SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 5, (2026), pp. 636-653
Open Access. Online & Print

DOI: 10.5281/zenodo.1250052

U-NET TO TRANSFORMERS: A COMPREHENSIVE REVIEW OF ENCODER-DECODER ARCHITECTURES FOR BRAIN TUMOR SEGMENTATION AND CLASSIFICATION

Palwinder Singh[1], Namit Gupta[2*]

[1] *Research Scholar Teerthanker Mahaveer University, Moradabad, India.*
[2] *Associate Professor Teerthanker Mahaveer University, Moradabad, India.*

Received: 01/12/2025
Accepted: 02/01/2026

Corresponding author: Namit Gupta (namit.k.gupta@gmail.com)

ABSTRACT

Brain tumor segmentation and classification are essential components of computer-aided diagnosis systems in neuro-oncology, supporting early detection, treatment planning, and outcome assessment. However, the complexity of tumor morphology and variability in imaging protocols make automated analysis challenging. In recent years, deep learning has appeared as a dominant approach for addressing these challenges, proving superior performance over traditional machine learning methods. This review provides a comprehensive overview of deep learning techniques for brain tumor segmentation and classification using medical imaging data. It covers key imaging modalities, preprocessing strategies, and highlights the role of U-Net-based encoder-decoder architectures as the foundation of modern segmentation models. Existing methods are systematically categorized based on architectural design, including CNN-based encoder-decoder models, attention-based approaches, multi-scale and cascaded networks, and transformer-based frameworks. The review further analyzes commonly used datasets and evaluation strategies and summarizes the strengths and limitations of current approaches. Finally, it outlines major research challenges and future directions, including model generalization, interpretability, computational efficiency, and clinical applicability.

KEYWORDS: Brain tumor segmentation, Encoder-decoder network, Deep learning, Transformer Multi scale networks, U-Net architecture.

1. INTRODUCTION

Neurons make up a network of cells that provide the human body with information processing and enable responses to internal and external stimulation. The nervous system consists of an array of interconnected neurons and consists of two divisions, the central nervous system (CNS), and the peripheral nervous system (PNS) [1]. The CNS consists of the brain and spinal cord and is protected by the skull and vertebral column. The brain is the main control center of the body and coordinates all cognitive functions, interprets sensory information, and controls voluntary and involuntary motor functions [2].

Brain tumours are abnormal cell growth within the tissues of the brain caused by uncontrolled division of the cells [3]. Brain tumours can be classified as either binine or malignan. Binine tumours are made up of non-cancerous cells that normally grow slowly and remain within one area of the brain [4]. Malignan tumours, on the other hand, grow much faster, have a much higher likelihood of recurring after treatment, and may invade into other portions of the brain. Furthermore, brain tumours can be categorised as either primary or secondary based on their source of origin. Primary tumours originate from scar tissue in the brain and are normally localised [5]. Secondary tumours, on the other hand, originate from other cancers in the body and are spread to the brain via blood circulation [6].

Brain Tumours are regarded as a significant health concern for patients who have been diagnosed with this type of cancer because of the high mortality rate and the complex treatments involved and ultimately, the impact of brain tumours on quality of life [5]. According to GLOBOCAN 2020, brain tumours represent approximately 1.9% (19th most common cancer) of total cancers; in addition, brain tumours account for approximately 2.5% (12th most common cause of cancer death) of total cancer deaths, indicating that brain tumours are extremely aggressive and have a poor prognosis overall [7]. Therefore, early detection and accurate diagnosis of brain tumours are important to improve treatment success rates and reduce mortality rates.

Clinicians primarily use non-invasive imaging techniques, with MRI providing the best soft tissue contrast which enables accurate delineation of the tumour itself as well as surrounding edema [8]. Nevertheless, the manual interpretation of MRI scans involves a process that is extremely time-consuming, labour-intensive and can suffer from variability between observers (inter-observer), impacting on the accuracy of the diagnosis. Due to the rapid proliferation of medical imaging data and therefore huge burden on clinicians, the development of automated diagnostic systems will provide the foundation for accurate, reliable, and timely delivery of results regardless of the clinical situation [9].

CAD systems are utilized to assist the clinician in tumor detection, identification and classification. CAD systems allow for the combination of new imaging technologies and the use of advanced machine learning techniques to provide a greater degree of accuracy in diagnosis while also reducing the workload and subjectivity of the physician.

Recent advances with deep learning methods, particularly CNNs and encoder-decoder architectures, have shown significant advances in the automation, accuracy and efficiency of brain tumor analysis. This article reviews encoder-decoder architectures with an emphasis on the advancements of U-Net based models that provide a significant advancement of the ability to segment and classify brain tumours using images obtained from imaging modalities.

1.1. The contributions listed here are:

- This article is a complete review of the many different encoder-decoder models currently available to brain tumor segmenting and categorizing including the U-Net, V-Net, Attention U-Net, and U-Net++.
- A new taxonomy of encoder-decoder architectures will be presented based on design principles.
- Present a comparative analysis of commonly used datasets e.g., BRATS, TCIA and evaluation metrics.
- Identifies key challenges, including data scarcity, domain adaptation, and model interpretability.
- Discusses emerging trends including hybrid CNN-Transformer networks, lightweight 3D models, and self-supervised learning.

This article will be organized as follows, A brief introduction into the motivation and clinical background of performing brain tumor imaging and categorization will be presented in Section 1, discussion of foundational concepts will be provided in Section 2, including the various types of imaging modality, preprocessing methods, and the U-Net architecture. Existing work will be categorized in Section 3 based on the taxonomy of encoder-decoder architecture, Section 4 will provide information about datasets used, evaluation metrics, and performance comparisons, Section 5 will provide discussion regarding current issues and future research opportunities, while Section 6 will provide a conclusion.

2. BACKGROUND

The section delivers a fundamental understanding of how deep-learning algorithms can be used for segmentation and classification of tumours within an image of a patient's brain with respect to their radiological appearances. The most common imaging modalities that are used within the area of neuro-oncology may have some form of preprocessing applied to them prior to using them for classification and a simple U-Net architecture serve as the foundation for many encoder-decoder-based segmentation frameworks that exist today.

2.1 Imaging Modalities for Brain Tumor Analysis

Brain tumours are diagnosed, characterized, and treated using medical images. Each imaging modality offers its own unique complementary anatomy and function, thereby enhancing automated segmentation and classification.

2.1.1 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) is a common imaging technique used to evaluate brain tumours due to its superior contrast between soft tissues, ability to obtain images in multiple planes, and non-invasive nature [10]. Multiple imaging sequences can be obtained using MRI including T1-weighted images for anatomic structure, contrast-enhanced (T1c) images to define areas of tumor activity, T2-weighted images to visualize area of fluid accumulation, and fluid-attenuated inversion recovery (FLAIR) images to visualize edema around the tumor [11]. Multiple deep learning models combine these sequences into a multi-channel input to increase feature representation and performance of any segmentation method. In addition, MRI is subjected to multiple factors including intensity inhomogeneity, variability between MRI scanners, and noise which creates domain shifts in the data across datasets [12,13].

2.1.2 Computed Tomography (CT)

Computed Tomography (CT) uses X-ray attenuation to produce cross-sectional images of the brain and is typically used for emergency and trauma cases because of its fast acquisition and readiness for use and is useful for detecting the presence of hemorrhage, calcifications, and bone aberrations that may be associated with brain tumours [14].

However, due to lower contrast of soft tissues with CT than with MRI, exact identification of tumor borders is more difficult using CT sequences. As a result, deep-learning techniques that use CT for brain tumor segmentation and classification tend to be used less and are also typically less effective compared to similar CT imaging approaches [15,16].

2.1.3 Positron Emission Tomography (PET)

Positron Emission Tomography is a functional imaging modality that measures metabolic activity in brain tissues through radiotracer uptake. It is useful for tumor grading and distinguishing recurrence from treatment effects such as radiation necrosis [17]. PET is often fused with MRI in deep learning studies to combine metabolic and anatomical information, improving tumor analysis [18,19]. However, PET has low spatial resolution, excessive cost, and radiation exposure. Given below Table 1 summarizes the advantages, limitations, and roles of commonly used imaging modalities.

Table 1: Comparison of Imaging Modalities for Brain Tumor Analysis

Modality	Advantages	Limitations	Mitigation strategies	Role
MRI	High soft-tissue contrast, multi-sequence	Intensity inhomogeneity, noise	Denoising, intensity normalization	Primary modality
CT	Fast, widely available	Poor soft-tissue contrast	Enhancement, artifact reduction	Auxiliary imaging
PET	Functional and metabolic insight	Low resolution, radiation	Reconstruction, smoothing, fusion	Grading

2.2 Preprocessing Techniques

Preprocessing removes noise, artifacts, and irrelevant structures from medical images. It improves data consistency and enhances model performance. Table 2 summarizes the aims and impacts of commonly used preprocessing techniques in brain tumor analysis.

