

Deep Learning Based Methods for Brain Tumor Segmentation In MRI: A Review

Mohammed Mikael Rashid, Shaurya Singh , Arbaz Ahmed

Department of Computer Science, Sharda University, Greater Noida, India

2022526514.mohammedmikael@ug.sharda.ac.in, 2022006789.shaurya@ug.sharda.ac.in,

2022006308.arbaz@ug.sharda.ac.in

ABSTRACT

Brain tumor segmentation is an essential task in medical image analysis, since precise identification and depiction of tumors directly assist in the study and treatment planning. Conventional segmentation as well as machine learning based techniques are generally troubled with drawbacks like unsatisfactory generalization, dependency on custom-made features, and susceptibility to imaging artifacts. To overcome these weaknesses, deep learning-based methods have shown to be strong alternatives across various measures in the task of brain tumor segmentation in medical scans. In this review, we examine the latest advancements in deep learning-based approaches, with special emphasis on convolutional neural networks (CNNs), encoder–decoder models, attention mechanisms, and hybrid architectures. Additionally, we explore several benchmark datasets and give a correlative performance of models to give an integrated overview of ultramodern methods. Lastly, we address the problems, such as data imbalance, clarity, and computational cost, as well as possible avenues of future research. The objective of this review is to perform as a starting point and help future research in the area of brain tumor segmentation with deep learning.

Keywords: *convolution, neural networks, deep learning, architectures, segmentation*

1. Introduction

Image segmentation is the process of dividing an image into various distinct regions of interest. In contrast to classification and detection, which are carried out at the object or image level, segmentation involves categorization at the pixel level where each pixel is mapped into a particular, well-defined class. This provides precise representation of boundaries in the image, which is important for its analysis and further following processes [1]. Segmentation plays an important role in the field of medical image analysis because it makes it possible to outline anatomical structures and disease areas with high accuracy. In the context of our discussion, segmentation of MRI brain scans makes it possible to detect tumors, analyze its size and shape, and follow up the development of the disease [2]. Medical Resonance Imaging (MRI) has been the imaging modality most frequently utilized for analysis of brain tumors because it yields excellent soft-tissue contrast and various imaging sequences like T1, T2, FLAIR, and post-contrast scans.

Conventional brain tumor segmentation techniques include thresholding, region growing, clustering, edge-based techniques, and deformable models. Although these techniques offered early advances, they are very sensitive to noise, intensity inhomogeneities, and non-uniform tumor borders frequently found in MRI images. Limited automatic parameter setting and inability to cope with multi-modal 3D data severely restrict their performance [3]. Machine learning (ML) techniques provided advancements through learning from manually designed features, but generalizability across tumor types and datasets was still restricted. ML methods also encounter complication in dealing with complex tumor diversity since features are hand-engineered and cannot potentially represent the in-depth intensity and structure nuances of MRI scans. Deep Learning (DL) methods have reformed the field of medical image segmentation by outperforming conventional and ML-based methods in precision and adaptability [4]. DL models learn selective traits end-to-end from raw input data automatically, reducing the demand

for manual feature engineering. DL networks not only handle large multi-modal datasets but also firmly grab related and spatial interdependencies among pixels. This is especially beneficial for brain tumor segmentation because tumors tend to be irregularly shaped, varied in sizes, and infiltrative into adjacent tissues. Modeling both local details and global information, DL-based methods drastically improve segmentation robustness. Among several models, U-Net and its extensions have been the most widely used, being effective in 2D as well as 3D MRI segmentation. Encoder–decoder architectures, skip connections, and attention mechanisms also enhance these designs by enhancing localization accuracy and boundary refinement [5].

The major contributions of this paper are summarized as follows:

- This paper provides a comprehensive review of deep learning-based methods applied specifically to brain tumor segmentation using MRI scans.
- The paper highlights widely used architectures, significant benchmark datasets, and current challenges encountered in brain tumor segmentation tasks.
- The review aims to facilitate research efforts by summarizing recent progress and outlining possible future research directions in this domain.

The organization of this paper is as follows: Section 2 discusses various DL architectures commonly used for brain tumor segmentation. Section 3 provides a comparative analysis of various DL models on standard benchmarks followed by section 4 containing various datasets used for segmentation of brain tumors. Section 5 contains challenges and corresponding future research directions and section 6 provides a concise summary of recent advancements in the field. Section 7 provides the conclusion of this review.

