IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 31 December 2023, accepted 30 January 2024, date of publication 9 February 2024, date of current version 23 February 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3365048*

TOPICAL REVIEW

# Developments in Brain Tumor Segmentation Using MRI: Deep Learning Insights and Future Perspectives

**SHAHID KARIM[1,2,3,4], GENG TONG[1,2,3,4], YITING YU[1,2,3,4,5], ASIF ALI LAGHARI[6,7], ABDULLAH AYUB KHAN[6], MUHAMMAD IBRAR[7], AND FAISAL MEHMOOD[8]**

[1]Research and Development Institute, Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China
[2]College of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China
[3]Key Laboratory of Micro/Nano Systems for Aerospace, Ministry of Education, Northwestern Polytechnical University, Xi'an 710072, China
[4]Shaanxi Province Key Laboratory of Micro and Nano Electro-Mechanical Systems, Northwestern Polytechnical University, Xi'an 701172, China
[5]Research Center of Smart Sensing Chips, Ningbo Institute of Northwestern Polytechnical University, Ningbo 315103, China
[6]Department of Computer Science, Sindh Madressatul Islam University, Karachi, Sindh 74000, Pakistan
[7]Software College, Shenyang Normal University, Shenyang 110034, China
[8]School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen 518055, China

Corresponding author: Yiting Yu (yyt@nwpu.edu.cn)

This work was supported in part by the Shenzhen Science and Technology Innovation Committee under Grant YFJGJS1.0; and in part by the Science and Development Program of Local Lead by Central Government, Shenzhen Science and Technology Innovation Committee (2021Szvup112).

**ABSTRACT** The human brain is an incredible and wonderful organ that governs all body actions. Due to its great importance, any defect in the shape of its regions should be reported quickly to reduce the death rate. The abnormal region segmentation helps to plan and monitor the treatment. The most critical procedure is isolating normal and abnormal tissues from each other. So far, remarkable imaging modalities are being used to diagnose abnormalities at their early stages, and magnetic resonance imaging (MRI) is renowned and noninvasive among those modalities. This paper investigates the current landscape of brain tumor segmentation (BTS) by exploring emerging deep learning (DL) methods for brain MRI analysis. The findings offer a comprehensive comparison of recent DL approaches, emphasizing their effectiveness in handling diverse tumor types while addressing limitations associated with data scarcity and robust validation. DL has shown a vital improvement for BTS, so our primary focus is to include significant DL robust models to analyze the brain MRI. However, DL outperforms traditional methods; still, there are several limitations, especially related to the diverse tumor types, lack of datasets, and weak validations. The future perspectives of DL-based BTS present significant potential for revolutionizing the diagnosis and treatment of brain tumors.

**INDEX TERMS** Brain tumor segmentation, deep learning, medical imaging, MRI.

## I. INTRODUCTION

Image segmentation is one of computer vision's basic and challenging tasks [1]. In recent decades, its research has been increasingly innovative and the development process continues to advance with the emergence of new technologies. Several high-level computer vision applications need positioning, recognition, detection, and segmentation functions. The processing of medical images needs to detect objects, and we often use the way of detecting specific semantic objects in digital images and locating them by their bounding boxes to obtain the results. Image segmentation aims to determine whether there are characteristic objects to be detected in the image. If so, the coordinate pixels and area size of each object detected can be attained [2].

With the continuous development and popularization of medical imaging, the tumor detection rate is increasing [3], [4]. Radiation and laser therapy greatly rely on accurate

The associate editor coordinating the review of this manuscript and approving it for publication was Roberta Palmeri.

tumor segmentation. Otherwise, it can be harmful to several sensitive nerves [5]. Medical treatment first requires doctors to give diagnostic results while observing the lesions. This process will cause a great burden on the clinical medical system for a large number of patient groups. Moreover, the subjective factors of finding between different doctors will cause trouble in clinical diagnosis and treatment. Therefore, a system that can comprehend automatic medical image segmentation will have great clinical significance. Glioma is the most frequent brain tumor in adult brain diseases [6]. It will cause irreversible damage to the brain by infiltrating the surrounding tissues. Accurately segmenting brain gliomas is very time-consuming in the process of clinical diagnosis and planning of treatment. If medical resources can be liberated through automatic segmentation, many medical resources will be saved for research work that needs more human intervention. Such resources promote the development of medical image processing-related technologies. Medical image processing is a necessary and technical means for diagnosing and planning by analyzing the lesions in the brain tumor focus area. BTS aims to classify the pixels of the lesion area in an image and obtain the detailed distribution of the lesion area, which can help doctors better understand the symptoms and diagnosis. Traditional image segmentation approaches can be divided into graph theory segmentation methods, watershed image segmentation algorithms, and so on [7]. These algorithms usually rely on the feature extraction of the image itself and its color, texture and other shallow features. Brain gliomas usually infiltrate the surrounding tissues. Their shallow features are not as obvious as those in natural images, which can affect the performance of traditional segmentation methods for BTS. It can affect the segmentation accuracy, and create other problems to prohibit the final segmented image to put into use, resulting in the decline of the significance of introducing image segmentation into clinical medical auxiliary diagnosis.

Various problems, such as environmental pollution, irregular life, and population aging, are becoming critical. Because of these factors, the incidence rate of brain diseases such as epilepsy, Parkinson's disease, brain tumors, and cerebrovascular diseases is increasing yearly [8]. It seriously threatens human health. Especially brain tumors and cerebrovascular diseases are becoming the first cause of death in China due to their high disability rate, high mortality, and high recurrence rate. The medical diagnosis of the brain, especially brain tumors and cerebrovascular diseases, often depends on medical imaging such as computed tomography (CT) [9], X-rays and magnetic resonance imaging (MRI) [10], [11]. In addition, these medical imaging techniques are also helpful for surgical planning and postoperative treatment effect evaluation. Therefore, an accurate computer-aided system can help doctors quickly locate lesions, assist in recognition, and improve doctors' efficiency and accuracy. With the development and large-scale application of machine learning, computer-aided medical image analysis based on machine learning has become one of the research hotspots in medical image analysis and machine learning. Although the current level of medical technology has made significant progress compared with the past, there is still a lack of effective treatment for brain tumors. Patients can only prolong their lives through various conservative, comprehensive treatments and operations with high-risk coefficients. If the brain tumor can be found at the early stage of its growth then the patients can get the first opportunity for treatment, and the survival probability will be greatly improved. The technology based on medical image segmentation has been applied to early brain tumor detection, but the traditional manual segmentation method requires much professional labor. It needs doctors with rich clinical experience and much professional knowledge to divide manually. Due to the lack of medical resources, there is no doubt that this method cannot be popularized and applied on a large scale.

The core problem of BTS based on deep learning (DL) is to extract, fuse and classify the different levels of brain tumor image feature information through DL technology within acceptable consumption and to improve the final accuracy of the model.

Most research works mainly concentrate on traditional structures, with little creativity in creating new neural network architectures that are customized for the complexities of brain tumor segmentation (BTS). Closing this gap requires a move toward the creation of complex models that can efficiently utilize the complementing information present in multimodal data. Furthermore, the evaluation and comparison of current approaches frequently lack critical information, which makes it difficult to establish a benchmark for assessing the actual usefulness of various strategies. Many studies using image-guided surgery use a variety of datasets that display a range of imaging modalities, including MRI, CT, and intraoperative datasets [12]. Typically, these datasets include annotated tumor area examples, making algorithmic training and validation easier. This review paper tries to summarize the available literature on image-guided surgery, focusing on DL's critical role in improving precision. The most recent approaches are briefly reviewed and demonstrate advances in recurrent architectures, attention mechanisms, and convolutional neural networks (CNNs), demonstrating the ongoing development of DL methods for precise and effective BTS in the context of image-guided surgical interventions. This paper mainly focuses on the recent developments and challenges for BTS. The key points for this contribution are:

(a) The paper introduces and evaluates cutting-edge DL methods for BTS in MRI during 2016-2023, showcasing a significant leap forward in the methodology employed for accurate analysis.

(b) The findings offer a thorough comparison of recent DL approaches, emphasizing their effectiveness in handling diverse tumor types. The paper addresses data scarcity and validation robustness limitations, providing valuable insights for researchers.

(c) The open challenges for BTS based on DL and traditional approaches have been highlighted. The advanced DL impacts on BTS and future recommendations are comprehensively described.

This paper consists of seven sections. Section II describes the background of the study. Section III illustrates the basic theory of BTS. Sections IV and V present the traditional and DL-based BTS methods, respectively. Section VI comprises important types of image-guided surgery for brain tumors. Section VII covers discussion and future recommendations. Finally, section VIII concludes this review paper.

## II. BACKGROUND

The research fields of computer vision include remote sensing applications, vehicle and pedestrian monitoring, medical image segmentation and recognition, map satellite navigation, tire defect detection, object positioning and recognition [13], scene classification and segmentation [14], [15], etc. It has also been widely used in various fields of daily life, such as military territorial air defense, medical disease analysis, intelligent video monitoring [16], and remarkable success in specific objects such as intelligent home machine control, surveillance video recognition, and medical surgery object positioning. Many object detection achievements proposed by scientific researchers for a long time have gradually increased, mainly including traditional and DL-based object detection methods [17], [18].

DL is a new technology that has been rising recently, and it is a new branch under the machine learning sub-directory [19], [20]. Its concept comes from scientists' deeper research on artificial neural networks (ANNs) [21], such as multilayer perceptron (MLP) from the perspective of structure [22], which is a simple DL structure. The network contains hidden layers, which is a typical feature of the DL structure. The DL model mainly uses a mathematical means (convolution) to extract the deep hidden feature representation in the input data. Based on the extracted feature representation, the network model can fit the input data well without manual preset rules like other methods, which is the unique learning-fitting ability of the DL network. In recent years, the best DL method in medical image processing originated from LeCun, known as CNN [23]. Later generations have made many simple and effective improvements from different angles, further enhancing CNN's data representation ability. The CNN training is more stable and simpler, which is suitable for various task scenarios. Although CNN was proposed long ago, it is still promising for semantic segmentation and object detection. The models with extraordinary performance are basically improved based on CNN variations. The number of published papers during 2016-2023 are extracted from PubMed with the included expressions of "Brain Tumor", "Deep Learning", "Artificial Intelligence", "CNN" and excluded expressions "case study", "overall surgery", "Model training for information learning" and "Not accessible" as shown in Figure 1. This analysis shows that numerous articles have been published on BTS during 2016-2023 which reveals the importance of this research.

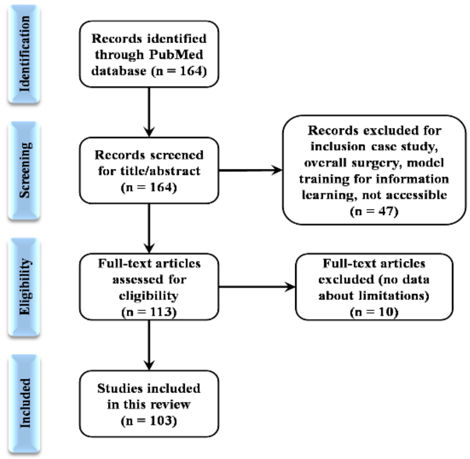



**FIGURE 1.** **PRISMA diagram of systematic review of BTS from 2016 to 2023.**

At present, partial resection of brain tumors through surgery is the best treatment method. Accurate BTS images can help doctors quickly view the lesions and implement treatment to reduce patients' pain. However, this segmentation task for brain tumors is not easy, not only because the size, shape and location of gliomas are quite different among patients, but also affected by many factors, which will lead to inaccurate segmentation, which greatly limits the availability of glioma segmentation information [24]. Big data analysis and preprocessing techniques such as cleansing, profiling, enrichment, and transformation play critical roles in early BTS [25].

In addition, the tumor mass effect will change the arrangement of surrounding normal tissues. All medical imaging modalities have some shortcomings and planar X-ray can be used to visualize the skull but has no diagnostic value for brain pathologies. For example, intensity heterogeneity or different intensity ranges between acquisition scanners [26]. The diversity of MRI acquisition parameters and sequences leads to great differences, increasing glioma variability between different patients. Therefore, it is very important to help patients find tumors as soon as possible and carry out relevant diagnosis and treatment [27]. At present, the tumor area is manually marked and segmented by radiologists. However, due to the changes caused by the appearance and shape of the glioma, the process is very time-consuming, and the consistency between evaluators is low. Therefore, automatic segmentation is attractive because it can describe relevant tumor parameters faster, more objectively, and accurately. However, due to the irregular nature of tumors, developing algorithms that can segment tumors efficiently and accurately is still challenging [28]. This inherent heterogeneity of glioma is also reflected in its imaging phenotype. Its sub-region is generated by multimodal MRI, which reflects

**TABLE 1.** Comparison of our review with existing reviews.

| Ref. No. | Scope: Conventional Approaches | Deep Learning Approaches | Image Guided Surgery Techniques | Drawbacks/ Limitations | Future Perspectives | Data/ Repository Links |
|---|---|---|---|---|---|---|
| [24] | ✓ | ✓ | | ✓ | | |
| [26] | | ✓ | | | ✓ | |
| [27] | ✓ | ✓ | | ✓ | ✓ | |
| [28] | | | ✓ | | | |
| [29] | ✓ | ✓ | | ✓ | ✓ | |
| [36] | | ✓ | | ✓ | ✓ | ✓ |
| [37] | ✓ | ✓ | | ✓ | ✓ | |
| [38] | | ✓ | | ✓ | ✓ | ✓ |
| Our paper | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

different biological characteristics of tumors according to different intensity distributions [29].

