concatenate Layer	(None, 2, 240, 240, 64)	-	Concatenate
	ConvLSTM2D Layer	(None, 240, 240, 32)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 240, 240, 64)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 240, 240, 32)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 240, 240, 16)	3×3, strides=1	ReLU
	Convolutional Layer	(None, 240, 240, 8)	3×3, strides=1	ReLU
	Output Layer	(None, 240, 240, 4)	1×1, strides=1	Softmax

while still capturing the complex features of tumor sub-regions. Importantly, we focus on balancing performance and computational efficiency, ensuring that our model achieves high accuracy without excessive requirements. By bridging the gap between research and clinical practice, we hope to enhance diagnosis and patient care while making the model more accessible for real-world applications.

III. PROPOSED METHOD: M-BCDU-NET FUSION APPROACH FOR BRAIN TUMOR SEGMENTATION

Algorithm 1 Pseudo-code for the M-BCDU-Net framework

```

1: Input: MRI image dataset (e.g., BraTS 2020/2021)
2: Output: Segmented brain tumor regions
3: Load and preprocess the MRI image dataset
4: for each image in the dataset do
5:     Normalize the image intensity values
6:     Coding Path (Encoder):
7:         Pass the preprocessed image through the Bi-
            Directional ConvLSTM layers
8:         Extract spatiotemporal features using forward
            and backward ConvLSTM units
9:         Apply downsampling (e.g., strided convolutions or
            pooling) to reduce spatial dimensions
10:        Increase the number of feature maps to capture
            high-level features
11:    Decoding Path (Decoder):
12:        Pass the feature maps through the decoder, using
            upsampling layers (e.g., transposed convolutions)
13:        Reduce the number of feature maps while increas-
            ing the spatial resolution
14:        Use skip connections to combine high-resolution
            features from the encoder and decoder
15:        Apply convolutions to refine the segmentation
16:        Output the final segmented brain tumor regions
17: end for
18: Evaluate the performance using relevant metrics (e.g.,
    Dice coefficient, IoU)

```

The M-BCDU-Net approach is inspired by the influential architectures U-Net [23], BConvLSTM [24], and DCC [25]. It combines the strengths of these architectures and hopes to return leading results on the challenging task of brain tumor identification and localization within 3D medical images.

The segmentation process diagram in Figure 1 illustrates the workflow of extracting meaningful regions from medical images using a deep learning model. It begins with the input MRI scan, which undergoes pre-normalization and resizing normalization to enhance model performance. The image is then fed into an M-BCDU-Net architecture, as shown in figure 2, consisting of a contracting path (encoder) that captures spatial features through convolutional and max-pooling layers, followed by a bottleneck layer that learns high-level representations. The expanding path (decoder) reconstructs the segmented regions using upsampling layers and skip connections to retain spatial details.

A. ENCODING PATH

The encoding path generates high-dimensional feature representations with significant semantic information. It takes advantage of four steps of execution and captures the context through successive layers: The model initially processes input data through convolutional layers, each convolving the input with learnable filters, detailed in Table 1, to extract increasingly abstract features. These convolutional operations are interspersed with max-pooling layers to reduce spatial dimensions while retaining essential information. Finally, the ReLU activation function is used to introduce non-linearity into the network, enabling it to learn complex patterns in the data. These steps enable the network to learn hierarchical abstract features that are essential in understanding brain structures.

As the encoding phase progresses, the model's capacity to capture intricate features is augmented through the stacking of convolutional layers, which allow the network to progressively capture more complex image features with reduced dimensionality. Effectively, the first, second, and third steps are composed of 2 convolutional layers using, respectively, 64, 128, and 256 [3×3] filters followed by a 2×2 max pooling operation and ReLU activation. Dropout layers are strategically inserted to regularize the model to prevent overfitting by randomly dropping units during training and to promote the emergence of robust features: X_c^i with $i = [0, \dots, 2]$ copied feature maps extracted from several convolution steps. The concatenation of feature maps from different layers enriches the representation with diverse spatial information, enhancing the model's discriminative capabilities. In the final step, the sequence of convolutional layers used in the original U-Net process is replaced by DCC [25] to enhance the reuse of features and the flow of gradients through the model, enabling it to learn more discriminative representations. This subsequently strongly enhances network potential for finer information capture, which is necessary for brain tumor classification. Considered DCC, employed at the final convolutional layer of the encoding path, is composed of N blocks. Each block is a succession of 2 convolutions. The input of each block receives the concatenated feature maps of all preceding blocks to ensure maximal information flow and mitigate the vanishing gradient problem. The obtained output of the encoding path is noted $X_d = [X_d^1, \dots, X_d^N]$. This method improves feature propagation and reuse within the network and enhances the capability to capture complicated and varied structures within the brain at the same time. Therefore, the DCC will considerably enhance the representational power and, consequently, the performance of the tasks of brain segmentation.