2. Brain Tumor Segmentation Architectures

There are several architectures that lay the foundation of DL models proficient in Brain tumor segmentation. A few prominent architectures have been discussed in the following subsections and Table 1 contains their summary.

2.1. Encoder-Decoder based Architectures

Pixel-level labeling for image segmentation was first explored using SegNet [6], which employs a fully trained encoder–decoder framework to produce smooth segmentation predictions. The accuracy of segmentation improves with network depth, allowing a broader contextual understanding of MRI features. In SegNet, the encoder compresses the input image into feature maps, while the decoder progressively reconstructs these maps back to the original resolution to delineate tumor regions.

Among encoder–decoder designs, the U-Net architecture is the most widely adopted in brain tumor segmentation tasks. Its U-shaped structure connects encoder and decoder pathways through skip connections, preserving fine spatial information while capturing high-level contextual features. The encoder path extracts meaningful representations from multi-modal MRI scans (T1, T1c, T2, FLAIR) using convolution and pooling operations, while the decoder upsamples these features to produce detailed segmentation maps of tumor subregions such as whole tumor (WT), tumor core (TC), and enhancing tumor (ET). These skip connections bridge corresponding layers of the encoder and decoder, improving boundary precision and overall segmentation quality.

2.2. Hybrid Architectures

The Cascaded 3D U-Net design was introduced in 2018 and is structured in two sequential stages. The first network provides a rough outline of the region of interest, while the second stage refines this result into a more precise segmentation. To enhance computational efficiency, the search area is restricted to specific regions, termed C1 for stage one and C2 for stage two. The final mapping of features to the correct output classes is achieved through a $1 \times 1 \times 1$ convolution layer.

A similar approach can be seen in cascaded deep CNN frameworks applied to brain MRI segmentation [7], where both local structural features and overall contextual information are captured using a two-pathway mechanism. To address the significant class imbalance present in brain tumor labels, a two-step training process is used. When evaluated on the 2013 BraTS dataset, this approach demonstrated nearly 30 times faster performance compared to several contemporary methods.

2.3. Multiscale Architectures

Segmentation with these architectures involves dividing the input MRI scan into multiple homogeneous regions across different scales. This design improves high-level segmentation performance by combining feature information extracted at various resolutions for accurate pixel classification.

U-Net++, presented in 2018, made use of feature maps at a single scale. This enhanced architecture applies Dense Block connections, which allows more effective utilization of intermediate characteristic illustrations. Unlike U-Net, U-Net++ adds convolutional layers between skip pathways connecting the encoder and decoder parts [8]. Such extra layers improve characteristic alignment and refinement, leading to better representation of tumor margins. In addition, the architecture uses deep surveillance, which allows intermediate layers to be utilized in the training process and enhancing overall segmentation accuracy for tumor subregions in MRI scans.

2.4. Attention-based Architectures

The application of attention mechanisms in segmentation improves the ability of networks to identify necessary features and suppress unnecessary information. The mechanisms provide weights to areas of the feature maps so that the model may particularly attend to areas of the tumor in MRI scans.

The Attention U-Net is one of the most used models in this group. It modifies the skip connections of the baseline U-Net by embedding attention gates that screen out non-relevant responses and send only the most important features to the decoder [9]. This enhances segmentation through sharpening of tumor contours and false positive reduction.

Another significant model is the Dual Attention Network. It incorporates both spatial and channel attention modules. Spatial attention marks tumor areas in the image and channel attention enhances the contribution of beneficial feature channels. This combined mechanism enhances the overall representation of brain tumors.

Self-attention models have also been used for brain tumor segmentation. They are able to capture long-range relationships within the image by connecting faraway pixels. This aspect

assists in the detection of anomalous tumor morphology and enhances the consistency of segmentation results over different MRI modalities.

2.5. Transformer-Based Architectures

The self-attention concept was first introduced with transformer-based models. The image is divided into patches which are converted into a sequence of embeddings. Each embedding interacts with others through multi-head attention, allowing the model to capture long-range dependencies across the entire image.