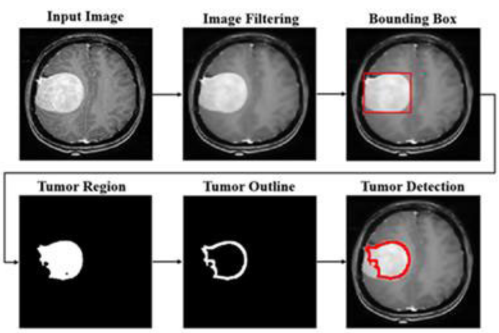



**FIGURE 2.** BTS and detection using MRI.

The object detection algorithm is also applied to the field of medical images. Tumor detection using MRI has been attained by implementing preprocessing approaches such as image filtering, thresholding, morphological operations, eroding, and subtraction, as shown in Figure 2. We have perceptibly described the contribution of this review com-pared to existing reviews in Table 1. Early tumor detection tasks are usually based on some traditional image algorithms [30]. Karkanis et al. proposed a method of tumor detection assisted by color wavelet features [31]. This method is based on a novel color feature extraction to depict different regions in the sequence and on the covariance of the second-order texture measure. A feature selection algorithm is proposed to determine the image regions along the video frame using linear discriminant analysis. Bercoff et al. proposed a tumor detection method using transient elastic imaging [32]. This method detects tumors by tracking the propagation of extremely low-frequency shear waves generated by the vibration system on the body surface in soft tissue. It has been developed to detect and quantify soft tissue and hard lesions. Wu et al. proposed a color-based tumor detection method using K-means clustering technology to observe tumor objects in MRI [33]. Its core idea is to convert a given gray MR image into a color space image and then use K-means clustering. Iftekharuddin et al. proposed a fractal-based multimodal MRI tumor detection method, which integrates two novel texture features and intensity in multimodal MRI images [34]. The texture features involve the segmented triangular prism surface area (TPSA) technique for fractal features. Meanwhile, the other texture feature uses the Brownian motion method, which integrates wavelet and fractal analysis for fractal feature extraction. Mustaqeem et al. proposed a method for tumor detection based on watershed and threshold segmentation [35]. By improving the quality of scanned images, morphological operators are used to detect tumors in scanned images. For the traditional tumor detection methods, on the one hand, due to the variety of tumor shape, size and appearance, accurate measurement is challenging. The tumor can grow suddenly, leading to the defects of adjacent tissues, providing an overall abnormal structure for healthy tissues. On the other hand, applying a tumor detection algorithm is complicated due to the particularity of medical images.

## III. THEORY OF BRAIN TUMOR SEGMENTATION

In the middle of the last century, Warren McCulloch et al. proposed the mathematical model of ANN, which laid the theoretical foundation of neural network structure in future generations [39]. Frank proposed a machine to simulate human perception and named it perceptron [40]. The structure of the perceptron is relatively simple. It is a single-layer network composed of threshold and linear units. Although single-layer perceptron can only classify linear tasks, its emergence represents the prototype of modern CNN structure. The brain tumor illustration has been presented in Figure 3, which describes the importance of different modalities. The recent developments and basic theory have been de-scribed in the following sections.

### A. ARTIFICIAL NEURAL NETWORK (ANN)

At the beginning of ANN, the computing power of the first generation of computers could not match it, which limited the role of neural networks, so it did not enter the stage of rapid development. The current computer can give full play to the characteristics of neural networks and promote the rapid development of artificial intelligence. In essence, ANN uses mathematical methods to imitate the complete process of biological visual perception [41]. In this bionic process, the data input into ANN can be regarded as the electrical stimulation signal obtained by visual cells. The network's data processing corresponds to the stimulation signal's transmission process in the biological visual perception system. Finally, the network simulates the performance of organisms for different stimuli and outputs different results. Artificial neurons play an absolute core role in the components of ANN. Multiple neurons can be stacked and connected in a certain way to form ANN.

The neurons in the same layer of ANN are not connected. The neurons between adjacent layers are connected through the weight matrix. The connection mode can be customized. Any neuron in the hidden layer is connected with all neurons in the adjacent layer. The network adopts the forward calculation method: the output of the previous layer is the input of the next layer, which is calculated layer by layer, and finally, the output is obtained. To learn the parameters in the network, Rumelhart et al. proposed an error backpropagation algorithm (BP) [42]. The algorithm first calculates the error between the network output and the true value, then backpropagates the error from back to front, adjusts the parameters of the network according to the error, and finally iterates continuously to make the network model converge to the global optimization or local optimization. From this intuitive process, we can understand the whole operation process of neurons. If a huge number of neurons are connected somehow, the ANN can disclose complex nonlinear tasks. Multilayer Perceptron (MLP) is a typical complex network that combines multiple neurons in a predetermined way. MLP is composed of layers of neurons and nodes. The layer closest to the input data is often called the input layer neuron, whose main function is to read the data into the shallow layer of the network model. The next series of neurons are called hidden layer neurons. In these layers, the network is mainly used to extract the multi-level features of input data, which is an important part of the network. The number of hidden layers is generally unlimited, and the appropriate number can be selected according to the task objectives. Developers need to make a trade-off between efficiency and accuracy. The last layer of neurons is called the output layer. After passing through the output layer, the network will provide the analyzed results of specific tasks. From this process, we can see that there is mutual propagation between multiple neurons. Based on this relationship, the characteristics of the input data move forward layer by layer in the neural network and finally form the output result of the network.

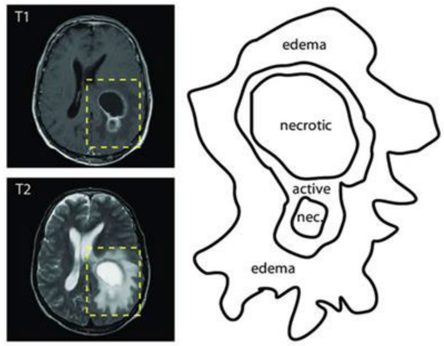



**FIGURE 3. Example of a brain tumor showing the importance of the diverse modalities (T1 with contrast and T2) [43].**

### B. CONVOLUTIONAL NEURAL NETWORK (CNN)

The most significant change brought by CNN to deep neural networks is the introduction of convolutional structures in the process of network construction. Applying this structure in deep neural networks can effectively reduce the memory and parameters occupied by the deep network calculation process in the training process, accelerate the training process, and make the network easier to converge. Compared with the traditional ANN, CNN reduces the training complexity of the network model by setting local connections in the model and sharing the weight value of each neuron in the calculation process. The number of parameters in the network is effectively reduced, and the possible overfitting phenomenon is alleviated. Another special feature of CNN is that it is very suitable for processing two-dimensional images. The convolution structure can effectively extract various features of the image (such as approximate shape, texture detail information, color information, etc.) in the network. These features help the network model understand the internal information in the image and improve the ability of the network to process image data. Because the CNN network has good translation invariance, it can still accurately extract and understand the image's content even if the image's shape changes. Because of these characteristics of CNN, there are more and more new ideas and methods in computer vision.

LeCun et al. [44] proposed CNN in handwritten character recognition and achieved great success. With the great success of AlexNet [45] in classifying ImageNet dataset, CNN has been widely used in image processing. The traditional ANN often adopts the way of full connection, and the network parameters increase exponentially with the increase of layers, which makes the network difficult to train. Even the shallow network is easy to overfit. In order to alleviate this problem, CNN introduced strategies such as local receptive field and weight sharing [46], which not only reduced the network parameters but also the weight sharing strategy can automatically extract image features from the layers. With the development of CNN, various excellent basic CNN models have emerged, such as VGG [47], GoogleNet [48], ResNet

[49], DenseNet [50], etc., but all networks are built on a series of functional layers.

## IV. BRAIN TUMOR SEGMENTATION METHODS

Among the advanced technological developments, two main factors are limiting BTS; one is medical imaging technology, and the other is segmentation methods. MRI has become a key technology suitable for brain imaging with the continuous development of medical imaging technology. MRI completes the imaging process of complex brain structures based on the magnetic field and radio waves. MRI is a promising imaging method that does not damage the human body and can generate multimodal data. The tissue characteristics presented by different modal data are different, and the multimodal data complement each other to ameliorate the accuracy of tumor segmentation. After solving the imaging problem, the key point of developing BTS is the research of segmentation methods. In recent years, the development of deep neural networks has led to the segmentation accuracy of segmentation methods based on DL has far exceeded the traditional segmentation methods. The significant segmentation methods based on supervised learning are random forest, decision tree, particle swarm optimization (PSO), dynamic sparse field (DSF); and based on unsupervised methods are K-means, C-means, mean shift, self-organizing map (SOM), markov random field (MRF), gray-level cooccurrence matrix (GLCM), multi-context fuzzy clustering (MCFC), and extreme learning machine (ELM). Some BTS methods based on traditional and DL approaches are divided into groups, as shown in Figure 4.

In BTS, the region to be segmented is not an independent region but multiple overlapping regions, such as the complete tumor region, core tumor region, and enhanced tumor region. The complete tumor region includes all tumors, and the enhanced tumor will be included in the core tumor region. Due to this characteristic, some scholars proposed that the BTS task can be divided into multiple subtasks for training to reduce the complexity of directly segmenting all the regions. At the same time, optimizing the single-segmented region is easier, and improving the overall segmentation accuracy through the fusion of the final segmentation results. Based on this idea of sub-task segmentation, Wang et al. proposed a model including three cascaded segmentation networks for BTS [51]. The three segmentation networks segment different tumor regions, respectively. The overall segmentation process is divided into three steps, and each step is re-segmented based on the previous step. Zhou and others proposed a one-step multi-task segmentation model for learning shared features, which divides the brain tumor into different segmentation tasks, integrates these tasks into one model, strengthens the interaction between tasks through shared parameters in the training process, and can complete multi-task prediction in one step directly in the prediction stage [52]. Shen et al. proposed a multi-task segmentation network, which distinguishes the regions through the up-sampling operation of different paths, then uses the designed loss function for binary classification, and finally fuses the different segmentation results to form a complete region segmentation [53].

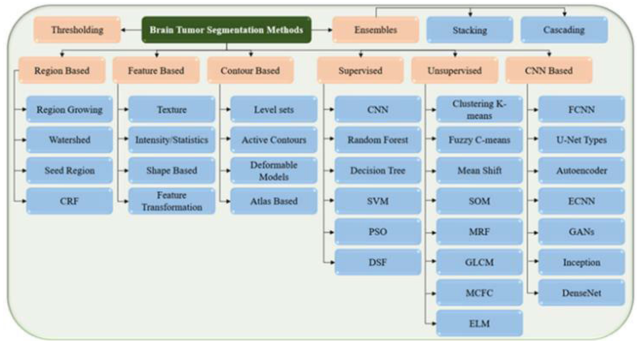



**FIGURE 4.** Classification of BTS methods.

### A. REGION-BASED SEGMENTATION

The process of grouping pixels into larger regions is known as region growing which has been utilized in region-based image segmentation algorithms [54]. Initially, scan the adjacent pixels and evaluate the merging of adjacent pixels into the region. Each pixel in the same attribute set will be used to allocate pixels for the growth process of the region. The shape of the region grows according to the intensity to complete the morphological edge detection of the input image and reconstruct the input image based on expansion and corrosion to enhance the image. The center point selects the pixels within a certain gray range and the pixels are evenly placed in the divided area. The receiving area starts from the accurate position of the first pixel, and then the seed point grows according to the set rules. An example of region-based segmentation is shown in Figure 5.

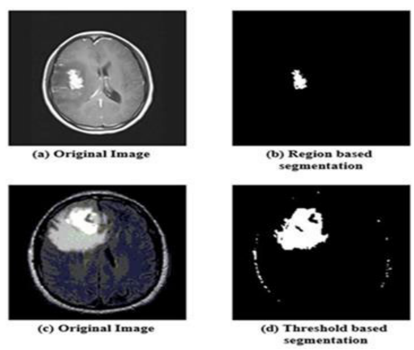



**FIGURE 5.** Results on MR brain image. (a, c) Original image (b) the result using the region growing model (d) threshold-based segmentation for brain MRI [55], [56].

### B. THRESHOLD-BASED SEGMENTATION

Threshold-based segmentation is one of the simplest image segmentation methods because it consumes less computation and has high efficiency, so it plays an effective and significant role in image segmentation. The gray image is converted into a binary image output corresponding to its region using a threshold. This method traverses all image pixels. If the pixel value of the image is higher than the determined threshold, the pixel is set to the maximum value of the proportion used; otherwise, it is the minimum value. Threshold selection techniques can be divided into two categories: two-level and multilevel. In the former case, only one threshold is needed to separate two objects of the image: one represents the object, and the other represents the background. When different objects are depicted in a given scene, multiple thresholds must be selected for correct segmentation, usually called multilevel threshold processing [57]. As seen in Figure 5, thresholding is used to process the brain MRI. Automatic image segmentation is proposed using cranium removal, morphological reconstruction, thresholding, and subtraction to get a binary mask for segmentation [58]. The segmentation results using active contour and graph cut theory have been shown in Figure 6.

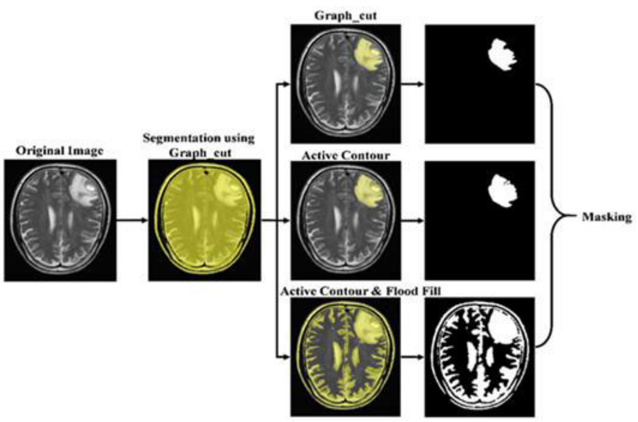



**FIGURE 6. Examples of segmentation using graph cut, active contour, and flood fill algorithms.**

### C. WATERSHED SEGMENTATION

A watershed algorithm is a mathematical morphology segmentation algorithm based on topology theory. It was introduced and improved by Lantuéjoul and Beucher [61], [62]. Subsequently, in terms of definition and implementation, several algorithms of watershed transformation with different variants are developed to reduce the computational complexity [63], [64]. The watershed algorithm realizes segmentation by constructing regions. The image to be segmented is regarded as a topographic map in geography. The gray value of pixels is used to represent the height of the terrain [65]. Mountains and valleys correspond to the gray value, respectively. When we inject a drop of water into the valley, the water will eventually converge to the local lowest point due to the action of gravity. There may be a minimum plane in which all points are minimum points. Pierce the valley, the bottom begins to fill with water, and the water level rises at a constant rate. As the water level rises, the water becomes more and more, and finally covers the whole surface. The

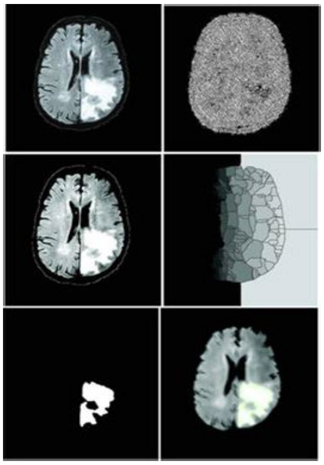


**FIGURE 7. Result of watershed segmentation for brain MRI from top to bottom (a) artifacts removal (b) watershed transform and (c) edge smoothing [59].**

watershed segmentation and accuracy assessments of conventional methods for BTS are shown in Figure 7 and 8 respectively.

