

A review of deep learning models (U-Net architectures) for segmenting brain tumors

Mawj Abdul-Ameer Al-Murshidawy¹, Omran Al-Shamma²

¹Informatics Institute for Postgraduate Studies, Iraqi Commission for Computers and Informatics, Baghdad, Iraq

²Department of Scientific Affairs, University of Information Technology and Communications, Baghdad, Iraq

Article Info

Article history:

Received Feb 14, 2023

Revised May 4, 2023

Accepted Jun 4, 2023

Keywords:

Brain tumor segmentation
BRAST dataset
Deep learning
Magnetic resonance imaging
images
U-Net model

ABSTRACT

Highly accurate tumor segmentation and classification are required to treat the brain tumor appropriately. Brain tumor segmentation (BTS) approaches can be categorized into manual, semi-automated, and full-automated. The deep learning (DL) approach has been broadly deployed to automate tumor segmentation in therapy, treatment planning, and diagnosing evaluation. It is mainly based on the U-Net model that has recently attained state-of-the-art performances for multimodal BTS. This paper demonstrates a literature review for BTS using U-Net models. Additionally, it represents a common way to design a novel U-Net model for segmenting brain tumors. The steps of this DL way are described to obtain the required model. They include gathering the dataset, pre-processing, augmenting the images (optional), designing/selecting the model architecture, and applying transfer learning (optional). The model architecture and the performance accuracy are the two most important metrics used to review the works of literature. This review concluded that the model accuracy is proportional to its architectural complexity, and the future challenge is to obtain higher accuracy with low-complexity architecture. Challenges, alternatives, and future trends are also presented.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Omran Al-Shamma

Department of Scientific Affairs, University of Information Technology and Communications

Al-Nidhal Street, Baghdad, Iraq

Email: o.al_shamma@uoitc.edu.iq

1. INTRODUCTION

Highly accurate tumor segmentation and classification are required to treat a brain tumor appropriately. The tumor is classified as either benign or malignant [1]. It has different types, but gliomas are the most dangerous [2]. A glioma is classified based on four levels of microscopic magnetic resonance imaging (MRI) image and tumor characteristics. Grades I and II are referred to as low-grade glioma (LGG), which are benign and slow-growing, while Grades III and IV are referred to as high-grade glioma (HGG), which are malignant and active. The current treatment consists of surgery, radiation, and chemotherapy [3]. However, the earlier a brain tumor is detected, the greater the patient's probability of survival. MRI is one of the numerous imaging modalities clinicians use to assess the presence or absence of tumors. MRI is a non-intrusive, high-significant information provided by contrast imaging on brain tumors' shape, size, and location. Today, many MRI sequences are used in clinical imaging for better diagnosis and tumor delineation. Among these are T1-weighted MRI (T1w), T1-weighted MRI with contrast enhancement (T1wc), T2-weighted MRI (T2w), and fluid-attenuated inversion recovery (FLAIR) [4]. After an MRI diagnosis of a brain tumor, tumor segmentation is undertaken.

Brain tumor segmentation (BTS) approaches can be classified into manual, semi-automatic, and fully automatic. Manual segmentation necessitates the radiologist's utilization of multi-modality information offered by MRI scans and knowledge of anatomy and physiology acquired through practice and training. The approach entails the radiologist combing over numerous slices of photographs one after the other, diagnosing the tumor, and precisely outlining the tumor areas. A trained radiologist often performs this procedure by drawing a circle around the tumor's region of interest (ROI). This method has the advantage of permitting expert experience, but it also has certain drawbacks, such as slow tumor detection times, inter-observer variability, and inconsistent manual segmentation of MRI images by specialists. In addition, this method suffers from the consumption in time and considers difficult work. Thus, to overcome this problem, semi and fully-automatic approaches are started. These methods save time and give reliable results [5].

More specifically, certain problems with the manual segmentation method are attempted to be resolved by semi-automatic segmentation. The user's effort and time investment can be decreased using algorithms to facilitate performing the segmentation, such as spreading to eliminate the requirement for slice-by-slice segmentation to extend segmentation to other slices or increase segmentation throughout an area. There are several semi-automatic approaches; for instance, algorithms might be applied during (or before) segmentation, while others might assign after segmentation is complete. Semi-automatic techniques provide an initial segmentation that is "objective" to reduce inter-observer uncertainty. Because the manual segmentation and algorithm settings impact the results, inter-observer variability will continue to exist [6]. Semi-automatic segmentation is still prone to inter-rater user error even if it takes less time and produces more reliable results than manual approaches variability. Thus, most current techniques for segmenting brain tumors are entirely automated.

Segmentation automatically is considered the best method because it gives accurate results. In addition, it does not take much time to give results compared to other methods. There are two types of fully automatic segmentation: unsupervised and supervised brain tumor segmentation, where both types do not require human intervention. In the unsupervised approach, a dataset containing ground truth labels is unnecessary for unsupervised learning techniques. These methods are based on symmetry, color, geographical position, and other distinctive qualities of the brain tumor. Non-learning techniques often focus on a single application and accomplish segmentation using the images and illness features. These non-learning techniques need to construct a distinct model for each segmentation task because they are application-specific. Since the tumors may have a description identical to other parts of the brain, non-learning techniques like fuzzy C-means or slic are not especially helpful. Therefore, when the tumor spots are close to other locations with similar pixel values, this approach has trouble differentiating white matter from grey matter [7].

In contrast, the supervised approach uses labeled datasets (ground truth). These datasets can be used to train or "supervise" computers to recognize data and predict the future. Explicitly described inputs and outputs can be used to validate the model's accuracy, and it can develop over time. Unlabeled data sets are inspected and clustered using machine learning (ML) techniques. These algorithms find hidden data patterns without human intervention. The goal of supervised learning is to give people the ability to make precise predictions based on current facts. However, with unsupervised learning techniques, significant knowledge is gleaned from enormous amounts of recent data. In addition, computers may identify what is unusual or interesting in data sets using ML [8].

Convolutional neural networks (CNNs) have significantly influenced image analysis and comprehension, particularly in image segmentation, classification, and analysis [9]. Recently, deep learning (DL) as a subset of ML employs hierarchical constructions for learning complex abstractions from data. DL is a novel method commonly used in traditional artificial intelligence areas, like computer vision. DL's current boom may be linked to three primary factors: heavy-load chip processing capacities (e.g., graphical processing unit (GPU)), drastically decreased computer hardware prices, and significant breakthroughs in ML methods. DL networks with many layers may extract many previously unattainable characteristics.