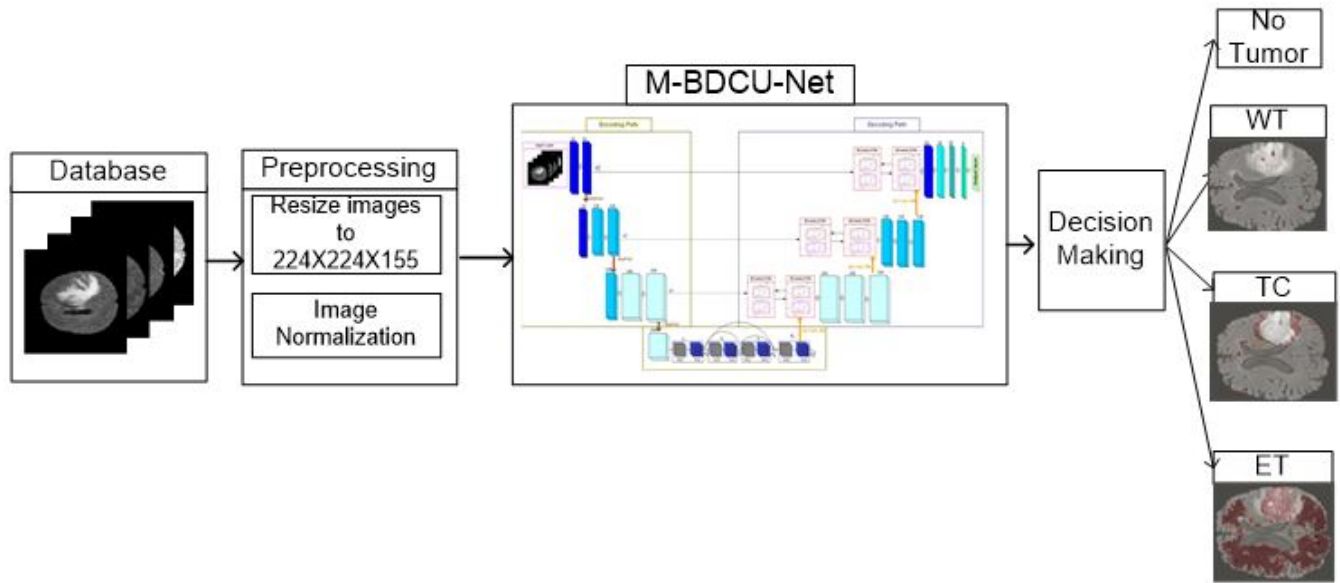


FIGURE 1: Brain Tumor Segmentation Process Using M-BDCU-Net

B. DECODING PATH

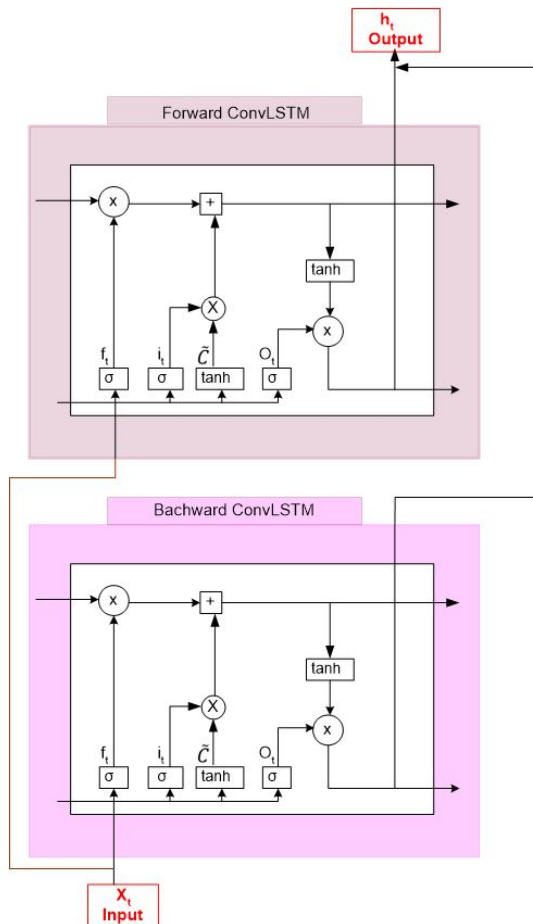


FIGURE 3: Architecture of the BConvLSTM

The decoder path restores the spatial resolution through up-sampling operations. Effectively, transposed convolutional layers reverse the downsampling performed during encoding, facilitating the reconstruction of higher-resolution feature maps. Reshape and concatenation layers prepare the data for the subsequent BConvLSTM layer. Instead of cropping the feature-copying X_c feature map to the decoder path, as used in the basic U-Net architecture, we employ the BConvLSTM algorithm to process X_c and X_d feature maps, capturing temporal dependencies and refining the representation further. X_d first goes through an upsampling function, then batch normalization and activation functions to keep the network stable and nonlinear. This step ensures the fidelity of the reconstructed features and increases the stability of the neuronal network. High-resolution features from the contracting path are added to the decoding path through BConvLSTM connections. This system keeps track of previous and subsequent contexts during any decision-making process. The U-Net combines context and spatial information by joining these features together to make accurate and detailed segmentation maps of brain MRIs. The structure of the BConvLSTM used in our algorithm is illustrated in Figure 3. The BConvLSTM consists of two ConvLSTM layers: a forward ConvLSTM layer that processes the input sequence in chronological order, while the backward layer processes it in reverse order. The final output is a combination of the hidden states from both directions. The equations governing the operations of the BConvLSTM are as follows:

