A concise review of U-Net-based deep learning models for brain tumor segmentation

Alaa Waleed Salman
Ministry of Education
General Directorate of Education of Al-Rusafa II
Specialization: Master of Science in Computer Science
School: First High School for Outstanding Students

Abstract.

A brain tumor requires highly discriminative segmentation and classification of the tumor in order to provide effective treatment. Three categories of brain tumor segmentation (BTS) exist fully automated, semi-automatic and manual. Both tumor segmentation in therapy, treatment planning, and diagnostic assessment have frequently relied on the deep learning (DL) approach to automate the process. It is founded on the U-Net paradigm that has recently demonstrated a superb performance in multimodal BTS. The study retrieves a literature review with the application of U-Net models to the BTS. It is a general method of training a new U-Net model on brain tumor segmentation. The procedures of this DL method have been explained to allow the derivation of the desired model. These involves collection of the dataset, pre-processing, selection/ or development of structure of model, transfer learning (optional) and image enhancement (optional). The performance and structure of the model are the two most essential measures that are used to assess the literature. Based on the findings of the review, there is a direct relationship between model accuracy and the architectural complexity of that model; therefore, it will be a future challenge to produce increased accuracy using a less complex design. There are also futures trends, alternatives, and challenges.

Keywords: - Brain tumor segmentation, Deep learning, U-Net, Medical classification

1- Introduction

When dealing with a brain tumor, extremely precise classification and categorization are necessary. There are two classifications for the tumor: benign and malignant [1]. Gliomas are the most hazardous of their several forms [2]. Four stages of micro-MRI (magnetic resonance imaging) images and tumor features are used to classify gliomas. Grades I and II, characterized as benign and slow-growing, are referred to as (LGG) low-grade gliomas, whereas Grades III and IV, which are malignant and aggressive, are classified

as (HGG) high-grade gliomas. Which are active and malignant.

Chemotherapy, radiation, and surgery are the current forms of treatment [3]. Nonetheless, a patient's chances of surviving a brain tumor increase with its early detection. Clinicians utilize a variety of imaging techniques, including magnetic resonance imaging (MRI), to determine if malignancies are present or not. An MRI is a non-invasive test that uses contrast imaging to provide highly significant information on the location, size, and shape of brain tumors. To improve tumor delineation and diagnosis, clinical imaging uses a range of MRI sequences. These encompass fluid-attenuated inversion recovery (FLAIR), T1-weighted MRI (T1w), contrast-enhanced T1-weighted MRI (T1wc), and T2-weighted MRI (T2w). [4]. Tumor segmentation is done following a brain tumor's MRI diagnosis. There are three types of brain tumor segmentation (BTS) methods: fully automatic, semi-automatic, and manual. The radiologist must use the multimodality information provided by MRI scans as well as their experience and training-acquired anatomy and physiology knowledge in order to perform manual segmentation. In order to diagnose the tumor and accurately delineate the tumor locations, the radiologist must go through multiple slices of pictures one after the other. This procedure is often performed by a certified radiologist who delineates the tumor's (ROI) region of interest. The above method facilitates expert input; nonetheless, it has various drawbacks, include inconsistent manual segmentation of MRI data by experts, protracted tumor identification durations, and inter-observer variability. This strategy is both time-consuming and considers a demanding task. Consequently, semi-automatic and fully-automatic techniques are being devised to resolve this problem. These strategies provide reliable outcomes and conserve time. [5].

Specifically, semi-automatic segmentation is an attempt to address several issues with the manual segmentation method. Algorithms that facilitate segmentation can reduce the amount of time and effort required by the user. For example, spreading can increase segmentation throughout an area or remove the need for slice-by-slice segmentation to extend segmentation to neighboring slices.

Several semi-automatic methods exist; for example, some algorithms may be used during (or prior to) segmentation, while others may be assigned following the completion of segmentation. Semi-automatic methods offer a "objective" initial segmentation to lower inter-observer uncertainty. Inter-observer variability will persist since the algorithm settings and manual

segmentation affect the outcomes [6]. Even though semi-automatic segmentation is faster and yields more accurate results than manual techniques, it is still susceptible to inter-rater user error. As a result, the majority of brain tumor segmentation methods used today are fully automated. Automatic segmentation is seen to be the best approach since it produces precise results. Furthermore, compared to other approaches, it produces results quickly. Unsupervised and supervised brain tumor segmentation are the two forms of completely automated segmentation that don't need human involvement. In the unsupervised method, unsupervised learning techniques do not require a dataset with ground truth labels. These techniques are predicated on the brain tumor's symmetry, color, location, and other distinguishing characteristics. Non-learning methods often concentrate on a single application and use the photos and disease characteristics to achieve segmentation. Since these non-learning methods are application-specific, they must create a unique model for every segmentation task. Non-learning methods like fuzzy C-means or slic are not particularly useful because the tumors may have a description that is the same as that of other brain regions. Consequently, our method struggles to distinguish between white matter and grey matter when the tumor areas are near other places with comparable pixel values [7]. The supervised method, on the other hand, makes use of labeled datasets, or ground truth. These datasets can be used to "supervise" or teach computers to identify patterns and make predictions. The accuracy of the model may be verified using explicitly stated inputs and outputs, and it can evolve over time. Machine learning (ML) approaches are used to examine and cluster unlabeled data sets. Without human assistance, without assistance from humans, these algorithms find hidden data patterns. Giving humans the ability to make accurate predictions based on current information is the aim of supervised learning. On the other hand, a great deal of current data can yield important knowledge when using unsupervised learning approaches. Additionally, ML may be used by computers to determine what in data sets is odd or intriguing [8]. image analysis and understanding have been greatly impacted by convolutional neural networks (CNNs), especially in the areas of picture segmentation, classification, and analysis [9]. In order to learn complicated abstractions from data, deep learning (DL), a subset of machine learning (ML), recently uses hierarchical constructs. DL is a cutting-edge technique frequently applied in more conventional domains of artificial intelligence, such as computer vision.

Heavy-load chip processing capabilities (such as graphics processing units, or GPUs), sharply declining computer hardware costs, and important advances in machine learning techniques are the three main causes of DL's present surge. Multi-layer DL networks may extract a lot of previously unreachable features.

Automatic BTS is applicable in locating brain tumors, in treatment planning, and in evaluation of the effectiveness of work therapy. Over the recent past, convolutional neural networks (CNNs) have performed rather well in multimodal behavior tracking systems (BTS). The machine learning approach required a set of training MRI labeled images [10]. With the availability of much more computing power in the form of GPUs, we are able to construct deep neural networks with many layers that can learn a bigger and deeper set of features that were previously inaccessible. CNN for segmentation can be classified into many groups according on the convolutional kernel's dimension. 2D CNNs with 2D convolutional kernels are used to predict the segmentation map for a single slice. Segmentation maps are predicted for the full volume by generating predictions for each slice. The context throughout the slice's height and width can be used by the 2D convolutional kernels to generate predictions.

However, because 2D CNNs only take one slice as input, they are unable to leverage following slices' context by default. Voxel information from nearby slices could be useful for segmentation map prediction [11]. Using 3D convolutional kernels, a volumetric scan patch was segmented using the 3D CNN. Inter-slice context can enhance performance but attack a computational cost because these CNNs employ a significant number of parameters [12].

The paper is a review for brain tumor segmentation using deep learning (DL). technique is presented in this work. As such, it is a typical method for modeling a DL approach to brain tumor segmentation. The DL procedures to create the necessary model are explained in Section 2. Gathering the dataset, preprocessing, and, if necessary, image augmentation were the first phases in this procedure. While Section 4 offers alternatives and challenges, Section 3 examines the literature based on the model architecture, including 2D, 3D, and hybrid. Section 5 discusses upcoming trends. The conclusion reached is finally summarized in section 6.

2- Deep Learning Approach

As a general method for creating a DL model, it consists of four major steps [13]. The first step is to enter the dataset photographs. We got the dataset from a website that has datasets and cleaned it up by getting rid of the images we didn't want, which made it less likely that we would produce a good model. The second step is pre-processing and adding more data. This step is optional and depends on the quantity and how useful the images in the dataset are. Pre-processing includes reducing noise, changing the format of images to fit the model's input specifications, and adjusting the size of the dimensions. Data augmentation techniques—such as mirroring, rotation, flipping, and cropping—are employed to synthetically expand and diversify the training dataset, thereby improving model generalization To expand the dataset, data augmentation is required because the DL model is image-hungry. The DL model is the core component of the segmenting system. It is either a well-known (standard) model or one that was designed. The BTS of the entire tumor (edema) is the model's output. The final step is the transfer learning block. It can be used in place of the data augmentation procedure to improve the model's accuracy.