The TransUNet is the most notable architecture in this category. It combines a CNN encoder with a transformer module placed in the bottleneck. The transformer processes the image patches and models global context while the decoder reconstructs the segmentation map [10]. This structure retains local detail while also incorporating global information.

Another important model is the Swin-UNet. It makes use of shifted window attention which restricts computation to smaller windows while still maintaining connections between them [11]. The architecture is hierarchical and processes features at multiple scales, making it suitable for detecting tumors of varying sizes.

Transformers have also been applied to 3D segmentation tasks. UNETR is a representative architecture where the transformer acts as the main encoder. Volumetric patches from the MRI scan are processed directly by transformer blocks and gradually decoded into voxel-wise predictions. This results in accurate tumor segmentation with consistency across slices.

Table 1: Summary of architectures

Architectures	Key Highlights
Encoder-Decoder based Architectures	<ul style="list-style-type: none"> Pixel-level labeling for image segmentation was first explored using SegNet. The U-Net architecture is the most widely adopted in brain tumor segmentation tasks.
Hybrid Architectures	<ul style="list-style-type: none"> The Cascaded 3D U-Net employs two sequential stages. The first network provides a rough outline of the region of interest, while the second stage refines this result into a more precise segmentation Cascaded deep CNN frameworks use a two-step training process to address class imbalance in tumor labels
Multi-scale Architectures	<ul style="list-style-type: none"> These architectures segment the input image by dividing it into multiple homogeneous regions across different scales. U-Net++ utilized feature maps only at a single scale. This enhanced design incorporates Dense Block connections, allowing more efficient use of intermediate feature representations.
Attention-based Architectures	<ul style="list-style-type: none"> Attention mechanisms assign weights to regions of the feature maps, allowing the model to selectively focus on tumor regions in MRI scans. Attention U-Net modifies the skip connections of the original U-Net by placing attention gates that filter irrelevant responses and pass only the most significant features to the decoder.
Transformer-based Architectures	<ul style="list-style-type: none"> Concept of self-attention , the image is divided into patches which are converted into a sequence of embeddings. TransUNet combines a CNN encoder with a transformer module placed in the bottleneck. Swin-UNet makes use of shifted window attention which restricts computation to smaller windows while still maintaining connections between them

3. Comparative Analysis of DL Models for Brain Tumor Segmentation

In Section 2 we examined architectural families and their suitability for MRI-based brain tumor segmentation. This section summarizes how representative implementations of those architectures actually perform on standard benchmarks. Table 2 lists selected models that are widely reported in the literature and gives representative Dice similarity scores for the three BraTS evaluation regions: whole tumor (WT), tumor core (TC) and enhancing tumor (ET). The values below are drawn from publicly available challenge results and peer-reviewed reports to provide a fair, reproducible comparison.

Table 2: Comparison of models

Model	Dataset (split)	Dice (WT)	Dice (TC)	Dice (ET)	Source
3D-UNET (baseline)	BraTS 2018 (validation)	0.871	0.771	0.706	[12]
Two-stage cascaded U-Net	BraTS 2019 (test)	0.88796	0.83697	0.83267	[13]
Attention U-Net	BraTS 2021 (validation)	0.879	0.819	0.793	[14]
U-Net++	BraTS 2019 (validation)	0.8712	0.7817	0.7192	[15]
U-Net	BraTS 2017 (validation)	0.896	0.797	0.732	[16]

The comparative results clearly demonstrate that architectural enhancements such as dense skip connections, attention mechanisms, and hybrid feature extraction strategies significantly improve segmentation accuracy across tumor subregions. U-Net maintains its relevance as a baseline model, while U-Net++ offers refined boundary delineation through feature re-use and deep supervision. Attention U-Net further improves tumor localization by selectively weighting critical features and suppressing irrelevant background information, leading to better delineation of heterogeneous tumor structures. Transformer-based approaches like TransUNet and UNETR, though not included in this table due to dataset-specific benchmarks, have also shown strong potential in modeling long-range dependencies in volumetric MRI data. These advancements indicate a clear progression from conventional convolution-based designs toward architectures integrating attention and global context modeling, pushing brain tumor segmentation closer to clinical viability.