Automatic BTS is useful in the diagnosis, treatment planning, and therapy evaluation of brain tumors. In recent years, CNNs have attained state-of-the-art performance for multimodal BTS. They required a collection of annotated training MRI images for learning as an ML approach [10]. Because of the availability of high computational power (mostly based on GPUs), it is now possible to build deep neural networks with a large number of layers that can extract a large number of previously unattainable features. Based on the dimension of the convolutional kernel used, CNN for segmentation can be divided into different categories. The segmentation map for a single slice is predicted using 2D CNNs with 2D convolutional kernels. By making predictions for each slice, segmentation maps are projected for the entire volume. The 2D convolutional kernels can use the context throughout the height and width of the slice to create predictions. However, 2D CNNs only accept one slice as input, so they cannot use the context of subsequent slices by default. For the prediction of segmentation maps, voxel information from neighboring slices may be helpful

[11]. The 3D CNN was employed to segment a volumetric scan patch using 3D convolutional kernels. Because these CNNs use a sizeable number of parameters, inter-slice context can improve performance but attack a computational cost [12].

This paper demonstrates a literature review for brain tumor segmentation using a DL approach. Consequently, it represents a common way to model a DL approach for segmenting brain tumors. Section 2 describes the DL steps to design the required model. These steps started with gathering the dataset, pre-processing, and augmenting the images if needed. Section 3 reviews the literature based on the model architecture, including 2D, 3D, and hybrid, whereas section 4 presents challenges and alternatives. Future trends are described in section 5. Finally, section 6 summarizes the conclusion obtained.

2. DEEP LEARNING APPROACH

Figure 1 shows the typical block diagram of a BTS model. It comprises four main stages as a general approach in designing a DL model [13]. Inputting the dataset images is the first stage. The dataset was downloaded from a dataset website and was refined to remove the unwanted images that lessen the opportunity of getting a desirable model. Pre-processing and data augmentation is the second stage. It is an optional stage based on the number of images and its quality in the dataset. Pre-processing includes filtering the noisy images, formatting the image to suit the model used, and resizing the image dimensions, while mirroring, rotating, flipping, and cropping are some functions of data augmentation process. Since DL model is hungry for images, data augmentation is needed to increase the dataset. The core of the segmenting system is the DL model. It is either a designed or a famous (standard) model. The model output is BTS of the whole tumor (edema). Transfer learning block is the last stage. It is optionally used to enhance the model accuracy instead of the data augmentation process.

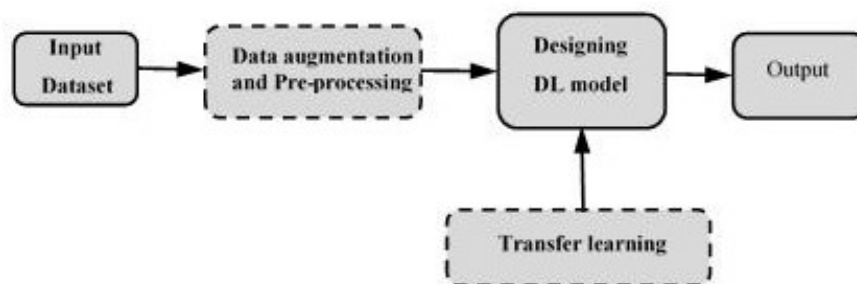


Figure 1. General block diagram for designing a DL model

2.1. Input dataset

The computing of medical images and intervention with the help of the computer for the first time launched a multimedia brain tumor challenge in 2012. Data groups have been updated annually since then to face new problems. This free data collection can be accessed online to create new ways to classify brain tumors. This dataset aims to evaluate BTS as a criterion of today's latest mechanical division of the brain. In addition, the photography data set was manually diagnosed by human specialists. Since the data set is suitable for all requirements, the proposed form classifies brain tumors automatically. Therefore, low-degree and high-grade-class tumors (HGG) appear in an MRI scan that forms the data set stored as Neuroimaging Informatics Technology Initiative (NIFTI) files (.nii.gz). The available data set is characterized by the site (BRATS). It is processed by the author when it is devoid of the skull [4]. Also, each patient contains four models (T1, T2, FLAIR, and T1C) represented by the enhanced tumor (T1), edema surrounding the tumor (T2), and tumor pulp (FLAIR) and non-augmented tumor (T1C) are the three sub-areas of the tumor shown. Three overlapping sub-areas were created from the explanatory comments: total tumor (WT), tumor pulp (TC), and augmented tumor (ET). The dimension of each MRI is 155×240×240 (Axial, Coronal, and Sagittal). Thus, the location complexity of the difference in shape, size, and density, the dataset has been divided into three parts based on its sub-regions, as shown in Figure 2 [14]. Two different data collections were provided for testing and checking health that lacks the basic labels of truth [15]. The BRATS dataset is now the most widely used open MRI dataset for objectively segmenting brain tumors. The open dataset for segmenting brain tumors is shown in Table 1.

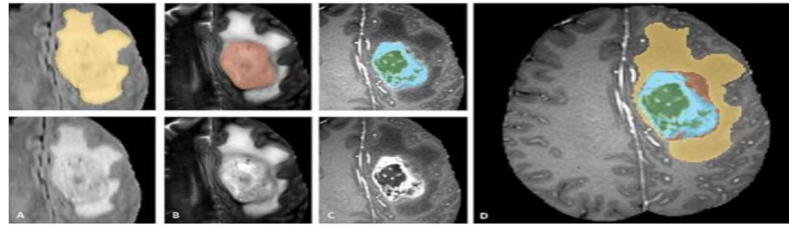


Figure 2. All regions are combined to obtain the last sub-regions tumor labels enhancing core (blue), necrotic/cystic core (green), non-enhancing solid core (red), and edema (yellow) [14]

Table 1. Summary of datasets and their descriptions

Author(s)/Year	Dataset	Description	link
Pereira <i>et al.</i> [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 $10 \times 155 \times 5 = 7550$ for LGG and 15500 for HGG.	www2.imm.dtu.dk/projects/BRATS2012/
Dvořák <i>et al.</i> [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 <i>et al.</i> [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 <i>et al.</i> [19]/2016	BRATS 2016	It is similar to BRATS 2015.	www2.imm.dtu.dk/projects/BRATS2012
Rezaei <i>et al.</i> [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 <i>et al.</i> [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 <i>et al.</i> [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 <i>et al.</i> [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 <i>et al.</i> [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 <i>et al.</i> [25]/2015	IBSR20	20 T1-w scans ($256 \times 63 \times 256$) 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 <i>et al.</i> [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 pre-processing