2.2.1 Skull Stripping

Removing tissues outside of the brain i.e., skull, scalp, eye from a brain image through the process of skull stripping is done to remove any irrelevant background information [20]. Automated techniques can isolate brain tissue areas to improve the accuracy of segmentation and model performance. Consistent brain boundaries among modalities are crucial for multimodal MRI analysis [21].

2.2.2 Intensity Normalization

Voxel intensity variations that arise due to the difference in scanners and acquisition protocols can be mitigated through intensity normalization. This is conducted by using normalizing methods, such as z-score normalization, histogram matching, and percentile scaling, to standardize intensity distributions. This will help deep learning models in concentrating on tumor-related characteristics, thereby enhancing model strength and generalizability across the various datasets [22].

2.2.3 Noise Removal

There are many distinct types of noise and artifacts in medical imaging that negatively affect the quality of the medical image. Noise reduction or denoising techniques, like median filtering, anisotropic diffusion, and wavelet-based denoising methods, can be used to remove or reduce the impact of noise and preserve the structural detail of the image leading to more accurate detection of tumor boundaries and fewer false positives when the image is segmented [23].

2.2.4 Data Augmentation

Data augmentation increases dataset size and diversity using transformations such as rotation, flipping, scaling, intensity changes, and elastic deformation. It reduces overfitting and helps address limited and imbalanced medical imaging datasets [24].

Table 2: Role of Preprocessing Techniques in Brain Tumor Analysis

Technique	Objective	Impact
Skull stripping	Remove non-brain tissue	Reduces background bias
Intensity normalization	Standardize intensities	Improves generalization
Noise removal	Suppress artifacts	Enhances boundary detection
Data augmentation	Increase data diversity	Reduces overfitting

2.3 Basic U-Net Architecture / Overview and Motivation of U-Net

The architecture we call U-Net is a type of deep learning model designed to perform segmentation of medical images. Its main goal is to provide correct localisation of pixels at the level of detail and support the use of context-semantics when doing so. U-Net performs efficiently with a small amount of labelled data making it the de facto model for research into the segmentation of brain tumours [25]. The U-Net architecture includes many different types of components, as illustrated in Fig 1. It is important to understand these primary components before trying to understand more complex segmentation methods.

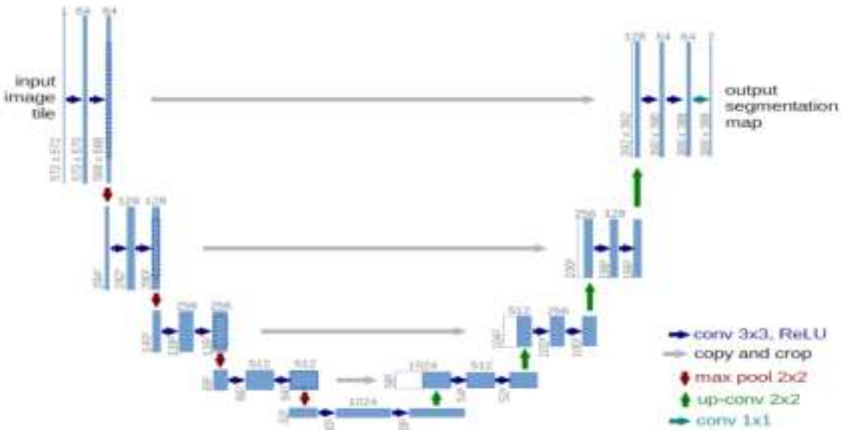

Figure 1: Illustration of the U-Net architecture (32 × 32 pixels at the lowest resolution). Blue boxes correspond to multi-channel feature maps, with channel numbers shown above and spatial dimensions indicated at the lower left. White boxes represent copied feature maps, while arrows denote different operations

2.3.1 Encoder (Contracting Path)

An input image tile enters the U-Net framework and is processed through what is traditionally referred to as an encoder the contracting path, generating hierarchical feature representations at multiple levels of depth. The encoder levels consist of two successive layers of 3 × 3 convolution operations followed by ReLU activations. This allows increasingly discriminative feature levels to be generated through the convolutional operations [26].

With every convolution, we apply a 2 × 2 max-pooling operation with a stride of 2, thus lowering the output spatial resolution of the convolution, while also increasing the size of the overall receptive field of that convolution's output. At each stage of encoding, we repeat encoding at multiple locations in the encoder, continually reducing the spatial dimensions of each output feature while simultaneously increasing the number of channels per layer. The encoder will accumulate both local features and global context because of this hierarchical process, which are both important components for segmenting images [27].

2.3.2 Bottleneck

The U-Net's final layer is the bottleneck, the bottom-most layer, is the most compressed spatially and has the most feature channels. This layer applies repeated convolution layers all using ReLU activations to each feature map to encode global context and create a highly abstract, compact representation, which is the key connection between the encoder and decoder.

2.3.3 Decoder (Expanding Path)

The decoding part is also used to restore the spatial resolution, and to make the segmentation predictions more accurate. Accomplish this, the decoder creates an upsampled feature map at every level of the encoder's feature maps using a 2x2 transposed convolution that reduces the number of channels in the upsampled feature map by half. After this upsampling, each of the decoder's upsampled feature maps is concatenated with the corresponding encoder's feature map through skip connections [28].

Next, two distinct consecutively applied 3x3 convolutional layers using the ReLU activation function are used to sharpen the fused representations of the encoder and decoder features. This entire process is repeated until the original resolution of the input image is restored. Then a single layer 1x1 convolution is used to convert the decoder's feature maps into the number of classes that will be used to represent the output segmentation [29].

2.3.4 Skip Connections and Feature Fusion

Skip connections connect corresponding encoding and decoding layer channels together which will combine early low-level spatial features with later high-level semantic information about them through this process. The use of skip connections will reduce the loss of information by downsampling, resulting in a large improvement in segmentation precision [30]. The fusion of features via skip connections is especially advantageous when dealing with the segmentation of small tumours and difficult to define boundaries present in brain tumours [31].

2.3.5 U-Net as the Structural Foundation for Advanced Encoder-Decoder Variants

Numerous advanced segmentation models that build on U-Net's architecture. Examples are Attention U-Net, U-Net++, Residual U-Net, and a hybrid CNN-Transformer model. Each of these variants maintains the original encoder-decoder structure of U-Net, but they also add attention mechanisms, dense connections, residual learning, and global context modeling to the segmentation model to improve performance [32]. Because of this versatility and ability to support future improvements, U-Net is considered a significant research framework for developing deep learning-based classification techniques for brain tumour pathology. Given below Table 3 summarizes the evolution of U-Net-based encoder-decoder architectures and their key enhancements.

Table 3: Evolution of U-Net-Based Encoder-Decoder Architectures

Variant	Key Enhancement	Purpose
U-Net	Encoder-decoder with skip connections	Baseline segmentation
Attention U-Net	Attention gates	Focus on tumor regions
U-Net++	Dense skip connections	Improved feature fusion
Hybrid CNN-Transformer	Global context modeling	Long-range dependency learning

Fig 2. summarizes the deep learning pipeline discussed in Section 2 and motivates the architectural taxonomy in Section 3.

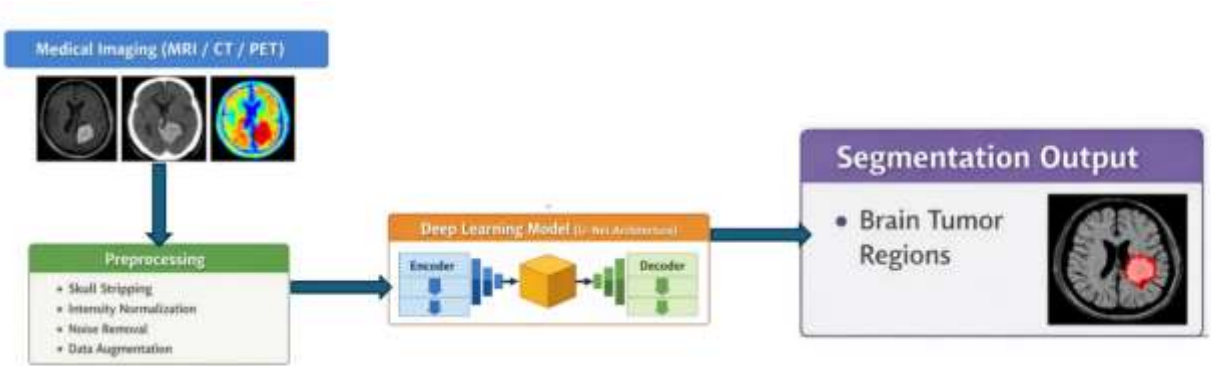

Figure 2: Deep Learning Pipeline for Brain Tumor Analysis

3. LITERATURE REVIEW

This section reviews deep learning-based approaches for brain tumor analysis, focusing on segmentation methods and their role in supporting tumor classification. To provide a structured overview of the reviewed studies, Fig 3 presents proposed taxonomy of deep learning-based approaches for brain tumor analysis. The literature is broadly categorized into deep learning brain tumor segmentation and deep learning-based brain tumor classification approaches, which are further organized based on architectural design principles.