4. Datasets

Brain Tumor segmentation models are trained on high quality datasets. These datasets are required to be well labelled, organized and cross-verified by field experts. The performance of the models will always directly depend on the dataset they are trained on. A few prominent segmentation datasets have been discussed below and summarized as well in Table 3.

4.1 BRISC (2025)

The BRISC dataset consists of 6,000 contrast-enhanced T1-weighted MRI scans annotated by certified radiologists. It includes three key tumor types—glioma, meningioma, pituitary—and non-tumorous cases. Datasets span axial, sagittal, and coronal planes, enabling cross-view

generalization and multi-angle learning. It offers high-resolution, high-quality labels aimed at both segmentation and classification tasks

4.2 TextBraTs (2025)

TextBraTS is the first multimodal dataset combining volumetric MRI scans with paired radiological text reports. Built on the BraTS2020 dataset, it facilitates exploration of text-guided segmentation, enabling models to fuse visual and linguistic data. This innovation supports development of vision-language fusion algorithms for enhanced tumor boundary delineation.

4.3 BraTS-MEN-RT (2024)

The BraTS-MEN-RT challenge introduces a large multi-institutional MRI dataset specifically for meningioma radiotherapy planning. It includes post-contrast T1-weighted 3D MRIs annotated with gross tumor volumes (GTV) and risk-margin labels, covering both intact and postoperative cases. Segmentation is evaluated using lesion-wise Dice and HD95 metrics, enhancing automated radiotherapy preparation.

4.4 RHUH-GBM (2023)

The Río Hortega University Hospital Glioblastoma Dataset (RHUH-GBM) compiles multiparametric MRI scans (including preoperative, early postoperative, and recurrence images), volumetric assessments, molecular data, and survival outcomes in glioblastoma patients. Expert-corrected tumor segmentations accompany the imaging, making it valuable for studying recurrence and post-treatment tumor progression.

4.5 Brain Metastasis 3D Segmentation Dataset (2023)

This dataset provides 3D MRI scans (T1 pre/post-contrast, T2, FLAIR) for 200 brain metastasis patients, along with detailed segmentations of enhancing lesions, necrotic cores, and peritumoral edema. It includes 975 contrast-enhancing lesions, many sub-centimeter in size, paired with clinical metadata to support robust AI model development for metastasis detection and segmentation.

Table 3: Summary of Datasets

Dataset	Year	Images/Patients	Key Highlights
BRISC	2025	6,000 images	It includes three key tumor types—glioma, meningioma, pituitary
TextBraTS	2025	369 individual cases	First multi-modal dataset combining volumetric MRI scans with paired radiological text reports
BraTS-MEN-RT	2024	750 scans	Segmentation is evaluated using lesion-wise Dice and HD95 metrics, enhancing automated radiotherapy preparation
RHUH-GBM	2023	40 patients	Compiles multiparametric MRI scans
Brain Metastasis 3D Segmentation Dataset	2023	200 patients	This dataset provides 3D MRI scans along with detailed segmentations of enhancing lesions, necrotic cores, and peritumoral edema.

5. Challenges and Future Research Directions

While DL has contributed significantly to the advancement of brain tumor segmentation, several challenges remain that open up avenues for future research. This section discusses some of the prominent challenges in DL based segmentation methods and also points out corresponding future research directions.

5.1 Limited Generalization Across Datasets

Most deep learning models show promising results on specific benchmark datasets, yet their performance drops significantly when tested on unseen clinical datasets. Developing robust architectures that generalize well across diverse patient populations, imaging protocols, and tumor types remains an open challenge.

5.2 Scarcity of Annotated Data

Manual annotation of brain tumors in MRI scans is time-consuming and requires expert radiologists, which restricts the availability of large-scale labeled datasets. Future work should explore semi-supervised, weakly supervised, and synthetic data generation approaches to reduce dependency on extensive manual labeling.

5.3 Handling Tumor Variability and Class Imbalance

Brain tumors exhibit high heterogeneity in size, shape, and intensity, which makes consistent segmentation difficult. Moreover, the class imbalance between healthy tissues and tumor regions hampers training. Advanced loss functions and adaptive sampling strategies can help mitigate these issues.

5.4 High Computational and Memory Requirements

State-of-the-art models often demand substantial computational power and memory, making them less suitable for real-time clinical deployment. Future research should focus on lightweight architectures, model compression, and hardware-aware optimization to enable practical applicability.