The first step after gathering the dataset is data augmentation. If the dataset is limited, then data augmentation is necessary to achieve the DL requirements to enlarge the dataset. Transfer learning is another solution. Thus, data augmentation is an optional step. Image pre-processing is the second step, and it is also an optional task. It makes minor improvements to the dataset images if required. However, for pre-processing data, a variety of tools and techniques are employed [27], including the following:

- From a large population of data, sampling chooses a representative subset;
- A single input is created through transformation, which modifies raw data;
- Denoising removes noise from data;
- For missing values, imputation creates a composite of statistically significant data;

- e. Feature extraction extracts a subset of pertinent features that are important in a specific situation;
- f. Masking removal is the first task of pre-processing to remove mask images from the entire data set, if any, to make the image clearer;
- g. Normalization means that the mean is divided by the density standard deviation inside the slice, using bias correlation to remove noise by applying the N4ITK algorithm;
- h. Using 2D slicing to convert the image dimensions from 3D to 2D before the image is entered into the model and using data augmentation to enlarge the number of images in the dataset [28];
- i. Also, one of the pre-processing methods is resized images, a major limitation of CNN, which is the requirement to resize the image in the dataset to a fixed dimension.

2.3. Designing deep learning model (U-Net model)

As mentioned earlier, image segmentation means dividing an image into several smaller segments, which can be achieved using U-Net architecture. Although it was introduced in 2015, it had a fabulous performance then. Since then, several successful U-Net versions have been developed. The original implementation of the U-Net architecture is presented in this section. It consists of three main parts: convolutional process (down-sampling), flattening, and de-convolutional process (up-sampling). The layers that down-sample the input is considered a component of the encoder, while the layers that up-sample the input are regarded as a decoder component, as illustrated in Figure 3 [29]. This network includes three layers: input, hidden, and output layers. The input layers include entering data (images), while the hidden layers extract features from the input data using the convolution blocks and pooling layers since each block contains a convolutional layer, rectified linear unit (ReLU activation function and batch normalization). The result represents the feature map [30].

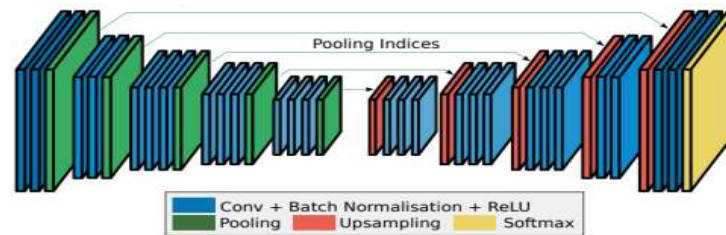
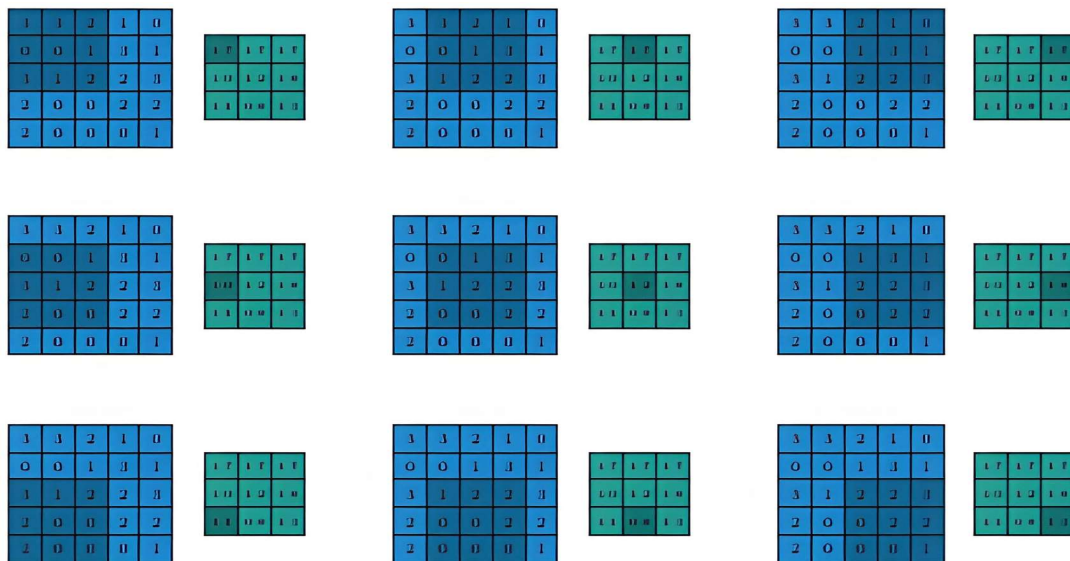
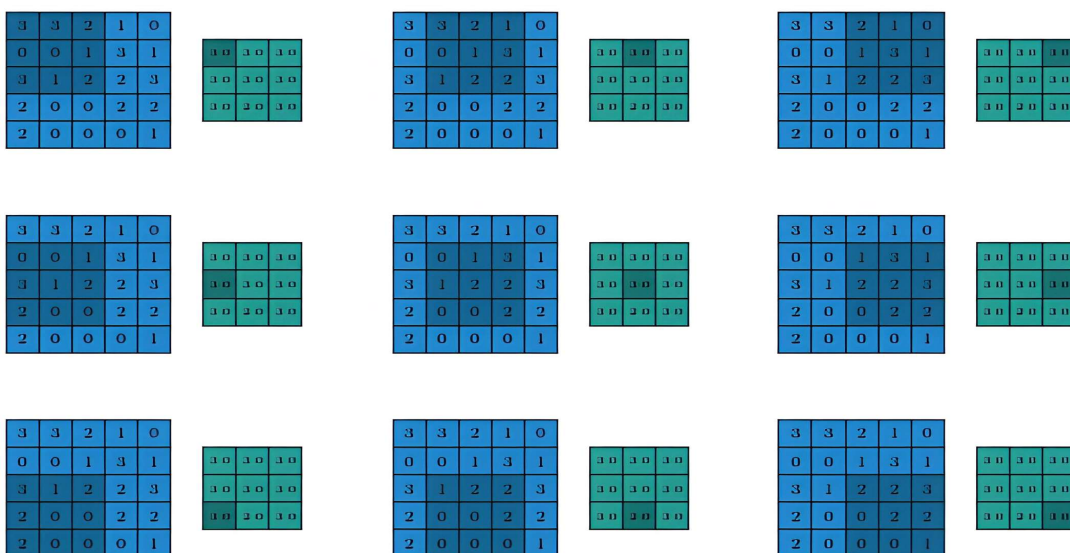


Figure 3. U-Net with downsampling and upsampling [29]

The convolution layers are considered the main layer in the CNN. The activations from the previous layers are convolved with several small parameterized filters, often of size 3×3 . This layer scans each input data (images) by using the number of kernels to produce the features map. Since features that show in one area of the image are likely also present in adjacent areas. It is possible to detect horizontal lines, for example, wherever they exist, using a filter. Note that in making a convolution, the image size will be reduced, and the number of pixels will decrease, so what is known as (padding) is done in this layer. The padding adds pixels (including columns and rows of zero for each side) to the image to keep its dimension unchanged. Moreover, instead of the filter making a complete scan of the original matrix and moving horizontally and vertically step by step, it can move two or more steps instead of one step, known as (stride convolution) [31].