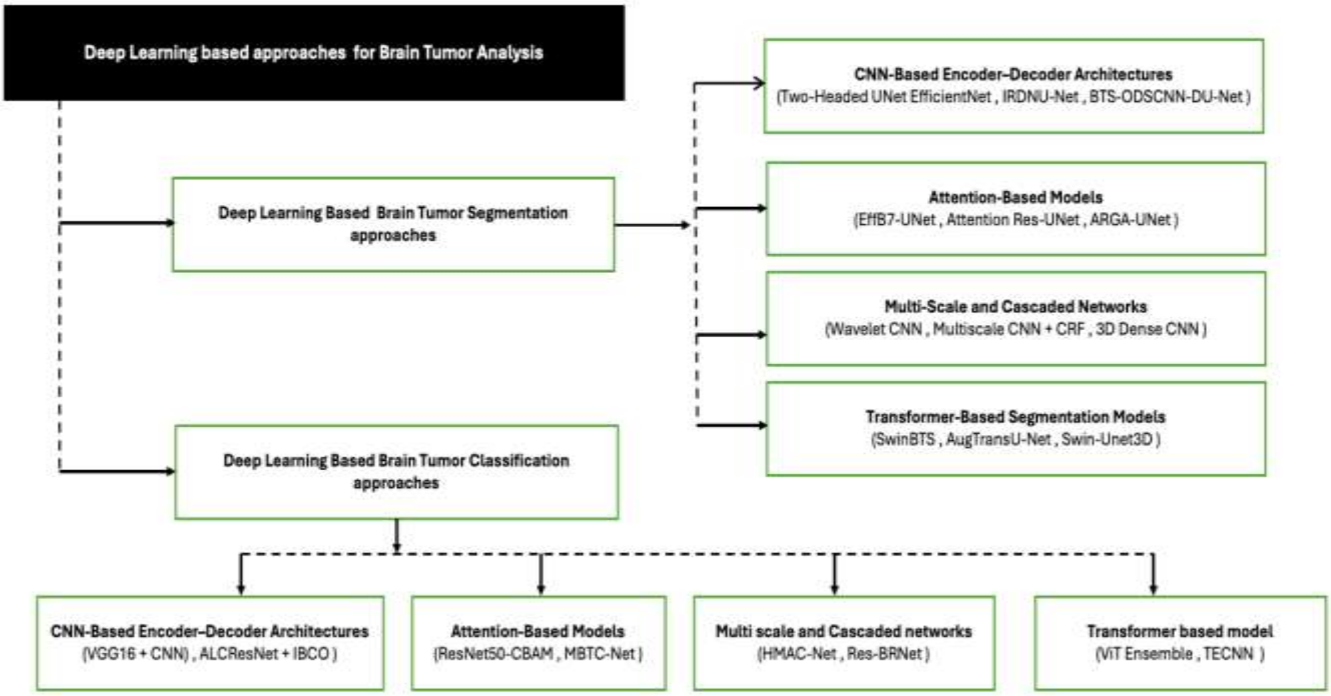

Figure 3: Proposed Taxonomy of deep learning-based approaches for brain tumor analysis

3.1 Deep Learning Based Brain Tumor Segmentation approaches

Deep learning-based segmentation methods aim to generate pixel-level or voxel-level tumor masks that accurately delineate tumor regions and substructures in brain MRI scans [33]. The information provided by segmentation, segregating different anatomical structures, or segments, to diagnose and treat patients is critical for clinical decision-making in all three areas. In addition, segmentation outputs are frequently required as input to downstream tasks such as identifying tumor types, determining their grade, predicting likelihood of the tumor progressing [34,35].

Please refer to Figure 3. Segmentation approaches based on deep learning can be divided into four primary categories CNN-based encoder-decoder networks, attention-based models, multi scale and cascade networks, and transformer-based architectures.

3.1.1 CNN-Based Encoder-Decoder Architectures

Brain tumor segmentation is a common application of the encoder-decoder CNN architecture, which is highly localizing, has good computational efficiency, and is well-suited for working with limited amounts of annotated data [36]. Several research groups have improved the segmentation performance of segmentation algorithms based on CNN encoder-decoder architectures without adding too much complexity to the models through various means, including the improvement of convolutional blocks, optimization of decoding strategies, and the development of new architectural innovations [37-39].

In this context [40] optimized Dense U-Net structure depth-wise separable convolutions for better segmentation quality and lower computation cost [41] presented IRDNU-Net inception and residual structures in a densely nested encoder-decoder manner that achieves better feature representation with fewer parameters. Localising using ROIs together with U-Net-based segmentation was shown [42], to provide reduced background interference and greater accuracy in defining tumour boundaries.

A new segmentation framework based on the residual Net architecture to improve the flow of information in the model with connection of previous features with each new cycle of the model this should provide a solution for vanishing gradients [43]. It also added to the segmented versions of networks by developing a series of encoder-decoder networks with hybrid dilated convolutional layering techniques for identifying edge-aware features more robustly [44]. Lastly [45] added a feature pyramid network to their U-NET models, which solved for fusing features at multiple resolutions to provide more robust performance regardless of the overall size of the respective tumours. Selected CNN-based segmentation models are discussed in Table 4 with respect to the main contributions, limitations, and metrics used to evaluate their performance.

Table 4: CNN-Based Encoder-Decoder Architectures for Brain Tumor Segmentation

Reference	Method name	Contribution	Limitations	Evaluation Metric
36	Two-Headed UNetEfficientNet	Parallel segmentation and classification using EfficientNet backbone integrated with U-Net	Increased model complexity and higher computational cost	DICE 94.03, Jaccard 91.20
37	PG-DBCWMF + HV Region + CTSIFT	Hybrid preprocessing, segmentation, and feature extraction pipeline	Relies heavily on handcrafted preprocessing techniques	Higher PSNR and improved accuracy
38	GoogLeNet CNN	GoogLeNet with attention modules for improved feature extraction	Primarily focused on classification rather than precise segmentation	overall accuracy 97.62
39	Dual-Module Detection Framework	MRI enhancement followed by neural network segmentation and SVM classification	Multi-stage pipeline increases processing time	DSC 0.981, sensitivity 0.99
40	BTS-ODSCNN-DU-Net	Depth-wise separable CNN integrated with Dense U-Net and optimization algorithm	Optimization algorithms increase training complexity	Higher dice index and accuracy
41	IRDNU-Net	Combines inception modules, residual connections, and dense skip pathways	Deep architecture increases training time	Dice similarity index 0.888
42	ROI-Aided Localization and Segmentation U-Net	Two-stage architecture using 2D U-Net for localization and 3D U-Net for segmentation	Two-stage framework increases computational overhead	Dice similarity index 0.876
43	Improved ResNet Segmentation Model	Enhanced residual network with improved shortcut connections	Performance depends on large training datasets	DICE Score 0.864 , Sensitivity 0.736
44	Improved U-Net with Hybrid Dilated Convolution	Serial encoder-decoder structure with hybrid dilated convolution modules	Increased model complexity and memory usage	MIoU 86.8%
45	Improved FPN-U-Net	Integrates Feature Pyramid Network with U-Net for multi-scale feature extraction	Limited evaluation on large multi-institution datasets	DICE rating 92%

3.1.2 Attention-Based Models

Attention-based models allow the segmentation of brain tumours with greater accuracy by directing the network's attention onto areas where a tumor is and directing its attention away from non-tumor related areas. This is important to accurately identify small tumours as well as tumours of different shapes and sizes on MRI scans [46]. A lightweight 3D UKAN model based on patch embedding and tokenization established [47], which enhances segmentation efficiency [48] present BT-Net, a multi-task framework for tumor classification, segmentation, and localization.