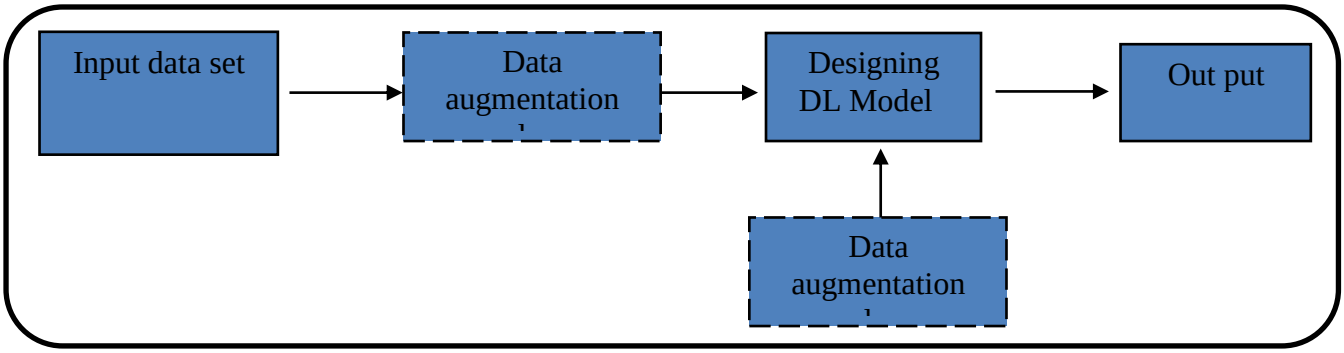

Figure 1. General schematic representation of the Deep learning (DL) Model design

2.1 Input dataset

The integration of computer-assisted techniques in medical image analysis led to the initiation of the first Multimodal Brain Tumor Segmentation Challenge (BTS) in 2012. Since its inception, the associated datasets have been updated annually to address evolving research challenges. This publicly accessible resource supports the development and validation of advanced methodologies for brain tumor classification. The primary objective of the dataset is to serve as a standardized benchmark for evaluating the performance of contemporary brain tumor segmentation (BTS) algorithms. Furthermore, human experts manually diagnosed the photography data set. The suggested form automatically categorizes brain tumors because the data set satisfies all conditions. So, an MRI image that makes up the data set saved as part of the Neuroimaging Informatics Technology (NIFTI) files (. nii.gz) Initiative shows both low-degree and high-grade-class tumors (HGG). The dataset used in this study is sourced from the BRATS website. It has been preprocessed by the author to remove the skull component [4] Each patient's data includes four imaging modalities: T1, T2, FLAIR, and T1C. These modalities highlight different tumor characteristics—T1 shows the enhanced tumor, T2 visualizes the surrounding edema, FLAIR reveals the tumor core, and T1C represents the non-enhancing tumor. Based on these modalities, three overlapping tumor sub-regions are identified: the whole tumor, the tumor core (TC), and the enhancing tumor (ET) The dimensions of each MRI scan are 155×240×240, corresponding to the axial, coronal, and sagittal planes. Due to the complexity associated with variations in tumor shape, size, and intensity, the dataset was divided into three sub-regions, as illustrated in Figure 2 [14]. Additionally, two distinct data subsets—lacking ground truth labels—were provided for evaluation and validation purposes [15]. Currently, the BRATS dataset represents the most widely utilized open-access MRI dataset for the objective segmentation of brain tumors. An overview of the dataset used for brain tumor segmentation is presented in Table 1

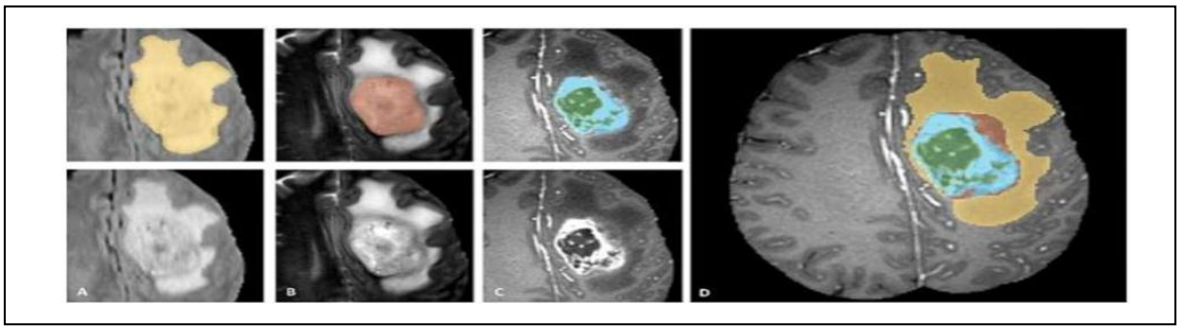

Figure 2 : Schematic representation of the integration of all annotated regions to derive the final tumor sub-region labels: enhancing tumor core (blue), necrotic/cystic core (green), non-enhancing solid core (red), and peritumoral edema (yellow) [14]

Table 1. Summary and Description of the Datasets

Author(s)/Year	Dataset	Description	link
Pereira *et al.* [16]/2016	BRATS 2013	10 LGG and 20 HGG make up the training set, and manual segmentations are accessible. The total images in this dataset are 10×155×5=7550 for LGG and 15500 for HGG.	www2.imm.dtu.dk/projects/BRATS2012/
Dvořák *et al.* [17]/2015	BRATS 2014	The dataset includes patient volumes for 252 HGG and 57 LGG glioma cases. 195300 images for HGG and 44175 images for LGG.	www2.imm.dtu.dk/projects/BRATS2012
Li *et al.* [18]/2019	BRATS 2015	The training dataset contains 54 patients who have LGG and 220 patients who have HGG, with expert segmentations provided as ground truth. The total images for HGG are 170500, and 41850 for LGG images.	www2.imm.dtu.dk/projects/BRATS2012
Zhao *et al.* [19]/2016	BRATS 2016	It is similar to BRATS 2015.	www2.imm.dtu.dk/projects/BRATS2012
Rezaei *et al.* [20]/2017	BRATS 2017	It has 220 HGG, and 108 LGG MRI scans made up of the segmentation training dataset. 170500 images for HGG and 83700 for LGG.	www2.imm.dtu.dk/projects/BRATS2012
Weninger *et al.* [21]/2018	BRATS 2018	Its training dataset comprises 75 scans for LGG and 210 for HGG. The total images for HGG are 162750 and 58125 for LGG.	www2.imm.dtu.dk/projects/BRATS2012
Jiang *et al.* [22]/2019	BRATS 2019	The training dataset consists of 76 cases of LGG and 259 cases of HGG. The total images are 200725 for HGG and 58900 for LGG.	www2.imm.dtu.dk/projects/BRATS2012
Mehta *et al.* [23]/2022	BRATS 2020	125 patient cases with diffuse gliomas make up the validation dataset. The total number of images in this dataset is 96875.	www2.imm.dtu.dk/projects/BRATS2012
Fidon *et al.* [24]/2022	BRATS 2021	There are 1251 cases in the training dataset and 219 cases in the validation dataset. The total number of images in this dataset is 169725 images.	www2.imm.dtu.dk/projects/BRATS2012
Valverde *et al.* [25]/2015	IBSR20	20 T1-w scans (256×63×256) make up the IBSR20 image set. Additionally, the authors offer signal intensity histograms, labeled volumes, and key tissue annotations (GM, WM, and CSF) for evaluation based on qualified experts using signal intensity histograms and a semi-automated intensity contour mapping technique. The order of these images reflects their difficulty. The most difficult scans have significant acquisition irregularities and artifacts.	www.nitrc.org/projects/ibsr
Jiang *et al.* [26]/2018	Brain Web	This simulated brain database has a set of realistic MRI data volumes delivered by an MRI simulator that was commonly applied for evaluating the performance of denoising approaches.	brainweb.bic.mni.mcgill.ca/brainweb/

2.2. Data Augmentation and Preprocessing

following dataset collection, the initial step typically involves data augmentation, particularly when the available dataset is limited. This process is crucial to meet the data volume requirements of deep learning (DL) models by artificially increasing dataset size and variability. Alternatively, transfer learning may be employed to mitigate the challenges associated with limited data, rendering data augmentation an optional, though beneficial, procedure. The subsequent step involves image pre-processing, which, while also optional, serves to enhance image quality through minor adjustments when necessary. A range of techniques and tools are available for effective implementation of the pre-processing stage [27], including the following:

a. **Sampling** involves selecting a representative subset from a larger population of data to facilitate efficient analysis while preserving the statistical characteristics of the original dataset.

b. **Transformation** refers to the process of converting raw data into a structured input format by applying specific modifications, thereby enhancing data compatibility with analytical models.

c. **Denoising** is employed to eliminate noise or irrelevant variations from the dataset, thus improving data quality and the reliability of downstream analysis.

d. **Imputation** addresses the issue of missing values by estimating and replacing them with statistically meaningful values, often derived from patterns observed within the existing data.

e. **feature extraction** involves identifying and isolating a subset of relevant features that are most informative for the specific analytical task.

f. **Masking removal**, as an initial pre-processing step, eliminates any superimposed mask artifacts from the dataset to enhance image clarity.

g. **Normalization** standardizes image intensity by dividing an average by the standard deviation within each slice, often employing bias field correction techniques such as the N4ITK algorithm to reduce noise and improve data consistency

h. To prepare volumetric data for model input, **2D slicing** is applied to convert three-dimensional images into two-dimensional representations. Additionally, **data augmentation** techniques are employed to increase the number of training samples and enhance dataset diversity, thereby improving model generalization [28];

i. Reducing images is another pre-processing method. This is a major issue with CNN because it needs the images in the dataset to be the same size.