5.5 Integration of Multi-modal MRI Data

Although multi-modal MRI provides complementary information, effectively combining these modalities for accurate segmentation remains a challenge. Novel fusion strategies and attention-based mechanisms are required to leverage the full potential of multi-modal imaging.

5.6 Clinical Translation and Interpretability

Despite achieving high segmentation accuracy in experimental settings, clinical adoption remains restricted due to the black-box nature of deep learning models and limited trust from medical practitioners. Incorporating Explainable AI (XAI) techniques can enhance transparency by providing interpretable insights into model predictions. Future research should focus on integrating XAI frameworks with segmentation models and validating their reliability through large-scale, multi-institutional clinical studies.

5.7 Federated Learning for Collaborative Model Training

Strict privacy regulations and institutional policies often prevent direct sharing of medical imaging data, limiting access to diverse, multi-institutional MRI datasets. Federated Learning (FL) addresses this challenge by enabling collaborative model training across different centers while keeping patient data localized. However, issues such as non-uniform data distribution, communication overhead, and maintaining performance across heterogeneous datasets remain

unresolved. Future research should explore optimized FL frameworks and privacy-preserving techniques to improve generalization while ensuring compliance with data security standards.

6. Recent Advances

DL has shown immense potential in the domain of brain tumor segmentation. It is a mainstream domain that consistently sees new advancements. Some important recent developments are detailed in this section.

6.1 CNN and Residual/Encoder–Decoder Based Architectures

- Bhardwaj et al. [18] presented a ResUNet50 CS model for better classification and segmentation of brain tumors from MRI scans. The model combines ResNet50's deep residual connections with U-Net's encoder–decoder architecture to improve the feature propagation while retaining the spatial details. Experimental analysis showed that the proposed Viola deep learning method had better performance than traditional deep learning methods with an accuracy of 97% for classification and reliable segmentation results. These results demonstrate the ability of ResUNet50 to manage the complex tumor features and aid accurate and clinically relevant analysis.
- EffUNet++, a deep learning structure framework tailored for accurate brain tumor segmentation in FLAIR MRI images was introduced by Yadav et al. [18]. This model ranges the standard UNet++ model by including Efficient-NetB7 as the encoder and using pre-trained weights of the AdvProp dataset to get improved feature illustration. Skip connections from the decoder allowed multi-scale feature fusion and mask modification. The Kaggle lower-grade gliomas (LGG) dataset displayed outstanding performance in experiments, with a mean IoU of 0.9123 and a Dice coefficient of 0.9387. With these findings, EffUNet++ successfully increases the accuracy of tumour boundary description and segmentation, which makes it perfect for clinical use.

6.2 Attention and Inception Enhanced Architectures

- Rutoh et al. [19] introduced ABI-Net, a variant of U-Net that integrates attention-based inception modules for the segmentation of brain tumours. The beginning blocks pull out spatial features from different scales, and the consideration mechanism highlights areas of interest in multimodal 3D MRI scans that are specific to tumours. The model was trained on the BraTS 2020 dataset and got Dice scores of 0.8354 (ET), 0.8505 (TC), and 0.8782 (WT), which were better than any other method. ABI-Net sustained to do well on the validation set of 125 patients, with scores of 0.8189, 0.8401, and 0.8673. The architecture looks like it could be valuable in the clinic for diagnosis, planning treatment, and keeping an eye on patients.
- Echine and Darouichi [20] introduced a 3D AGSE-Res-UNet architecture specifically engineered for accurate brain tumour segmentation in multimodal MRI scans. The model is founded on Res-UNet and adds three important features: Atrous Spatial Pyramid Pooling (ASPP) for mining features at different scales, Attention Gates to highlight important tumour areas, and Squeeze-and-Excitation modules to improve channel-wise feature representation. The model was examined on the BraTS 2020 dataset and got Dice scores of 0.8262 for the whole tumour (WT), 0.7699 for the tumour core (TC), and 0.7430 for the enhancing tumour (ET). A relative analysis displayed that this method worked better than many other cutting-edge methods. This shows how actual it is to combine attention, multi-scale processing, and channel recalibration in a residual learning framework.