A cascaded CNN with a distance-wise attention mechanism to further improve feature learning for tumor-centric tasks [49,50] created MBANet, which is a 3D CNN using multi-branch shuffle attention to learn the spatial and channel relationship of a given input. The study in [51] introduced Attention Res-UNet with attention gates and guided decoding to enhance feature representation.

Furthermore [52], integrated a joint spatial pyramid module with attention to capture multi-scale contextual information for improved glioblastoma subregion segmentation, while [53] proposed an improved U-Net combining residual grouped convolutions and convolutional block attention modules. In addition [54], proposed M-Net, an ensemble-based architecture for segmentation and classification, whereas [55] introduced a hybrid Res2Net-U-Net model integrating multi-scale feature extraction with channel-spatial attention for improved segmentation accuracy. Table 5 summarizes attention-based models for brain tumor segmentation, highlighting their performance and key limitations.

Table 5: Attention-Based Deep Learning Models for Brain Tumor Segmentation

Reference	Method name	Contribution	Limitations	Evaluation metric
46	EffB7-UNet (BrainView framework)	Hybrid architecture integrating EfficientNetB7 encoder with U-Net decoder for tumor segmentation and DeepBrainNet for classification with cloud-based deployment	Limited evaluation on multi-institution MRI datasets	Dice coefficient 92.734 %
47	3D UKAN	Lightweight 3D architecture using patch embedding and tokenization to improve segmentation efficiency	Performance may degrade for extremely small tumor regions	Dice coefficient: 92%, Jaccard index: 86%
48	BT-Net	Multi-task framework performing tumor classification, segmentation, and localization simultaneously	increased computational complexity	DICE similarity score 0.86
49	Cascaded CNN with Distance-Wise Attention	Cascaded architecture with attention mechanism to enhance tumor-focused feature learning	Cascaded design increases training time	Whole tumor dice score 0.9203
50	MBANet	Multi-branch shuffle attention integrating spatial and channel attention	Requires high GPU memory	DICE of ET 80.18%
51	Attention Res-UNet	Attention gates combined with residual learning and guided decoder supervision	Increased model complexity	Mean IoU 0.838
52	Attention Spatial Pyramid Network	Joint spatial pyramid module with attention for multi-scale contextual feature extraction	Sensitive to hyperparameter tuning	DICE of ET and WT is 79.90 %, 89.63 %
53	ARGA-UNet	Improved U-Net using Residual Grouped Convolution (RGCM), CBAM attention, and bilinear interpolation upsampling	Evaluated on FLAIR MRI sequence dataset	DICE score 97.581
54	M-Net	Ensemble-based encoder-decoder architecture for segmentation and classification	High computational cost	Accuracy 99%
55	Res2Net-U-Net with Attention	Hybrid architecture integrating multi-scale feature extraction and channel-spatial attention	Increased architectural complexity	DICE WT 0.9630, DICE ET 0.9513

3.1.3 Multi-Scale and Cascaded Networks

Multi-scale and cascaded architectures improve brain tumor segmentation by capturing hierarchical contextual features at different spatial resolutions and refining segmentation outputs [56]. A multiscale cascaded multitask network that performs simultaneous tumor segmentation and classification using hierarchical feature aggregation was introduced [57]. In [58], a multiscale CNN combined with Conditional Random Fields to improve segmentation accuracy and reduce false positives in MRI images.

A proposed AG-MS3D-CNN [59], a multiscale attention-guided 3D CNN for robust tumor segmentation, while [60] introduced the MM-MSCA-AF framework that integrates multi-modal MRI with multi-scale contextual aggregation and attention fusion. Studies [61-63] further improved segmentation performance by integrating multi-scale feature interaction, densely connected CNN architectures, and cascaded residual multi-scale convolution frameworks within U-Net-based models. Table 6 summarizes multi-scale and cascaded segmentation models, highlighting their performance and limitations.

Table 6: Multi-Scale and Cascaded Segmentation Models

Reference	Method name	Contribution	Limitations	Evaluation metrics
56	Wavelet CNN	Multi-resolution wavelet feature extraction	Preprocessing dependency	98.8% detection accuracy
57	MCMTN	Multiscale cascaded multitask network	High model complexity	DCS 96.2 and mean IoU 95.88
58	Multiscale CNN + CRF	CRF refinement for segmentation	Additional post-processing	Sensitivity 99.86%
59	AG-MS3D-CNN	Multiscale attention-guided 3D CNN	High computational cost	High dice score
60	MM-MSCA-AF	Multimodal multi-scale contextual aggregation	Requires multimodal MRI	DICE score 0.8158
61	MVSI-Net	Multi-view attention with multi-scale features	Model complexity	DSC whole tumor 0.876
62	3D Dense CNN	Dense connectivity with multi-scale receptive fields	Memory intensive	DSC WT is 0.95
63	MambaBTS	Cascaded residual multi-scale convolution	Complex training	Dice score 0.856 for WT

3.1.4 Transformer-Based Segmentation Models

In recent years, transformer architectures have been used to segment brain tumours because they are able to recognise long-range connections and global context in Magnetic Resonance Imaging (MRI) data [64]. In [65], a combination of Swin Transformers with an encoder–decoder architecture, termed SwinBTS, was proposed for multimodal MRI segmentation. Similarly, in [66], an architecture termed AugTransU-Net was proposed, incorporating augmented transformer modules and paired attention mechanisms to enhance feature interaction and improve segmentation accuracy in MRI images.

There have been many studies that investigated the combination of transformers with hybrid structure to enhance segmentation accuracy [67]. Dual Vision Transformer-DSUNET was developed by Zakariah et al., who developed a dual transformer module and integrated feature fusing for precise segmentation of tumours [68]. TransResUNet, a transformer-enhanced residual U-Net, was proposed in [69] to enhance contextual feature learning. In [70], the 3D Medical Axial Transformer (MAT), a lightweight transformer utilizing axial attention, was developed for efficient 3D segmentation. Finally, in [71], MWG-UNet++, a hybrid transformer-based U-Net integrated with generative adversarial networks (GANs), was proposed to enhance segmentation performance. Table 7 summarizes transformer-based segmentation models, highlighting their performance and limitations.

Table 7: Transformer-Based Brain Tumor Segmentation Models

Reference	Method name	Contribution	Limitations	Evaluation metric
64	Swin-Unet3D	Hybrid CNN-Transformer architecture	High computational cost	Dice coefficient of 0.840 on ET, 0.87 on TC
65	SwinBTS	Swin transformer for multimodal MRI segmentation	Memory intensive	Average dice score 81.5
66	AugTransU-Net	Augmented transformer U-Net architecture	Complex architecture	Dice value of 89.7 for WT, 78.2 for ET
67	CKD-TransBTS	Cross-attention multimodal transformer	High computational cost	High dice score
68	Dual Vision Transformer-DSUNET	Transformer feature fusion model	Large parameters count	Cumulative dice score 91.29%
69	TransResUNet	Transformer-enhanced residual U-Net	Training complexity	Accuracy 98%
70	MAT	Lightweight axial transformer	Limited validation	Improve performance
71	MWG-UNet++	Transformer U-Net with GAN augmentation	GAN training complexity	Average evaluation metric is 0.89

Overall, existing deep learning segmentation models employ hierarchical feature extraction, attention mechanisms, multi-scale learning, and global self-attention to improve tumor localization and contextual understanding. However, many architectures introduce increased computational complexity and may require large, annotated datasets for robust training.

3.2 Deep Learning Based Brain Tumor Classification Methods

Brain tumor classification aims to assign clinically meaningful labels, such as tumor type or grade, from MRI images. Classification can be performed either on whole images or on tumor-localized regions obtained through segmentation. The following subsections review classification approaches using the same architectural taxonomy adopted for segmentation models.

3.2.1 CNN-Based Encoder-Decoder Architectures

CNN-based architectures are widely used for brain tumor classification due to their ability to automatically learn discriminative spatial features from MRI images. In [72], a dual CNN framework combining VGG16 with a custom CNN architecture was proposed to enhance multi-class tumor classification, while [73] introduced a dense CNN architecture integrating transfer learning models for improved performance. Hybrid CNN-based models such as M-C&M-BL [74] and NeXtBrain [75] further enhance classification by combining CNN-based feature extraction with advanced deep learning modules. Additionally, in [76], an IBCO-optimized ALCResNet framework was developed, integrating GAN-based augmentation, DeepLabV3 segmentation, and CNN-based classification to improve tumor classification and grading accuracy.