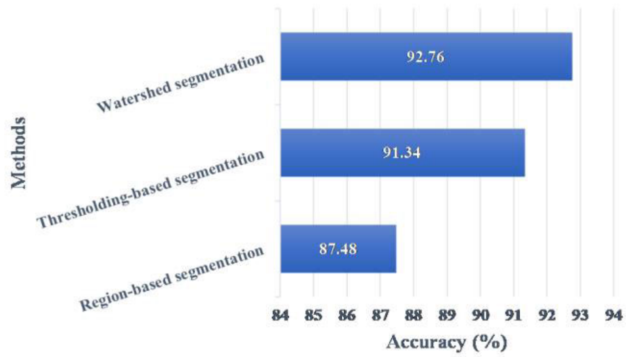



**FIGURE 8. Accuracy assessment of the traditional segmentation methods [60].**

### D. CLUSTERING-BASED SEGMENTATION

The collection of similar data into clusters depends on homogeneity, known as clustering. Among clustering algorithms, K-means clustering [66] is the base for image segmentation. The similar components in the dataset belong to the same cluster in this algorithm. On the other hand, fuzzy C-means (FCM) clustering resolves the multi-grouping issue by assigning each pixel to a separate class. It means one pixel could associate with more clusters, although each pixel reveals a distinct similarity value for every cluster. The optimization function of the C-means algorithm affects the accuracy of results. Among clustering algorithms, K-means improved K-means [67], and C-means [68] are most commonly used in segmentation [69]. There are several improved methods based on the FCM fundamental approach, such as kernelized fuzzy C-means (KFCM) [70], generalized fuzzy C-means (GFCM) [71], fast generalized FCM algorithm

(FGFCM) [72], enhanced FCM (EnFCM) (..) [73] and so on. Similarly, the K-means algorithm has been widely used for image segmentation [74], [75], [76], [77], [78]. The brain image segmentation using K-means clustering at different K levels is shown in Figure 9.

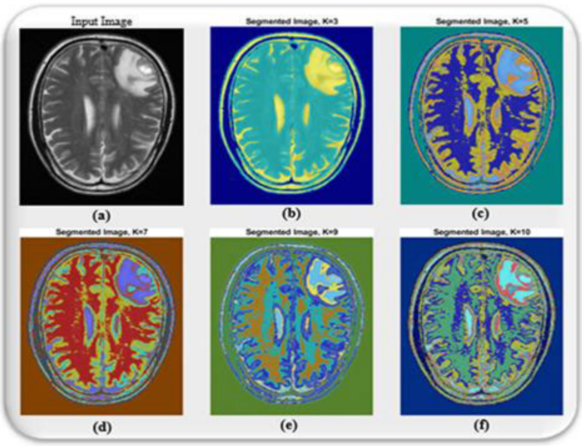



**FIGURE 9. Brain image segmentation using K-means clustering algorithm at different K values.**

## V. NEURAL NETWORK-BASED BRAIN TUMOR SEGMENTATION

In the basic theory of BTS, we will introduce the related basis of neural networks. Firstly, the related principle of ANN, the structure and application of common and advanced CNNs are introduced. Then several common segmentation frameworks based on CNN will be discussed in the following sections. Image segmentation is one of the widely used directions of the CNN model. According to the input and output, image segmentation methods based on CNN can be divided into two categories: image block-based segmentation and end-to-end segmentation.

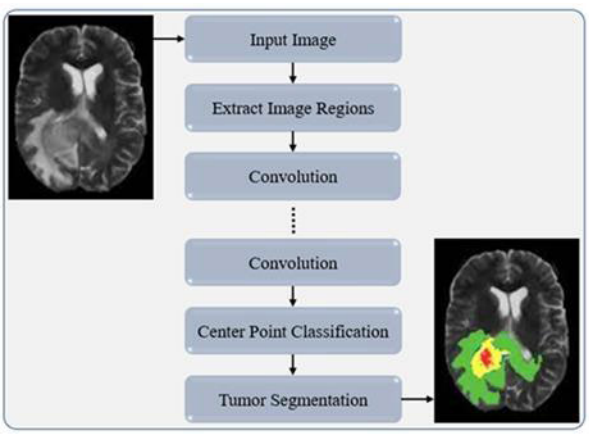



**FIGURE 10. BTS using region extraction and CNN.**

### A. IMAGE REGION-BASED SEGMENTATION METHODS

The success of AlexNet in image classification promotes the application of CNN in image segmentation. The traditional CNN model includes three functional layers: convolutional layer, pooling layer, and fully connected layer. It can automatically extract the features of the image and classify the image by using the image features. Therefore, the segmentation method based on image block maps a fixed-size image block in the image to the category using CNN's powerful fitting ability. Figure 10 shows the segmentation method framework based on regions. Firstly, the regions are extracted. Since the center points of the regions are classified, it is necessary to use the sliding window method to obtain the fixed-size blocks in the image. Then, CNN is used to extract features and classify the center points of regions. Finally, the classification of all pixels is the segmentation result of the whole image.

Because the image segmentation task is transformed into the classification of region center points, the commonly used CNN classification models, such as AlexNet, GoogleNet, ResNet, etc., can be used for image segmentation. However, unlike image classification, the size of regions often greatly impacts the network's results. Another point is that different types of regions can be selected through certain strategies in network training. Havaei et al. developed a CNN-based tumor segmentation method by exploiting global and local features simultaneously, exhibiting promising results [79].

### B. END-TO-END SEGMENTATION METHODS

The resolution of the feature image extracted by the traditional CNN model decreases with the increase of the number of network layers, so the classification center point is used to evaluate image segmentation. However, with the upsampling and deconvolution layers, the low-resolution feature map can be generated into a high-resolution feature map to analyze the end-to-end segmentation method. Common end-to-end segmentation frameworks include FCN [80], SegNet [81], U-Net [82], DeepLab [83], [84], [85], PSPNET [86], etc. the following will introduce three classic segmentation networks: FCN, U-Net, and DeepLab, of which FCN and U-Net are commonly used in medical image segmentation. Nevertheless, the overall review of end-to-end tumor segmentation methods describes the tumor tissues, which include necrosis, active tumor, and peritumoral edema [26]. It also clearly discloses brain anatomy, end-to-end method scheme, classifications and comparisons.

#### 1) GENERATIVE ADVERSARIAL NETWORKS (GANS)

The GAN model is a generalized basic unsupervised learning model, and G and D are used as generative and discriminative models. The success of the GAN model in unsupervised image generation has driven the application of the model in other image processing fields, including unsupervised image domain conversion [87], [88], text to image [89], image fusion [90], etc. In addition, the GAN model is also widely used in medical image processing, such as image super-resolution reconstruction [91], region of interest location [92], image segmentation [93], etc. To obtain image features, DCGAN uses a deconvolution network as the generation network and a convolution network as the discrimination network, which enhances the expression ability of the network and can obtain high-quality generated images [94]. Similarly, LapGAN generates a series of images with different resolutions by sampling images and then uses multiple DCGANs to generate high-quality images from low resolution to high resolution [95]. However, LapGAN needs too many networks and is difficult to train. ProGAN generates low-resolution images to high-resolution images by iteratively increasing the network depth [96]. Compared with LapGAN, it generates high-quality images with many parameters reduced.

The proposed direct PET image reconstruction network by wasserstein GAN (WGAN) model increases the image quality. Experiments were performed on mouse and patient data using a different method, and the results show that DPIR-NET enhances the reconstructed PET image quality more than DeepPET [97]. Generally, the image quality decreases due to a decrease in the radiotracer dose of positron emission tomography (PET) imaging. To enable the GAN model to generate images according to the given category, the conditional GAN (cGAN) model adds category conditions to the network when generating the network and determining the network input [98], [99]. To enable the discriminant network to learn the category features better, the proposed auxiliary classifier GAN (ACGAN) removed the conditions in the input of the discriminant network and added an output in the image category [100]. InfoGAN uses information entropy to comprehend unsupervised feature extraction of different types of images [101]. KL divergence is often used in GAN to measure the distance between the real data distribution and the generated data distribution, which leads to the discontinuity of the loss function at some points and the problem of mode collapse. To alleviate this problem, Arjovsky et al. proposed WGAN, which uses the Earth Mover distance instead of KL divergence to measure the difference between the two distributions [102]. Gulrajani et al. adopted a gradient penalty term instead of weight clipping to achieve Lipschitz restriction to solve the problem of parameter centralization caused by weight clipping of WGAN [103]. In addition, Mao et al. adopted the least square loss function, and WGAN can be considered a special case [104]. Lim et al. used hinge loss (HL) to measure the distance between two distributions [105]. HL is an extension of EM distance, which can make the GAN model training more stable.

### 2) FULLY CONVOLUTION NEURAL NETWORK (FCN)

In recent years, there have been more and more excellent models in the field of image segmentation. Long et al. proposed a fully convolution neural network (FCN), which is undoubtedly a classic network [80]. Unlike the traditional CNN, which uses the full connection layer to obtain the probability of a one-dimensional vector corresponding to many categories, FCN replaces the full connection layer with the convolution layer, outputs an image, and realizes end-to-end training. Moreover, FCN can input images of any size, and the size of input and output images is consistent through the upsampling operation. Using a jump connection structure, the low-level and high-level feature information is added and fused point by point to ensure that more features can be learned.

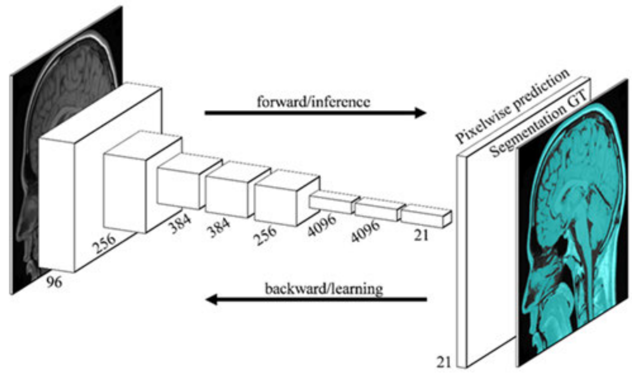



**FIGURE 11.** FCN framework for pixel-wise prediction and semantic segmentation.

Based on the upsampling technology and multi-scale fusion method, the middle layer can be divided into FCN-32s, FCN-16s, and FCN-8s. FCN-32s directly outputs an image with the same size as the original image through the upsampling operation with a step size of 32, but loses a lot of detail information; FCN-16s performs two upsampling operations. First, the last layer is upsampled, and then a convolution operation of $1 \times 1$ is added to the fourth pool layer. Then, the convolution results are fused with the upsampling results, and an upsampling operation is performed to restore the original image size. The network retains more detailed features; FCN-8s also samples up the output of the last layer, then samples up the fourth pool layer, and then integrates the results of the two upsampling with the third pool layer before upsampling, to obtain better results. However, FCN still has poor extraction of image detail information, and the processing of pixels is independent and lacks spatial consistency [106]. The authors have proposed a feature extraction and concatenation method, which would help the early detection of BT [107]. DensNet201 and Inception-v3 models are used to evaluate brain tumor detection and classification, resulting in 99.34% and 99.51% accuracies and achieving the highest performance in brain tumor detection. Figure 11 is the structural diagram of FCN.

### 3) SEGNET

SegNet is an improved model based on FCN, including encoding and decoding structures [81]. Compared with the pooling layer in other network structures, the advantage of SegNet is that the pooling layer has an index function, which is utilized to record the position of the max pooling result associated with the pooling core. The pool index connects the pooled layer output to the corresponding upper sampling layer. Since the network exploits the symmetrical structure, the first pooling layer links to the last upsampling. In the encoding phase, each pooling operation will save the relative position of the weight selected by maximizing the $2 \times 2$ filter. Then, in the decoding phase, the pooling operation is used to process these stored indexes and their corresponding feature maps. The decoding process uses zero filling volume, mainly used to fill the feature map information sampled on the sampling without generating parameters to be learned; it saves memory space and generates a sparse feature map. These feature maps are then convoluted to produce dense feature maps. Finally, the classifier outputs the maximum values of different categories through softmax to obtain the final segmentation graph. Salma et al. proposed FCNN SegNet to automatically segment brain tumors and parts of tumors by applying four different imaging modalities of MRI (i.e., T1, T1ce, T2, and Flair) [108]. The basic architecture of SegNet is shown in Figure 12.

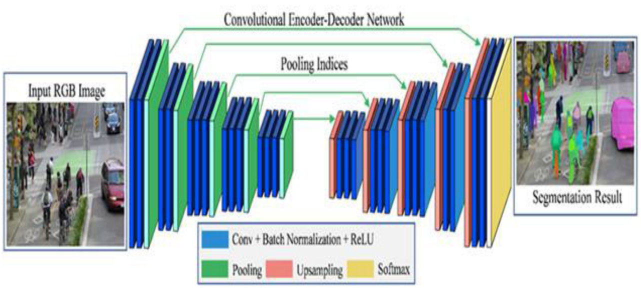



**FIGURE 12. SegNet architecture for pixelwise segmentation [81].**

### 4) U-NET

Unlike natural images, for medical images such as brain tumor imaging, even small marginal segmentation errors will significantly reduce the experience of DL automatic segmentation. The main reason for limiting the model to accurately segment this part of the difficult areas such as irregular contour, discontinuous organization, and fine-grained organization of the target is the ability of the model to make full use of shallow information and capture more details. To solve this problem, 3D U-net introduced a dual channel attention mechanism and applied it to the channel dimension and spatial dimension of the feature map through the channel attention and spatial attention branch, respectively, and embedded it into the layer hopping connection to alleviate the semantic gap caused by the direct integration of low-level features and high-level features, also to utilize the efficiency of low-level features.

The U-Net model adopts the encoding-decoding method, as shown in Figure 13 [82]. The encoding method is to continuously obtain the characteristics of the input image through the convolution layer and the lower sampling layer. With the increase in the number of layers, the image features extracted by the coding layer are more and more abstract, and the feature resolution is decreasing. Unlike FCN, which uses these features for prediction results, U-Net uses decoding to generate high-resolution (fine-grained) features from low-resolution (coarse-grained) features. Finally, U-Net classifies high-resolution features to obtain segmentation results with the same input resolution. As shown in Figure 13, U-Net adopts a multilayer downsampling layer during encoding. However, the downsampling layer will lose image information, which leads to the lack of image information for the features after upsampling during decoding. In order to resolve this problem, u-net uses a horizontal connection to connect the encoded features with the decoded features. In U-Net, the upper sampling layer is used to upgrade the dimension of features. It can be seen that the deconvolution layer also has the same ability and is more flexible than the upper sampling layer. Therefore, the deconvolution layer can replace the upper sampling layer.