- Fan et al. [21] tackled the problem of incomplete MRI modalities in brain tumour segmentation by suggesting a multimodal reversible feature learning framework. The recommended method uses a reversible feature interaction module to fill in missing modality features by looking for similarities and differences between the presented modalities. To get an accurate segmentation outputs a dual-scale attention module helps in recovering fine details and cutting down on noise during reconstruction. The results of experimenting on BraTS dataset showed that these method works better than other methods, especially when imaging data is missing.

6.3 Hybrid and Multi-Stage Approaches

- A YOLO-driven adaptive deep neural architecture for analysing brain tumours using MRI was suggested by Vinayakan et al. [22]. Using transfer learning, the method improves a cutting-edge YOLO model to discover meningioma, glioma, and pituitary tumours. The network precisely pinpoints tumours and demonstrates good results on small lesions. While processing segmentation on areas that have been found, YOLO detector is a fast front-end that narrows the examination space. The study shows that MRI scans can be constantly detected and located, which is useful for clinical triage and as a pre-segmentation module in a full pipeline.
- Revathi et al. [23] suggested a hybrid method that associates fuzzy C-means (FCM) segmentation with Support Vector Machine (SVM) classification to detect brain tumours from MRI images. The SVM classifier further separates tumour types based on the segmented features, while the FCM algorithm uses unsupervised clustering to separate tumour regions from healthy tissue. An accuracy of 98% was shown in the experimental results, representing the dependability of combining machine learning-based classification with traditional segmentation techniques which enhances diagnostic performance. This study demonstrates how hybrid methods can improve brain tumour analysis prediction accuracy.
- To perform a detailed assessment of deep learning techniques for brain tumour segmentation and MGMT promoter classification Amor et al. [24] used the BraTS 2021 and BraTS 2022 datasets. The study concluded numerous cutting-edge methods, emphasising that hybrid frameworks that combined various deep learning techniques outperformed single-model approaches in terms of segmentation performance. Also, five different methods were evaluated for predicting the methylation status of the MGMT promoter, a crucial biomarker for the treatment planning of glioblastoma. According to the findings, combining segmentation and molecular classification can greatly improve diagnostic precision and facilitate individualised treatment plans.

6.4 Graph and Non-CNN Based Architectures

- A multi-modal graph M2GCNet, a complex network that was designed for exact brain tumour segmentation across several MRI arrangements, was given by Zhou [25]. To successfully extract both local and global related information, the framework presents a Multi-Modal Graph Convolution Module (M2GCM), which examines the image pixels as nodes and hidden relationships as edges. There are two specific parts included in M2GCM: a channel-wise graph convolution module (CGCM) to model inter-channel relationships and a spatial-wise graph convolution module (SGCM) to get spatial dependencies across tumour areas. To leverage nonlinear connections between MRI modalities, a multimodal correlation loss function is also employed. M2GCNet also improved modality fusion and clinical understandability while beating current

techniques in segmentation accuracy, according to the benchmark dataset of experimental evaluation.

6.5 Optimization and Training Strategies

- Ahmad and Chen [26] presented a method that improves brain tumor segmentation in 3D multimodal MRI scans by integrating an optimized Gabor filter. This method improved feature extraction which were frequently overlooked by conventional preprocessing, this was implemented by capturing subtle structural details and spatial variations. The study significantly improved brain tumour segmentation by implementing a dual pre-processing strategy for its deep learning model. A Gabor filter was firstly applied to improve the discriminative power of the input features. Alongside, the multichannel MRI data was removed and rescaled to reduce the computational workload. This approach led to a significant increase in performance, with accuracy growing from 0.982 to 0.991 and the mean Connection over Union (mIoU) jumping from 0.714 to 0.804. These results effectively proved that integrating sophisticated filtering techniques with deep learning can lead to much more accurate tumour enlightenment in volumetric datasets.
- Susanto et al. [27] suggested a progressively trained extremely administered network for brain tumour segmentation which overcome the high computational cost of training large scale deep learning models and this was implemented by breaking the network up into smaller, sequentially trained segments, the method maintains competitive performance while drastically lowering GPU memory requirements by up to 40%. The BraTS 2021 dataset was used to assess the architecture, which combines a ConvNeXt encoder with a UNet3+ decoder. The technique obtained Dice scores of 90.80 for the entire tumour (WT), 87.57 for the tumour core (TC), and 82.39 for the enhancing tumour (ET). Progressive training is ideal for research settings with constrained computational capacity because it enables effective hyperparameter tuning and resource management, despite a slight increase in total training time capacity