3.2.2 Attention-Based Models

Attention-based models enhance brain tumor classification by enabling networks to focus on tumor-relevant regions while suppressing irrelevant background information. In [77], an attention-guided multipath CNN was proposed to emphasize critical tumor features and enhance classification accuracy in MRI images. Similarly, in [78], a convolutional block attention module was integrated with ResNet50 to improve feature representation and classification performance. New techniques which improve classification accuracy and interpretability in diagnosis of brain tumor now use more advanced forms of attention, such as the Ensemble Attention Framework described by [79], The Multi-Head Attention Model MBTC-Net described by [80], and the Hybrid CNN-Transformer Attention Architecture described by [81].

3.2.3 Multi-Scale and Cascaded Networks

The author [82] developed a Multi-Scale CNN (MSCNN) approach that uses several different convolution kernels to allow for better classification of multi-class brain tumor scans and lower overall computing costs. Similarly, [83] introduced a hierarchical multi-scale feature fusion network (HMAC-Net) that combines global and local features using attention-based feature fusion to improve classification performance. In addition,[84] proposed MAProtoNet, which integrates multi-scale modules with attention mechanisms to generate interpretable tumor classification maps. Recent studies [85,86] further improved classification accuracy through multi-path CNN architectures and deep residual regional CNN frameworks that capture heterogeneous tumor patterns in MRI images.

3.2.4 Transformer-Based classification Models

Transformer-based models improve brain tumor classification by capturing long-range dependencies and global contextual information in MRI images. In [87], an ensemble of Vision Transformer (ViT) models was employed for multi-class tumor classification, achieving high accuracy on MRI datasets. Similarly, studies [88-91] proposed hybrid CNN–transformer, fine-tuned ViT, data-efficient transformer, and cross-transformer architectures to enhance classification performance in brain tumor diagnosis. These models demonstrate strong potential for improving automated medical image analysis by effectively capturing both local and global tumor features. Table 8 summarizes deep learning-based brain tumor classification models across different architectural categories, highlighting their performance and limitations.

Table 8: Deep learning-based brain tumor classification methods

Reference	Architecture type	Method name	Contribution	Limitations	Evaluation metric
72	CNN Based Encoder decoder	Dual CNN (VGG16 + CNN)	Combines VGG16 with a custom CNN architecture to improve multi-class brain tumor classification.	Higher computational complexity due to dual networks.	Training accuracy 100 % and testing accuracy 99 %
73		Modified Dense CNN	integrates DenseNet and VGG16 transfer learning to enhance classification performance.	Transfer learning models increase training time.	Average precision 0.96, average f1 score 0.950
74		M-C&M-BL	Hybrid model combining multiple CNN layers with BiLSTM to improve feature representation for tumor classification.	Hybrid architecture increases model complexity.	Accuracy 99.93%, f1 score 99.31%
75		NeXtBrain	CNN-based architecture with improved convolution blocks for enhanced tumor classification accuracy.	Requires significant computational resources.	Accuracy 99.78%
76		ALCResNet + IBCO	optimized ResNet for automated tumor classification and grading.	Multi-stage pipeline increases system complexity.	Improved Accuracy
77	Attention based model	Attention-Guided CNN	Multipath CNN with attention mechanism to highlight tumor-relevant features	Limited dataset diversity	Accuracy 98.61%
78		ResNet50-CBAM	ResNet50 integrated with convolutional block attention module	Increased model parameters	Accuracy 99.43 %, AUC 99.25%
79		Ensemble Attention Model	Co-attention mechanism using MobileNetV3 and EfficientNetB7 for improved tumor classification	High computational complexity	Accuracy on fig share dataset 98.94%
80		MBTC-Net	EfficientNetV2B0 with multi-head attention for multimodal brain tumor classification	Requires multimodal data	Accuracy 97.54% on 15 classes
81		DenseTransformer	DenseNet201 with self-attention and transformer modules for tumor classification	Hybrid architecture complexity	Accuracy 99.41%

82	Multi scale and Cascaded networks	MSCNN	Multi-scale CNN using multiple convolution kernels for MRI tumor classification	Sensitive to image noise	Accuracy 91.2 %, f1 score 91%
83		HMAC-Net	Hierarchical multi-scale feature fusion with attention mechanisms	High model complexity	Accuracy 0.921, precision 0.960
84		MAProtoNet	Multi-scale attentive prototypical network for interpretable tumor classification	Computational overhead	Improved BAC score
85		Multi-Path CNN	Multi-path CNN with diverse convolution kernels for feature extraction	Kernel selection sensitivity	Accuracy 92.25 % with all features
86		Res-BRNet	Deep residual and regional CNN capturing local and global tumor features	Requires large training data	F1-score 0.9841
87	Transformer based model	ViT Ensemble	Ensemble of Vision Transformer models for multi-class tumor classification	Requires large training data	Overall testing accuracy 98.7%
88		TECNN	Hybrid CNN-Transformer model combining local and global feature learning	High computational complexity	Accuracy on Figshare 96.75 %
89		FTVT	Fine-tuned Vision Transformer models for multi-class MRI tumor classification	Requires large dataset for training	Accuracy 98.70%
90		LCDEiT	Linear-complexity data-efficient transformer for MRI tumor classification	Model architecture complexity	Accuracy 98.11 % and f1 score 97.86 % on figshare
91		Cross-Transformer	Transformer-based framework for tumor classification and detection	Complex training pipeline	More than 97% accuracy

4. PERFORMANCE REVIEW

Performance in brain tumor segmentation and classification depends on the datasets used, evaluation metrics, and variations in experimental protocols such as preprocessing and train-test splits. Consequently, fair comparison of reported results requires careful interpretation and consistent benchmarking across studies. In addition to model architecture, factors such as dataset composition, annotation quality, class imbalance, and modality availability also influence model performance.

4.1 Publicly Available Brain Tumor Datasets

Public datasets are essential for reproducible benchmarking in brain tumor analysis. Segmenting studies often use multiple modalities of MRI data while usually, classifications are made using curated sets of MRI data. Within the segmentation domain, the BraTS dataset family remains one of the most frequently used benchmarks to segment using standardised multi-modal MRIs with expert annotated tumours, and many datasets available from the TCIA or TCGA ecosystems are also often used for brain tumour studies including, TCGA-LGG, or sets to classify brain tumours available from sites such as Figshare or BRISC. Recent benchmark datasets of the BraTS-GOAT 2024 have provided an opportunity for additional robustness evaluations across multiple different clinical imaging situations. Table 9 presents commonly used benchmark datasets for brain tumor segmentation and classification, along with their modalities and characteristics.

Table 9: Benchmark Datasets Used in Brain Tumor Segmentation and Classification

Dataset / Source	Task	Modalities	Notes for benchmarking
BraTS family	Segmentation	T1, T1c, T2, FLAIR MRI	Standardized benchmark; widely compared across models
BraTS-GOAT 2024 [92]	Segmentation	Multi-modal MRI	Introduced updated evaluation and robustness checks
TCGA-LGG (via TCIA) [93]	Segmentation / Classification	Multi-modal MRI + Clinical	Frequently used for segmentation-assisted classification pipelines
BRISC [94]	Segmentation / Classification	Single-modal MRI (T1c)	Annotated masks; balanced dataset

4.2 Comparative Analysis and Discussion

Due to variations in datasets, imaging modalities, preprocessing methods, and evaluation metrics, direct comparison of deep learning approaches for brain tumor segmentation and classification is challenging, and reported performance often depends more on data and evaluation strategy than architecture alone [1] [52] [72] [75]. Among existing methods, CNN-based encoder–decoder models, particularly U-Net and its variants, remain the most reliable baselines due to their strong localization capability, computational efficiency, and robustness with limited data, often achieving improved results with enhanced architectures and optimized designs [13] [18] [44] [47].

Table 10 consists of several representative design categories for architectural paradigms deployed in brain tumor segmentation and classification, as well as their respective strengths, weaknesses, and common use scenarios, with a view towards providing a structured overview of these paradigms.