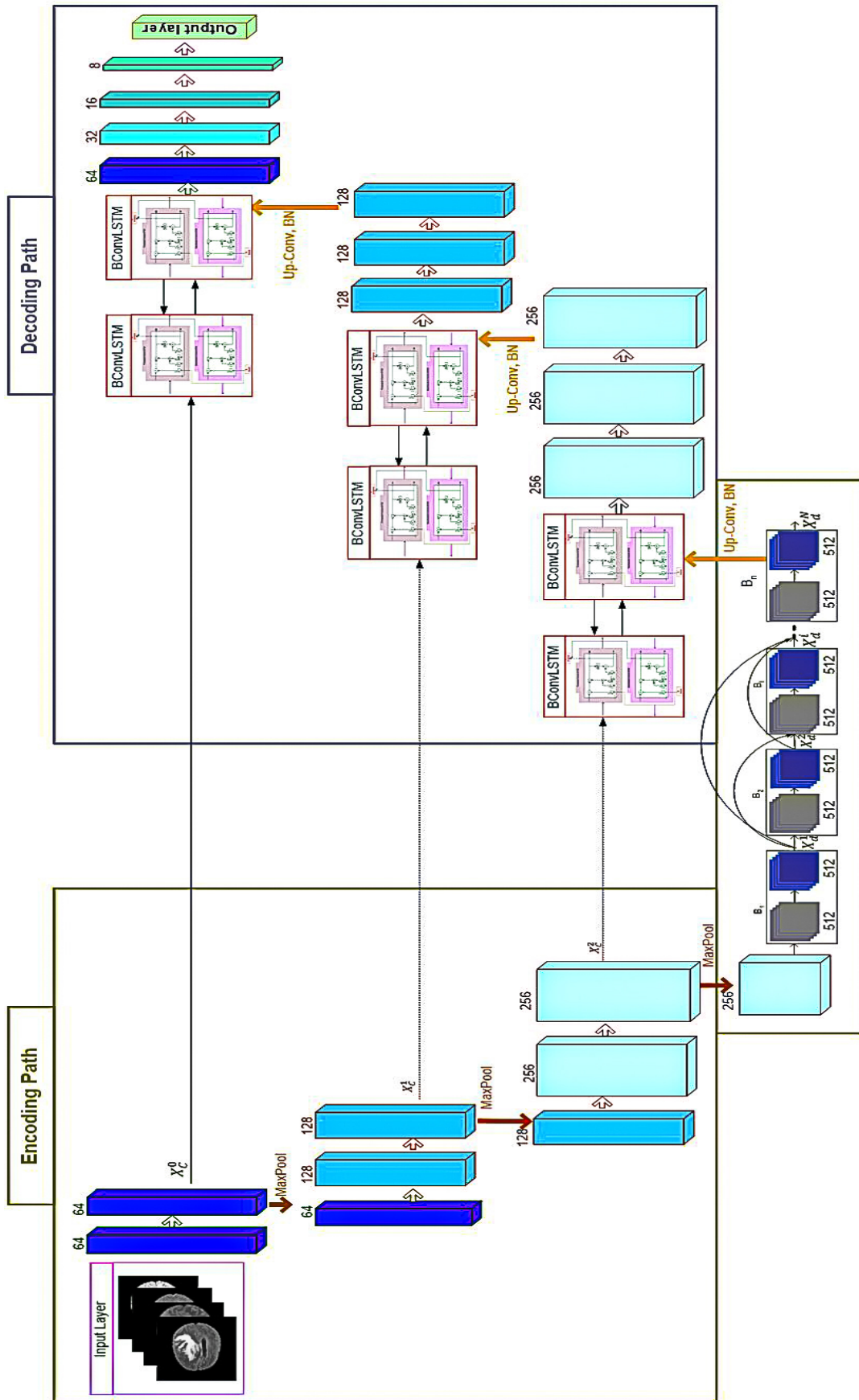


FIGURE 2: M-BDCU-NET architecture block diagram

$$\begin{aligned} i_t &= \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + b_i), \\ f_t &= \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + b_f), \\ o_t &= \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + b_o), \\ \tilde{C}_t &= \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c), \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \\ H_t &= o_t \odot \tanh(C_t), \end{aligned}$$

where:

- X_t is the input at time t ,
- h_t is the hidden state at time t ,
- C_t is the cell state at time t ,
- i_t , f_t , and o_t are the input, forget, and output gates, respectively,
- \tilde{C}_t is the candidate cell state,
- σ is the sigmoid activation function,
- $*$ denotes the convolution operation,
- \odot denotes element-wise multiplication,
- W and b are the weights and biases, respectively.

The final output of the BConvLSTM is a combination of the hidden states from both the forward and backward ConvLSTM layers. Using the BConvLSTM, our model accurately captures temporal dependencies going both forward and backward. This technique makes it easier to see sequential MRI slices and improves the accuracy of segmentation. The model culminates in convolutional layers to generate the final segmentation mask outputs to process the classification step. Table 1 encapsulates, in detail, the intricate interplay of convolution, pooling, normalization, and activation operations. These details elucidate the network's hierarchical processing of input data and offer a comprehensive understanding of its functionality and expressive power.

C. DATASETS DESCRIPTIONS

In the context of brain tumor segmentation, the labels WT (Whole Tumor), ET (Enhancing Tumor), and TC (Tumor Core) hold significant clinical importance and play a crucial role in understanding the different regions of interest within a brain MRI. Accurate segmentation based on these categories helps clinicians effectively tailor treatments.

- **WT (Whole Tumor):**
This label refers to the entire tumor region, encompassing all abnormal tissues. It includes the necrotic core (the dead part of the tumor), the surrounding edema (fluid accumulation caused by inflammation or damage from the tumor), and the enhancing tumor area. WT provides a comprehensive view of the tumor's total extent, which is critical for evaluating tumor progression and overall size.
- **ET (Enhancing Tumor):**
This label specifically identifies the active tumor area that appears enhanced on MRI images after the injection of a contrast agent. This region is associated with aggressive and proliferative tumor cells, as the contrast

agent highlights areas with high tumor vascularization (often a sign of actively growing tumors). ET is, therefore, a critical indicator of malignancy and tumor activity.

- **TC (Tumor Core):**
This label includes the necrotic (dead) tumor tissue and the enhancing tumor but excludes the surrounding edema. It focuses on the tumor's central and most critical regions. TC is important for identifying the active core or central parts of the tumor, which often require special attention during treatment planning or surgical interventions.











The comparison between the BraTS datasets from 2018 to 2021, as shown in Table 2, highlights significant advancements over the years. The table demonstrates that all the datasets share the same modalities and segmentation labels. However, the more recent datasets, BraTS 2020 and BraTS 2021, include significant enhancements, such as larger dataset sizes, advanced data augmentation techniques, and improved annotations. These advancements make BraTS 2020 and BraTS 2021 more suitable for evaluating the robustness and generalizability of our approach. Therefore, we chose to test our model on these two datasets to ensure alignment with the latest standards and to leverage their improved features. BraTS 2021 substantially increased the dataset size, offering more training, validation, and testing cases, which can significantly improve the training process and model performance due to the larger data pool. To assess the efficacy of our proposed architecture, we leveraged two distinct datasets within our study. The first dataset, BraTS2020, encompasses a collection of 369 training images, 125 validation images, and 166 testing images. Correspondingly, each image within the BraTS2020 dataset shares the same dimensions as those within BraTS2021, specifically $240 \times 240 \times 155$. The second dataset utilized is BraTS2021, a contemporary publicly accessible repository comprising 1251 training images and 219 validation images. Each image within this dataset is characterized by dimensions measuring $240 \times 240 \times 155$. Both datasets provide the same image modalities (Flair, t1, t1ce, and t2) and segmentation labels, ensuring consistency in the type of data provided. Notably, we visually represented a sample slice from each modality within these datasets (Table 3). Each 3D scan comprises a total of 155 slices, with each slice possessing dimensions of 240×240 pixels. The introduction of advanced data augmentation techniques in BraTS2021, such as elastic deformations and intensity shifts, marks a critical enhancement over the basic augmentations used in BraTS2020. This can lead to better generalization and robustness of models trained on BraTS2021 data.

Additionally, BraTS2021 benefits from a combination of manual and semi-automated annotations, which potentially increases the accuracy and consistency of the ground truth labels. BraTS2021 has expanded the benchmark metrics to include the Normalized Surface Dice, offering a more comprehensive evaluation of segmentation performance.

TABLE 2: Comparative Analysis of BraTS Datasets

Aspect	BraTS2018	BraTS2019	BraTS2020	BraTS2021
Dataset Size	285 training cases, 66 validation cases, 191 testing cases	335 training cases, 125 validation cases, 166 testing cases	369 training cases, 125 validation cases, 166 testing cases	1251 training cases, 219 validation cases, 530 testing cases
Image Modalities	Flair, t1, t1ce, t2			
Segmentation Labels	Enhancing tumor (ET), tumor core (TC), whole tumor (WT)			
Challenge Tasks	Tumor segmentation, Survival prediction	Tumor segmentation, Survival prediction	Tumor segmentation, Survival prediction, Uncertainty estimation	Tumor segmentation, Survival prediction, Uncertainty estimation
Data Augmentation	Basic augmentations: rotations, flips			Advanced augmentations: rotations, flips, elastic deformations, intensity shifts
File Format	NIfTI (.nii.gz)			
Data Annotations	Manual annotations by expert radiologists			Manual and semi-automated annotations by expert radiologists
Benchmark Metrics	Dice score, Hausdorff distance			Dice score, Hausdorff distance, Normalized Surface Dice
Notable Improvements	Introduction of survival prediction task	Increased dataset size, improved annotations	Enhanced evaluation metrics, uncertainty estimation	Largest dataset, advanced augmentation strategies, normalized surface dice metric