When a window is slid across an input, the content is fed to a pooling function, which then performs pooling. On the other hand, the pooling layers apply a function to summarize the sub-regions, like choosing the maximum or the average value. In addition, pooling procedures reduce the feature map size. The pooling functions are somewhat similar to a discrete convolution. Figure 4 [32] illustrates the average pooling, while maximum pooling is illustrated in Figure 5 [32].

Batch normalization, which enables an increased learning rate and prevents overfitting when training deep networks, is crucial for enhancing convergence and generalization. Moreover, it is being used to quicken the training process. Small batches are processed to facilitate and speed up training rather than processing the entire dataset. Thus, batch normalization is frequently added after each convolution layer [33]. In contrast, the activation functions play a most important task in successfully training DL models. The choice of activation functions substantially impacts deep network training dynamics and task performance. There are multiple activation functions, but they do not produce identical results due to various statistical designs. The most common activation functions are leaky ReLU, ReLU, Hyperbolic Tangent function, Sigmoid, and SoftMax [4].

Figure 4. The computations of applying a 3×3 average pooling on a 5×5 input with a 1×1 stride [32]Figure 5. The computations of applying a 3×3 max pooling on a 5×5 input with a 1×1 stride [32]

2.4. Post-processing

The brain MRI image is segmented, and then the picture is subjected to several post-processing techniques to identify the tumor region in the brain. These post-processing operations include the morphological erosion applied to the segmented image using 3×3 structuring. In addition, a binary tumor-masked window is produced for the segmentation purpose. The procedure's main goal is to display the image region where the tumor is more intense and larger. Additionally, tumor tissue is more intense than the tissues around it. Thus, the tumor mask is applied to the dilated picture to produce the final image that detects brain tumors [4].

2.5. Performance measures

Generally, the performance of various segmentation techniques is evaluated using ground truth images, whereas manual segmentation, done by experts (radiologists), is used as the ground truth images. The most popular metric score to evaluate cutting-edge image segmentation algorithms is the dice score (DS) [34]. In their narrative review, Das *et al.* [1] declare that 83% of studies utilized the DS to evaluate the BTS

performance, whereas only 22% used the Hausdorff distance. In addition, they analyzed various measuring techniques in various medical imaging like lung, coronary lesion, and carotid lesion segmentation to state that the DS is the foremost measure for lesion segmentation. Hausdorff distance can be an effective measure in other domains regarding sizes and shapes, like other boundary measures [35]. However, no single measure can completely convey important information, instead of the significant criteria are task reliant on. When contrasting the ground truth photos with a predicted image, standardized assessment criteria like accuracy, specificity, and precision are used in addition to the DS.

The DS is often employed in MRI segmentation, a scale that rates how much the segmented picture and the real image overlap. It can be determined as in (1):

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

Where $|A|$ represents the predicted model pixels as tumors, and $|B|$ represents the classified pixels as tumors in the ground truth. In addition, the confusion matrix calculates the dice coefficient using (2):

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

Other performance measures like accuracy, precision, specificity, and recall are often used with DL models but rarely in segmentation.

3. DEEP LEARNING SEGMENTATION ARCHITECTURES

Initially, detecting cancerous brain tumors in the early stages is very important. Because there is no ideal method for tumor segmentation, various attempts and architectures have been proposed. Researchers have developed different DL models using MRI images to segment brain tumors automatically. These DL architectures were classified as 2D, 3D, and hybrid (2D+3D models) according to their dimensionality.

3.1. 2D architecture

Considering 2D architecture is quite common in biomedical image segmentation. It consists of two sides: an encoder, a decoder, and a link using a skip connection. Pereira *et al.* [16] developed a method for automatically segmenting data using a CNN model and tiny 3x3 kernels. In addition to being effective against overfitting, which results in fewer weights, the use of small kernels enables the creation of deeper architectures, as seen in Figure 6 [16]. They applied data augmentation on the datasets BRATS 2013 and BRATS 2015 and used intensity normalization as a pre-processing step. They improved the CNN model by improving the dice to 0.88 for BRATS 2013 and 0.78 for BRATS 2015.

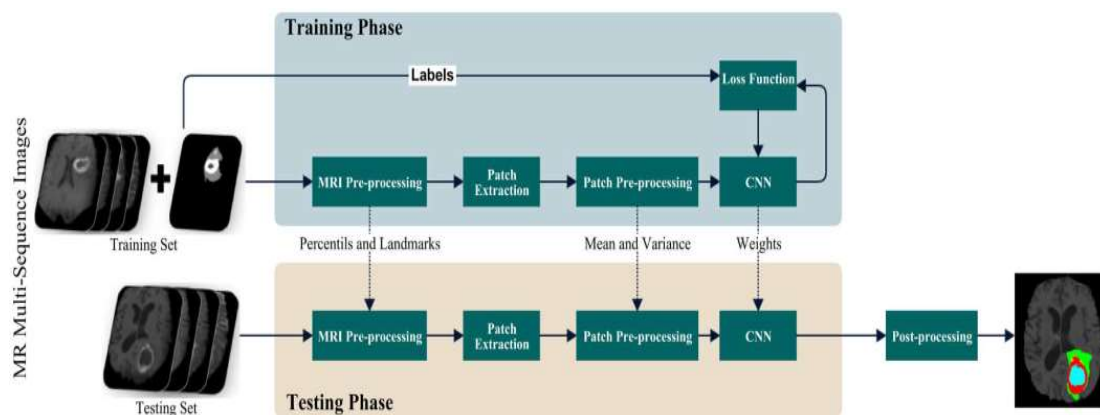


Figure 6. Overview of the proposed method in [16]

A novel U-Net architecture-based 2D fully-convoluted segmentation model is proposed by Dong *et al.* [36]. It comprises encoder and decoder paths, as depicted in Figure 7 [36]. They performed simple transformations such as zoom, shift, rotation, and flipping to augment the dataset images. In addition, they achieved their results using the dataset BRATS 2015 with WT 0.86 dice. The framework has not had

better efficiency but necessitates fewer clinical inputs because multimodal MRI data are frequently unavailable owing to constrained acquisition time and patient symptoms.