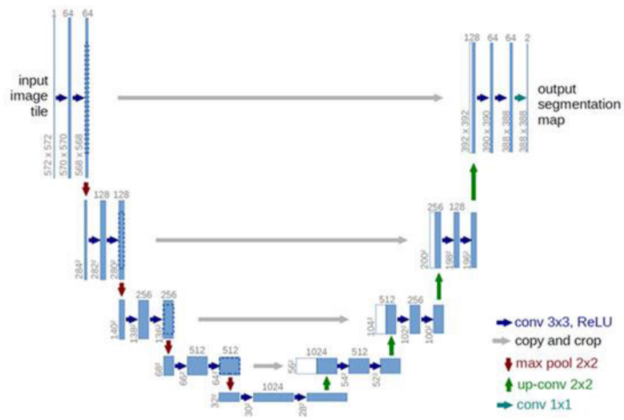



**FIGURE 13. U-Net architecture [82].**

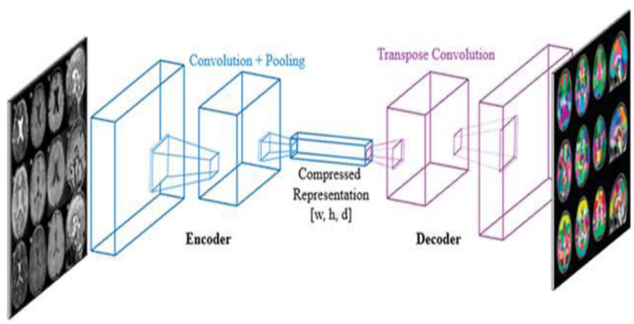



**FIGURE 14. DeepLab architecture for image segmentation.**

### 5) DEEPLAB MODELS

The DeepLab model is also an encoding-decoding method. Its coding method also adopts convolutional down sampling. DeepLab also adds a residual layer to better extract features and build a deeper network. In the network, too many downsampling layers will lead to image information loss. However, the downsampling layer can increase the receptive field of the network and can obtain global features. In order to ensure that the network has a larger receptive field without reducing the feature resolution, DeepLab cancels the last two lower sampling layers and uses the Atrous Spatial Pyramid Pooling (ASPP) module to segment targets at different scales. The ASPP module uses cavity convolution to obtain larger receptive fields. At the same time, it uses cavity convolution with different expansion rates to obtain receptive fields of different sizes and then combines these features to obtain pyramid features. The decoding method of the DeepLab model also adopts the upsampling plus convolution method and introduces the horizontal connection.

However, unlike the U-Net method, DeepLab finds that the 4-fold upsampling is better [109]. Later on, the authors proposed DeepLab v1, which introduced the variation in pooling stride and padding size, and this model was based on the VGG network [83]. In DeepLab v1, conditional random fields (CRFs) are introduced to enhance the segmentation accuracy [110]. Consequently, DeepLab v2 was proposed to resolve the segmentation problem caused by the same object scale variations in the same image [109]. DeepLab v3 utilized ResNet-101 with a cascaded Atrous module to get promising image segmentation at multi-scales [84]. The authors have utilized the cascaded network of two DeepLab v3+ [85] with ResNet-50 to get improved segmentation results [111]. The DeepLab architecture for image segmentation is shown in Figure 14.

## VI. IMAGE GUIDED SURGERY FOR BRAIN TUMORS

### A. STEREOTACTIC SURGERY

In the form of image-guided surgery (IGS), stereotactic surgery uses several imaging modalities, including CT and MRI, to produce a precise map of the brain [142], [143]. The surgeon then used this map to precisely direct a surgical tool, such as a drill or biopsy needle, to the tumor's position. This method seeks to precisely target the tumor while causing the least amount of harm possible to nearby healthy brain tissue. In order to keep the patient's head motionless throughout the treatment, the patient is first placed in a unique head-holding equipment [144], [145]. Next, a CT or MRI scan of the patient's head can be performed to get precise images of the patient's brain. A 3-dimensional map of the brain can be made using MRI scans which helps the surgery. A frameless stereotactic device used by the surgeon throughout the process, confers to the head-holding device and guides the surgical tools to the exact position of the tumor. The surgeon can sample or remove the tumor using the instruments. Diverse types of brain tumors, including benign and malignant tumors, can be treated with IGS. It is particularly beneficial for brain tumors that are located in hard-to-reach regions, like deep-seated.

### B. FLUORESCENCE-GUIDED SURGERY

A specific dye can be placed into the patient's blood circulation during fluorescence-guided surgery (FGS), which uses IGS to target the tumor only [28], [146]. The fluorescence can be detected by specific cameras, which enable the surgeon to see the tumor and surrounding structures during the surgery. The risk of harming the surrounding healthy brain tissue can be reduced using this method though increasing the precision and accuracy of brain tumor surgery. The process initiates with controlling a specific dye to the patient. Tumor cells favorably absorb this dye, which causes them to glow when struck by light. An exclusive camera detects the fluorescence scans and develops detailed brain images. The surgical procedure can be planned using these images. The specialist uses the fluorescence images to lead the surgical instruments to the precise spot of the tumor. This aids in reducing the chance of destruction of healthy brain tissue by extending the specialist's capability.

Gliomas, a form of brain tumor known to invade surrounding brain tissue, have been found to respond particularly well to FGS [147]. These tumors can be challenging to remove with conventional surgical methods. The probability of leaving any tumor cells behind after surgery can be reduced by FGS. It has been discovered that the method is efficient and safe, with a high degree of precision.

### C. INTRAOPERATIVE MRI (IMRI)

Intraoperative MRI (iMRI) is a type of IGS that uses an MRI machine in the operating room to give the physician real-time brain imaging while executing surgery [148]. This technique intends to improve the accuracy and safety of brain tumor surgery by giving the physician instant access to the location and size of the tumor. Using MRI, detailed brain scans can be taken prior to the treatment. These images guide the surgical tools according to the procedure. The physician can confirm the surgery's position and extent using real-time images.

Cancers located in difficult-to-reach parts of the brain, such as deep-seated tumors or tumors close to critical functioning areas, have been found to benefit particularly from iMRI [149]. This systematic procedure has proven to be quite accurate and rarely problematic. It also enables the physician to confirm that the tumor has been completely removed. The primary benefit of iMRI is that it provides the physician with real-time imaging during the procedure, to enhance the accuracy and safety. iMRI is not yet widely used because it is comparatively a new technique.

### D. IMAGE-GUIDED RADIATION THERAPY (IGRT)

Using imaging tools like CT or MRI, image-guided radiation treatment correctly targets and carries radiation to the brain tumor while limiting exposure to the surrounding tissues [150], [151]. This method intents to enhance the precision and efficacy of radiation therapy for brain malignancies. Initially, the patient undergoes imaging, such as CT or MRI scan, to get accurate brain images. During therapy, the patient is positioned on the treatment table, and the position of their body is tracked by specialized imaging tools. A linear accelerator generates high-energy x-rays or particles and can be utilized to guide the radiation beams at the tumor. Imaging equipment can be employed to improve the accuracy.

**TABLE 2.** Comparison of DL methods for BTS.

| DL Method | Model Architecture | Segmentation Task | Advantages | Disadvantages |
|---|---|---|---|---|
| Attention-Gated Network (AGN) [112][113][114][115] | U-Net with attention gates | Whole tumor, tumor core, enhancing tumor | High accuracy, robust to noise | Requires large amounts of data |
| Hybrid Attention Network (HAN) [116][117][118][119] | U-Net with hybrid attention mechanism | Whole tumor, tumor core, enhancing tumor | Efficient training handles complex tumor shapes | Limited ability to generalize to other datasets |
| Multi-Scale Attention U-Net (MSAU-Net) [120][121][122][123] | U-Net with multi-scale attention modules | Whole tumor, tumor core, enhancing tumor | High accuracy, captures long-range dependencies | Increased computational complexity |
| Conditional Random Field (CRF) with U-Net [124][125][126][127][10 ] | U-Net with CRF post-processing | Whole tumor, tumor core, enhancing tumor | Improved spatial consistency, reduces segmentation errors | Requires additional training for the CRF |
| DeepLabV3+ and ASPP [85][128][129][130] | DeepLabV3+ and atrous spatial pyramid pooling | Whole tumor, tumor core, enhancing tumor | Efficient for large datasets, handles different tumor sizes | Less accurate than some U-Net based methods |
| Swin Transformer [131][132][133][134][13 5] | Swin Transformer with multi-level attention | Whole tumor, tumor core, enhancing tumor | State-of-the-art performance handles complex tumor shapes and pathologies | High computational cost, requires large amounts of data |
| Generative Adversarial Network (GAN) [136][137][138][139][14 0] | GAN-based model for generating synthetic brain tumor images | Whole tumor, tumor core, enhancing tumor | Improved data augmentation, generates realistic synthetic data | Requires careful training to avoid artifacts, limited control over specific tumor characteristics |
| SegNet [81][108][141] | SegNet with encoders and decoders | Whole tumor, tumor core, enhancing tumor | Efficient architecture, easy to implement | Lower performance compared to newer methods |

It has been discovered that IGRT is particularly helpful for brain cancers which are located in hard-to-reach regions of the brain. This method is highly accurate and it enables the tumor to get a larger radiation dose while reducing exposure to the surrounding healthy regions. The benefit of IGRT is that it enables the radiation oncologist to have an eye on the tumor's location and change the radiation beams as necessary to ensure that the tumor receives the right amount of radiation while minimizing exposure to the surrounding tissues [152]. Radiation oncologists, technologists, and radiologists have been trained in the tools and image interpretations are needed as part of a specialized team.

### *E. AUGMENTED REALITY-BASED SURGERY*

Augmented reality (AR) is being employed more and more in the field of neurosurgery to remove brain tumors [153], [154]. Real-time, 3-dimensional (3D) scans of a patient's anatomy can be sent to the physician using AR. As the brain is a fragile, complex organ that can be exciting to explore, AR can be exclusively helpful. The physician's ability to view the tumor and surrounding anatomy in 3D can increase surgical precision and accuracy. Since it can be difficult for the physician to see the tumor and surrounding tissues with normal 2D imaging techniques.

AR can be used to aid surgical planning and reduce human errors. For instance, the physician can utilize AR to construct a virtual map of a patient's brain [155]. It may be beneficial to reduce the risk of destruction and improve surgical accuracy. AR can help to lessen the need for numerous surgical passes, which can shorten the surgery's duration. There are already a variety of systems available for AR-based brain tumor surgery, each with exceptional high-tech capabilities [156]. AR is progressively being used in neurosurgery, and the advantages of employing AR in brain tumor surgery include increased precision and accuracy, better surgical planning, and a shorter anesthetic stay. AR will play a bigger role in the treatment of brain tumors and other neurological diseases as technology advances.

### *F. ULTRASOUND-GUIDED SURGERY*

Brain tumors can be found and removed with the help of ultrasound-guided surgery (UGS) [157]. High-frequency sound waves are used in this procedure. They are sent to the tumor and then returned to a computer screen. The physician can use this information to lead the surgical tools precisely to the tumor's location. Real-time brain imaging is the primary benefit of UGS. Previously, physicians were forced to rely on preoperative imaging in a standard tumor surgery, such as MRI or CT scans, which could be less accurate and could fail to pinpoint the tumor's specific location. The benefit of UGS being less intrusive than conventional brain surgery. UGS is more precise and with less harm to the surrounding tissues. This phenomenon may result in fewer difficulties and quicker recovery times. Transcranial ultrasonography, or TCUS, is the ultrasound technology used in this procedure [158]. With the use of this technology, deep-seated lesions like brain tumors can be seen in real-time. A computer analyzes and compiles the echoes formed by the tumor.

UGS is continuously being studied as a relatively new procedure [159]. However, early research proposes it may help treat brain malignancies. In one study, researchers discovered that when compared to conventional brain surgery, UGS resulted in a very high rate of total tumor removal [160]. Patients who underwent surgery with UGS also experienced a shorter hospital stay. Moreover, there are several restrictions with UGS. For larger tumors or cancers found in specific regions of the brain, for instance, it might not be suitable.

## VII. DISCUSSION

The multimodal MRI brain tumor image segmentation is to segment the whole tumor region, core tumor region, and enhanced tumor region from normal brain tissue, using the image data of different imaging modalities. Accurate segmentation of multimodal MRI brain tumor images with AI has great significance in clinical diagnosis. It can reduce much time for doctors to manually divide brain tumor areas, make doctors pay more attention to treating patients and saving people, and improve doctors' medical intellectual level. Therefore, it has always been an important topic in medical image processing. With the continuous development of modern medical imaging, the traditional manually labeled MRI image segmentation is gradually replaced by computer-aided diagnosis. Accurate and rapid diagnosis not only helps doctors make judgments in a short time but also reduces the error of manual markings. As a common malignant tumor, early non-invasive diagnosis is very helpful for treating patients. However, due to the differences in the appearance and shape of brain tumors among different patients, evaluating tumor areas is very time-consuming. Therefore, the automatic segmentation based on DL is very attractive in this case because it can describe the relevant tumor parameters more objectively and accurately. Databases with complete details provide easy and quick access to specific and significant studies [29], [36]. Similarly, the pros and cons of overall categorized methods into thresholding, traditional machine learning, region-based, DL variants, and hybrid approaches can be found in [37], [161], [162], and [163]. Suppose the multi-scale information helps the network segment some easily confused areas at the edge of brain tumors. In that case, two branches are added to the original single-scale confrontation network to form a parallel multi-scale segmentation network. Stacked residual blocks and attention mechanisms are introduced to improve the segmentation accuracy of the model. The residual module can alleviate the problems of gradient dispersion and network failure in the process of deepening the network depth, accelerate the convergence process of the network, and allow the network to extract the characteristic information of data at different scales simultaneously in combination with the multi-path structure. The feature information is used by multi-scale fusion. It focuses limited attention on the details of the lesion area, which helps to improve the sensitivity of the model and improve the accuracy of fine segmentation of brain tumor areas.