3.2 Designing the Deep Learning Model (U-Net Architecture)

Image segmentation involves partitioning an image into multiple meaningful regions, a task effectively addressed by the U-Net architecture. Originally introduced in 2015, U-Net demonstrated exceptional performance and has since inspired numerous enhanced variants. This section outlines the standard U-Net design, which is composed of three primary components: a convolutional path (down-sampling), a bottleneck or flattening layer, and a deconvolutional path (up-sampling). The down-sampling layers form the encoder, which captures context by progressively reducing spatial dimensions, while the up-sampling layers constitute the decoder, responsible for reconstructing spatial resolution and enabling precise localization in Figure 3 [29]. The network architecture comprises three fundamental layers: input**,** hidden**,** and output layers. The input layer receives the raw image data, while the hidden **layers** are responsible for feature extraction through a series of convolutional blocks and pooling operations. Each convolutional block typically consists of a convolutional layer, followed by a Rectified Linear Unit (ReLU**)** activation function and batch normalization, resulting in the generation of a refined feature map that captures essential spatial patterns from the input [30].

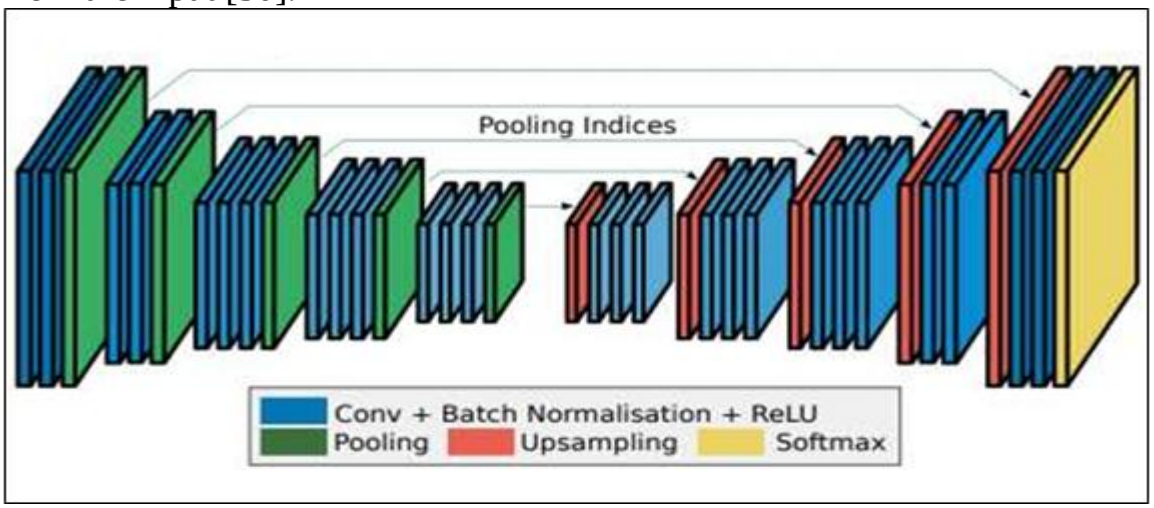

Figure 3: Schematic representation of U-Net with down sampling and up sampling

convolutional layers are the core components of CNNs, applying multiple small (typically 3×3) learnable filters to input images to generate feature

maps. These filters detect spatial patterns, such as edges, across the image. To preserve the input dimensions despite size reduction during convolution, **padding**—adding zeros around image borders—is employed. Additionally, **stride** controls the filter's step size during scanning, allowing it to move multiple pixels per step, which affects the output resolution [31].

Pooling layers slide a window over the input and apply a summarizing function, such as maximum or average pooling, to each sub-region. This process reduces the spatial dimensions of the feature maps, helping to decrease computational complexity and control overfitting. Pooling operations resemble discrete convolutions. Figures 4 and 5 [32] illustrate average and maximum pooling, respectively

Patch normalization enhances training by enabling higher learning rates, accelerating convergence, and reducing overfitting, typically applied after convolutional layers [33]. It processes small data batches to improve efficiency. Activation functions are critical for deep learning performance, influencing training dynamics and outcomes. Common activation functions include ReLU, leaky ReLU, Sigmoid, Hyperbolic Tangent, and SoftMax, each with distinct properties affecting model behavior [4]

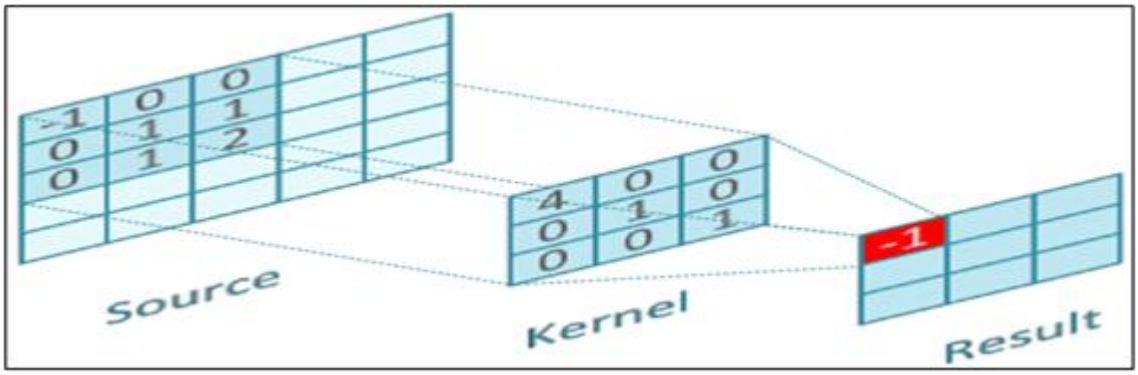

Figure4: Schematic representation The calculations back applying a 1×1 stride and a 3×3 average pooling on a 5×5 input

2.4 processing

To determine the location of the brain tumor, the brain MRI image is first segmented and then put through a number of post-processing processes. Using 3×3 structuring, morphological erosion is one of these post-processing techniques used to the segmented image. Additionally, for segmentation purposes, a binary tumor-masked window is created. The primary objective of the process is to show the area of the image where the tumor is larger and

more intense. Furthermore, the intensity of tumor tissue is higher than that of the surrounding tissues. In order to create in the final image displaying brain tumors, the tumor mask is put on top of the expanded image . Also, for segmentation purposes, a binary tumor masked window is generated. The primary objective of the process is to show the area of the image where the tumor is larger and more intense. Furthermore, the intensity of tumor tissue is higher than that of surrounding tissues. In order to create the final image that identifies brain tumors, the tumor mask is placed to the dilated image [4].

2.5 Performance measures

Ground truth images are annotated manually by individuals who are experts in radiology in order to evaluate the effectiveness of segmentation methods. Dice Score (DS) is a widely used measure and is likely the most powerful in sophisticated image segregating algorithms. According to Das et al. [1], eighty-three percent of the researches in their narrative review applied DS to evaluate BTS performance, as opposed to 22 percent of the studies applied Hausdorff Distance (HD). The analysis of between-applications evaluation measures in Das et al. paper dealt with medical imaging applications including lung, coronary, and carotid lesion, segmentations and identified that DS remains a primary measure of lesion segmentation. HD and other measurements based on boundaries would, however, work better in regions that are peculiarly concerned with differences on shapes in similarity of sizes [35].

However, vital information cannot be fully conveyed by a single metric; rather, the task depends on the significant criteria. In addition to the DS, established evaluation criteria including accuracy, specificity, and precision are applied when comparing the ground truth images with a forecast image. The Dice Score (DS) is commonly used in MRI segmentation to quantify the overlap between the segmented output and the ground truth image. It is calculated as shown in Equation (1):

$$Dice\ score(A,B) = \frac{2|A \cap B|}{|A|+|B|} \qquad (1)$$

Here, |A| denotes the number of pixels predicted as tumor by the model, while |B| represents the tumor pixels in the ground truth. Additionally, the Dice Coefficient can also be computed from the confusion matrix as shown in Equation (2):

$$Dice\ score = \frac{2TP}{FP + 2TP + FN} \qquad (2)$$

While metrics such as accuracy, precision, specificity, and recall are commonly used in deep learning models, they are less frequently applied in the context of image segmentation

3. Deep Learning Architectures for Image Segmentation

Firstly, early detection of malignant brain tumors is crucial. Numerous methods and architectures have been put forth because there is no one best way to segment tumors. In order to automatically segment brain tumors, researchers have created various DL models utilizing MRI images. Based on their dimensionality, these DL architectures were divided into three categories: 2D, 3D, and hybrid (2D+3D models).

3.1. 2D architecture

2D architectures are widely used in biomedical image segmentation, typically comprising an encoder, a decoder, and skip connections. Pereira et al. [16] proposed a CNN-based method utilizing small 3×3 kernels, which reduce overfitting by limiting the number of parameters and support deeper network design. They applied this model to the BRATS 2013 and BRATS 2015 datasets, incorporating intensity normalization and data augmentation as preprocessing steps. Their approach achieved Dice Scores of 0.88 and 0.78 for BRATS 2013 and BRATS 2015, respectively (Figure 6).