TABLE 3: Multimodal MRI of a Single Brain Slice from BraTS2020 and BraTS2021 Dataset

	Flair	t1	t1ce	t2	Mask
BraTS2020					
BraTS2021					

D. PREPROCESSING

In the domain of medical image analysis, labeled data is often limited. Consequently, data augmentation techniques are used to enhance the generalization capability of machine learning models. In this study, we implemented this crucial technique as part of the data generation process to train a deep learning model for brain tumor segmentation. Several transformations were applied to the input images, including random rotations, translations, shearing, zooming, and horizontal flipping. These transformations were carefully chosen to simulate variations in image acquisition, such as different angles, patient positioning, and magnification levels, while maintaining spatial coherence across different MRI modalities. Additionally, appropriate strategies were used to fill in newly created pixels after transformations to ensure no information loss. The augmented images were resized to the target dimensions and normalized to facilitate efficient training. This comprehensive augmentation process led to a significant improvement in the model's performance, as evidenced by higher dice coefficients and mean intersection over union (IoU) metrics on the validation set, highlighting the importance of data augmentation in training robust and generalizable models in medical image segmentation.

IV. RESULTS AND COMPARATIVE ANALYSIS

A. EXPERIMENTAL SETUP AND PARAMETERS

All experiments were carried out on a server equipped with dual Intel Xeon Gold 6326 CPUs (totaling 32 cores and 64 threads, running at 2.9-3.5 GHz), 512 GB of RAM, and a storage array of three 4 TB NVMe SSDs configured in RAID0. The server also featured six Nvidia A40 GPUs, each with 48 GB of VRAM. The proposed model was developed using TensorFlow and Keras as the backend frameworks. The Adam optimizer was employed for optimal performance, as it is recognized as one of the most effective optimizers. Image pixel intensity levels were normalized to mitigate potential complexity.

B. EXPERIMENTAL SETTINGS: ASSESSING MODEL PERFORMANCE

To evaluate our model's performance, we employ metrics tailored to this task: Dice Similarity Coefficient (DSC) and accuracy. DSC: T is a robust measure of accuracy in MRI scans. This specific metric quantifies the overlap between predicted (P) and ground truth (Q) sets using the formula:

$$DSC(P, Q) = \frac{2 \times |\text{intersection}(P, Q)|}{|P| + |Q|} \quad (1)$$

Accuracy: This metric gauges the alignment between calculated results and actual data, computed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Here, TP signifies True Positives, TN denotes True Negatives, FP represents the count of real positive cases, and FN signifies the count of real negative cases in the dataset.

Sensitivity:

Sensitivity measures the proportion of actual positives that

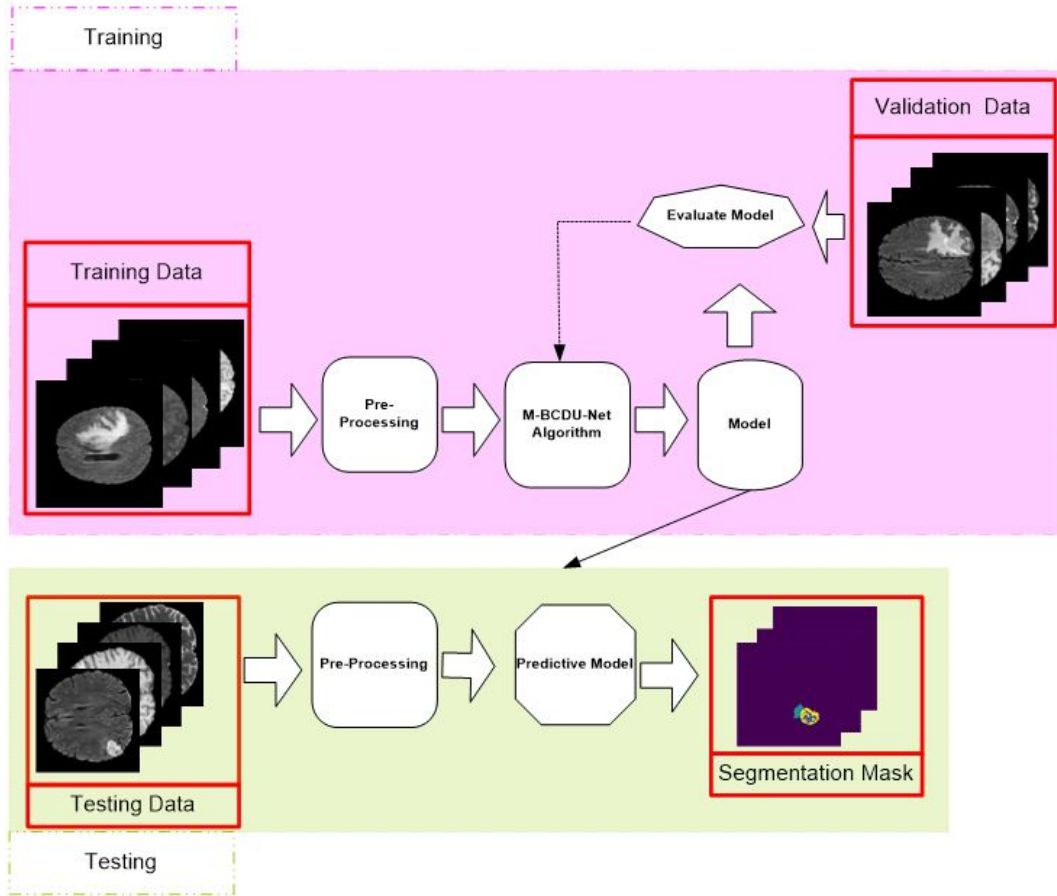


FIGURE 4: Steps of experimentation process

TABLE 4: Performance Metrics of the Segmentation Model on BraTS2020 and BraTS2021 Datasets

Models	Dataset	Accuracy	Precision	Sensitivity	Mean_IOU	DSC
U-net	BraTS2020	0.92	0.91	0.90	0.82	0.78
	BraTS2021	0.90	0.89	0.88	0.80	0.76
U-Net + BConvLSTM	BraTS2020	0.94	0.93	0.92	0.85	0.81
	BraTS2021	0.92	0.91	0.90	0.83	0.79
M-BDCU-NET	BraTS2020	0.98	0.99	0.98	0.94	0.80
	BraTS2021	0.95	0.94	0.94	0.81	0.82

are correctly identified by the model. It is calculated as:

$$\text{Sensitivity} = \frac{TN}{TP + FN} \quad (3)$$

Precision:

Precision measures the proportion of predicted positives that are correct. It is calculated as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

In this work, we employ a combined loss function for multiclass segmentation that integrates Categorical Cross-Entropy Loss(CCE Loss) with Dice Loss.