The progress of semantic segmentation—from traditional approaches to sophisticated DL techniques—is highlighted, with a particular emphasis on pixel-level classification [164]. The authors cover both supervised and unsupervised learning algorithms, emphasizing the vital significance of DL in tackling important issues. It also offers an extensive examination of surveys devoted to semantic segmentation in the context of DL. BTS has made rapid developments with the development of DL. The major advantage of CNN is that its hidden layer automatically learns image features and iterates features through the connection between layers. However, CNN has the problem of learning blindness. The design of network depth and structure is very important for the speed and accuracy of feature extraction. Several DL methods have been discussed, which are the most significant and outperformers compared to all earlier models, as shown in Table 2. The dice scores are calculated in the literature based on three categories: tumor core, whole tumor, and enhancing tumor, and we have quoted the mean values of the whole tumor from test data. A couple of articles have provided accuracy (i.e., given in %) instead of dice scores, which can be observed in Table 2. Their associated dice scores cannot predict the performance of the given methods because these values vary with respect to test and training data, and it is recommended that the complete article is read for a better understanding.

### A. LIMITATIONS AND FUTURE PERSPECTIVES

For BTS, DL methods have been frequently used for the last decade, which involve multiple layers and many steps in the computer vision algorithms to comprehend the intensity and symmetry-related information. All these properties are combinedly and used for the classification of different regions of tumors such as necrosis, oedemic, gliomas, enhancing or non-enhancing tumors, etc. Frequent use of AI still needs many validations, specifically clinical and biological validations [165]. Firstly, there are some limitations to getting such validations in which the data is a key limitation. In several computer vision scenarios, we are improving the application outcomes by utilizing our models and with the help of non-related datasets. Therefore, various techniques have been developed to overcome this limitation, such as data augmentation, transfer learning, etc. Logically, data augmentation techniques should be avoided for sensitive and deadly scenarios. As we know, several types of tumors are different in shape, location, and size. Boundaries of tumors are irregular, discontinuous, and unclear, so it is better to develop a cumbersome dataset exclusively for one specific disease, which can better assist computer vision approaches. Nevertheless, BraTS development has lessened this limitation dramatically by introducing multiple imaging for BTS [166]. Secondly, biological validation reveals that images cannot interpret the biological structures because tumor lesions can

**TABLE 3.** Description of DL models with their dice scores and training data.

| Sr. No. | Model Name/Backbone | Papers Description (code link) | Dice Score | Extra Training Data | Year |
|---|---|---|---|---|---|
| 1 | CNN [161] | Automated segmentation based on CNN that explores small 3×3 kernels. The small kernels let scheming the deeper network, in addition to prohibiting overfitting, and adapt to the smaller amount of weights. Intensity normalization is utilized at the pre-processing stage. (NA) | 0.88 | No | 2016 |
| 2 | Cascaded CNN [79] | Automatic BTS is conferred by DNN and utilizes local and global features. (https://github.com/jadevaibhav/Brain-Tumor-Segmentation-using-Deep-Neural-networks) | 0.88 | No | 2017 |
| 3 | 3D CNN + CRF [167] | It is based on 11-layers deep structure, multi-scale, 3D CNN architecture and a robust training process. (https://github.com/deepmedic/deepmedic) | 0.85 | No | 2017 |
| 4 | CNN + 3D filters [168] | A CNN with 3D filters is used for brain MRI and also applied to hand bones. (https://github.com/BRML/CNNbasedMedicalSegmentation) | 0.85 | Yes | 2017 |
| 5 | Cascaded Anisotropic CNNs [51] | A cascade of FCN is developed and introduces a hierarchy of subregions (i.e. tumor categories). (https://github.com/hellopipu/brats17) | 0.87 | No | 2017 |
| 6 | U-Net + more filters + data augmentation + dice-loss [169] | A U-Net based model and a dice loss function are employed to improve the performance and overcome overfitting. (https://github.com/shalabh147/Brain-Tumor-Segmentation-and-Survival-Prediction-using-Deep-Neural-Networks) | 0.85 | No | 2018 |
| 7 | AFN-6 [170] | An autofocus convolutional layer with the capability of multi-scale processing generates strong features that work adaptively for semantic segmentation. (https://github.com/yaq007/Autofocus-Layer) | 0.84 | No | 2018 |
| 8 | Att U-Net [113] | The attention gate model is proposed which observes the varying shapes. It eliminates the modules for organ localization in cascaded CNN. (https://github.com/sfczekalski/attention_unet) | 0.84 | No | 2018 |
| 9 | NVDLMED [171] | An encoder-decoder-based semantic segmentation network is proposed for subregion segmentation in 3D MRI. (https://github.com/IAmSuyogJadhav/3d-mri-brain-tumor-segmentation-using-autoencoder-regularization) | 0.87 | No | 2018 |
| 10 | ModelGenesis [172] | This model named Models Genesis is created using self-supervision, without labeling and application-specific models for segmentation and classification problems. (https://github.com/MrGiovanni/ModelsGenesis) | 0.925 | No | 2019 |
| 11 | KMFCM [173] | A hybrid energy-efficient uses active contours K-means and fuzzy C-means to segment tumors. (NA) | 0.82 | Yes | 2019 |
| 12 | ECNN [174] | An enhanced CNN model with an optimized loss function is developed for automatic segmentation. (NA) | 0.90 | No | 2019 |
| 13 | DCNN [175] | It uses DL-based 3D CNN and multimodal MRIs for BTS. The potent features are selected from segmented regions using a cross-validation decision tree for survival prediction. (https://github.com/shalabh147/Brain-Tumor-Segmentation-and-Survival-Prediction-using-Deep-Neural-Networks) | 0.90 | No | 2019 |
| 14 | Bag of tricks [176] | It categorizes tumor segmentation tricks into semi-supervised, patch-size training, and data sampling. It also elaborates on optimization methods. (NA) | 0.88 | No | 2019 |
| 15 | OM-Net + CGAp [177] | The authors proposed a light weight deep network to resolve class imbalance compared to cascaded models. It integrates joint and discriminative features into one model. (https://github.com/chenhong-zhou/OM-Net) | 0.92 | No | 2020 |
| 16 | MS-Dual-Guided [178] | This model overcomes the drawbacks of the encoder-decoder network by utilizing contextual dependencies and self-attention. (https://github.com/sinAshish/Multi-Scale-Attention) | 0.80 | No | 2020 |
| 17 | Semantic Genesis [179] | To semantically enrich information, self-classification, self-discovery, and self-restoration are introduced to improve segmentation. (https://github.com/fhaghighi/SemanticGenesis) | 68.8% | No | 2020 |
| 18 | Feature Concatenation [107] | It introduces extraction concatenation of multi-level features to diagnose tumors at an early stage. The DensNet201 and Inception-v3 pretrained models are used for validation. (NA) | 99.51% | No | 2020 |

**TABLE 3.** ***(Continued.)*** **Description of DL models with their dice scores and training data.**

| | | | | | |
|---|---|---|---|---|---|
| 19 | ELM and RELM-LOO [180] | The images have been enhanced by fuzzy median filters which help to segment accurately using fuzzy set approach. (NA) | 0.96, 0.87 | No | 2020 |
| 20 | DeepSeg [181] | A decoupling framework is developed in an encoder-decoder way. A CNN provides spatial information in the encoder part. (https://github.com/razeineldin/DeepSeg) | 0.84 | No | 2020 |
| 21 | U-Net and CNN based Model [182] | The authors employed U-Net to segment tumors; pretrained Vgg16 using transfer learning and a classifier are used to grade tumors. (NA) | 0.84 | Yes | 2020 |
| 22 | nnU-Net [183] | Several modifications are introduced in BraTS to improve segmentation such as postprocessing, data augmentation, region-based training and minor variations to nnU-Net. (https://github.com/MIC-DKFZ/nnUNet) | 0.89 | No | 2020 |
| 23 | Extension of nnU-Net [184] | 3D U-Net based model is trained by utilizing patch strategies and normalization for BTS. (https://github.com/woodywff/brats) | 0.89 | No | 2020 |
| 24 | Segtran (i3d) [185] | A transformer-based model is developed to overcome the limitation of receptive field which is not addressed in the earlier models. The squeezed attention can regularize the self-attention and the expansion block can learn diversified representations. (https://github.com/askerlee/segtran) | 0.82 | No | 2021 |
| 25 | TransBTS [186] | The transformer is utilized in 3D CNN to segment brain tumors. This framework is analogous to the encoder-decoder network. (https://github.com/Wenxuan-1119/TransBTS) | 0.90 | No | 2021 |
| 26 | DL and attention mechanism [187] | The region-oriented preprocessing is proposed in this paper. It ends overfitting and decreases computational time. (NA) | 0.92 | No | 2021 |
| 27 | CMFT and CMFF [188] | The cross-modality feature transition and fusion are introduced in the DL framework to enhance the feature representation. (NA) | 0.90 | No | 2021 |
| 28 | MRA-UNet [126] | Multiscale residual attention-UNet preserves sequential information. It exploits cascaded type multiscale learning to employ adaptive ROIs to segment tumor regions precisely. (NA) | 0.90 | No | 2022 |
| 29 | DCNN [189] | A DCNN based model is proposed which introduces a dynamic loss function to significantly consider tumor sub-regions. Morphological relationship and symmetric attention are utilized to lessen boundary constraints and highlight locations of tumors, respectively. (NA) | 0.916 | No | 2022 |
| 30 | MFF-DNet [190] | A dual path framework based on multimodal feature fusion is proposed which uses kernel multiplexing to comprehend perceptual domain and nonlinear mapping features (i.e. to enhance coherence). (NA) | 0.92 | No | 2022 |
| 31 | SwinBTS [191] | SwinBTS blends edge and semantic data to produce precise segmentation. The method, which uses multimodal MRI scans to segment brain tumors, makes use of 3D context and performs at the cutting-edge level on benchmark datasets. (NA) | 0.89 | No | 2022 |
| 32 | TBraTS [192] | A trusted BTS network that offers trustworthy uncertainty estimations and strong segmentation outcomes. It does so without placing an undue strain on the backbone network's computing resources. (https://github.com/Cocofeat/TBraTS) | -- | No | 2022 |
| 33 | $D^2$-Net [193] | D2-Net with missing modalities consists of a modality disentanglement stage (MD-Stage) and a tumor-region disentanglement stage (TD-Stage). The MD-Stage utilizes a spatial-frequency joint modality contrastive learning scheme to separate modality-specific information from MRI data. (https://github.com/CityU-AIM-Group/D2Net) | 0.72 | No | 2022 |
| 34 | Transformer [194] | It consists of an edge detection module using CNNs, an edge spatial attention block (ESAB) for feature improvement, and a semantic segmentation module using the Swin Transformer and a shifted patch tokenization technique. (https://github.com/HXY-99/brats) | 0.91 | No | 2023 |
| 35 | DPAFNet [195] | Dual path convolution is used by DPAFNet to increase network scale and residual connections are included to prevent degradation. At the channel level, it introduces an attention fusion module that integrates global and local information. DPAFNet improves the semantic information in the collected features by combining feature maps of various scales. (NA) | 0.90 | No | 2023 |

**TABLE 3.** ***(Continued.)*** **Description of DL models with their dice scores and training data.**

| | | | | | |
|---|---|---|---|---|---|
| 36 | U-Net [196] | U-Net-based network, redundant filters are found using a genetic approach. Filter pruning is viewed as a multiobjective optimization issue that concurrently seeks to optimize performance and inference time. The compressed network is then used for the segmentation of brain tumors. (https://github.com/FBehrad/Evo_conv) | 0.87 | No | 2023 |
| 37 | CKD-TransBTS [197] | Based on MRI imaging principles, it divides input modalities into two categories, utilizing radiologists' experience in identifying brain cancers from a variety of modalities. The model combines the advantages of the Transformer and CNN to pinpoint precise lesion borders and collect long-range features in 3D volumetric images. (NA) | 0.93 | No | 2023 |
| 38 | Federated Learning [198] | Federated learning-based model permits collaborative learning by training the model with distributed data. Due to the scalability of this proposed model, it can be applicable to multiple medical imaging scenarios. | 0.89 | No | 2023 |
| 39 | Handcrafted features + CNN [199] | The proposed approach utilizes handcrafted features and CNN network to improve BTS performance. The features are based on intensity, shape, and texture from MRI scans. | 0.91 | No | 2023 |

grow differently after removing them, and tumor invasions cannot be observed by AI or imaging techniques. Some limitations in the field of BTS are listed below:

- Lack of training datasets for DL methods
- Sophisticated segmentation techniques are mandatory when applying for annotations or changing structured labels
- 3D segmentation models have been implemented using 2D segmentation [200].
- The training of the DL system on BraTS dataset needs consideration to handle uncertainty and noise.
- The small receptive field for a large dataset, then deep models are not worthwhile. Training a network has various constraints, such as limited memory, GPU, and bandwidth.
- Fixed-size for a kernel for image slicing may damage some valuable information.
- Data augmentation (i.e., scaling and rotation) and normalization approaches have been used to develop new lesions of brain tumors, which may generate class imbalance [161].

In medical image segmentation, especially the end-to-end image segmentation algorithm, there is a critical problem of data imbalance. Although scholars have proposed many methods to alleviate the data imbalance problem, such as using data enhancement, reducing image size, changing the loss function of network training, etc., these methods cannot solve the problem. In the future, we need to explore effective methods to make it easier for the network to learn the characteristics of a small number of samples in brain medical image segmentation and improve the segmentation accuracy of the network. Currently, the main algorithm of medical image segmentation is still supervised. Still, the supervised algorithm has high requirements for input data, and the acquisition of label data requires a lot of human resources, which leads to high-cost requirements for this algorithm and makes it unsuitable for the current situation in the medical field. Using weakly supervised or unsupervised algorithms to realize medical image analysis is an important direction in the future. The commonly used brain medical image segmentation algorithms based on neural networks are derived from natural image processing algorithms. However, brain medical images differ from natural images, and there is a large number of prior knowledge of medical anatomy in brain medical images. This prior knowledge helps to enhance the segmentation performance of the network, but the existing brain medical image segmentation based on the neural network often ignores this point. Therefore, in the future, it is necessary to combine neural networks with medical anatomy knowledge to enhance the segmentation performance of the network.