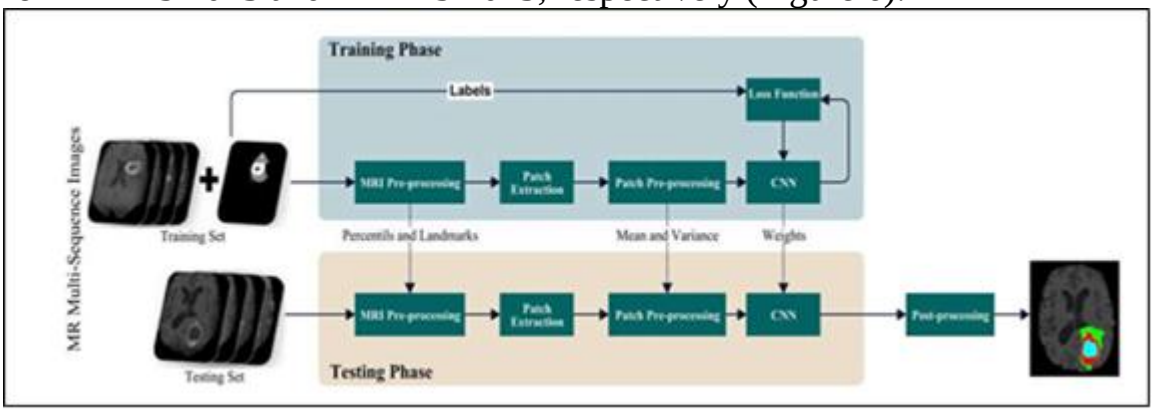

Figure5: Schematic representation Overview of the proposed method [16]

Dong et al. [36] offer a unique 2D fully-convoluted segmentation model based on the U-Net architecture. As shown in Figure 7, it consists of encoder and decoder routes [36]. To enhance the dataset photographs, they applied basic adjustments like flipping, rotation, shift, and zoom. Additionally, they

used the BRATS 2015 dataset using WT 0.86 dice to get their results. The framework has not achieved improved effectiveness due to frequently unavailability of multimodal MRI data, which is typically constrained by limited acquisition time and patient symptoms, however it necessitates less clinical inputs.

Isensee et al. [37] They proposed a 2D U-Net model for brain tumor segmentation utilizing MRI data, which includes a context aggregation pathway. This design enables the encoder to progressively abstract input features as it moves deeper into the network A localization pathway that recombines shallower features with these representations is then sought in order to precisely localize the structures of interest. They used normalization in the pre-processing step, rescaling the generated images to [0, 1], with the non-brain region set to 0, and clipping them at [−5, 5] to eliminate outliers. In contrast, their model achieved a Dice Score of 0.858 for whole tumor segmentation on the BRATS 2015 dataset. However, further improvement is needed, particularly in accurately predicting tumor location relative to critical brain structures such as the ventricles, optic nerves, and other essential pathways.

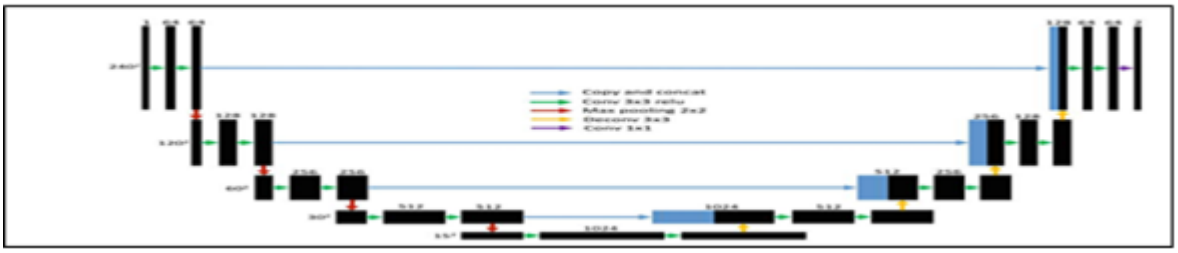

Figure 6: Schematic representation Architecture of the proposed U-Net model for brain tumor segmentation [36]

Kermi et al. [38] developed a fully automated brain tumor segmentation model based on the 2D U-Net architecture, achieving accurate intra-tumor and whole-brain segmentation. The model was trained to segment both LGG and HGG tumors using the BRATS 2018 dataset, with normalization and data augmentation applied during preprocessing. The model attained Dice Scores of 0.805 for the tumor core (TC), 0.868 for the whole tumor, and 0.783 for the enhancing tumor. The authors recommended upgrading GPU resources to further accelerate and enhance the training process.

Venu et al. [39] proposed a 2D U-Net model for brain tumor segmentation using the BRATS 2015 dataset. Image normalization, bias field correction,

and patch extraction were used in the preprocessing pipeline. A total accuracy of 86.3 was achieved by the model on the total tumor segmentation. Iqbal et al. [40] introduced three deep learning structures, Interpolated Network, SeNet and SkipNet, in an attempt to enhance the effectiveness of segmentation. Each model has four sub-blocks whose design comprises of an encoder decoder structure. The preprocessing pipeline performed normalization, bias correction, and 2D slicing of 3D images in order to make them 2D, and cropping to accelerate training and increase the number of image samples Noori et al. [41] proposed a low-parameter 2D U-Net-based network utilizing two distinct approaches. The first approach integrates high-level and low-level features with an attention mechanism to enhance segmentation performance. This attention-based approach reduces model confusion by adaptively weighting individual channels. The second method involves a multi-view fusion technique, which, despite using a 2D model, leverages 3D contextual information from the input images. As part of the preprocessing, the N4ITK bias field correction algorithm is applied to each MRI modality to correct image inhomogeneity. Prior to normalization, the top and bottom 1% of intensities in each modality were clipped to enhance robustness, followed by standardization to achieve zero mean and unit variance. Using the BRATS 2017 and 2018 datasets, the model achieved Dice scores of 0.776 for enhancing tumor (ET), 0.888 for whole tumor (WT), and 0.821 for tumor core (TC). However, segmentation performance was hindered by the model's bias toward larger tumor regions and the computational cost involved. However, SegNet utilizes an encoder-decoder architecture completed by a final layer for pixel-wise categorization. Four distinct 3D MRI modalities (FLAIR, T1, T1c, and T2) were utilized individually to train the SegNet model, as stated by Alqazzaz et al. [42] then combined the findings in the post-processing step. A pair of encoders (downsampling) with 13 convolutional layers with 3x3 filters and max-pooling layers made up the architecture, as did the decoder (upsampling), which also had 13 convolutional layers. Pre-processing procedures included matching, normalizing, and bias field correction. Using the BRATS 2017 dataset, they obtained dice scores of 0.81, 0.85, and 0.79 for TC, entire tumor, and enhancing tumor, respectively. Among the model's drawbacks are a laborious training phase problem and the potential for future accuracy improvements through improved post-processing methods. Rehman et al. [43] introduced the BU-Net architecture for brain tumor segmentation (BTS),

achieving Dice scores of 0.901, 0.837, and 0.788 for whole tumor, tumor core, and enhancing tumor (ET), respectively, on the BRATS 2017 dataset. To enhance performance, residual connections (RES) were integrated between the encoder and decoder. Additionally, N4ITK bias correction and intensity normalization were applied to standardize the data to zero mean and unit variance. Lin et al. [44] proposed the Path Aggregation U-Net (PAU-Net) for brain tumor segmentation using multi-modality MRI. The model has a bottom-up Path Aggregation (PA) encoder, which aims to lessen the semantic gap between deep features and output layer and help lessen noise. A long sort code (constant) decoder holds in-depth details. Moreover, an effective feature pyramid (EFP) is used to enhance mask prediction that takes fewer computations. Preprocessing involved normalization of four MRI modalities and merging them into a four-channel input array, followed by data shuffling. On the BRATS 2018 test set, the model achieved Dice scores of 0.8563 (WT), 0.6751 (TC), and 0.6002 (ET), while on the BRATS 2017 test set, it reached 0.9000 (WT), 0.7095 (TC), and 0.6357 (ET). Despite promising results, the approach was limited by weak supervision. Ebied et al. [45] proposed a modified U-Net architecture, structurally similar to the standard U-Net but with specific enhancements. The encoding path consists of six convolutional layers, each block comprising two 5×5 filters with a stride of two, resulting in 2048 additional feature maps. The decoding path includes six deconvolutional layers with matching filter sizes and strides, reducing the number of feature maps. The model was evaluated on three datasets: TCIA, BRATS 2019, and FIGSHARE. Batch normalization was applied to the BRATS 2019 and FIGSHARE datasets, while medium and soft filters were used for the TCIA dataset. The model achieved Dice scores of 85.02 on BRATS 2019, 91.96 on FIGSHARE, and 86.68 on TCIA. Table 2 summarizes 2D U-Net-based studies for brain tumor segmentation

3.2. 3D architecture

3D models often face optimization challenges due to the high number of trainable parameters in 3D kernels, which complicates maintaining stable forward and backward signal propagation, as illustrated in Table 3.