CCE Loss measures the difference between the true class labels and the predicted probabilities. For each pixel i , it is given by:

$$\text{CCE Loss} = - \sum_{c=1}^C y_{i,c} \log(p_{i,c}), \quad (5)$$

where C is the number of classes, $y_{i,c}$ is the binary indicator for class c , and $p_{i,c}$ is the predicted probability for class c .

Dice Loss assesses the overlap between the predicted and actual binary masks. For each class k , the Dice Coefficient is calculated as:

$$\text{Dice Coefficient}_k = \frac{2 \sum_{i=1}^N p_{i,k} \cdot y_{i,k} + \text{smooth}}{\sum_{i=1}^N p_{i,k} + \sum_{i=1}^N y_{i,k} + \text{smooth}}, \quad (6)$$

where N is the number of pixels, $p_{i,k}$ is the predicted probability for pixel i of class k , and $y_{i,k}$ is the ground truth label. The Dice Loss is then:

$$\text{Dice Loss} = 1 - \frac{1}{K} \sum_{k=1}^K \text{Dice Coefficient}_k, \quad (7)$$

Where K is the number of classes.

The combined loss function is the sum of the Categorical Cross-Entropy Loss and the Dice Loss:

$$\text{Combined Loss} = \text{CCE Loss} + \text{Dice Loss}. \quad (8)$$

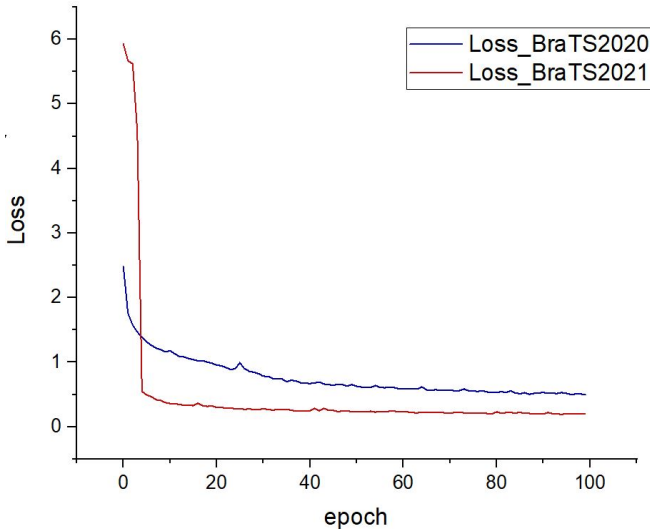


FIGURE 5: Training Loss Curves for BraTS2020 and BraTS2021 Datasets

This combined approach helps balance pixel classification and segmentation accuracy, which is particularly useful for handling class imbalances. Figure 5 shows the training loss curves for two different datasets, BraTS2020 and BraTS2021. The curve shows the training loss decreasing towards zero over 100 epochs, demonstrating that the training process is effective and the model is successfully improving.

C. RESULTS AND DISCUSSION

In this section, the performance of the proposed model is evaluated, and a detailed comparative analysis is conducted. As detailed in Fig. 4, the experimentation process involves training and testing steps. The M-BDCU-NET algorithm initially receives pre-processed images to train the model through cross-validation. The trained model uses testing data to perform predictions once the training process is complete.

Table 4 highlights the performance of the proposed segmentation model on the BraTS2020 and BraTS2021 datasets, evaluated across multiple metrics: accuracy, precision, sensitivity, Mean IoU, and Dice Score. We did an ablation study to find out what each part of our proposed M-BDCU-Net architecture does. We compared the baseline U-Net, U-Net with Bi-Directional ConvLSTM, and the full M-BDCU-Net.

The baseline U-Net achieves moderate performance, with a Dice Score (DSC) of 0.78 on BraTS2020 and 0.76 on BraTS2021. While it demonstrates reasonable accuracy (0.92 on BraTS2020 and 0.90 on BraTS2021), its limitations in capturing complex tumor features are evident, particularly in the lower Mean IoU (0.82 on BraTS2020 and 0.80 on BraTS2021). This highlights the need for additional components to improve feature learning and segmentation robustness. Adding the BConvLSTM module improves temporal feature learning, as reflected in the increased Dice Score (0.81 on BraTS2020 and 0.79 on BraTS2021). The model also shows improvements in Sensitivity (0.92 on BraTS2020 and 0.90 on BraTS2021) and Mean IoU (0.85 on BraTS2020 and 0.83 on BraTS2021). These results demonstrate the importance of modeling temporal dependencies in MRI slices for better segmentation performance. The full proposed model, which integrates U-Net, BConvLSTM, and DCC, achieves the highest performance. The model demonstrates high accuracy on both datasets, achieving 0.98 on BraTS2020 and 0.95 on BraTS2021. The BraTS2021 dataset shows a slight decrease in accuracy. This variation is possibly due to the increased complexity of the newer dataset. Precision is also consistently strong, with values of 0.99 for BraTS2020 and 0.94 for BraTS2021. This observation indicates the model's effectiveness in correctly identifying positive cases. Sensitivity is similarly high, at 0.98 for BraTS 2020 and 0.94 for BraTS 2021, reflecting the model's ability to detect true positives. However, there is a noticeable drop in Mean IoU, from 0.94 on BraTS 2020 to 0.81 on BraTS2021. The model might struggle with certain aspects of overlap in the newer dataset. Interestingly, the Dice Score goes up from 0.80 for BraTS2020 to 0.82 for BraTS2021, which shows that the model does a little better at segmenting the new dataset. So, even though other metrics show that the model may not be as good at segmenting with the BraTS 2021 data, its agreement with the ground truth gets better. This improvement could be attributed to the model's enhanced ability to manage class imbalances or tackle challenging cases more effectively in the newer dataset.

For each dataset, we present a comprehensive set of images to evaluate our model's performance in figure 6. The display includes the original image alongside the FLAIR sequence, which enhances the visibility of specific structures. We also show the ground truth annotations, which provide the definitive reference for all classes. The model's predictions are visualized for comparison, highlighting three key classes TC, EC, and WT. This comparative visualization enables a thorough assessment of the model's segmentation accuracy and its ability to correctly identify and classify different tissue types or abnormalities.