## VIII. CONCLUSION

This paper delineates the significant contributions of recent studies for BTS. The key role of this article is to provide an understanding of selecting robust methods for competent segmentation and future studies, which will help doctors get concrete disease identification. However, we have also grouped the BTS methods with respect to their characteristics. Several limitations related to datasets, algorithms, state of the arts, etc., have been discussed briefly to fascinate researchers for deep studies. Furthermore, we have presented various optimal methods with their dice scores, training data, and published time with references. Neurologists persist in discussing that the datasets are limited for several diseases individually, which can be helpful for computer-based diagnosis systems, and that can be overcome by centralizing the systems of hospitals. Several hospitals around the globe might have limited datasets individually, but we can devise and assist in benchmarking the datasets in the medical field, which can be a good prospect. DL approaches outperform conventional methods due to their limitations. The recent models can be further improved with ensemble and data augmentation procedures.

## DATA AVAILABILITY

The data are available from the corresponding author upon reasonable request.

## CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## ACKNOWLEDGMENT

*(Shahid Karim, Geng Tong, and Yiting Yu contributed equally to this work.)*

## REFERENCES

[1] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, Feb. 2020.

[2] T. S. Yoo, *Insight Into Images: Principles and Practice for Segmentation, Registration, and Image Analysis*. Natick, MA, USA: AK Peters Ltd, 2004.

[3] S. Metlek and H. Çetiner, "ResUNet+: A new convolutional and attention block-based approach for brain tumor segmentation," *IEEE Access*, vol. 11, pp. 69884–69902, 2023.

[4] Z. Liu, L. Tong, L. Chen, Z. Jiang, F. Zhou, Q. Zhang, X. Zhang, Y. Jin, and H. Zhou, "Deep learning based brain tumor segmentation: A survey," *Complex Intell. Syst.*, vol. 9, no. 1, pp. 1001–1026, Feb. 2023.

[5] I. Hasan, S. Roy, B. Guo, S. Du, W. Tao, and C. Chang, "Recent progress in nanomedicines for imaging and therapy of brain tumors," *Biomater. Sci.*, vol. 11, no. 4, pp. 1270–1310, 2023.

[6] E. C. Holland, "Progenitor cells and glioma formation," *Current Opinion Neurol.*, vol. 14, no. 6, pp. 683–688, Dec. 2001.

[7] K.-S. Fu and J. K. Mui, "A survey on image segmentation," *Pattern Recognit.*, vol. 13, no. 1, pp. 3–16, 1981.

[8] P. Picone and D. Nuzzo, "Biofabrication of nanovesicles for brain diseases," *Neural Regeneration Res.*, vol. 18, no. 3, pp. 525–526, 2023.

[9] H. Jiang, T. Shi, Z. Bai, and L. Huang, "AHCNet: An application of attention mechanism and hybrid connection for liver tumor segmentation in CT volumes," *IEEE Access*, vol. 7, pp. 24898–24909, 2019.

[10] Y. Ding, F. Chen, Y. Zhao, Z. Wu, C. Zhang, and D. Wu, "A stacked multi-connection simple reducing net for brain tumor segmentation," *IEEE Access*, vol. 7, pp. 104011–104024, 2019.

[11] S. Rajendran, S. K. Rajagopal, T. Thanarajan, K. Shankar, S. Kumar, N. M. Alsubaie, M. K. Ishak, and S. M. Mostafa, "Automated segmentation of brain tumor MRI images using deep learning," *IEEE Access*, vol. 11, pp. 64758–64768, 2023.

[12] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, "A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned," *Magn. Reson. Imag.*, vol. 61, pp. 300–318, Sep. 2019.

[13] C. Cao, B. Wang, W. Zhang, X. Zeng, X. Yan, Z. Feng, Y. Liu, and Z. Wu, "An improved faster R-CNN for small object detection," *IEEE Access*, vol. 7, pp. 106838–106846, 2019.

[14] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state of the art," *Proc. IEEE*, vol. 105, no. 10, pp. 1865–1883, Oct. 2017.

[15] X. Hu, L. Jing, and U. Sehar, "Joint pyramid attention network for real-time semantic segmentation of urban scenes," *Appl. Intell.*, vol. 52, no. 1, pp. 580–594, Jan. 2022.

[16] A. Ben Mabrouk and E. Zagrouba, "Abnormal behavior recognition for intelligent video surveillance systems: A review," *Exp. Syst. Appl.*, vol. 91, pp. 480–491, Jan. 2018.

[17] W. Guo, H. Ye, and F. Cao, "Feature-grouped network with spectral–spatial connected attention for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2021, Art. no. 5500413.

[18] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277–281, Feb. 2020.

[19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[20] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.

[21] D. Anderson and G. McNeill, "Artificial neural networks technology," *Kaman Sci. Corp.*, vol. 258, no. 6, pp. 1–83, 1992.

[22] M. W. Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences," *Atmos. Environ.*, vol. 32, nos. 14–15, pp. 2627–2636, Aug. 1998.

[23] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[24] M. I. Sharif, J. P. Li, J. Naz, and I. Rashid, "A comprehensive review on multi-organs tumor detection based on machine learning," *Pattern Recognit. Lett.*, vol. 131, pp. 30–37, Mar. 2020.

[25] F. Ullah, A. Salam, M. Abrar, and F. Amin, "Brain tumor segmentation using a patch-based convolutional neural network: A big data analysis approach," *Mathematics*, vol. 11, no. 7, p. 1635, Mar. 2023.

[26] S. Rajput and M. S. Raval, "A review on end-to-end methods for brain tumor segmentation and overall survival prediction," 2020, *arXiv:2006.01632*.

[27] 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," *Comput. Biol. Med.*, vol. 143, Apr. 2022, Art. no. 105273.

[28] A. J. Schupper, M. Rao, N. Mohammadi, R. Baron, J. Y. K. Lee, F. Acerbi, and C. G. Hadjipanayis, "Fluorescence-guided surgery: A review on timing and use in brain tumor surgery," *Frontiers Neurol.*, vol. 12, Jun. 2021, Art. no. 682151.

[29] T. A. Soomro, L. Zheng, A. J. Afifi, A. Ali, S. Soomro, M. Yin, and J. Gao, "Image segmentation for MR brain tumor detection using machine learning: A review," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 70–90, 2023.

[30] G. Latif, "Brain MR image classification for glioma tumor detection using deep convolutional neural network features," *Current Med. Imag. Rev.*, vol. 17, no. 1, pp. 56–63, 2021.

[31] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, and M. Tzivras, "Computer-aided tumor detection in endoscopic video using color wavelet features," *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 3, pp. 141–152, Sep. 2003.

[32] J. Bercoff, M. Tanter, and M. Fink, "Supersonic shear imaging: A new technique for soft tissue elasticity mapping," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 51, no. 4, pp. 396–409, Apr. 2004.

[33] M.-N. Wu, C.-C. Lin, and C.-C. Chang, "Brain tumor detection using color-based K-means clustering segmentation," in *Proc. 3rd Int. Conf. Intell. Inf. Hiding Multimedia Signal Process. (IIH-MSP)*, Nov. 2007, pp. 245–250.

[34] K. M. Iftekharuddin, J. Zheng, M. A. Islam, and R. J. Ogg, "Fractal-based brain tumor detection in multimodal MRI," *Appl. Math. Comput.*, vol. 207, no. 1, pp. 23–41, Jan. 2009.

[35] A. Mustaqeem, A. Javed, and T. Fatima, "An efficient brain tumor detection algorithm using watershed & thresholding based segmentation," *Int. J. Image, Graph. Signal Process.*, vol. 4, no. 10, pp. 34–39, Sep. 2012.

[36] M. W. Nadeem, M. A. A. Ghamdi, M. Hussain, M. A. Khan, K. M. Khan, S. H. Almotiri, and S. A. Butt, "Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges," *Brain Sci.*, vol. 10, no. 2, p. 118, Feb. 2020.

[37] A. Fawzi, A. Achuthan, and B. Belaton, "Brain image segmentation in recent years: A narrative review," *Brain Sci.*, vol. 11, no. 8, p. 1055, Aug. 2021.

[38] M. K. Balwant, "A review on convolutional neural networks for brain tumor segmentation: Methods, datasets, libraries, and future directions," *IRBM*, vol. 43, no. 6, pp. 521–537, Dec. 2022.

[39] W. S. Mcculloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.

[40] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958.

[41] J. Yang, N. Kudulaiti, Z. Chen, L. Gao, N. U. F. Hameed, R. Feng, and S. Lu, "Within and beyond the visual cortex: Brain tumors induce highly sensitive plasticity of visual processing in whole-brain neural functional networks," *Cerebral Cortex*, vol. 32, no. 20, pp. 4422–4435, Oct. 2022.

[42] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.
[43] J. J. Corso, E. Sharon, S. Dube, S. El-Saden, U. Sinha, and A. Yuille, "Efficient multilevel brain tumor segmentation with integrated Bayesian model classification," *IEEE Trans. Med. Imag.*, vol. 27, no. 5, pp. 629–640, May 2008.
[44] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[45] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[46] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J Physiol.*, vol. 195, pp. 215–243, Mar. 1968.
[47] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[48] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[50] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
[51] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2017, pp. 178–190.
[52] C. Zhou, C. Ding, Z. Lu, X. Wang, and D. Tao, "One-pass multi-task convolutional neural networks for efficient brain tumor segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2018, pp. 637–645.
[53] H. Shen, R. Wang, J. Zhang, and S. McKenna, "Multi-task fully convolutional network for brain tumour segmentation," in *Proc. Annu. Conf. Med. Image Understand. Anal.* Cham, Switzerland: Springer, 2017, pp. 239–248.
[54] R. Anastácio, L. R. O. Mamere, P. C. Carneiro, T. A. A. Macedo, and A. C. Patrocinio, "Preprocess enhancement of CT image for liver segmentation with region growing algorithm," in *Proc. 6th Eur. Conf. Int. Fed. Med. Biol. Eng.* Cham, Switzerland: Springer, 2015, pp. 134–137.
[55] W. Deng, W. Xiao, H. Deng, and J. Liu, "MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve," in *Proc. 3rd Int. Conf. Biomed. Eng. Informat.*, vol. 1, Oct. 2010, pp. 393–396.
[56] G. P. Anandgaonkar and G. S. Sable, "Detection and identification of brain tumor in brain MR images using fuzzy C-means segmentation," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 2, no. 10, pp. 3964–3967, 2013.
[57] T. Yang, H. W. Tian, and X. M. Liu, "Research on image segmentation algorithm based on edge detection and Otsu," *Comput. Eng.*, vol. 42, no. 11, pp. 255–266, 2016.
[58] Z. Bobotov and W. S. Bene, "Segmentation of brain tumors from magnetic resonance images using adaptive thresholding and graph cut algorithm," in *Proc. 20th Central Eur. Seminar Comput. Graph.*, Slovakia, U.K., 2016, pp. 1–8.
[59] S. M. K. Hasan and M. Ahmad, "Two-step verification of brain tumor segmentation using watershed-matching algorithm," *Brain Informat.*, vol. 5, no. 2, pp. 1–11, Dec. 2018.
[60] N. Baraiya and H. Modi, "Comparative study of different methods for brain tumor extraction from MRI images using image processing," *Indian J. Sci. Technol.*, vol. 9, no. 4, pp. 1–5, Jan. 2016.
[61] C. Lantuejoul and S. Beucher, "On the use of the geodesic metric in image analysis," *J. Microsc.*, vol. 121, no. 1, pp. 39–49, Jan. 1981.
[62] S. Beucher, "Watersheds of functions and picture segmentation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 1982, pp. 1928–1931.
[63] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," in *Mathematical Morphology in Image Processing*. Boca Raton, FL, USA: CRC Press, 2018, pp. 433–481.
[64] J. Serra, "Morphological filtering: An overview," *Signal Process.*, vol. 38, no. 1, pp. 3–11, Jul. 1994.
[65] H. Meddeber and B. Yagoubi, "A new parallel method for medical image segmentation using watershed algorithm and an improved gradient vector flow," in *Proc. Int. Conf. Eur. Middle East North Afr. Inf. Syst. Technol. Support Learn.* Cham, Switzerland: Springer, 2018, pp. 641–651.
[66] A. Likas, N. Vlassis, and J. J. Verbeek, "The global K-means clustering algorithm," *Pattern Recognit.*, vol. 36, no. 2, pp. 451–461, Feb. 2003.
[67] Z. Zhang, J. Zhang, and H. Xue, "Improved K-means clustering algorithm," in *Proc. Congr. Image Signal Process.*, May 2008, pp. 169–172.
[68] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy *c*-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, Jan. 1984.
[69] G. Latif, J. Alghazo, F. N. Sibai, D. N. F. A. Iskandar, and A. H. Khan, "Recent advancements in fuzzy C-means based techniques for brain MRI segmentation," *Current Med. Imag.*, vol. 17, no. 8, pp. 917–930, Aug. 2021.
[70] D.-Q. Zhang and S.-C. Chen, "A novel kernelized fuzzy C-means algorithm with application in medical image segmentation," *Artif. Intell. Med.*, vol. 32, no. 1, pp. 37–50, Sep. 2004.
[71] L. Zhu, F.-L. Chung, and S. Wang, "Generalized fuzzy c-means clustering algorithm with improved fuzzy partitions," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 3, pp. 578–591, Jun. 2009.
[72] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognit.*, vol. 40, no. 3, pp. 825–838, Mar. 2007.
[73] L. Szilagyi, Z. Benyo, S. M. Szilagyi, and H. S. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2003, pp. 724–726.
[74] P. M. Patel, B. N. Shah, and V. Shah, "Image segmentation using K-mean clustering for finding tumor in medical application," *Int. J. Comput. Trends Technol.*, vol. 4, pp. 1239–1242, May 2013.
[75] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using K-means clustering algorithm and subtractive clustering algorithm," *Proc. Comput. Sci.*, vol. 54, pp. 764–771, Jan. 2015.
[76] X. Zheng, Q. Lei, R. Yao, Y. Gong, and Q. Yin, "Image segmentation based on adaptive K-means algorithm," *EURASIP J. Image Video Process.*, vol. 2018, no. 1, pp. 1–10, Dec. 2018.
[77] S. Dalmiya, A. Dasgupta, and S. Kanti Datta, "Application of wavelet based K-means algorithm in mammogram segmentation," *Int. J. Comput. Appl.*, vol. 52, no. 15, pp. 15–19, Aug. 2012.
[78] A. S. Bin Samma and R. A. Salam, "Adaptation of K-means algorithm for image segmentation," *World Acad. Sci. Eng. Technol.*, vol. 50, pp. 58–62, Oct. 2009.
[79] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.
[80] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
[81] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[82] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
[83] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected CRFs," 2014, *arXiv:1412.7062*.
[84] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*.
[85] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder–decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 801–818.
[86] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2881–2890.