Table 2. 2D U-Net architecture

Author(s)/Year	Segmentation approach	Dataset	Results
Pereira *et al.* 2016/[16]	Automated segmentation technique based on CNN with kernels of size 3×3.	BRATS 2013 BRATS 2015	WT 0.88 dice for BRATS 2013 WT 0.78 dice for BRATS 2015
Dong *et al.* 2017/[36]	The U-Net structure was used to create a distinctive 2D fully convolutional segmentation model.	BRATS 2015	WT 0.86 dice
Isensee *et al.* 2017/[37]	U-Net based on 2D CNN and its effectiveness in segmenting brain tumors was carefully tuned.	BRATS 2015	WT 0.858 dice
Kermi *et al.* 2018/[38]	Modified U-net architecture-based 2D Deep CNNs.	BRATS 2018	WT 0.868 dice
Venu *et al.* 2018/[39]	An algorithm for DL based on CNN.	BRATS 2015	WT 86.3% dice
Iqbal *et al.* 2018/[40]	It is proposed to use CNN with three main network architectures: SE-Net, Skip-Net, and interpolated network.	BRATS 2015	IntNet WT 0.90 dice, SkipNet WT 0.87 dice, and SENate WT 0.88 dice.
Noori *et al.* 2019/[41]	2D U-Net for automated segmentation	BRATS 2018	WT 0.89 dice
Alqazzaz *et al.* 2019/[42]	SegNet approach.	BRATS 2017	WT 0.85 dice
Rehman *et al.* 2020/[43]	They suggested using a 2D BU-net to automatically segment an image of a brain tumor.	BRATS 2017	WC 0.90 dice
Lin *et al.* 2021/[44]	They proposed the path aggregation U-Net (PAU-Net) model of neural networks (MRI).	BRATS 2018 and BRATS 2017	WT 0.85 dice for BRATS 2018 and WT 0.90 dice for BRATS 2017
Ebied *et al.* 2022/[45]	They suggest that they can use a modified 2D U-Net network to show how to segment the brain tumor.	BRATS 2019, FIGSHARE, and TCIA datasets	WT85.02 dice for BRATS 2019, WT 91.96 dice for FIGSHARE, and WT 86.68 dice for TCIA

Kamnitsas et al. [46] addressed lesion segmentation using a 3D architecture enhanced with residual connections. Intensity normalization was performed individually for each scan by dividing the mean by the standard deviation. Additionally, training data augmentation was implemented by mirroring images across the mid-sagittal plane using the BRATS 2015 dataset. Erden et al. [47] proposed a novel 3D U-Net-based model optimized against a loss function during training. The architecture was designed to be relatively shallow and narrow, yielding satisfactory performance. As a preprocessing step, a bounding box was applied to the dataset, although this introduced additional complexity in model handling. The model achieved a Dice score of 0.71 on the BRATS 2017 dataset. To reduce computational cost and training

time, the study utilized only a single CT modality instead of combining all four, allowing for broader experimentation with different architectures. Lachinov et al. [48] proposed a cascaded 3D U-Net approach for automated brain tumor segmentation using multimodal MRI inputs. The 3D U-Net architecture was modified to efficiently handle multiple modalities, and segmentation quality was enhanced by incorporating contextual information from models with similar topologies trained on downscaled data. Preprocessing involved data normalization, assigning zeros to background voxels, scaling brain voxel intensities to the 0–10 range, and filtering out noise and outliers by restricting values within the −5 to 5 range. The proposed method achieved Dice scores of 0.720 (enhancing tumor), 0.878 (whole tumor), and 0.785 (tumor core) on the BRATS 2018 dataset. Mehta et al. [49] introduced an enhanced 3D U-Net model for brain tumor segmentation using multimodal MRI volumes. As shown in Figure 8 [49], the architecture builds upon the standard 3D U-Net framework. Preprocessing involved rescaling voxel intensities to a 0–1 range through mean subtraction and division by the standard deviation, followed by cropping volumes to 184×200×152. On the BRATS 2018 dataset, the model achieved Dice scores of 0.771 (tumor core), 0.871 (whole tumor), and 0.706 (enhancing tumor). However, performance declined in the tumor core and enhancing tumor categories on the test dataset.

Table 3. 3D U-Net architecture

Chen et al. concurrently presented an S3D U-Net framework for BTS [50]. They suggest a novel separable 3D convolution that divides each 3D convolution into three parallel branches—axial, sagittal, and coronal—in order to fully use 3D volumes. They also provide a detachable 3D block of

Author(s)/Year	Segmentation Approach	Dataset	Results
Kamnitsas *et al.* 2016/[46]	3D CNN architecture was presented. Which they further improve by adding residual connections.	BRATS 2015	WT 89.6 dice
Erden *et al.* 2017/[47]	Three-dimensional FCNN moreover, they used a U-NET architecture.	BRATS 2017	WT 0.71
Lachinov *et al.* 2018/[48]	They offer a deep cascaded approach for automatically segmenting brain tumors (3D U-Net architecture).	BRATS 2018	WT 0.87
Mehta *et al.* 2018 [49]	3D CNN.	BRATS 2018	WT 0.871
Chen *et al.* 2018/[50]	They suggest a novel separable 3D convolution with separable 3D U-Net architecture.	BRATS 2018	WT 0.89
Ali *et al.* 2020/[51]	3D CNN and a U-Net.	BRATS 2019	WT 0.906
Baid *et al.* 2020/[52]	The author developed a novel 3D U-Net network for segmenting various brain tumors.	BRATS 2018	WT 0.88
Bukhari *et al.* 2021/[53]	Using E1D3 U-Net.	BRATS 2018 BRATS 2021	WT 91.0 for BRATS 2018 and WT 91.9 for BRATS 2021

state-of-the-art residual inception architecture. The bias field caused by the inhomogeneity of the magnetic field and the small motions during scanning is first removed using the N4ITK bias correction method. Furthermore, normalization is essential when performing multi-mode scanning with a single algorithm. Ultimately, they achieved the following results during the testing phase: dice scores of 0.83093, 0.89353, and 0.74932 for the enhancing tumor, tumor core, and total tumor, respectively.

Ali et al. [51] proposed an ensemble approach combining a 3D CNN and a U-Net architecture to enhance segmentation accuracy. The ensemble leveraged the complementary strengths of both models to improve prediction precision. Various data augmentation techniques, including mirroring, rotation, and cropping, were employed to enhance model robustness during training. The proposed ensemble achieved Dice scores of 0.846 for tumor core, 0.906 for whole tumor, and 0.750 for enhancing tumor on the validation set. However, the approach has several limitations. First, evaluation was limited to a single metric using the challenge's official validation set. Second, the study lacked comprehensive pre- and post-processing of the dataset and results. Finally, the method's generalizability remains uncertain, as it was not tested on independent MRI datasets beyond the challenge data

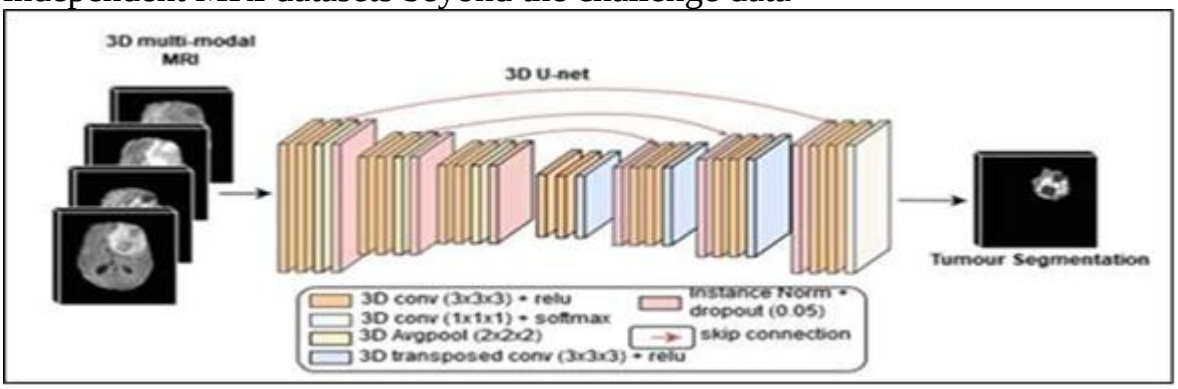

Figure 7: Schematic representation Using four input images and a 3D U-Net architecture to segment a multi-class tumor [49]

Baid et al. [52] proposed a modified 3D U-Net architecture for brain tumor segmentation, incorporating a weighted patch extraction strategy focused on tumor borders and employing fewer network levels with increased filter counts per level. Preprocessing involved normalization and N4ITK bias correction. Evaluated on the BRATS 2018 dataset, the method achieved mean Dice scores of 0.75 (ET), 0.88 (WT), and 0.83 (TC). However, the model's extended training time remains a notable limitation. Bukhari et al.

[53] introduced the E1D3 U-Net, an enhanced variant of the standard 3D U-Net architecture for brain tumor segmentation, featuring a single encoder and three decoders, as shown in Figure 9 [53]. The architecture extends the baseline encoder-decoder design by incorporating two additional decoders, structurally similar to the original. Each of the three decoders receives feature maps directly from the single encoder and independently generates a segmentation output.

Prior to training and evaluation, each 3D MRI volume was normalized within the whole-brain region to zero mean and unit variance. Using the BRATS 2018 dataset, the model achieved Dice scores of 91.0 (WT), 86.0 (TC), and 80.2 (ET), while on BRATS 2021, it attained 91.9, 86.5, and 82.0, respectively.