Table 10: Comparative Analysis of Deep Learning Architectures for Brain Tumor Segmentation and Classification

Architecture Category	Representative Models from literature	Primary Strengths	Key Limitations	Typical Use Case
CNN Encoder-Decoder	U-Net, V-Net, U-Net++, ResUNet	Efficient training, strong localization, works with limited data	Limited global context modeling	Baseline segmentation, clinical pipelines
Attention-Based	Attention U-Net, ResUNet+, attention-CNNs	Focus on tumor regions, improved boundary precision	Increased model complexity, tuning sensitivity	Small or heterogeneous tumours
Multi-Scale & Cascaded	Multi-scale CNNs, cascaded U-Net pipelines	Captures global + local context, improved subregion segmentation	Error propagation, longer inference	Tumor subregion analysis, grading
Transformer-Based	Hybrid CNN-Transformer, ViT-based models	Long-range dependency modeling, global context	Data-hungry, high compute cost	Large datasets, research-focused systems

Segmentation is greatly enhanced with the use of attention mechanisms as they enable the focus to be on important features while minimizing noise, which is especially beneficial for small and complicated tumours. The Dice scores associated with the use of attention models based on the U-Net architecture have shown to be improved, and the number of false positive predictions is also reduced while still having higher complexity and sensitivity to tuning [13], [15], [70]. Additionally, Multi-scale, and cascaded architectures have enhanced segmentation performance through the ability to capture features at multiple scales, and through the ability to perform additional refinements, which allows for a more in-depth hierarchy-level analysis of tumours; however, these architectures are characterized by greater computational overheads and more opportunities for the propagation of errors [39], [45], [58], [95]. Segmentation-based classification techniques also improve accuracy in diagnostics by incorporating the spatial location of the tumour, these methods often outperform direct methods using images at the level of classification emphasizing the requirement of successful segmentation for clinical applications [39], [55], [60], [86]. Transformer-based techniques and hybrid CNN – Transformer technologies demonstrate improved global context modelling and greater robustness in complicated conditions, yet they demand an extensive amount of database material, high computational requirements, and can be adversely impacted by domain displacement [7], [22], [65], [72].

The following information was compiled from various comparative study results and survey results to generate Table 11, which compares and summarizes segmentation quality, classification accuracy, generalizability, and clinical applicability across different architectural paradigms.

Table 11: Performance Trend Summary Across Architectural Paradigms

Design Paradigm	Segmentation Performance Trend	Classification Impact	Generalization	Clinical Readiness
CNN Encoder-Decoder	Stable, strong baseline	Moderate-High	Good	High
Attention-Based	Improved boundary accuracy	High	Moderate	Medium
Multi-Scale or Cascaded	Strong for complex tumours	High	Moderate	Medium
Transformer-Based	Potentially high	Promising	Uncertain	Low-Medium

In comparing these paradigms, we can see that the performance improvements due to architectural innovation can be large. There is no single best performing paradigm across all datasets and clinical scenarios. The architectures that are currently most validated or deemed to have inconsistent performance include the CNN encoder-decoder architectures, while the attention multi-scale and transformer-based architectures have provided increased performance for specific improvements in complex scenarios but are still ongoing areas of research.

5. CONCLUSION

In this review, we provide a comprehensive review of deep learning methods that can be used for brain tumor segmentation/classification based on architectural design principles, imaging modalities and preprocessing methods. We find that encoder-decoder CNN architectures, such as U-Net and its variants, continue to be the most effective and most used due to their ability to localize accurately, their computational efficiency, and how well they can work with a limited number of labelled training examples. Attention mechanisms and multi-scale or cascading models can also help to improve the accuracy of brain tumor segmentation by allowing the model to focus on tumor-relevant regions and capture their hierarchical context. Although transformer-based and hybrid CNN-transformer models show potential in accurately modelling global dependencies, they can require more computation and data than other methods. This article is intended to raise awareness when comparing model performance in different studies because of different evaluation protocols and data sets. Also, segmentation assisted classification methods in general provide better interpretability and diagnostic accuracy than direct classification methods. As such, it is important for future progress to include better data availability, better generalizability to different clinical situations, better interpretability and more robust models for aiding with clinical decisions in the real world.

Acknowledgments

Authors of this paper wish to thank all researchers and groups who have contributed to the body of evidence documented herein pertaining to deep learning applications for the segmentation and classification of brain tumours. The authors consider that the vast body of literature in medical imaging provides significant insights into the direction and development of this review article. Authors also acknowledge the contributions of the people responsible for developing and maintaining publicly available datasets used in scientific research in this area, such as the Brain Tumor Segmentation (BraTS) and The Cancer Imaging Archive (TCIA). These two datasets have been instrumental in helping promote and advance research in this field. The authors confirm that no specific funding was received for this study.

REFERENCES

1. Townsend K. L. (2024). One Nervous System: Critical Links Between Central and Peripheral Nervous System Health and Implications for Obesity and Diabetes. Diabetes, 73(12), 1967–1975. https://doi.org/10.2337/dbi24-0004
2. Hussain, S. S., Wani, N. A., Kaur, J., Ahmad, N., & Ahmad, S. (2025). Next-generation automation in neuro-oncology: advanced neural networks for MRI-based brain tumor segmentation and classification. IEEE Access. DOI:10.1109/ACCESS.2025.3547796
3. Pratap Joshi, K., Gowda, V. B., Bidare Divakarachari, P., Siddappa Parameshwarappa, P., & Patra, R. K. (2025). VSA-GCNN: Attention Guided Graph Neural Networks for Brain Tumor Segmentation and Classification. Big Data and Cognitive Computing, 9(2), 29. https://doi.org/10.3390/bdcc9020029
4. Nancy, A. M., & Maheswari, R. (2025). Brain tumor segmentation and classification using transfer learning based CNN model with model agnostic concept interpretation. Multimedia Tools and Applications, 84(5), 2509-2538. DOI:10.1007/s11042-024-20353-1
5. Renugadevi, M., Narasimhan, K., Ramkumar, K., & Raju, N. (2025). A novel hybrid vision UNet architecture for brain tumor segmentation and classification. Scientific Reports, 15(1), 23742. DOI:10.1038/s41598-025-09833-y
6. Gurunathan, P., Srinivasan, P.S. & S, R. Deep transfer learning based feature fusion model with Bonobo optimization algorithm for enhanced brain tumor segmentation and classification through biomedical imaging. Sci Rep 15, 34030 (2025). https://doi.org/10.1038/s41598-025-12973-w
7. Ilic, I., & Ilic, M. (2023). International patterns and trends in the brain cancer incidence and mortality: An observational study based on the global burden of disease. Heliyon, 9(7), e18222. https://doi.org/10.1016/j.heliyon.2023.e18222
8. Satushe, V., Vyas, V., Metkar, S., & Paul Singh, D. (2025). Advanced CNN Architecture for Brain Tumor Segmentation and Classification using BraTS-GOAT 2024 Dataset. Current medical imaging, 21, e15734056344235. https://doi.org/10.2174/0115734056344235241217155930
9. Ali, A., Li, X., Mashwani, W. K., Abiad, M., Karim, F. K., & Mostafa, S. M. (2025). Multi-class brain tumor MRI segmentation and classification using deep learning and machine learning approaches.