Isensee *et al.* [37] suggested a 2D U-Net model for segmenting brain tumors using MRI images. They have a context aggregation pathway as part of their architecture, which allows the encoder to gradually abstract the input representations when it proceeds deeper into the model. In order to precisely localize the interested structures, a localization pathway is then pursued that recombines shallower features with these representations. In the pre-processing step, they applied normalization, clipping the resulting images at $[-5, 5]$ to remove outliers and rescaling them to $[0, 1]$, with the non-brain region set to 0. In contrast, they achieved the result of 0.858 dice for the whole tumor on the BRATS 2015 datasets. However, the work needs to improve the prediction task by looking at the tumor's location in other brain structures, such as the ventricles, optic nerves, or other important pathways.

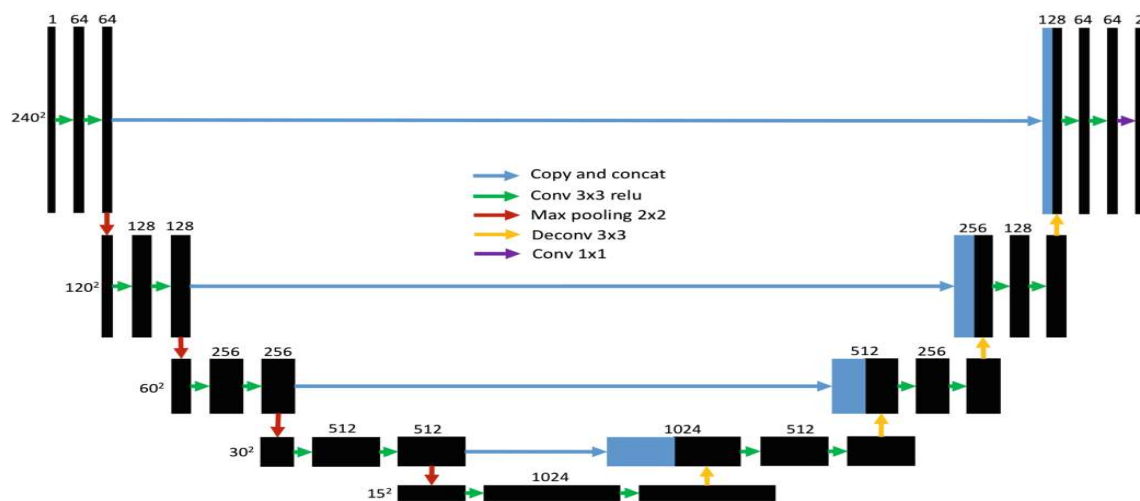


Figure 7. The architecture of the developed U-Net [36]

Kermi *et al.* [38] proposed a fully automated model using the well-known 2D U-Net with an accurate approach for intra-tumor areas, and whole-brain tumor segmentation has been developed. The designed model was trained for segmenting the LGG and HGG. Normalization and augmentation were used on BRATS 2018 dataset. They achieved a tumor core (TC), whole tumor, and enhancing tumor of 0.805, 0.868, and 0.783 dice scores, respectively. They planned and recommended enhancing the GPU for a more powerful and further accelerated model learning phase. At the same time, Venu *et al.* [39] introduced another 2D U-Net model to segment brain tumors problem. Image normalization, bias field correction, and patch extraction were employed using the BRATS 2015 dataset. They obtained an accuracy of 86.3% for the whole tumor. Simultaneously, Iqbal *et al.* [40] presented three different DL architectures to boost the model performance, including interpolated network, SeNet, and SkipNet. These models are composed of four sub-blocks with encoder and decoder architectures. In the pre-processing step, normalization, bias correction, 2D slicing (for converting the images from 3D to 2D images), and cropping (to speed the training process and feed more sets from images). The BRATS 2015 dataset was used to validate each model. For complete tumor portions, the obtained dice score was 0.9 for IntNet, 0.87 for SkipNet, and 0.88 for SENate.

Noori *et al.* [41] introduced a 2D U-Net model-based low-parameter network using two different methods. After combining high-level and low-level features, a mechanism for paying attention is used as the initial technique. This method avoids model confusion by adaptively weighing each of the channels. The multi-view fusion technique is the second method. While still using a 2D model, they can utilize this method to take advantage of the 3D contextual data of the input images. As a bias-field correction algorithm, the N4ITK is initially applied to each MRI modality for pre-processing the data to rectify the image inhomogeneity. The modality top and bottom intensities are subtracted by 1% before applying normalization to get a mean of zero and a variance of one. BRATS 2018 and 2017 datasets achieved a dice of ET 0.776, WT 0.888, and TC 0.821. Thus, the model's attempt to learn the larger classes and segmentation performance suffered from a time-consuming problem.

In contrast, the SegNet has an encoder-decoder network, afterward, a final layer of pixel-wise classification. Alqazzaz *et al.* [42] presented the SegNet model trained using four 3D MRI modalities (flair, T1, T1ce, and T2) individually before combining the results in the post-processing stage. The architecture consisted of a pair of encoders (downsampling), which has 13 convolutional layers with 3×3 filters and is flowed by max-pooling layers, and the decoder (upsampling), which has the same 13 convolutional layers as the encoder. Matching, normalizing, and bias field correction steps were applied for pre-processing. They achieved a dice score of 0.81, 0.85, and 0.79 for TC, whole tumor, and enhancing tumor, respectively, using the BRATS 2017 dataset. Model limitations include a time-consuming issue in the training phase, and the accuracy can improve in the future by using better post-processing techniques.

Rehman *et al.* [43] proposed a BU-Net architecture for BTS. Using the dataset BRATS 2017, they achieved a dice score of 0.901, 0.837, and 0.788 for whole, core, and enhancing tumors (ET), respectively. The author used RES between the encoder and decoder to get better performance. In addition, as a bias correction technique, the N4ITK algorithm was applied with normalizing to achieve a mean of zero with unity variance.