[87] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, ''StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8789–8797.

[88] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, ''Unpaired image-to-image translation using cycle-consistent adversarial networks,'' in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.

[89] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, ''Generative adversarial text to image synthesis,'' in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1060–1069.

[90] S. Karim, G. Tong, J. Li, A. Qadir, U. Farooq, and Y. Yu, ''Current advances and future perspectives of image fusion: A comprehensive review,'' *Inf. Fusion*, vol. 90, pp. 185–217, Feb. 2023.

[91] J. Zhu, G. Yang, and P. Lio, ''How can we make GAN perform better in single medical image super-resolution? A lesion focused multi-scale approach,'' in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1669–1673.

[92] M. M. R. Siddiquee, Z. Zhou, N. Tajbakhsh, R. Feng, M. Gotway, Y. Bengio, and J. Liang, ''Learning fixed points in generative adversarial networks: From image-to-image translation to disease detection and localization,'' in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 191–200.

[93] Y. Xue, T. Xu, H. Zhang, L. R. Long, and X. Huang, ''SegAN: Adversarial network with multi-scale $L_1$ loss for medical image segmentation,'' *Neuroinformatics*, vol. 16, no. 3, pp. 383–392, 2018.

[94] A. Radford, L. Metz, and S. Chintala, ''Unsupervised representation learning with deep convolutional generative adversarial networks,'' 2015, *arXiv:1511.06434*.

[95] E. L. Denton, S. Chintala, and R. Fergus, ''Deep generative image models using a Laplacian pyramid of adversarial networks,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 1–9.

[96] T. Karras, T. Aila, S. Laine, and J. Lehtinen, ''Progressive growing of GANs for improved quality, stability, and variation,'' 2017, *arXiv:1710.10196*.

[97] Z. Hu, H. Xue, Q. Zhang, J. Gao, N. Zhang, S. Zou, Y. Teng, X. Liu, Y. Yang, D. Liang, X. Zhu, and H. Zheng, ''DPIR-Net: Direct PET image reconstruction based on the Wasserstein generative adversarial network,'' *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 1, pp. 35–43, Jan. 2021.

[98] M. Mirza and S. Osindero, ''Conditional generative adversarial nets,'' 2014, *arXiv:1411.1784*.

[99] M. Zhang and X. Hu, ''Translation-aware semantic segmentation via conditional least-square generative adversarial networks,'' *J. Appl. Remote Sens.*, vol. 11, no. 4, p. 42622, Dec. 2017.

[100] A. Odena, C. Olah, and J. Shlens, ''Conditional image synthesis with auxiliary classifier GANs,'' in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2642–2651.

[101] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, ''InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.

[102] M. Arjovsky, S. Chintala, and L. Bottou, ''Wasserstein generative adversarial networks,'' in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 214–223.

[103] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, ''Improved training of Wasserstein GANs,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[104] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, ''Least squares generative adversarial networks,'' in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2794–2802.

[105] J. H. Lim and J. C. Ye, ''Geometric GAN,'' 2017, *arXiv:1705.02894*.

[106] M. S. Smirnov, T. R. Garrett, and R. Yasuda, ''An open-source tool for analysis and automatic identification of dendritic spines using machine learning,'' *PLoS One*, vol. 13, no. 7, Jul. 2018, Art. no. e0199589.

[107] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, ''A deep learning model based on concatenation approach for the diagnosis of brain tumor,'' *IEEE Access*, vol. 8, pp. 55135–55144, 2020, doi: 10.1109/ACCESS.2020.2978629.

[108] S. Alqazzaz, X. Sun, X. Yang, and L. Nokes, ''Automated brain tumor segmentation on multi-modal MR image using SegNet,'' *Comput. Vis. Media*, vol. 5, no. 2, pp. 209–219, Jun. 2019.

[109] L. C. Chen, G. Papandreou, and I. Kokkinos, ''DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Jun. 2017.

[110] P. Krähenbühl and V. Koltun, ''Efficient inference in fully connected CRFs with Gaussian edge potentials,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 24, 2011, pp. 1–9.

[111] J. Czajkowska, P. Badura, S. Korzekwa, and A. Płatkowska-Szczerek, ''Automated segmentation of epidermis in high-frequency ultrasound of pathological skin using a cascade of DeepLab v3+ networks and fuzzy connectedness,'' *Computerized Med. Imag. Graph.*, vol. 95, Jan. 2022, Art. no. 102023.

[112] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, and D. Rueckert, ''Attention gated networks: Learning to leverage salient regions in medical images,'' *Med. Image Anal.*, vol. 53, pp. 197–207, Apr. 2019.

[113] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y Hammerla, B. Kainz, B. Glocker, and D. Rueckert, ''Attention U-Net: Learning where to look for the pancreas,'' 2018, *arXiv:1804.03999*.

[114] J. Hu, L. Shen, and G. Sun, ''Squeeze-and-excitation networks,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

[115] M. Yeung, L. Rundo, E. Sala, C.-B. Schönlieb, and G. Yang, ''Focal attention networks: Optimising attention for biomedical image segmentation,'' in *Proc. IEEE 19th Int. Symp. Biomed. Imag. (ISBI)*, Mar. 2022, pp. 1–5.

[116] N. Ilyas, Y. Song, A. Raja, and B. Lee, ''Hybrid-DANet: An encoder–decoder based hybrid weights alignment with multi-dilated attention network for automatic brain tumor segmentation,'' *IEEE Access*, vol. 10, pp. 122658–122669, 2022.

[117] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, ''CBAM: Convolutional block attention module,'' in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.

[118] J. Cheng, S. Tian, L. Yu, C. Gao, X. Kang, X. Ma, W. Wu, S. Liu, and H. Lu, ''ResGANet: Residual group attention network for medical image classification and segmentation,'' *Med. Image Anal.*, vol. 76, Feb. 2022, Art. no. 102313.

[119] J. M. J. Valanarasu, P. Oza, I. Hacihaliloglu, and V. M. Patel, ''Medical Transformer: Gated axial-attention for medical image segmentation,'' in *Proc. 24th Int. Conf. Med. Image Comput. Comput. Assist. Intervent.*, Strasbourg, France. Cham, Switzerland: Springer, 2021, pp. 36–46.

[120] A. Amer, T. Lambrou, and X. Ye, ''MDA-Unet: A multi-scale dilated attention U-Net for medical image segmentation,'' *Appl. Sci.*, vol. 12, no. 7, p. 3676, Apr. 2022.

[121] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, ''Attention gate ResU-Net for automatic MRI brain tumor segmentation,'' *IEEE Access*, vol. 8, pp. 58533–58545, 2020.

[122] J. Chen, C. Ma, Y. Li, S. Fan, R. Shi, and X. Yan, ''HAU-Net: Hybrid attention U-Net for retinal blood vessels image segmentation,'' *Proc. SPIE*, vol. 12782, Aug. 2023, Art. no. 127821H.

[123] F. Zhang, L. Wu, Y. Wang, Y. Yang, M. Li, J. Li, and Y. Xu, ''A multi-scale brain tumor segmentation method based on U-Net network,'' *J. Phys., Conf. Ser.*, vol. 2289, no. 1, Jun. 2022, Art. no. 012028.

[124] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, ''VoxelMorph: A learning framework for deformable medical image registration,'' *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019.

[125] 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 *Proc. 2nd Int. Workshop Brainlesion, Glioma, Multiple Sclerosis, Stroke Traumatic Brain Injuries*, Athens, Greece. Cham, Switzerland: Springer, Oct. 2016, pp. 75–87.

[126] Z. Ullah, M. Usman, M. Jeon, and J. Gwak, ''Cascade multiscale residual attention CNNs with adaptive ROI for automatic brain tumor segmentation,'' *Inf. Sci.*, vol. 608, pp. 1541–1556, Aug. 2022.

[127] Z. Wang, Y. Zou, and P. X. Liu, ''Hybrid dilation and attention residual U-Net for medical image segmentation,'' *Comput. Biol. Med.*, vol. 134, Jul. 2021, Art. no. 104449.

[128] A. R. Choudhury, R. Vanguri, S. R. Jambawalikar, and P. Kumar, ''Segmentation of brain tumors using DeepLabv3+,'' in *Proc. 4th Int. Workshop Brainlesion, Glioma, Multiple Sclerosis, Stroke Traumatic Brain Injuries*, Granada, Spain. Cham, Switzerland: Springer, Sep. 2018, pp. 154–167.

[129] F. K. Shoushtari, S. Sina, and A. N. V. Dehkordi, ''Automatic segmentation of glioblastoma multiform brain tumor in MRI images: Using Deeplabv3+ with pre-trained ResNet18 weights,'' *Phys. Medica*, vol. 100, pp. 51–63, Aug. 2022.

[130] N. M. AboElenein, S. Piao, and Z. Zhang, ''Encoder–decoder network with depthwise atrous spatial pyramid pooling for automatic brain tumor segmentation,'' *Neural Process. Lett.*, vol. 55, no. 2, pp. 1697–1713, Apr. 2023.
[131] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, ''Swin Transformer: Hierarchical vision transformer using shifted windows,'' in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 10012–10022.
[132] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, ''Swin-Unet: Unet-like pure transformer for medical image segmentation,'' in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2022, pp. 205–218.
[133] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, ''Swin UNETR: Swin Transformers for semantic segmentation of brain tumors in MRI images,'' in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2021, pp. 272–284.
[134] B. Chen, Y. Liu, Z. Zhang, G. Lu, and A. W. K. Kong, ''TransAttUnet: Multi-level attention-guided U-Net with transformer for medical image segmentation,'' *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 8, no. 1, pp. 55–68, Feb. 2024.
[135] P. Wang, S. Liu, and J. Peng, ''AST-Net: Lightweight hybrid transformer for multimodal brain tumor segmentation,'' in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 4623–4629.
[136] C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda, S. Muramatsu, Y. Furukawa, G. Mauri, and H. Nakayama, ''GAN-based synthetic brain MR image generation,'' in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 734–738.
[137] Y. Lei, R. L. J. Qiu, T. Wang, W. J. Curran, T. Liu, and X. Yang, ''Generative adversarial networks for medical image synthesis,'' in *Biomedical Image Synthesis and Simulation*. Amsterdam, The Netherlands: Elsevier, 2022, pp. 105–128.
[138] Y. Zheng, B. Wang, and Q. Hong, ''UGAN: Semi-supervised medical image segmentation using generative adversarial network,'' in *Proc. 15th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Nov. 2022, pp. 1–6.
[139] R. Sille, T. Choudhury, A. Sharma, P. Chauhan, R. Tomar, and D. Sharma, ''A novel generative adversarial network-based approach for automated brain tumour segmentation,'' *Medicina*, vol. 59, no. 1, p. 119, Jan. 2023.
[140] Y. Ding, C. Zhang, M. Cao, Y. Wang, D. Chen, N. Zhang, and Z. Qin, ''ToStaGAN: An end-to-end two-stage generative adversarial network for brain tumor segmentation,'' *Neurocomputing*, vol. 462, pp. 141–153, Oct. 2021.
[141] S. Cui, M. Wei, C. Liu, and J. Jiang, ''GAN-segNet: A deep generative adversarial segmentation network for brain tumor semantic segmentation,'' *Int. J. Imag. Syst. Technol.*, vol. 32, no. 3, pp. 857–868, May 2022.
[142] W. E. L. Grimson, R. Kikinis, F. A. Jolesz, and P. M. Black, ''Image-guided surgery,'' *Sci. Amer.*, vol. 280, no. 6, pp. 62–69, 1999.
[143] S. A. D. Franca, W. M. Tavares, A. S. M. Salinet, M. J. Teixeira, and W. S. Paiva, ''Laser interstitial thermal therapy as an adjunct therapy in brain tumors: A meta-analysis and comparison with stereotactic radiotherapy,'' *Surgical Neurol. Int.*, vol. 11, p. 360, Oct. 2020.
[144] D. G. Barone, T. A. Lawrie, and M. G. Hart, ''Image guided surgery for the resection of brain tumours,'' *Cochrane Database Systematic Rev.*, vol. 1, pp. 1–37, 2014.
[145] J. J. Graber, C. S. Cobbs, and J. J. Olson, ''Congress of neurological surgeons systematic review and evidence-based guidelines on the use of stereotactic radiosurgery in the treatment of adults with metastatic brain tumors,'' *Neurosurgery*, vol. 84, no. 3, pp. 168–170, 2019.
[146] S. S. Cho, R. Salinas, and J. Y. K. Lee, ''Indocyanine-green for fluorescence-guided surgery of brain tumors: Evidence, techniques, and practical experience,'' *Frontiers Surgery*, vol. 6, p. 11, Mar. 2019.
[147] R. Sun, H. Cuthbert, and C. Watts, ''Fluorescence-guided surgery in the surgical treatment of gliomas: Past, present and future,'' *Cancers*, vol. 13, no. 14, p. 3508, Jul. 2021.
[148] C. M. Rogers, P. S. Jones, and J. S. Weinberg, ''Intraoperative MRI for brain tumors,'' *J. Neuro-Oncol.*, vol. 151, no. 3, pp. 479–490, Feb. 2021.
[149] G. Fang, M. C. K. Chow, J. D. L. Ho, Z. He, K. Wang, T. C. Ng, J. K. H. Tsoi, P.-L. Chan, H.-C. Chang, D. T.-M. Chan, Y.-H. Liu, F. C. Holsinger, J. Y.-K. Chan, and K.-W. Kwok, ''Soft robotic manipulator for intraoperative MRI-guided transoral laser microsurgery,'' *Sci. Robot.*, vol. 6, no. 57, Aug. 2021, Art. no. eabg5575.
[150] A. Corroyer-Dulmont, S. Valable, J. Fantin, L. Chatre, J. Toutain, S. Teulier, C. Bazille, E. Letissier, J. Levallet, D. Divoux, M. Ibazizène, S. Guillouet, C. Perrio, L. Barré, S. Serres, N. R. Sibson, F. Chapon, G. Levallet, and M. Bernaudin, ''Multimodal evaluation of hypoxia in brain metastases of lung cancer and interest of hypoxia image-guided radiotherapy,'' *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, May 2021.
[151] S. Rudra, N. Jiang, S. A. Rosenberg, J. R. Olsen, M. C. Roach, L. Wan, L. Portelance, E. A. Mellon, A. Bruynzeel, F. Lagerwaard, M. F. Bassetti, P. J. Parikh, and P. P. Lee, ''Using adaptive magnetic resonance image-guided radiation therapy for treatment of inoperable pancreatic cancer,'' *Cancer Med.*, vol. 8, no. 5, pp. 2123–2132, May 2019.
[152] A. Duffton, W. Li, and E. Forde, ''The pivotal role of the therapeutic radiographer/radiation therapist in image-guided radiotherapy research and development,'' *Clin. Oncol.*, vol. 32, no. 12, pp. 852–860, Dec. 2020.
[153] M. Satoh, T. Nakajima, T. Yamaguchi, E. Watanabe, and K. Kawai, ''Evaluation of augmented-reality based navigation for brain tumor surgery,'' *J. Clin. Neurosci.*, vol. 94, pp. 305–314, Dec. 2021.
[154] W.-X. Si, X.-Y. Liao, Y.-L. Qian, H.-T. Sun, X.-D. Chen, Q. Wang, and P. A. Heng, ''Assessing performance of augmented reality-based neurosurgical training,'' *Vis. Comput. Ind., Biomed., Art*, vol. 2, no. 1, pp. 1–10, Dec. 2019.
[155] N. Montemurro, S. Condino, M. Carbone, N. Cattari, R. D'Amato, F. Cutolo, and V. Ferrari, ''Brain tumor and augmented reality: New technologies for the future,'' *Int. J. Environ. Res. Public Health*, vol. 19, no. 10, p. 6347, May 2022.
[156] C. Lee and G. K. C. Wong, ''Virtual reality and augmented reality in the management of intracranial tumors: A review,'' *J. Clin. Neurosci.*, vol. 62, pp. 14–20, Apr. 2019.
[157] G. Unsgård and F. Lindseth, ''3D ultrasound-guided resection of low-grade gliomas: Principles and clinical examples,'' *Neurosurgical Focus*, vol. 47, no. 6, p. 9, Dec. 2019.
[158] K.-T. Chen, W.-Y. Chai, Y.-J. Lin, C.-J. Lin, P.-Y. Chen, H.-C. Tsai, C.-Y. Huang, J. S. Kuo, H.-L. Liu, and K.-C. Wei, ''Neuronavigation-guided focused ultrasound for transcranial blood-brain barrier opening and immunostimulation in brain tumors,'' *Sci. Adv.*, vol. 7, no. 6, Feb. 2021, Art. no. eabd0772.
[159] F. Prada, M. Y. S. Kalani, K. Yagmurlu, P. Norat, M. Del Bene, F. DiMeco, and N. F. Kassell, ''Applications of focused ultrasound in cerebrovascular diseases and brain tumors,'' *Neurotherapeutics*, vol. 16, no. 1, pp. 67–87, Jan. 2019.
[160] D. Aleo, Z. Elshaer, A. Pfnür, P. J. Schuler, M. M. Fontanella, C. R. Wirtz, A. Pala, and J. Coburger, ''Evaluation of a navigated 3D ultrasound integration for brain tumor surgery: First results of an ongoing prospective study,'' *Current Oncol.*, vol. 29, no. 9, pp. 6594–6609, Sep. 2022.
[161] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, ''Brain tumor segmentation using convolutional neural networks in MRI images,'' *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016.
[162] T. Magadza and S. Viriri, ''Deep learning for brain tumor segmentation: A survey of state-of-the-art,'' *J. Imag.*, vol. 7, no. 2, p. 19, Jan. 2021.
[163] F. Ullah, M. Nadeem, M. Abrar, F. Amin, A. Salam, A. Alabrah, and H. AlSalman, ''Evolutionary model for brain cancer-grading and classification,'' *IEEE Access*, vol. 11, pp. 126182–126194, 2023.
[164] U. Sehar and M. L. Naseem, ''How deep learning is empowering semantic segmentation: Traditional and deep learning techniques for semantic segmentation: A comparison,'' *Multimedia Tools Appl.*, vol. 81, no. 21, pp. 30519–30544, Sep. 2022.
[165] N. Sauwen, M. Acou, S. van Cauter, D. M. Sima, J. Veraart, F. Maes, U. Himmelreich, E. Achten, and S. van Huffel, ''Comparison of unsupervised classification methods for brain tumor segmentation using multi-parametric MRI,'' *NeuroImage, Clin.*, vol. 12, pp. 753–764, Feb. 2016.
[166] B. H. Menze et al., ''The multimodal brain tumor image segmentation benchmark (BRATS),'' *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Dec. 2014.
[167] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, ''Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,'' *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.
[168] B. Kayalibay, G. Jensen, and P. van der Smagt, ''CNN-based segmentation of medical imaging data,'' 2017, *arXiv:1701.03056*.