Cancer imaging : the official publication of the International Cancer Imaging Society, 25(1), 131. https://doi.org/10.1186/s40644-025-00953-2
10. Anitha, V., & Murugavalli, S. J. I. C. V. (2016). Brain tumour classification using two-tier classifier with adaptive segmentation technique. IET computer vision, 10(1), 9-17. DOI:10.1049/iet-cvi.2014.0193
11. Shen, L., Hou, Y., Chen, J., Diao, L., & Duan, Y. (2024). Mbdres-u-net: Multi-scale lightweight brain tumor segmentation network. arXiv preprint arXiv:2411.01896. https://doi.org/10.48550/arXiv.2411.01896
12. Lu, N. H., Huang, Y. H., Liu, K. Y., & Chen, T. B. (2025). Deep learning-driven brain tumor classification and segmentation using non-contrast MRI. Scientific reports, 15(1), 27831. https://doi.org/10.1038/s41598-025-13591-2
13. R, P., M, J. P. P., & J S, N. (2025). Brain tumor segmentation using multi-scale attention U-Net with EfficientNetB4 encoder for enhanced MRI analysis. Scientific reports, 15(1), 9914. https://doi.org/10.1038/s41598-025-94267-9
14. R, P., M, J. P. P., & J S, N. (2025). Brain tumor segmentation using multi-scale attention U-Net with EfficientNetB4 encoder for enhanced MRI analysis. Scientific reports, 15(1), 9914. https://doi.org/10.1038/s41598-025-94267-9
15. Zarenia, E., Far, A. A., & Rezaee, K. (2025). Automated multi-class MRI brain tumor classification and segmentation using deformable attention and saliency mapping. Scientific reports, 15(1), 8114. https://doi.org/10.1038/s41598-025-92776-1
16. Jin, C., Ng, T. F., & Ibrahim, H. (2025). Advancements in Semi-Supervised Deep Learning for Brain Tumor Segmentation in MRI: A Literature Review. AI, 6(7), 153. https://doi.org/10.3390/ai6070153
17. Thaha, M. M., Kumar, K. P. M., Murugan, B. S., Dhanasekeran, S., Vijayakarthick, P., & Selvi, A. S. (2019). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. Journal of medical systems, 43(9), 294. https://doi.org/10.1007/s10916-019-1416-0
18. Tiwary, P. K., Johri, P., Katiyar, A., & Chhipa, M. K. (2025). Deep learning-based MRI brain tumor segmentation with EfficientNet-enhanced UNet. IEEE Access. DOI:10.1109/ACCESS.2025.3554405
19. Lakshmi, K., Amaran, S., Subbulakshmi, G., Padmini, S., Joshi, G. P., & Cho, W. (2025). Explainable artificial intelligence with UNet based segmentation and Bayesian machine learning for classification of brain tumours using MRI images. Scientific reports, 15(1), 690. https://doi.org/10.1038/s41598-024-84692-7
20. Saad, M. A. (2025). Automated Segmentation and Classification of Brain Tumor Using Deep Learning: Improving Healthcare Efficiency and Patient Outcomes. Int J Phys Med & Rehbtn, 2(1), 1-7. DOI:10.33564/IJEAST.2021.v06i01.057
21. Bahadure, N. B., Ray, A. K., & Thethi, H. P. (2018). Comparative Approach of MRI-Based Brain Tumor Segmentation and Classification Using Genetic Algorithm. Journal of digital imaging, 31(4), 477–489. https://doi.org/10.1007/s10278-018-0050-6
22. Tran, A. T., Zeevi, T., & Payabvash, S. (2025). Strategies to Improve the Robustness and Generalizability of Deep Learning Segmentation and Classification in Neuroimaging. BioMedInformatics, 5(2), 20. https://doi.org/10.3390/biomedinformatics5020020
23. Raghuwanshi, G., Vinu, D., Gupta, P., Goyal, M. K., Khaitan, S., & HS, M. (2025). Brain tumor segmentation in multi-modal MRI: A comparative study. Journal of Integrated Science and Technology, 13(7), 1153-1153. DOI:10.62110/sciencein.jist.2025.v13.1153
24. Wang, Y., Xu, J., Guan, Y., Ahmad, F., Mahmood, T., & Rehman, A. (2025). MSegNet: A multi-view coupled cross-modal attention model for enhanced MRI brain tumor segmentation. International Journal of Computational Intelligence Systems, 18(1), 63. DOI:10.1007/s44196-025-00787-7
25. Muhammad Faheem Khan, Arslan Iftikhar, Huzaifa Anwar, & Sadaqat Ali Ramay. (2024). Brain Tumor Segmentation and Classification using Optimized Deep Learning. Journal of Computing & Biomedical Informatics, 7(01), 632–640. Retrieved from https://www.jcbi.org/index.php/Main/article/view/528
26. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Cham: Springer international publishing. https://doi.org/10.48550/arXiv.1505.04597
27. Almufareh, M. F., Imran, M., Khan, A., Humayun, M., & Asim, M. (2024). Automated brain tumor

segmentation and classification in MRI using YOLO-based deep learning. IEEE Access, 12, 16189-16207. DOI:10.1109/ACCESS.2024.3359418

28. Sharif, M., Tanvir, U., Munir, E.U., Khan, M.A., & Yasmin, M. (2018). Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection. Journal of Ambient Intelligence and Humanized Computing, 1-20. DOI:10.1007/s12652-018-1075-x
29. Yoo, J. J., Namdar, K., Wagner, M. W., Yeom, K. W., Nobre, L. F., Tabori, U., Hawkins, C., Ertl-Wagner, B. B., & Khalvati, F. (2025). Generative AI for weakly supervised segmentation and downstream classification of brain tumours on MR images. Scientific reports, 15(1), 22160. https://doi.org/10.1038/s41598-025-06741-z
30. Batool, A., & Byun, Y. C. (2024). Brain tumor detection with integrating traditional and computational intelligence approaches across diverse imaging modalities - Challenges and future directions. Computers in biology and medicine, 175, 108412. https://doi.org/10.1016/j.compbiomed.2024.108412
31. Asiri, A. A., Shaf, A., Ali, T., Aamir, M., Irfan, M., Alqahtani, S., Mehdar, K. M., Halawani, H. T., Alghamdi, A. H., Alshamrani, A. F. A., & Alqhtani, S. M. (2023). Brain Tumor Detection and Classification Using Fine-Tuned CNN with ResNet50 and U-Net Model: A Study on TCGA-LGG and TCIA Dataset for MRI Applications. Life (Basel, Switzerland), 13(7), 1449. https://doi.org/10.3390/life13071449
32. Iqbal, S., Ghani, M. U., Saba, T., & Rehman, A. (2018). Brain tumor segmentation in multi-spectral MRI using convolutional neural networks (CNN). Microscopy research and technique, 81(4), 419–427. https://doi.org/10.1002/jemt.22994
33. Shreeharsha, J. (2024). Brain Tumor Segmentation and Classification Using Binomial Thresholding-Based Bidirectional-Long-Short Term Memory. International Journal of Intelligent Engineering & Systems, 17(3). DOI: 10.22266/ijies2024.0630.13
34. Dénes-Fazakas, L., Kovács, L., Eigner, G., & Szilágyi, L. (2024). Enhancing Brain Tumor Diagnosis with L-Net: A Novel Deep Learning Approach for MRI Image Segmentation and Classification. Biomedicines, 12(10), 2388. https://doi.org/10.3390/biomedicines12102388
35. Amin, J., Sharif, M., Yasmin, M., Saba, T., Anjum, M. A., & Fernandes, S. L. (2019). A New Approach for Brain Tumor Segmentation and Classification Based on Score Level Fusion Using Transfer Learning. Journal of medical systems, 43(11), 326. https://doi.org/10.1007/s10916-019-1453-8
36. Rai, H. M., Yoo, J., & Dashkevych, S. (2024). Two-headed UNetEfficientNets for parallel execution of segmentation and classification of brain tumours: incorporating postprocessing techniques with connected component labelling. Journal of cancer research and clinical oncology, 150(4), 220. https://doi.org/10.1007/s00432-024-05718-1
37. Phan, T. C. (2024). Tumor segmentation and classification using machine learning approaches. International Journal of Data Informatics and Intelligent Computing, 3(1), 1-11. DOI: https://doi.org/10.59461/ijdiic.v3i1.89
38. Subba, A. B., & Sunaniya, A. K. (2025). Computationally optimized brain tumor classification using attention based GoogLeNet-style CNN. Expert Systems with Applications, 260, 125443. https://doi.org/10.1016/j.eswa.2024.125443
39. Asiri, A. A., Soomro, T. A., Shah, A. A., Pogrebna, G., Irfan, M., & Alqahtani, S. (2024). Optimized brain tumor detection: a dual-module approach for mri image enhancement and tumor classification. IEEE access, 12, 42868-42887. DOI:10.1109/ACCESS.2024.3379136
40. Revathi, K. G., Shirley, C. P., & Sreethar, S. (2025). Brain tumor segmentation using optimized depth wise separable convolutional neural network with dense U-Net. Knowledge-Based Systems, 324, 113678. DOI:10.1016/j.knosys.2025.113678
41. AboElenein, N. M., Songhao, P., & Afifi, A. (2022). IRDNU-Net: Inception residual dense nested u-net for brain tumor segmentation. Multimedia Tools and Applications, 81(17), 24041-24057. DOI:10.1007/s11042-022-12586-9
42. Li, S., Liu, J., & Song, Z. (2022). Brain tumor segmentation based on region of interest-aided localization and segmentation U-Net. International journal of machine learning and cybernetics, 13(9), 2435–2445. https://doi.org/10.1007/s13042-022-01536-4
43. Aggarwal, M., Tiwari, A. K., Sarathi, M. P., & Bijalwan, A. (2023). An early detection and segmentation of Brain Tumor using Deep Neural Network. BMC medical informatics and decision making, 23(1), 78. https://doi.org/10.1186/s12911-023-02174-8