Recently, multi-modality MR imaging was used by Lin *et al.* [44]. They suggested a new model of neural networks dubbed the path aggregation U-Net (PAU-Net) model. A bottom-up path aggregation encoder (PA) was applied to decrease the distance between the deep features and the output layers to lessen the entry of noise. They employed the enhanced decoder (ED) to keep additional intact information. Moreover, the efficient feature pyramid (EFP) was implemented to enhance mask prediction while utilizing fewer resources. In the pre-processing stage, the images from the four modalities were first normalized. Then, the final multi-modality array was created by merging the images to create a four-channel array. The dataset was finally shuffled. The findings, which were compared with up-to-the-date techniques on the test set of the BRATS 2018 dataset, were dice WT 0.8563, TC 0.6751, and ET 0.6002. In addition, the findings using the testing set of the BRATS2017 dataset were dice WT 0.9000, TC 0.7095, and ET 0.6357. In contrast, the work was suffering from weak supervision.

Ebied *et al.* [45] suggested a modified U-Net model, whose architecture is identical to the standard U-Net, except that some components have been changed. Six layers of convolution in the sampling-down (encoding) path. There are two 5×5 filters and two strides in each block. Therefore, there are 2048 more feature maps than the previous one. In addition, six layers of deconvolution in the decoding process (sampling up). Each block includes two 5×5 filters and a stride of two deconvolution layers. As a result, the number of feature maps decreased. Three datasets were used: the cancer imaging archive (TCIA), the BRATS 2019 challenge, and the FIGSHARE database. They employed a batch normalization method in the FIGSHARE and BRATS 2019 datasets, while for the TCIA dataset, they used two filters: a medium filter and a soft filter. The findings with the BRATS2019 dataset were 85.02, the FIGSHARE dataset was 91.96, and the (TCIA) dataset was 86.68. Table 2 summarizes the studies that have used 2D U-Net for brain tumor segmentation.

Table 2. 2D U-Net architecture

Author(s)/Year	Segmentation approach	Dataset	Results
Pereira <i>et al.</i> 2016/[16]	Automated segmentation technique based on CNN with kernels of size 3×3 .	BRATS 2013	WT 0.88 dice for BRATS 2013
Dong <i>et al.</i> 2017/[36]	The U-Net structure was used to create a distinctive 2D fully convolutional segmentation model.	BRATS 2015	WT 0.78 dice for BRATS 2015
		BRATS 2015	WT 0.86 dice
Isensee <i>et al.</i> 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 <i>et al.</i> 2018/[38]	Modified U-net architecture-based 2D Deep CNNs.	BRATS 2018	WT 0.868 dice
Venu <i>et al.</i> 2018/[39]	An algorithm for DL based on CNN.	BRATS 2015	WT 86.3% dice
Iqbal <i>et al.</i> 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 <i>et al.</i> 2019/[41]	2D U-Net for automated segmentation	BRATS 2018	WT 0.89 dice
Alqazzaz <i>et al.</i> 2019/[42]	SegNet approach.	BRATS 2017	WT 0.85 dice
Rehman <i>et al.</i> 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 <i>et al.</i> 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 <i>et al.</i> 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

3.2. 3D architecture

The optimization complexity and maintaining the forward and backward propagated signal can be challenging in 3D models due to the huge number of trainable parameters in 3D kernels, as in Table 3. It was provided by Kamnitsas *et al.* [46] for segmenting of lesions, which was improved by adding residual connections. The intensities within the brain were further adjusted for each scan individually by dividing the mean by the standard deviation. Moreover, training data were enhanced by reflecting on the mid-sagittal plane of the BRATS 2015 dataset.

Erden *et al.* [47] presented a novel 3D model. An optimized network against a loss function was performed during the training stage. The authors utilized a U-Net architecture to obtain satisfactory findings with a moderately shallow and narrow network. They added a bounding box over the dataset as a pre-processing step, which created a lot more frustration with model syntax. On the BRATS 2017 dataset, the model obtained a dice score of 0.71. Limited the dataset from the start to utilize a single CT scan reading rather than the four used to establish the various tumor parts. They selected this approach due to the costly training time of a 3D model with several layers and because they desired to test completely many different architectures.

Lachinov *et al.* [48] presented a cascade approach for automated segmentation using 3D U-Net architecture. The 3D U-NET was modified to handle multimodal MRI input efficiently. In addition, they presented ways to improve segmentation quality with context derived from models of the same topology operating on downsampled data. Also, data normalization is carried out. At the final step, they assigned zeros to the background and shifted brain voxels to the range of 0 to 10. Then, they removed noise and outliers by rearranging each value between the limits 5 and -5. The findings of the proposed strategy on the BRATS 2018 dataset earned a dice score of 0.720/0.878/0.785 for improving the segmentation of tumors, complete tumors, and TCs, respectively.

Concurrently, Mehta *et al.* [49] proposed a 3D U-Net model from multimodal brain MR volumes. The model was an improved type of the general 3D U-Net architecture shown in Figure 8 [49]. In the pre-processing step, the volume intensities were rescaled from 0 to 1 using mean subtraction, divided by the standard deviation, and cropped to 184×200×152 on dataset BRATS 2018 with a dice score of 0.771, 0.871, and 0.706 for TC, whole tumor, and enhancing tumor, respectively. Conversely, this method degraded performance on the test dataset in the TC and ET categories.

Table 3. 3D U-Net architecture

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

Concurrently, an S3D U-Net framework for BTS was presented by Chen *et al.* [50]. To fully utilize 3D volumes, they propose a novel separable 3D convolution by splitting each 3D convolution into three parallel branches, coronal, sagittal, and axial. In addition, they offer a separable 3D block of cutting-edge residual inception architecture. The N4ITK bias correction algorithm is first applied to eliminate the bias field brought on by the magnetic field's inhomogeneity and the minor motions during scanning. Additionally, normalization is crucial in using a single algorithm to conduct multi-mode scanning. Finally, during the testing stage, they obtained the following outcomes: dice scores of 0.83093, 0.89353, and 0.74932 for tumor core, whole tumor, and enhancing tumor, respectively.

Later, an ensemble of two segmentation networks was introduced by Ali *et al.* [51]. A 3D CNN and a U-Net are combined effectively yet to produce better and more precise predictions. Several data augmentation techniques (mirroring, rotation, and cropping) are applied to train the model successfully. The proposed ensemble obtained dice scores of 0.846, 0.906, and 0.750 for tumor core, whole tumor, and

enhancing tumor, respectively, on the validation set. However, this work still has certain limitations. Firstly, only one metric (the challenge's official validation set) was utilized to evaluate the introduced segmentation ensemble. Secondly, neither the dataset nor the results were thoroughly pre- or post-processed. Finally, testing on different MRI data, irrespective of the challenge, can further evaluate the technique's validity.