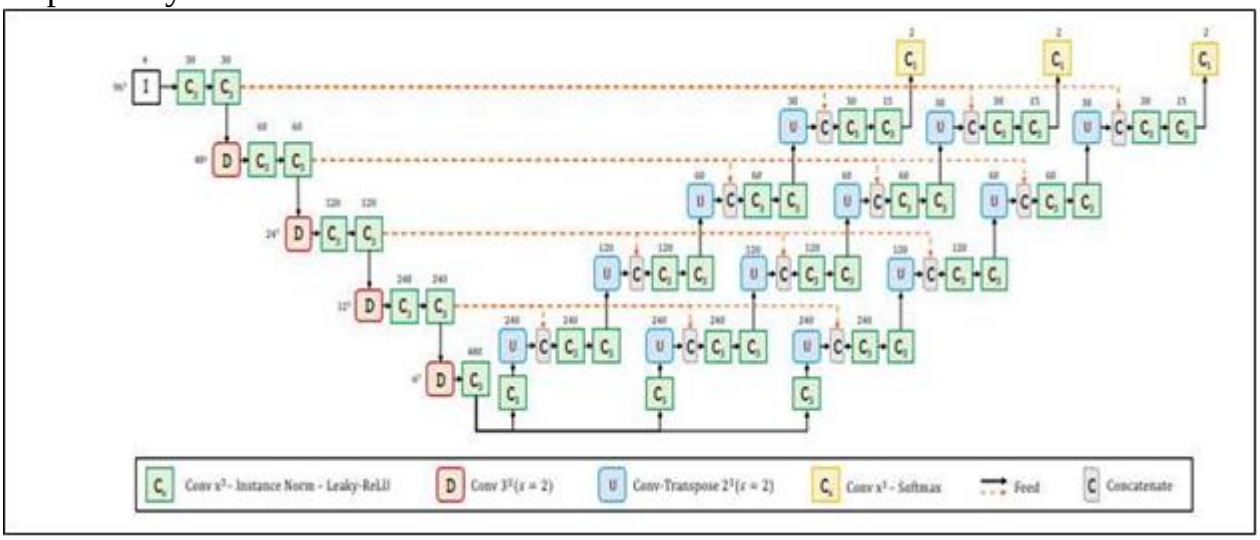

Figure 8: Schematic representation 3D U-Net design called E1D3 U-Net [53]

3.3. Hybrid architecture

Wang et al. [54] addressed brain tumor segmentation by integrating 2D and 3D U-Net architectures through a cascade of 2.5D models, balancing the computational efficiency of 2D CNNs with enhanced memory usage. Their model incorporated multi-scale predictions for deep supervision, employing three 3×3×1 convolutional layers at different network levels to generate intermediate outputs. Minimal preprocessing—limited to normalization—was applied, along with test-time augmentation to improve accuracy. Using the BRATS 2017 dataset, the model achieved Dice scores of 0.786 (tumor core), 0.905 (whole tumor), and 0.838 (enhancing tumor). The study highlights the

potential for improving CNN generalizability using partially or weakly annotated brain tumor images.

Mlynarski et al. [55] proposed a CNN-based model that combines long-range 2D and short-range 3D contextual information, enhancing segmentation performance. The architecture includes modality-specific sub-networks to improve robustness against missing MR sequences during training. Preprocessing involved advanced normalization techniques, including histogram matching. On the BRATS 2017 dataset, the model achieved median Dice scores of 0.918 (whole tumor), 0.883 (tumor core), and 0.854 (enhancing tumor). Despite recent GPU advancements, fully 3D models may still face limitations when processing large spatial volumes. This approach is summarized in Table 4.

Table 4. Summary of hybrid U-Net literature

Author(s)/Year	Segmentation approach	Dataset	Results
Wang *et al.* 2019/[54]	A cascade of 2.5D models balances the benefits of having a 2D CNN with memory efficiency and model complexity	BRATS 2017	WT 0.90
Mlynarski *et al.* 2019/[55]	They presented a CNN-based model that effectively integrates the benefits of both the long-range 2D context and the short-range 3D context	BRATS 2017	WT 0.91

4. CHALLENGES AND SOLUTIONS

Despite significant progress in brain tumor segmentation (BTS) using deep learning (DL) techniques, several challenges continue to limit their clinical applicability. One major issue is data imbalance, where certain tumor types or grades are underrepresented in datasets, leading to biased model predictions. Techniques such as data augmentation, synthetic data generation, and weighted loss functions can alleviate this problem by improving model exposure to minority classes. Another critical challenge is multi-center generalization. Models trained on data from a single institution often perform poorly on data from other centers due to variations in scanners, acquisition protocols, and patient populations. Approaches such as domain adaptation, transfer learning, and federated learning have been proposed to enhance cross-site robustness while maintaining data privacy. Additionally, the computational cost of training and deploying deep models remains high, especially for high-resolution 3D MRI volumes. Solutions include model compression, knowledge distillation, and lightweight architectures that reduce memory and processing requirements without significantly compromising accuracy. Addressing these limitations is essential to transition

U-Net–based BTS systems from research environments to reliable clinical tools capable of real-world application.

4.1 The appearance of brain tumors

Tumor lesions are generally diagnosed by gradient of intensities when compared to the immediate surrounding normal tissues. Nevertheless, the gradients still may be spoiled by the low-resolution images, bias field effects, or the partial volume effect. Due to variability in tumor location, size, shape, it is difficult to incorporate the prior knowledge about extent or position of tumor, as illustrated in Figure 10 [56].

The data on historical geography of normal brain tissue is not readily accessible because the normal, functioning brain tissue represents a wide range of lesions other than any tumor mass itself or cavity left behind after treatment. This shows the drawback of the sound brain modeling techniques, including that of an atlas, which presuppose that the location of healthy tissues remains constant. Incorporation of prior usage of data on tumor substructural characterization is complicated by a lack of predictable tumor MRI imaging caused by variation of morphology and invasiveness. High-grade gliomas share great enhancement in contrast and heterogeneity of the tumor, while low-grade gliomas do so by only about 60%. Increasing the depth of the model has worked to manage this inconsistency [57]. Also, using weighted loss functions where more weight is assigned to background labels between tissues that are adjacent improves segmentation performance[58]. Multi-modality approaches and super-pixel techniques also contribute to overcoming these challenges [59].

4.2. Dataset

Deep learning (DL) models require large volumes of labeled images for effective training, yet obtaining such datasets remains a costly, labor-intensive, and persistent challenge [60]. To mitigate this issue, data augmentation and transfer learning are widely adopted techniques [61]. Patch-wise training offers an alternative approach by dividing images into random or overlapping patches. Its effectiveness is influenced by factors such as patch size and the degree of overlap [62].

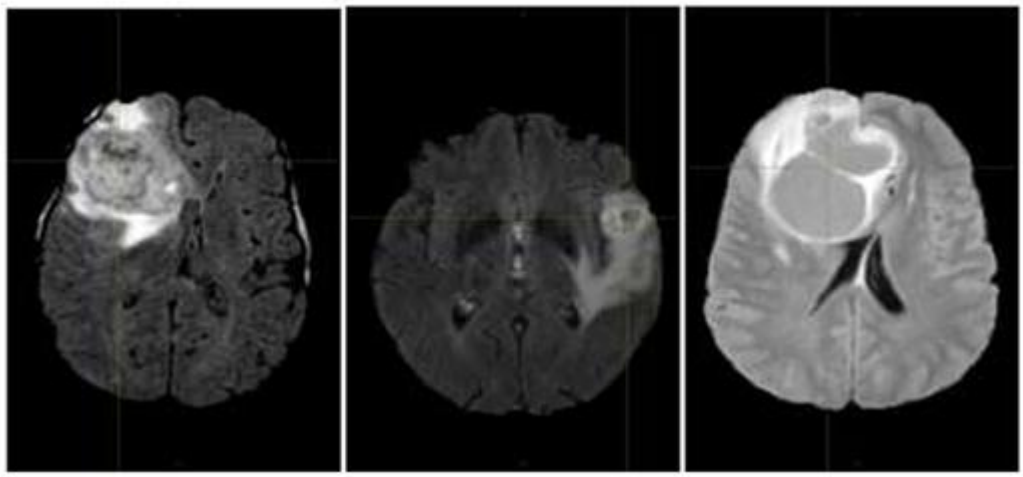

Figure 9: Schematic representation Axial slices of T2-FLAIR acquisitions of 3 different brains with tumors of variable grade [56]

4.3. Overfitting

Overfitting occurs when a model learns the training data too precisely, capturing noise and specific patterns that do not generalize well to unseen data. This leads to high accuracy on the training set but poor performance on validation or real-world datasets. The problem is particularly common in medical imaging tasks, where the number of annotated samples is limited.To mitigate overfitting, several regularization strategies are employed. Data augmentation expands the diversity of training data by applying random transformations such as rotation, flipping, or scaling. Dropout is another widely used technique, which randomly deactivates a subset of neurons during each training iteration to prevent co-adaptation and improve model generalization. Additionally, methods such as early stopping, L2 regularization, and cross-validation can further enhance the robustness of U-Net-based segmentation models against overfitting.

5. Future Trends

U-Net continues to serve as a cornerstone in medical image segmentation due to its encoder–decoder architecture and strong generalization capabilities. Future research directions are expected to focus on enhancing model scalability, robustness, and clinical applicability.

Emerging hybrid architectures that combine U-Net with transformer-based encoders have demonstrated superior performance in capturing long-range dependencies, thereby improving segmentation accuracy in complex medical datasets. Moreover, federated learning frameworks are gaining attention as a

privacy-preserving solution that enables multi-center collaboration without centralizing sensitive patient data—addressing one of the main barriers to real-world deployment.