7. Conclusions

In this review, we have explored and compared various DL models as well as datasets used for brain tumor segmentation in MRI scans. The applications of AI in healthcare are rapidly increasing and computer aided diagnosis greatly benefits from it. We began our discussion with various segmentation architectures that form the backbone of state-of-the-art models currently in use. However, there are some challenges associated with DL methods that include lack of annotated data, class imbalance, high computational requirements and clinical integration challenges. We also proposed several possible future research directions that include use of federated learning and explainable AI. There are numerous recent advancements in DL methods for brain tumor segmentation. We explored hybrid approaches including YOLO-driven adaptive architecture and combination of FCM with SVM classification. The recent usage of optimization strategies for DL models have also been explored. Deep learning has revolutionized the domain of brain tumor segmentation and has transformed clinical diagnosis and treatment planning.

Funding source

No funding was received for this study.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1]. Shu, X. et al. (2025) 'Variational methods with application to medical image segmentation: A survey', *Neurocomputing*, 639, p. 130260. doi:10.1016/j.neucom.2025.130260.
- [2]. Xia, Q. et al. (2025) 'A comprehensive review of Deep Learning for medical image segmentation', *Neurocomputing*, 613, p. 128740. doi:10.1016/j.neucom.2024.128740.
- [3]. Rayed, Md.E. et al. (2024) 'Deep learning for medical image segmentation: State-of-the-art advancements and challenges', *Informatics in Medicine Unlocked*, 47, p. 101504. doi:10.1016/j.imu.2024.101504
- [4]. S. Solanki, U. P. Singh, S. S. Chouhan and S. Jain, "Brain Tumor Detection and Classification Using Intelligence Techniques: An Overview," in *IEEE Access*, vol. 11, pp. 12870-12886, 2023, doi: 10.1109/ACCESS.2023.3242666.
- [5]. N. Siddique, S. Paheding, C. P. Elkin and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," in *IEEE Access*, vol. 9, pp. 82031-82057, 2021, doi: 10.1109/ACCESS.2021.3086020.
- [6]. V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, 1 Dec. 2017, doi: 10.1109/TPAMI.2016.2644615.
- [7]. S. Hussain, S. M. Anwar and M. Majid, "Brain tumor segmentation using cascaded deep convolutional neural network," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea (South), 2017, pp. 1998-2001, doi: 10.1109/EMBC.2017.8037243.
- [8]. Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh and J. Liang, "UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 1856-1867, June 2020, doi: 10.1109/TMI.2019.2959609.
- [9]. P. Baruah, B. Dutta, P. P. Dutta, H. Pallab Jyoti Dutta, B. Goswami and D. Sarma, "Brain Tumor Segmentation using Attention-based UNet," 2023 IEEE 3rd Applied Signal Processing Conference (ASPCON), India, 2023, pp. 282-286, doi: 10.1109/ASPCON59071.2023.10396058.
- [10]. J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "TransUNet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
- [11]. Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2023). Swin-Unet: Unet-like pure Transformer for medical image segmentation. *Lecture Notes in Computer Science*, 205–218. https://doi.org/10.1007/978-3-031-25066-8_9
- [12]. Mehta, R., & Arbel, T. (2019). 3D U-Net for brain tumour segmentation. *Lecture Notes in Computer Science*, 254–266. https://doi.org/10.1007/978-3-030-11726-9_23