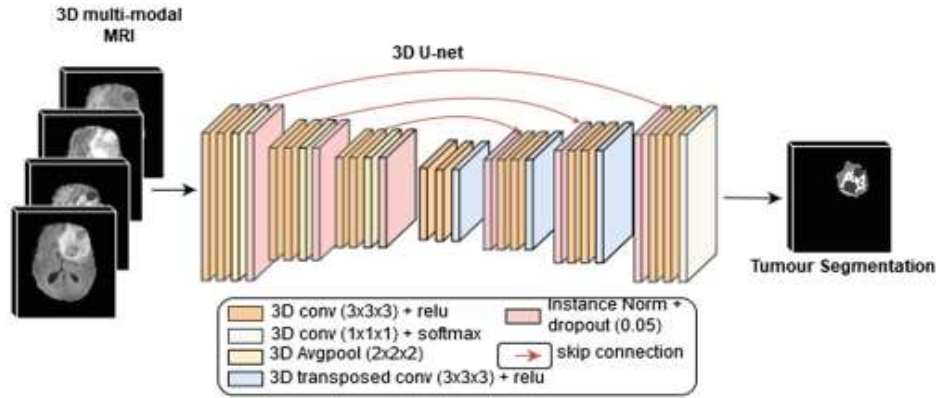


Figure 8. Multi-class tumor segmentation using four input images and 3D U-Net architecture [49]

Baid *et al.* [52] presented a novel 3D U-Net network for segmenting various brain tumors. The proposed method is unusual because it uses a weighted patch extraction method from the tumor's borders and creates a 3D U-net with fewer levels overall with more filters at each level. Normalization and the N4ITK tool were used for pre-processing step. The mean dice scores of 0.75, 0.88, and 0.83 for ET, WT, and TC, respectively, using the BRATS 2018 dataset. The authors used 3D U-Net, which needs more time for training. Thus, it can be considered a limitation.

Bukhari *et al.* [53] proposed forth an intriguing upgrade to the common 3D U-Net design that is tailored for segment brain tumors called E1D3 U-Net, which consists of one encoder and three decoders, as illustrated in Figure 9 [53]. The architecture has two extra decoders with a similar design to the original decoder to the baseline encoder-decoder architecture. One encoder and three decoders make up the resulting architecture, each of which gets feature maps from the encoder individually and produces a segmentation at the output. Before training and testing, they normalized each 3D MRI volume within the whole-brain area to zero mean and unit variance. Comparing the dice scores using the BRATS 2018 dataset with the BRATS 2021 dataset, the BRATS 2018 dice scores for WT, TC, and ET are 91.0, 86.0, and 80.2, while BRATS 2021 gives 91.9, 86.5, and 82.0, respectively. In contrast, this architecture lacks many commonly used elements, such as deep supervision and residual connections, which could considerably increase memory needs.

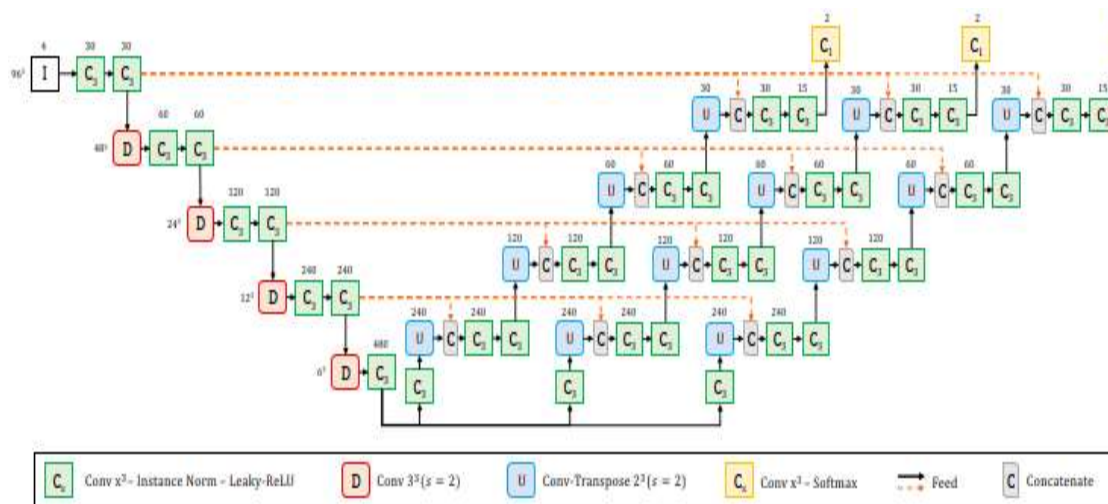


Figure 9. 3D U-Net design called E1D3 U-Net [53]

3.3. Hybrid architecture

Solving BTS involves combining the 2D and the 3D U-Net. Wang *et al.* [54] proposed a cascade of 2.5D models that balances the benefits of having a 2D CNN with memory enhancement. They suggested a CNN model that employed multiple prediction scales for deep supervision. Three prediction layers of $3 \times 3 \times 1$ convolution was utilized at various CNN levels to obtain several intermediate predictions. They used minimal pre-processing, which means they used only normalization, and to increase segmentation accuracy, they also used test-time augmentation. The quantitative assessment results using the BRATS 2017 were 0.786, 0.905, and 0.838 dice scores for tumor core, whole tumor, and enhanced tumor, respectively. It would be interesting to learn from brain tumor images that have only been partially or poorly tagged to increase the generalizability of the CNNs.

Mlynarski *et al.* [55] presented a CNN-based model that effectively integrates the benefits of the long-range 2D and short-range 3D contexts. Additionally, they recommended a model architecture with modality-specific sub-networks for additional resilience to the missing MR sequences issue throughout the training phase. They apply additional normalization for pre-processing techniques that have been offered, like histogram-matching techniques. With the median on the BRATS 2017 dataset dice scores of 0.918 (total tumor), 0.883 (tumor core), and 0.854, this technique generates precise segmentations (enhancing core). However, when the segmentation problem requires the study of a very vast spatial 3D environment, pure 3D techniques may be readily constrained by their processing requirements despite the significant recent advancements in GPUs. Table 4 lists this literature work.

Table 4. Summary of hybrid U-Net literature

Author(s)/Year	Segmentation approach	Dataset	Results
Wang <i>et al.</i> 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 <i>et al.</i> 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

Up-to-the-date DL techniques still experience many challenges and difficulties that need to be solved, although crucial advancement was made in BTS. These challenges are correlated to dataset images, training process, and performance accuracy, as in the following.