Further improvements are expected through the integration of semi-supervised and unsupervised learning strategies to better utilize unlabeled medical data, along with data augmentation and weighted loss functions to mitigate class imbalance. Additionally, explainable AI (XAI) techniques will play a vital role in increasing the interpretability and trustworthiness of U-Net predictions in clinical settings.

Ultimately, the combination of efficient architectures, privacy-preserving training paradigms, and explainable models will drive the next generation of U-Net-based segmentation systems toward broader and safer clinical adoption.

6. Conclusion

This review summarized the evolution of automated brain tumor segmentation (BTS) techniques using MRI datasets, emphasizing the pivotal role of U-Net architectures in advancing segmentation accuracy and efficiency. MRI remains the preferred imaging modality for BTS due to its superior soft tissue contrast and non-invasive nature. Despite substantial research progress, clinical adoption of automated BTS methods remains limited—largely due to challenges in communication and collaboration between clinicians and developers. While fully automated methods, particularly U-Net and its variants, have demonstrated remarkable accuracy compared to traditional and semi-automated approaches, several limitations persist. These include data imbalance, which affects model generalization; multi-center variability, which hinders cross-institutional robustness; and the high computational cost associated with training deep networks on large MRI datasets. Addressing these issues is essential for enabling real-world clinical deployment. Looking ahead, promising research directions include the development of transformer-based hybrid U-Net models capable of capturing global contextual information, and the adoption of federated learning frameworks to facilitate privacy-preserving, multi-center model training. Additionally, advances in explainable AI will enhance model transparency and trust among medical professionals. Together, these innovations are expected to bridge the gap between algorithmic performance and clinical

applicability, paving the way for broader integration of U-Net-based systems into medical imaging practice.

References

[1] S. Das, G. K. Nayak, L. Saba, M. Kalra, J. S. Suri, and S. Saxena, "An artificial intelligence framework and its bias for brain tumor segmentation: A narrative review," Computers in Biology and Medicine, vol. 143, p. 105273, Apr. 2022, doi: 10.1016/j.compbiomed.2022.105273.

[2] I. Mazumdar, "Automated brain tumour segmentation using deep fully residual convolutional neural networks," arXiv preprint arXiv:1908.04250, 2019

[3] Z. Liu et al., "Deep learning based brain tumor segmentation: a survey," Complex & Intelligent Systems, vol. 9, no. 1, pp. 1001 1026, Feb. 2023, doi: 10.1007/s40747-022-00815-5.

[4] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," Zeitschrift für Medizinische Physik, vol. 29, no. 2, pp. 102–127, May 2019, doi: 10.1016/j.zemedi.2018.11.002.

[5] A. Işın, C. Direkoğlu, and M. Şah, "Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods," Procedia Computer Science, vol. 102, pp. 317–324, 2016, doi: 10.1016/j.procs.2016.09.407.

[6] F. L. C. dos Santos, A. Joutsen, M. Terada, J. Salenius, and H. Eskola, "A Semi-Automatic Segmentation Method for the Structural Analysis of Carotid Atherosclerotic Plaques by Computed Tomography Angiography," Journal of Atherosclerosis and Thrombosis, vol. 21, no. 9, pp. 930–940, 2014, doi: 10.5551/jat.21279.

[7] W. Jiang, J. Chen, X. Ding, J. Wu, J. He, and G. Wang, "Review Summary Generation in Online Systems: Frameworks for Supervised and Unsupervised Scenarios," ACM Transactions on the Web, vol. 15, no. 3, pp. 1–33, Aug. 2021, doi: 10.1145/3448015.

[8] O. El Aissaoui, Y. E. A. El Madani, L. Oughdir, and Y. El Allioui, "Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles," Procedia Computer Science, vol. 148, pp. 87–96, 2019, doi: 10.1016/j.procs.2019.01.012.

[9] G. Litjens et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005

[10] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic Brain Tumor Segmentation Using Convolutional Neural Networks with Test-Time

Augmentation," in International MICCAI Brainlesion Workshop, 2019, pp. 61–72, doi: 10.1007/978-3-030 11726-9_6.
[11] M. Ghaffari, A. Sowmya, and R. Oliver, "Automated Brain Tumor Segmentation Using Multimodal Brain Scans: A Survey Based on Models Submitted to the BraTS 2012–2018 Challenges," IEEE Reviews in Biomedical Engineering, vol. 13, pp. 156–168, 2020, doi: 10.1109/RBME.2019.2946868
[12] H. Mzoughi et al., "Deep Multi-Scale 3D Convolutional Neural Network (CNN) for MRI Gliomas Brain Tumor Classification," Journal of Digital Imaging, vol. 33, no. 4, pp. 903–915, Aug. 2020, doi: 10.1007/s10278-020-00347-9.
[13] N. Q. Al-Ani and O. Al-Shamma, "A review on detecting brain tumors using deep learning and magnetic resonance images," International Journal of Electrical and Computer Engineering (IJECE), vol. 13, no. 4, pp. 4582–4593, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4582-4593.
[14] B. H. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.
[15] A. Myronenko, "3D MRI Brain Tumor Segmentation Using Autoencoder Regularization," in International MICCAI Brainlesion Workshop, Springer (LNIP), 2019, pp. 311–320, doi: 10.1007/978-3-030-11726-9_28.
[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] P. Dvořák and B. Menze, "Local Structure Prediction with Convolutional Neural Networks for Multimodal Brain Tumor Segmentation," in International MICCAI workshop on medical computer vision, Springer (LNIP), 2016, pp. 59–71, doi: 10.1007/978-3-319-42016-5_6.
[18] H. Li, A. Li, and M. Wang, "A novel end-to-end brain tumor segmentation method using improved fully convolutional networks," Computers in Biology and Medicine, vol. 108, pp. 150–160, May 2019, doi: 10.1016/j.compbiomed.2019.03.014.
[19] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Fan, and Y. Zhang, "Brain Tumor Segmentation Using a Fully Convolutional Neural Network with Conditional Random Fields," in International workshop on brainlesion: glioma, multiple

sclerosis, stroke and traumatic brain injuries, Springer (LNIP), 2016, pp. 75–87, doi: 10.1007/978-3-319-55524-9_8.
[20] M. Rezaei et al., "A Conditional Adversarial Network for Semantic Segmentation of Brain Tumor," in International MICCAI Brainlesion Workshop, Springer (LNIP), 2018, pp. 241–252, doi: 10.1007/978-3-319-75238-9_21.
[21] L. Weninger, O. Rippel, S. Koppers, and D. Merhof, "Segmentation of Brain Tumors and Patient Survival Prediction: Methods for the BraTS 2018 Challenge," in International MICCAI brainlesion workshop, Springer (LNIP), 2019, pp. 3–12, doi: 10.1007/978-3-030-11726-9_1.
[22] Z. Jiang, C. Ding, M. Liu, and D. Tao, "Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task," in International MICCAI brainlesion workshop, 2020, pp. 231–241, doi: 10.1007/978-3-030-46640-4_22.
[23] R. Mehta et al., "QU-BraTS: MICCAI BraTS 2020 Challenge on Quantifying Uncertainty in Brain Tumor Segmentation – Analysis of Ranking Scores and Benchmarking Results," Machine Learning for Biomedical Imaging, vol. 1, pp. 1–54, Aug. 2022, doi: 10.59275/j.melba.2022-354b.
[24] L. Fidon, S. Shit, I. Ezhov, J. C. Paetzold, S. Ourselin, and T. Vercauteren, "Generalized Wasserstein Dice Loss, Test-Time Augmentation, and Transformers for the BraTS 2021 Challenge," in International MICCAI Brainlesion Workshop, Springer (LNCS), 2022, pp. 187–196, doi: 10.1007/978-3-031-09002-8_17.
[25] S. Valverde, A. Oliver, M. Cabezas, E. Roura, and X. Lladó, "Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations," Journal of Magnetic Resonance Imaging, vol. 41, no. 1, pp. 93–101, Jan. 2015, doi: 10.1002/jmri.24517.
[26] D. Jiang, W. Dou, L. Vosters, X. Xu, Y. Sun, and T. Tan, "Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network," Japanese Journal of Radiology, vol. 36, no. 9, pp. 566–574, Sep. 2018, doi: 10.1007/s11604-018-0758-8.
[27] S. Ghosh, A. Chaki, and K. Santosh, "Improved U-Net architecture with VGG-16 for brain tumor segmentation," Physical and Engineering Sciences in Medicine, vol. 44, no. 3, pp. 703–712, Sep. 2021, doi: 10.1007/s13246-021-01019-w.