D. CROSS-VALIDATION OF THE PROPOSED MODEL

To rigorously assess the generalizability and robustness of the proposed model, we conducted a cross-dataset evaluation by training the model on one dataset and testing it on another. Specifically, we trained the model on the BraTS2020

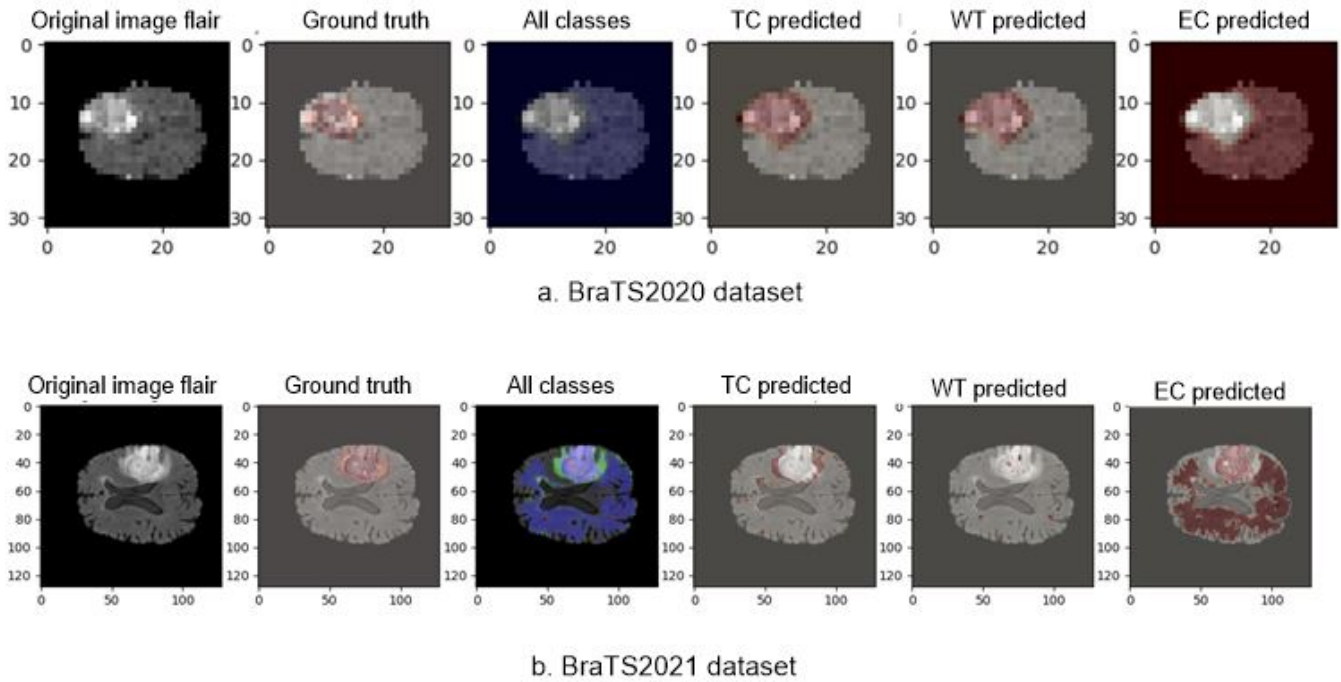


FIGURE 6: Prediction results for models trained on the BraTS2020 and BraTS2021 datasets.

dataset and evaluated its performance on the BraTS2021 dataset (First-Test (BraTS2020/BraTS2021)), and vice versa (Second-Test (BraTS2021/BraTS2020)). This approach helps identify how well the model can adapt to unseen data with different characteristics, which is critical for ensuring its practical applicability in real-world scenarios.

TABLE 5: Cross-Dataset Evaluation of the Proposed Model

Metrics	Tests	
	First-Test (BraTS2020/BraTS2021)	Second-Test (BraTS2021/BraTS2020)
Accuracy	0.90	0.97
Mean_IoU	0.81	0.70
DSC	0.72	0.81

The results, as shown in Table 5, reveal significant variations in performance between the two tests. In the first test, where the model trained on BraTS 2020 was evaluated, the model achieved an accuracy of 0.90, a MeanIoU of 0.81, and a DSC of 0.72. In contrast, the second test, where the model was trained on BraTS 2021, showed improved performance with an accuracy of 0.97, a Mean IoU of 0.70, and a Dice Score of 0.81. These differences show that the model trained on the more recent BraTS 2021 dataset is better able to handle tough cases, as shown by the higher Dice Score and accuracy, even though the Mean IoU is lower. This suggests that performance depends on which dataset was used for training.

E. COMPARISON WITH RELATED WORK METHODS

The results of the M-BDCU-NET model are presented in Table 6 for comparison across different datasets and method-

ologies. This table summarizes the DSC values for tumor types across various models and datasets, specifically BraTS 2020 and BraTS 2021. The M-BDCU-NET model shows relatively lower performance in WT tumor type segmentation on the BraTS 2020 dataset, with DSC values of 0.80 for TC, 0.82 for ET, and 0.78 for WT. This behavior may be attributed to the complexity of tumor delineation in the BraTS 2020 dataset, where other models, such as NestedFormer and Deep Residual U-Net, with more sophisticated architectures, achieve higher DSC scores in WT tumor prediction. These models likely benefit from enhanced feature extraction capabilities and deep learning mechanisms tailored to handle the complex variability in WT regions. However, the M-BDCU-NET model exhibits a notable improvement on the BraTS 2021 dataset, achieving DSC values of 0.81 for TC, 0.85 for ET, and 0.82 for WT. This suggests that the M-BDCU-NET model's design is more suitable for the characteristics of the BraTS 2021 dataset, which may involve differences in tumor size, image quality, or a more refined model architecture that is better at segmenting various tumor regions. The performance improvement is particularly significant in the segmentation of ET regions, where the M-BDCU-NET excels relative to other models, even surpassing the performance of some existing methodologies. This indicates that the model is better equipped to capture active tumor regions with high accuracy.