[169] 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 *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2017, pp. 287–297.

[170] Y. Qin, K. Kamnitsas, S. Ancha, J. Nanavati, G. Cottrell, A. Criminisi, and A. Nori, ‘‘Autofocus layer for semantic segmentation,’’ in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2018, pp. 603–611.

[171] A. Myronenko, ‘‘3D MRI brain tumor segmentation using autoencoder regularization,’’ in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2018, pp. 311–320.

[172] Z. Zhou, V. Sodha, M. M. R. Siddiquee, R. Feng, N. Tajbakhsh, M. B. Gotway, and J. Liang, ‘‘Models genesis: Generic autodidactic models for 3D medical image analysis,’’ in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2019, pp. 384–393.

[173] P. G. Rajan and C. Sundar, ‘‘Brain tumor detection and segmentation by intensity adjustment,’’ *J. Med. Syst.*, vol. 43, no. 8, pp. 1–13, Aug. 2019.

[174] M. M. Thaha, K. Kumar, B. S. Murugan, S. Dhanasekeran, P. Vijayakarthick, and A. S. Selvi, ‘‘Brain tumor segmentation using convolutional neural networks in MRI images,’’ *J. Med. Syst.*, vol. 43, no. 9, pp. 1–10, 2019.

[175] L. Sun, S. Zhang, H. Chen, and L. Luo, ‘‘Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning,’’ *Frontiers Neurosci.*, vol. 13, p. 810, Aug. 2019.

[176] Y.-X. Zhao, Y.-M. Zhang, and C.-L. Liu, ‘‘Bag of tricks for 3D MRI brain tumor segmentation,’’ in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2019, pp. 210–220.

[177] C. Zhou, C. Ding, X. Wang, Z. Lu, and D. Tao, ‘‘One-pass multi-task networks with cross-task guided attention for brain tumor segmentation,’’ *IEEE Trans. Image Process.*, vol. 29, pp. 4516–4529, 2020.

[178] A. Sinha and J. Dolz, ‘‘Multi-scale self-guided attention for medical image segmentation,’’ *IEEE J. Biomed. Health Informat.*, vol. 25, no. 1, pp. 121–130, Jan. 2021.

[179] F. Haghighi, M. R. H. Taher, Z. Zhou, M. B. Gotway, and J. Liang, ‘‘Learning semantics-enriched representation via self-discovery, self-classification, and self-restoration,’’ in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 137–147.

[180] M. Sharif, J. Amin, M. Raza, M. A. Anjum, H. Afzal, and S. A. Shad, ‘‘Brain tumor detection based on extreme learning,’’ *Neural Comput. Appl.*, vol. 32, no. 20, pp. 15975–15987, Oct. 2020.

[181] R. A. Zeineldin, M. E. Karar, J. Coburger, C. R. Wirtz, and O. Burgert, ‘‘DeepSeg: Deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images,’’ *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 6, pp. 909–920, Jun. 2020.

[182] M. A. Naser and M. J. Deen, ‘‘Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images,’’ *Comput. Biol. Med.*, vol. 121, Jun. 2020, Art. no. 103758.

[183] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, ‘‘nnU-Net for brain tumor segmentation,’’ in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2020, pp. 118–132.

[184] F. Wang, R. Jiang, L. Zheng, C. Meng, and B. Biswal, ‘‘3D U-Net based brain tumor segmentation and survival days prediction,’’ in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2019, pp. 131–141.

[185] S. Li, X. Sui, X. Luo, X. Xu, Y. Liu, and R. Goh, ‘‘Medical image segmentation using squeeze-and-expansion transformers,’’ 2021, *arXiv:2105.09511*.

[186] W. Wang, C. Chen, M. Ding, H. Yu, S. Zha, and J. Li, ‘‘TransBTS: Multimodal brain tumor segmentation using transformer,’’ in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2021, pp. 109–119.

[187] R. Ranjbarzadeh, A. B. Kasgari, S. J. Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, ‘‘Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images,’’ *Sci. Rep.*, vol. 11, no. 1, pp. 1–17, 2021.

[188] D. Zhang, G. Huang, Q. Zhang, J. Han, J. Han, and Y. Yu, ‘‘Cross-modality deep feature learning for brain tumor segmentation,’’ *Pattern Recognit.*, vol. 110, Feb. 2021, Art. no. 107562.

[189] P. Wang and A. C. S. Chung, ‘‘Relax and focus on brain tumor segmentation,’’ *Med. Image Anal.*, vol. 75, Jan. 2022, Art. no. 102259.

[190] L. Fang and X. Wang, ‘‘Brain tumor segmentation based on the dual-path network of multi-modal MRI images,’’ *Pattern Recognit.*, vol. 124, Apr. 2022, Art. no. 108434.

[191] Y. Jiang, Y. Zhang, X. Lin, J. Dong, T. Cheng, and J. Liang, ‘‘SwinBTS: A method for 3D multimodal brain tumor segmentation using Swin Transformer,’’ *Brain Sci.*, vol. 12, no. 6, p. 797, Jun. 2022.

[192] K. Zou, X. Yuan, X. Shen, M. Wang, and H. Fu, ‘‘TBraTS: Trusted brain tumor segmentation,’’ in *Proc. 25th Int. Conf. Med. Image Comput. Comput. Assist. Intervent.*, Singapore. Cham, Switzerland: Springer, 2022, pp. 503–513.

[193] Q. Yang, X. Guo, Z. Chen, P. Y. M. Woo, and Y. Yuan, ‘‘D2-Net: Dual disentanglement network for brain tumor segmentation with missing modalities,’’ *IEEE Trans. Med. Imag.*, vol. 41, no. 10, pp. 2953–2964, Oct. 2022.

[194] Z. Zhu, X. He, G. Qi, Y. Li, B. Cong, and Y. Liu, ‘‘Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI,’’ *Inf. Fusion*, vol. 91, pp. 376–387, Mar. 2023.

[195] Y. Chang, Z. Zheng, Y. Sun, M. Zhao, Y. Lu, and Y. Zhang, ‘‘DPAFNet: A residual dual-path attention-fusion convolutional neural network for multimodal brain tumor segmentation,’’ *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104037.

[196] F. Behrad and M. S. Abadeh, ‘‘Evolutionary convolutional neural network for efficient brain tumor segmentation and overall survival prediction,’’ *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 118996.

[197] J. Lin, J. Lin, C. Lu, H. Chen, H. Lin, B. Zhao, Z. Shi, B. Qiu, X. Pan, Z. Xu, B. Huang, C. Liang, G. Han, Z. Liu, and C. Han, ‘‘CKD-TransBTS: Clinical knowledge-driven hybrid transformer with modality-correlated cross-attention for brain tumor segmentation,’’ *IEEE Trans. Med. Imag.*, vol. 42, no. 8, pp. 2451–24561, Aug. 2023.

[198] F. Ullah, M. Nadeem, M. Abrar, F. Amin, A. Salam, and S. Khan, ‘‘Enhancing brain tumor segmentation accuracy through scalable federated learning with advanced data privacy and security measures,’’ *Mathematics*, vol. 11, no. 19, p. 4189, Oct. 2023.

[199] F. Ullah, M. Nadeem, M. Abrar, M. Al-Razgan, T. Alfakih, F. Amin, and A. Salam, ‘‘Brain tumor segmentation from MRI images using handcrafted convolutional neural network,’’ *Diagnostics*, vol. 13, no. 16, p. 2650, Aug. 2023.

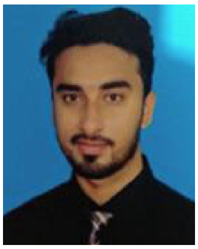
[200] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, ‘‘3D U-Net: Learning dense volumetric segmentation from sparse annotation,’’ in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2016, pp. 424–432.

**SHAHID KARIM** received the Ph.D. degree from the Department of Information and Communication Engineering, Harbin Institute of Technology, Harbin, China, in 2019. He is currently a Postdoctoral Fellow with Northwestern Polytechnical University, China. His current research interests include image processing, artificial intelligence, computer vision, and multidimensional imaging systems.

**GENG TONG** is currently pursuing the Ph.D. degree with the College of Mechanical Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi, China. His current research interests include image processing, machine learning, and deep learning.

**YITING YU** received the Ph.D. degree from Northwestern Polytechnical University, Xi'an, China, in 2010. From 2010 to 2012, he was a Postdoctoral Fellow with the University of Freiburg, Germany. In 2012, he joined Northwestern Polytechnical University, where he is currently a Professor with the School of Mechatronic Engineering. His research interests include MEMS devices, spectral imaging, polarization imaging, image processing, and machine vision.

**ASIF ALI LAGHARI** received the Ph.D. degree from the School of Computer Science and Technology, Harbin Institute of Technology. He is currently an Assistant Professor with Sindh Madressatul Islam University, Karachi, Pakistan, and Shenyang Normal University. He has published more than 100 technical papers in scientific journals and conference proceedings. His current research interests include machine learning, image and video analysis, cloud computing, the IoT, and multimedia QoE management.



**ABDULLAH AYUB KHAN** received the Ph.D. degree from the Department of Computer Science, Sindh Madressatul Islam University, Karachi. He has published around 40 research articles in well-reputed journals/publishers, such as IEEE ACCESS, IEEE TRANSACTIONS, MDPI, Elsevier, Spring, Wiley, and Hindawi, in the domain of digital forensics, blockchain, hyperledger technology, cloud computing, and artificial intelligence.

**MUHAMMAD IBRAR** is currently a Faculty Member with the Software College, Shenyang Normal University, China, who delves into the exciting fields of transportation, management information systems (MIS), emerging technologies, health care management, and artificial intelligence.

**FAISAL MEHMOOD** received the Ph.D. degree in control theory and control engineering from the Department of Automation, School of Electronics and Computer Science, University of Science and Technology of China (USTC), China, in 2019. He is currently a Postdoctoral Fellow with Peking University Shenzhen Graduate School, China. His current research interests include control theory, computer vision, and image processing.