- [13]. Jiang, Z. et al. (2020) 'Two-stage cascaded U-Net: 1st place solution to brats challenge 2019 segmentation task', Lecture Notes in Computer Science, pp. 231–241. doi:10.1007/978-3-030-46640-4_22.
- [14]. Wang, S., Li, L., Zhuang, X. (2022). AttU-NET: Attention U-Net for Brain Tumor Segmentation. In: Crimi, A., Bakas, S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021. Lecture Notes in Computer Science, vol 12963. Springer, Cham. https://doi.org/10.1007/978-3-031-09002-8_27
- [15]. N. Micallef, D. Seychell and C. J. Bajada, "Exploring the U-Net++ Model for Automatic Brain Tumor Segmentation," in IEEE Access, vol. 9, pp. 125523-125539, 2021, doi: 10.1109/ACCESS.2021.3111131.
- [16]. Isensee, F. et al. (2018) 'Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge', Lecture Notes in Computer Science, pp. 287–297. doi:10.1007/978-3-319-75238-9_25.
- [17]. A. Bhardwaj, P. Anand, D. Sharma and U. Gupta, "Improving Brain Tumor Classification and Segmentation Using ResUNet50," 2025 International Conference on Next Generation Information System Engineering (NGISE), Ghaziabad, Delhi (NCR), India, 2025, pp. 1-6, doi: 10.1109/NGISE64126.2025.11085229
- [18]. A. C. Yadav, M. H. Kolekar, Y. Sonawane, G. Kadam, S. Tiwarekar and D. R. Kalbande, "EffUNet++: A Novel Architecture for Brain Tumor Segmentation Using FLAIR MRI Images," in IEEE Access, vol. 12, pp. 152430-152443, 2024, doi: 10.1109/ACCESS.2024.3480271
- [19]. E. K. Rutoh, Q. ZhiGuang, J. C. Bore-Norton and N. Bahadar, "ABI-Net: Attention-Based Inception U-Net for Brain Tumor Segmentation From Multimodal MRI Images," in IEEE Access, vol. 13, pp. 134898-134916, 2025, doi: 10.1109/ACCESS.2025.3585926
- [20]. K. Echine and A. Darouichi, "3D AGSE-Res-UNet: An Efficient Deep Learning Approach based on Res-UNet for Brain Tumor Segmentation," 2024 International Conference on Connected Innovation and Technology (ICCITX), Rabat, Morocco, 2024, pp. 1-6, doi: 10.1109/ICCITX61791.2024.11070871.
- [21]. Y. Fan, L. Liu, X. Luan and W. Li, "Reversible Feature Learning for Brain Tumor Segmentation With Incomplete Modalities," in IEEE Signal Processing Letters, vol. 32, pp. 2419-2423, 2025, doi: 10.1109/LSP.2025.3572403.
- [22]. K. Vinayakan, G. Malleswari and A. S. Reddy, "YOLO-Driven Adaptive Deep Neural Architecture for Automated Brain Tumor Segmentation in MRI Imaging," 2025 3rd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India, 2025, pp. 830-834, doi: 10.1109/ICSSAS66150.2025.11081289.
- [23]. M. Revathi, P. Kalaiarasi, S. Jeevitha and G. Akilandasowmya, "Enhancing the Prediction Accuracy of Brain Tumor Detection by using Fuzzy C-Means Segmentation and Support Vector Machine," 2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS), Bengaluru, India, 2024, pp. 846-851, doi: 10.1109/ICICNIS64247.2024.10823224.
- [24]. F. Amor, H. Mzoughi, I. Njeh and M. B. Slima, "Review of MRI brain tumor segmentation and MGMT promoter classification methods on BraTs dataset based on Deep learning," 2024 IEEE 7th International Conference on Advanced Technologies,

- Signal and Image Processing (ATSIP), Sousse, Tunisia, 2024, pp. 249-254, doi: 10.1109/ATSIP62566.2024.10638990.
- [25]. T. Zhou, "M2GCNet: Multi-Modal Graph Convolution Network for Precise Brain Tumor Segmentation Across Multiple MRI Sequences," in *IEEE Transactions on Image Processing*, vol. 33, pp. 4896-4910, 2024, doi: 10.1109/TIP.2024.3451936
- [26]. N. Ahmad and Y. -T. Chen, "Enhanced Deep Learning Model Performance in 3D Multimodal Brain Tumor Segmentation with Gabor Filter," 2024 10th International Conference on Applied System Innovation (ICASI), Kyoto, Japan, 2024, pp. 406-408, doi: 10.1109/ICASI60819.2024.10547816.
- [27]. E. K. Susanto, H. Tjandrasa and C. Fatichah, "Progressively Trained Deeply Supervised Network for Brain Tumor Segmentation," 2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS), Bandung, Indonesia, 2025, pp. 1-5, doi: 10.1109/ICADEIS65852.2025.10933424