When comparing our results to existing models, we observe that while M-BDCU-NET shows competitive performance, especially on the BraTS 2021 dataset, it lags behind the best-performing models on BraTS 2020. For example, models such as 3D Deep Residual U-Net [18], NestedFormer

TABLE 6: Comparison of proposed model with state-of-the-art

Study	Architecture	Image Dimension	Dataset	Dice Score (DSC)			Model Complexity	
				TC	ET	WT	Parameters (M)	FLOPs (G)
[11]	SCAU-net	128×128×128	BraTS2018	0.85	0.80	0.90	33.24	1393.36
			BraTS2019	0.82	0.78	0.90		
			BraTS2020	0.82	0.78	0.90		
[13]	U-Net	240×240×155	BraTS2017	0.75	0.69	0.85	—	—
			BraTS2018	0.79	0.75	0.86		
			BraTS2019	0.74	0.69	0.86		
	AGU-Net	240×240×155	BraTS2017	0.77	0.72	0.87	—	—
			BraTS2018	0.79	0.77	0.87		
			BraTS2019	0.76	0.70	0.87		
	ResU-Net	240×240×155	BraTS2017	0.77	0.73	0.86	—	—
			BraTS2018	0.80	0.76	0.87		
			BraTS2019	0.76	0.70	0.86		
	AGResU-Net	240×240×155	BraTS2017	0.78	0.74	0.88	—	—
			BraTS2018	0.80	0.77	0.87		
			BraTS2019	0.77	0.70	0.87		
[14]	MVSI-Net	160×160×155	BraTS2019	0.88	0.80	0.86	47.6	73.2
			BraTS2020	0.87	0.81	0.87		
			BraTS2021	0.91	0.81	0.88		
[15]	Swin UNETR	128×128×128	BraTS2021	0.88	0.85	0.92	61.98	394.84
[17]	NestedFormer	240×240×155	BraTS2020	0.86	0.80	0.92	10.48	71.77
[18]	3D Deep Residual U-net	128×128×128	BraTS2020	0.83	0.80	0.86	30.47	374.04
Proposed Model	M-BDCU-NET	240×240×155	BraTS2020	0.80	0.82	0.78	8.81	60.34
			BraTS2021	0.81	0.85	0.82		

[17], MSVI-NET [14], and Swin UNTER [15] exhibit superior segmentation performance in the BraTS 2020 dataset's WT tumor type. This suggests that these models may incorporate advanced techniques, such as multi-scale feature learning or attention mechanisms, that improve WT tumor region segmentation. Despite this, the M-BDCU-NET model's balance of accuracy and efficiency stands out. It significantly reduces computational complexity while maintaining high segmentation quality, with only 8.81M parameters and 60.34 GFLOPs, compared to NestedFormer's 10.48M parameters and 71.77 GFLOPs, the 3D dResU-Net's 30.47M parameters and 374.04 GFLOPs, the MSVI-NET's 47.6M parameters and 73.2 GFLOPs, and Swin UNTER's 61.98M parameters and 394.84 GFLOPs. The SCAU-Net with its 33.24M parameters and 1393.36 GFLOPs, also demonstrates a much higher computational cost. In summary, while M-BDCU-NET may not outperform the leading models in certain areas, it presents a valuable trade-off, providing competitive DSC scores, particularly in ET detection, while being significantly lighter in terms of computational resources. This makes M-BDCU-NET an attractive option for applications requiring efficient and accurate tumor segmentation with relatively lower computational demands.

V. CONCLUSION

The importance of this work is underscored by the growing need for reliable, efficient, and automated tools for brain tumor diagnosis. For timely and effective treatment planning, it is important to accurately segment 3D MRI data, especially when dealing with the complexity of heterogeneous brain tumors. The proposed M-BDCU-Net provides a practical and robust solution to this challenge, making it a valuable tool for the medical imaging and scientific communities. The pro-

posed M-BDCU-Net combines Bi-Directional ConvLSTM layers and densely connected convolutions. The model offers a perfect balance between accuracy and computational efficiency. On the BraTS2021 dataset, it achieved competitive performance with DSC values of 0.81, 0.85, and 0.82, respectively, for TC, ET, and WT tumor regions. Additionally, the model demonstrates a significant reduction in complexity, with a 15% decrease in the worst case compared to state-of-the-art models. Furthermore, the model's efficiency and robustness make it well-suited for real-world applications. These extensions would further enhance the applicability of the proposed approach, ensuring its impact on both clinical practice and future AI-driven medical research.

REFERENCES

- [1] A., Javaria, S., Muhammad, RA., Mudassar, T. Saba, M. A. Anjum. Brain tumor detection using statistical and machine learning method. Computer methods and programs in biomedicine, vol. 177, p. 69-79, 2019.
- [2] G. Hemanth, M. Janardhan, and L. Sujihelen, "Design and implementing brain tumor detection using machine learning approach," in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1289–1294, 2019.
- [3] P. S. Chander, J. Soundarya, and R. Priyadharsini, "Brain tumour detection and classification using K-means clustering and SVM classifier," in RITA 2018: Proceedings of the 6th International Conference on Robot Intelligence Technology and Applications, pp. 49–63, 2020.
- [4] J. Seetha and S. S. Raja, "Brain tumor classification using convolutional neural networks," Biomedical & Pharmacology Journal, vol. 11, no. 3, p. 1457, 2018.
- [5] S. Kumar and D. P. Mankame, "Optimization driven deep convolution neural network for brain tumor classification," Biocybernetics and Biomedical Engineering, vol. 40, no. 3, pp. 1190-1204, 2020.
- [6] M. Havaei, H. Larochelle, P. Poulin, P. Jodoin, Within-brain classification for brain tumor segmentation. International journal of computer assisted radiology and surgery, 2016, vol. 11, p. 777-788, 2016
- [7] K. Mohammad Omid, S. Meenakshi, J. Vishal, J. M. Chatterjee, N. Z. Jhanjhi, A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. Irbm, 2022, vol. 43, no 4, p. 290-299.