4.1. Tumor appearance

Lesion sites can only be distinguished by intensity gradients relative to the normal tissue around them. These gradients may be masked or smoothed down by low-resolution acquisitions, bias field abnormalities, or partial volume effects. Because tumors can develop in any part of the brain and can be any size or shape, it is difficult to introduce previous knowledge about the tumor's location or extent, as shown in Figure 10 [56]. The quantity of previous spatial information about the healthy brain tissue is constrained by normal tissue displacement caused by the expanding tumor lesion (also known as the mass effect) or by a resection cavity following therapy. This problem reduces the efficiency of techniques that model a healthy brain to find the diseased regions while assuming the placement of healthy tissue, such as those that use a brain atlas. The heterogeneity of tumor appearance in MRIs, which reflects the range of tumor forms and their aggressiveness, makes it challenging to employ past knowledge about the relative appearance of tumor substructures. For instance, while contrast enhancement and tumor heterogeneity are highly evident in high-grade gliomas, contrast enhancement may be observed in only 60% of low-grade gliomas. However, increasing the model depth [57] is an effective solution. Another successful solution is to apply a weighted loss function alongside a higher weight assigned to the labels of the segmented background between contacting tissues [58]. Moreover, multi-modality-based approaches and super-pixel information can be useful in addressing this issue [59].

4.2. Dataset

To achieve the benefits of DL, the established models need a sizeable volume of labeled images to perform the training stage. Gathering such a volume of labeled images is hard, as it is very boring, costly, and still a challenge [60]. Data augmentation and transfer learning are common and widely used solutions [61]. Patch-wise training is another solution. This technique divides the image into randomly or overlapping multi-patch segments. Its efficiency relatively depends on the mini-patch size and the patch overlapping [62].

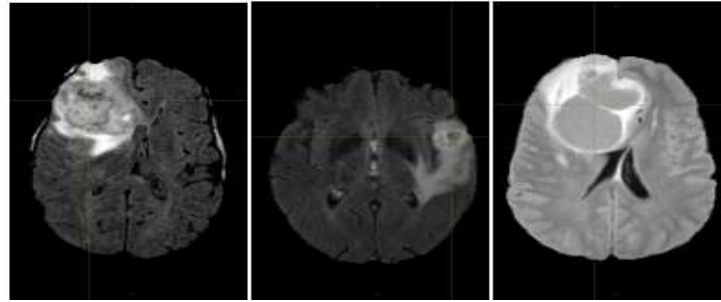


Figure 10. Axial slices of T2-FLAIR acquisitions of 3 different brains with tumors of variable grade [56]

4.3. Overfitting

In the training stage, overfitting occurs when the model can obtain the regularities and patterns with significantly higher accuracy than the unprocessed images [63]. The small number of training images is one reason for overfitting. Data augmentation is the solution. Another solution is to apply the dropout function to abandon a random neuron set outputs at each iteration during the training stage.

5. FUTURE TRENDS

As mentioned earlier, U-Net extracts the image features directly, while the traditional method extracts the hand-made features. BTS techniques will undoubtedly show great potential in the future, along with all the remarkable advances identified in this field. However, advanced tumor evaluation improvements, such as tumor volume estimation, future tumor progression estimation, and multiple tumor grading, will improve advances in current technologies. Some potential areas for future work in deep learning for segmentation include:

- Develop more efficient and effective architectures for DL models, such as those that can handle sizeable image sizes or additional classes.
- Incorporate techniques for handling class imbalance and overfitting, such as data augmentation, weighting loss functions, and regularization.
- Explore using semi-supervised and unsupervised learning techniques to use unlabeled data better.
- Investigate the use of other types of data, such as video and 3D data, for segmentation tasks.
- Develop more robust models using ensemble methods, multi-task learning, and domain adaptation.
- Develop explainable AI and interpretability techniques for DL segmentation models to make them more transparent and trustworthy.

In the specific case of U-Net, future work could include the following:

- Improving the computational efficiency of the architecture to make it more feasible to train and deploy on larger datasets and more complex tasks.
- Developing variants of U-Net better suited for handling larger image sizes, more classes, and other types of data, such as 3D data. Incorporating techniques for handling class imbalance and overfitting, such as data augmentation, weighting loss functions, and regularization.
- Investigate the use of other types of data, such as video and 3D data, for segmentation tasks.
- Develop more robust models using ensemble methods, multi-task learning, and domain adaptation. Develop explainable AI and interpretability techniques for U-Net to make it more transparent and trustworthy.

6. CONCLUSION

This article presents several techniques for automated BTS using MRI datasets. MRI-based methods are used more in BTS due to the good soft-tissue contrast and non-invasive MRI. However, the percentages of clinical applications of automated BTS methods are very low due to the lack of interaction between developers and clinicians. It is concluded that U-Net poses a very impressive architecture and provides significant results compared to traditional methods. Three methods are there to segment a brain tumor; manual, semi-automated, and full-automated. The best segmentation methods are fully automated, which has achieved high results and takes less time. Moreover, the paper concluded that architecture and performance accuracy are the most important metrics for evaluating model efficiency. Solutions to the current challenges associated with U-Net architecture were also introduced to facilitate the researchers' road in this field. Finally, future trends were suggested.

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. Direkçioğlu, 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 Alloui, "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 *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. Sasipriya, 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 U-Net 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 *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 *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.

BIOGRAPHIES OF AUTHORS



Mawj Abdul-Ameer Al-Murshidawy    received a B.Sc. degree in computer science from the University of Wasit, Kut, Iraq, in 2018. Since then, she has worked as a program developer using different computer languages and systems. She is seeking to obtain her M.Sc. degree in computer science/artificial intelligence at the Informatics Institute for Postgraduate Studies (IIPS), Baghdad, Iraq. Her field of interest includes artificial medical intelligence, machine, and deep learning applications, artificial database systems, and artificial inquiry systems. She can be contacted at email: mawjalquraishi@gmail.com.



Omran Al-Shamma    received a B.Sc. degree in control and systems engineering from the University of Technology (UoT), Baghdad, Iraq, in 1980 and M.Sc. degree in computer engineering from UoT, Baghdad, Iraq, in 1984. He received a Ph.D. in computer engineering from the University of Hertfordshire, Hatfield, UK, in 2013. Currently, he is an associate professor at the University of Information Technology and Communications, Baghdad, Iraq. His research interests include artificial medical intelligence, machine, and deep learning, computerized medical diagnosis, interactive computer design software, and preliminary aircraft design software. He can be contacted at email: o.al_shamma@uoitc.edu.iq.