[28] D. E. Cahall, G. Rasool, N. C. Bouaynaya, and H. M. Fathallah-Shaykh, "Dilated inception U-net (DIU-Net) for brain tumor segmentation," arXiv preprint arXiv:2108.06772, 2021.
[29] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Jun. 2015, pp. 3431–3440, doi: 10.1109/CVPR.2015.7298965.
[30] A. Syed and B. T. Morris, "SSeg-LSTM: Semantic Scene Segmentation for Trajectory Prediction," in 2019 IEEE Intelligent Vehicles Symposium (IV), IEEE, Jun. 2019, pp. 2504–2509, doi: 10.1109/IVS.2019.8813801.
[31] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.
[32] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," arXiv preprint arXiv:1603.07285, 2016.
[33] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," arXiv preprint arXiv:1710.05941, 2017.
[34] C. P. Behrenbruch, S. Petroudi, S. Bond, J. D. Declerck, F. J. Leong, and J. M. Brady, "Image filtering techniques for medical image post-processing: an overview," The British Journal of Radiology, vol. 77, no. suppl_2, pp. S126–S132, Dec. 2004, doi: 10.1259/bjr/17464219.
[35] X. Wu, "Cardiac CT Segmentation Based on Distance Regularized Level Set," in International Conference on Big Data Analytics for Cyber-Physical System in Smart City, Springer (LNDECT), vol. 103, pp. 123–131, 2022, doi: 10.1007/978-981-16-7469-3_13.
[36] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks," in Annual conference on medical image understanding and analysis, 2017, pp. 506–517, doi: 10.1007/978-3-319-60964-5_44.
[37] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge," in in International MICCAI Brainlesion Workshop, Springer (LNIP), 2018, pp. 287–297, doi: 10.1007/978-3-319-75238-9_25.
[38] A. Kermi, I. Mahmoudi, and M. T. Khadir, "Deep Convolutional Neural Networks Using U-Net for Automatic Brain Tumor Segmentation in

Multimodal MRI Volumes," in International MICCAI Brainlesion Workshop, Springer (LNIP), 2019, pp. 37–48, doi: 10.1007/978-3-030-11726-9_4.
[39] K. Venu, P. Natesan, N. Sasipriyaa, and S. Poorani, "Review on Brain Tumor Segmentation Methods using Convolution Neural Network for MRI Images," in 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW), IEEE, Dec. 2018, pp. 291–295, doi: 10.1109/I2C2SW45816.2018.8997387.
[40] S. Iqbal, M. U. Ghani, T. Saba, and A. Rehman, "Brain tumor segmentation in multi-spectral MRI using convolutional neural networks (CNN)," Microscopy Research and Technique, vol. 81, no. 4, pp. 419–427, Apr. 2018, doi: 10.1002/jemt.22994.
[41] M. Noori, A. Bahri, and K. Mohammadi, "Attention-Guided Version of 2D UNet for Automatic Brain Tumor Segmentation," in 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, Oct. 2019, pp. 269–275, doi: 10.1109/ICCKE48569.2019.8964956.
[42] S. Alqazzaz, X. Sun, X. Yang, and L. Nokes, "Automated brain tumor segmentation on multi-modal MR image using SegNet," Computational Visual Media, vol. 5, no. 2, pp. 209–219, Jun. 2019, doi: 10.1007/s41095-019-0139-y.
[43] M. U. Rehman, S. Cho, J. H. Kim, and K. T. Chong, "BU-Net: Brain Tumor Segmentation Using Modified U-Net Architecture," Electronics, vol. 9, no. 12, p. 2203, Dec. 2020, doi: 10.3390/electronics9122203.
[44] F. Lin, Q. Wu, J. Liu, D. Wang, and X. Kong, "Path aggregation U-Net model for brain tumor segmentation," Multimedia Tools and Applications, vol. 80, no. 15, pp. 22951–22964, Jun. 2021, doi: 10.1007/s11042-020-08795-9.
[45] H. M. Ebied, S. Amin, and M. Hassaan, "Brain Tumor Segmentation Using Modified U-Net," Research Square, 2022, doi: 10.21203/rs.3.rs-1653006/v2.
[46] K. Kamnitsas et al., "DeepMedic for Brain Tumor Segmentation," in International workshop on Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries, Springer (LNIP), 2016, pp. 138–149, doi: 10.1007/978-3-319-55524-9_14.
[47] B. Erden, N. Gamboa, and S. Wood, "3D convolutional neural network for brain tumor segmentation," in Computer Science, Stanford University, USA, Technical report, 2017.

[48] D. Lachinov, E. Vasiliev, and V. Turlapov, "Glioma Segmentation with Cascaded UNet," in International MICCAI Brainlesion Workshop, Springer (LNIP), 2019, pp. 189–198, doi: 10.1007/978-3-030-11726-9_17.
[49] R. Mehta and T. Arbel, "3D U-Net for Brain Tumour Segmentation," in International MICCAI Brainlesion Workshop, Springer (LNIP), 2019, pp. 254–266, doi: 10.1007/978-3-030-11726-9_23.
[50] W. Chen, B. Liu, S. Peng, J. Sun, and X. Qiao, "S3D-UNet: Separable 3D U-Net for Brain Tumor Segmentation," in in International MICCAI Brainlesion Workshop, Springer (LNIP), 2019, pp. 358–368, doi: 10.1007/978-3-030-11726-9_32.
[51] M. Ali, S. O. Gilani, A. Waris, K. Zafar, and M. Jamil, "Brain Tumour Image Segmentation Using Deep Networks," IEEE Access, vol. 8, pp. 153589–153598, 2020, doi: 10.1109/ACCESS.2020.3018160.
[52] U. Baid et al., "A Novel Approach for Fully Automatic Intra-Tumor Segmentation With 3D U-Net Architecture for Gliomas," Frontiers in Computational Neuroscience, vol. 14, p. 10, Feb. 2020, doi: 10.3389/fncom.2020.00010.
[53] S. T. Bukhari and H. Mohy-ud-Din, "E1D3 U-Net for Brain Tumor Segmentation: Submission to the RSNA-ASNR-MICCAI BraTS 2021 challenge," in in International MICCAI Brainlesion Workshop, Springer (LNCS, 2022, pp. 276–288, doi: 10.1007/978-3-031-09002-8_25.
[54] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks With Uncertainty Estimation," Frontiers in Computational Neuroscience, vol. 13, Aug. 2019, doi: 10.3389/fncom.2019.00056.
[55] P. Mlynarski, H. Delingette, A. Criminisi, and N. Ayache, "3D convolutional neural networks for tumor segmentation using long range 2D context," Computerized Medical Imaging and Graphics, vol. 73, pp. 60–72, Apr. 2019, doi: 10.1016/j.compmedimag.2019.02.001.
[56] S. Puch, "Multimodal brain tumor segmentation in magnetic resonance images with deep architectures," Ph.D. thesis, 2019.
[57] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," IEEE Transactions on Medical Imaging, vol. 36, no. 4, pp. 994–1004, Apr. 2017, doi: 10.1109/TMI.2016.2642839.
[58] J. Che, L. Yang, Y. Zhang, M. Alber, and D. Chen, "Combining fully convolutional and recurrent neural networks for 3D biomedical image

segmentation," in Advances in neural information processing systems, 2016, p. 29, doi: 10.5555/3157382.3157438.
[59] G. Zeng and G. Zheng, "Multi-stream 3D FCN with multi-scale deep supervision for multi-modality isointense infant brain MR image segmentation," in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, Apr. 2018, pp. 136–140, doi: 10.1109/ISBI.2018.8363540.
[60] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," NeuroImage, vol. 170, pp. 446–455, Apr. 2018, doi: 10.1016/j.neuroimage.2017.04.041.
[61] A. A. Abbood, Q. M. Shallal, and M. A. Fadhel, "Automated brain tumor classification using various deep learning models: a comparative study," Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 1, pp. 252–259, Apr. 2021, doi: 10.11591/ijeecs.v22.i1.pp252-259.
[62] F. Milletari et al., "Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound," Computer Vision and Image Understanding, vol. 164, pp. 92–102, Nov. 2017, doi: 10.1016/j.cviu.2017.04.002.
[63] L. Alzubaidi et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," Journal of Big Data, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

مراجعة لنماذج التعليم العميق لتقسيم اورام الدماغ بالاستناد على(U-Net)

م. م. الاء وليد سلمان
وزارة التربية/ المديرية العامة لتربية الرصافة الثانية
الاختصاص/ ماجستير علوم حاسبات
المدرسة / ثانوية المتفوقات الأولى

مستخلص البحث:

يتطلب علاج أورام الدماغ بشكل فعّال دقة عالية في عمليتي التقسيم (Segmentation) والتصنيف للورم. وتنقسم تقنيات تقسيم أورام الدماغ إلى ثلاث فئات: يدوية، شبه آلية، وآلية بالكامل. في السنوات الأخيرة، اعتمدت العديد من الدراسات على تقنيات التعلم العميق (DL) لأتمتة هذه العمليات، مستندة بشكل أساسي إلى نموذج U-Net الذي أثبت كفاءة عالية في التعامل مع البيانات متعددة الأنماط. يستعرض هذا البحث الأدبيات المتعلقة باستخدام تقنيات الذكاء الاصطناعي نماذج U-Netفي تقسيم أورام الدماغ، موضحًا الخطوات العامة لتدريب نموذج جديد وتشمل: جمع البيانات، المعالجة المسبقة، اختيار أو تصميم بنية النموذج، استخدام التعلم الانتقالي (Transfer Learning)وتحسين الصور (اختياريًا). تشير نتائج الدراسات إلى وجود علاقة طردية بين دقة النموذج وتعقيد بنيته المعمارية، مما يطرح تحديًا مستقبليًا يتمثل في تحقيق دقة أعلى باستخدام تصميمات أقل تعقيدًا. كما يناقش البحث الاتجاهات الحديثة، البدائل التقنية، والتحديات المرتقبة في هذا المجال.

الكلمات المفتاحية: قسيم أورام الدماغ ، التعلم العميق ، U-Net، التصنيف الطبي.