- [8] C. Zhang, H. Ge, S. Zhang, D. Liu, Z. Jiang, C. Lan, L. Li, H. Feng, R. Hu, Hematoma evacuation via image-guided para-corticospinal tract approach in patients with spontaneous intracerebral hemorrhage. *Neurology and Therapy*, Vol. 10, pp. 1001-1013, 2021.
- [9] L. Wang, X. Li, Z. Deng, Q. Cai, P. Lei, H. Xu, S. Zhu, T. Zhou, R. Luo, C. Zhang, Y. Yin, S. Zhang, N. Wu, H. Feng, R. Hu, Neuroendoscopic parafascicular evacuation of spontaneous intracerebral hemorrhage (NESICH technique): a multicenter technical experience with preliminary findings. *Neurology and Therapy*, vol. 13, no. 4, p. 1259-1271, 2024.
- [10] B. Jena, S. Jain, G. K. Nayak, and S. Saxena, "Analysis of depth variation of U-NET architecture for brain tumor segmentation," *Multimedia Tools and Applications*, vol. 82, no. 7, pp. 10723–10743, 2023.
- [11] D. Liu, N. Sheng, Y. Han, Y. Hou, B. Liu, J. Zhang, and Q. Zhang, "SCAU-net: 3D self-calibrated attention U-Net for brain tumor segmentation," *Neural Computing and Applications*, pp. 1–13, 2023.
- [12] A. M. Gab Allah, A. M. Sarhan, and N. M. Elshennawy, "Edge U-Net: Brain tumor segmentation using MRI based on deep U-Net model with boundary information," *Expert Systems with Applications*, vol. 213, p. 118833, 2023.
- [13] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, "Attention gate resU-Net for automatic MRI brain tumor segmentation," *IEEE Access*, vol. 8, pp. 58533–58545, 2020.
- [14] SUN, Junding, HU, Ming, WU, Xiaosheng, et al. MVSI-Net: Multi-view attention and multi-scale feature interaction for brain tumor segmentation. *Biomedical Signal Processing and Control*, 2024, vol. 95, p. 106484.
- [15] HATAMIZADEH, Ali, NATH, Vishwesh, TANG, Yucheng, et al. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In : *International MICCAI brainlesion workshop*. Cham : Springer International Publishing, 2021. p. 272-284.
- [16] Y. Jiang, Y. Zhang, X. Lin, J. Dong, T. Cheng, and J. Liang, "SwinBTS: A method for 3D multimodal brain tumor segmentation using swin transformer," *Brain Sciences*, vol. 12, no. 6, p. 797, 2022.
- [17] Z. Xing, L. Yu, L. Wan, T. Han, and L. Zhu, "NestedFormer: Nested modality-aware transformer for brain tumor segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 140–150, 2022.
- [18] R. Raza, U. I. Bajwa, Y. Mehmood, M. W. Anwar, and M. H. Jamal, "dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI," *Biomedical Signal Processing and Control*, vol. 79, p. 103861, 2023.
- [19] N. Ameri, N. Shoeibi, and M. Abrishami, "Segmentation of Hard Exudates in Retina Fundus Images Using BCDU-Net," in *2022 12th International Conference on Computer and Knowledge Engineering (ICCCKE)*, pp. 123–128, 2022.
- [20] B. M. S. Rani, V. R. Rajeev, V. P. Srinivasan, S. Thenmalar, and R. Kanimozhi, "Disease prediction based retinal segmentation using bi-directional ConvLSTMU-Net," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–10, 2021.
- [21] R. UN and G. MA, "BCDU-Net and chronological-AVO based ensemble learning for lung nodule segmentation and classification," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, no. 4, pp. 1491–1511, 2023.
- [22] K. K.L. Wong, W. Xu, M. Ayoub, Y. Fu, H. Xu, R. Shi, M. Zhang, F. Su, Z. Huang, W. Chen. Brain image segmentation of the corpus callosum by combining Bi-Directional Convolutional LSTM and U-Net using multi-slice CT and MRI. *Computer Methods and Programs in Biomedicine*, vol. 238, p. 107602, 2023.
- [23] G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, "Medical image segmentation based on U-Net: A review," *Journal of Imaging Science & Technology*, vol. 64, no. 2, 2020.
- [24] H. Song, W. Wang, S. Zhao, J. Shen, and K. M. Lam, "Pyramid dilated deeper convlstm for video salient object detection," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 715–731, 2018.
- [25] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, 2017.



crowwave circuits, artificial intelligence, Algorithm-Architecture Matching, and others.

MESHARI D. ALANAZI received the B.Sc. degree in electrical engineering (electronics and communications) from Jouf University, Saudi Arabia, the M.Sc. degree in electrical engineering from Bridgeport University, CT, USA, in 2015, and the Ph.D. degree in electrical engineering from the University of Sheffield, Sheffield, UK, in 2022. He is an Assistant Professor at the Department of Electrical Engineering at Jouf University. His research interests include antenna design, microwave circuits, artificial intelligence, Algorithm-Architecture Matching, and others.



AMNA MARAOUI received her Ph.D degree in Micro and Nano-electronics from the Faculty of Science of Monastir. Currently a member of Networked Objects, Control, and Communication Systems Laboratory(NOCCS). Her research interests include artificial intelligence, image and video processing, HEVC, and Future video coding (FVC) standards.



IMEN WERDA (Member, IEEE) received her Ph.D. degree in Electrical Engineering from the National School of Engineers of Sfax in 2009 and her HDR in 2023. She is currently a member of the Laboratory of Electronic and Information Technology. Her research includes artificial intelligence and video processing in graphics processors.



AHMED BEN ATALLAH received his Ph.D. degree in Electronics from the University of Bordeaux1 in 2007, France and the HDR Degree in 2013. He is currently an Associate Professor in Electronics at College of Engineering, Jouf University, Kingdom of Saudi Arabia. His main research activities are focused on image and Video Processing, Algorithm-Architecture Matching, Artificial Intelligence, SoC/SoPC architecture.



TURKI M. ALANAZI received the B.Sc. degree (Hons.) in electrical engineering from Bluefield State University, Bluefield, WV, USA, in 2010, the M.Sc. degree in electrical engineering from Gannon University, Erie, PA, USA, in 2011, and the Ph.D. degree in electrical engineering from the University of Dayton, Dayton, OH, USA, in 2018. Since 2020, he served as the Head of the Electrical Engineering Department, Vice Dean of Scientific Research and Dean of Scientific Research at Jouf University, Saudi Arabia. He is currently an Associate Professor at University of Hafr Al Batin, Saudi Arabia, a position he has held since 2023. His research interests include stochastic optimization, biomedical image, Computer vision, and signal processing. He was a recipient of several awards from the Saudi Arabian Cultural Mission, USA, and Bluefield State University.



MOHAMMED ALBEKAIRI is an expert in the domain of Electrical Engineering, specializing in areas such as Cybersecurity, Machine Learning, and Image Processing. Dr. Albekairi successfully completed his academic trajectory by achieving a highly esteemed PhD degree from the University of South Florida. He presently holds the position of assistant professor within the Department of Electrical Engineering at Jouf University in Saudi Arabia.



ANIS SAHBANI received the Electrical Engineer degree (specialization in industrial data processing) from National School of Engineers of Tunis, Tunisia, in 1998 and the M.S. degree in automation and computer engineering 1999. He then received the Ph.D. degree in robotics and computer science from University of Paul Sabatier—Robotics and Artificial Intelligence Group of the LAAS-CNRS, Toulouse, France, in 2003. He is an Associate Professor with Sorbonne university, Paris, France. His main research interests include Mechatronics, Robotics, Motion Planning, Optimization, Heuristics, and Intelligent Systems.



AMR YOUSEF (Member, IEEE) received the B.Sc. and M.Sc. degrees from the Engineering Mathematics Department and the Electrical Engineering Department, Alexandria University, in 2001 and 2006, respectively, and the Ph.D. degree in electrical and computer engineering from Old Dominion University (ODU), in May 2012. He was a Postdoctoral Research Associate with the Old Dominion Vision Laboratory, USA. He is currently an Assistant Professor with the Electrical Engineering Department, University of Business and Technology, Saudi Arabia. His research interests include optimization techniques, image processing/computer vision, and machine learning algorithms. He is a member of SPIE and OSA.

...