44. Zheng, P., Zhu, X., & Guo, W. (2022). Brain tumour segmentation based on an improved U-Net. BMC medical imaging, 22(1), 199. https://doi.org/10.1186/s12880-022-00931-1
45. Sun, H., Yang, S., Chen, L., Liao, P., Liu, X., Liu, Y., & Wang, N. (2023). Brain tumor image segmentation based on improved FPN. BMC medical imaging, 23(1), 172. https://doi.org/10.1186/s12880-023-01131-1
46. Ghose, P., & Jamil, H. M. (2026). BrainView: A cloud-based deep learning system for brain image segmentation, tumor detection and visualization. Biomedical journal, 49(1), 100871. https://doi.org/10.1016/j.bj.2025.100871
47. Shahid, A., Massod, S., Nasim, F., & Siddique, A. (2025). Comparative analysis of brain tumor segmentation: Nested UNet Vs Light UKAN. Journal of Innovative Computing and Emerging Technologies, 5(1). DOI:10.56536/jicet.v5i1.190
48. Rabby, S. F., Arafat, M. A., & Hasan, T. (2024). BT-Net: An end-to-end multi-task architecture for brain tumor classification, segmentation, and localization from MRI images. Array, 22, 100346. DOI:10.1016/j.array.2024.100346
49. Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S., Anari, S., Naseri, M., & Bendechache, M. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. Scientific reports, 11(1), 10930. https://doi.org/10.1038/s41598-021-90428-8
50. Cao, Y., Zhou, W., Zang, M., An, D., Feng, Y., & Yu, B. (2023). MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images. Biomedical Signal Processing and Control, 80, 104296. DOI:10.1016/j.bspc.2022.104296
51. Maji, D., Sigedar, P., & Singh, M. (2022). Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumours. Biomedical Signal Processing and Control, 71, 103077. DOI:10.1016/j.bspc.2021.103077
52. Liu, H., Huang, J., Li, Q., Guan, X., & Tseng, M. (2024). A deep convolutional neural network for the automatic segmentation of glioblastoma brain tumor: Joint spatial pyramid module and attention mechanism network. Artificial intelligence in medicine, 148, 102776. DOI: https://doi.org/10.1016/j.artmed.2024.102776
53. Siyi, X. U. N., Zhang, Y., Sixu, D. U. A. N., Mingwei, W. A. N. G., Jiangang, C. H. E. N., Tong, T. O. N. G., ... & Tao, T. A. N. (2024). Arga-unet: Advanced u-net segmentation model using residual grouped convolution and attention mechanism for brain tumor mri image segmentation. Virtual Reality & Intelligent Hardware, 6(3), 203-216. DOI: https://doi.org/10.1016/j.vrih.2023.05.001
54. Sreelakshmi, S., & Malu, G. (2023). M-Net: An encoder-decoder architecture for medical image analysis using ensemble learning. Results in Engineering, 17, 100927. DOI: https://doi.org/10.1016/j.rineng.2023.100927
55. Devi, C. N., Lawrence, T. S., & Subramaniam, P. R. (2025). A Novel Hybrid Res2Net-UNet Model for Accurate Brain Tumor Segmentation in MRI. International Journal of Cognitive Computing in Engineering. DOI:10.1016/j.ijcce.2025.11.005
56. Gokapay, D. K., & Mohanty, S. N. (2024). Enhanced MRI-based brain tumor segmentation and feature extraction using Berkeley wavelet transform and ETCCNN. Digital Health, 10, 20552076241305282. DOI: 10.1177/20552076241305282
57. Sobhaninia, Z., Karimi, N., Khadivi, P., & Samavi, S. (2023). Brain tumor segmentation by cascaded multiscale multitask learning framework based on feature aggregation. Biomedical Signal Processing and Control, 85, 104834. DOI:10.1016/j.bspc.2023.104834
58. Guennich, A., Othmani, M., & Ltifi, H. (2025). Advanced brain tumor segmentation with a multiscale CNN and conditional random fields. IEEE Access. DOI:10.1109/ACCESS.2025.3541183
59. Lilhore, U. K., Sunder, R., Simaiya, S., Alsafyani, M., Monish Khan, M. D., Alroobaea, R., ... & Baqasah, A. M. (2025). AG-MS3D-CNN multiscale attention guided 3D convolutional neural network for robust brain tumor segmentation across MRI protocols. Scientific Reports, 15(1), 24306. DOI:10.1038/s41598-025-09351-x
60. Aslam, W., Hussain, J., Aslam, M. Z., Jan, S., Riaz, T. B., Iqbal, A., ... & Khan, I. (2025). Enhanced brain tumor segmentation in medical imaging using multi-modal multi-scale contextual aggregation and attention fusion. Scientific Reports, 15(1), 37308. DOI:10.1038/s41598-025-21255-4

61. Sun, J., Hu, M., Wu, X., Tang, C., Lahza, H., Wang, S., & Zhang, Y. (2024). MVSI-Net: Multi-view attention and multi-scale feature interaction for brain tumor segmentation. Biomedical Signal Processing and Control, 95, 106484. DOI:10.1016/j.bspc.2024.106484
62. Adib, F., Amirmazlaghani, M., & Rahmati, M. (2025). 3D densely connected CNN with multi-scale receptive fields and hybrid loss for brain tumor segmentation. Discover Artificial Intelligence, 5(1), 213. DOI:10.1007/s44163-025-00279-9
63. Zhou, R., Wang, J., Xia, G., Xing, J., Shen, H., & Shen, X. (2024). Cascade residual multiscale convolution and mamba-structured unet for advanced brain tumor image segmentation. Entropy, 26(5), 385. DOI:10.3390/e26050385
64. Cai, Y., Long, Y., Han, Z., Liu, M., Zheng, Y., Yang, W., & Chen, L. (2023). Swin Unet3D: a three-dimensional medical image segmentation network combining vision transformer and convolution. BMC medical informatics and decision making, 23(1), 33. DOI:10.1186/s12911-023-02129-z
65. Jiang, Y., Zhang, Y., Lin, X., Dong, J., Cheng, T., & Liang, J. (2022). SwinBTS: A method for 3D multimodal brain tumor segmentation using swin transformer. Brain sciences, 12(6), 797. DOI:10.3390/brainsci12060797
66. Zhang, M., Liu, D., Sun, Q., Han, Y., Liu, B., Zhang, J., & Zhang, M. (2024). Augmented transformer network for MRI brain tumor segmentation. Journal of King Saud University-Computer and Information Sciences, 36(1), 101917. DOI:10.1016/j.jksuci.2024.101917
67. Lin, J., Lin, J., Lu, C., Chen, H., Lin, H., Zhao, B., ... & Han, C. (2023). CKD-TransBTS: clinical knowledge-driven hybrid transformer with modality-correlated cross-attention for brain tumor segmentation. IEEE transactions on medical imaging, 42(8), 2451-2461. DOI:10.1109/TMI.2023.3250474
68. Zakariah, M., Al-Razgan, M., & Alfakih, T. (2024). Dual vision Transformer-DSUNET with feature fusion for brain tumor segmentation. Heliyon, 10(18). DOI: https://doi.org/10.1016/j.heliyon.2024.e37804
69. Rasool, N., Bhat, J. I., Wani, N. A., Ahmad, N., & Alshara, M. (2024). TransResUNet: revolutionizing glioma brain tumor segmentation through transformer-enhanced residual UNet. IEEE access, 12, 72105-72116. DOI: 10.1109/ACCESS.2024.3402947
70. Liu, C., & Kiryu, H. (2024, January). 3D medical axial transformer: A lightweight transformer model for 3D brain tumor segmentation. In Medical imaging with deep learning (pp. 799-813). PMLR.
71. Lyu, Y., & Tian, X. (2025). MWG-UNet++: hybrid transformer U-Net model for brain tumor segmentation in MRI scans. Bioengineering, 12(2), 140. DOI: https://doi.org/10.3390/bioengineering12020140
72. Al-Zoghby, A. M., Al-Awadly, E. M. K., Moawad, A., Yehia, N., & Ebada, A. I. (2023). Dual Deep CNN for Tumor Brain Classification. Diagnostics 2023, 13, 2050. DOI:10.3390/diagnostics13122050
73. Özkaraca, O., Bağrıaçık, O. İ., Gürüler, H., Khan, F., Hussain, J., Khan, J., & Laila, U. E. (2023). Multiple Brain Tumor Classification with Dense CNN Architecture Using Brain MRI Images. Life (Basel, Switzerland), 13(2), 349. https://doi.org/10.3390/life13020349
74. Anand, V., Khajuria, A., Pachauri, R. K., & Gupta, V. (2026). Multi-class classification of brain tumours using optimized CNN and transfer learning techniques. Scientific reports, 16(1), 4709. https://doi.org/10.1038/s41598-025-34806-6
75. Pacal, I., Akhan, O., Deveci, R. T., & Deveci, M. (2025). NeXtBrain: Combining local and global feature learning for brain tumor classification. Brain research, 1863, 149762. https://doi.org/10.1016/j.brainres.2025.149762
76. Yaqub, M., Jinchao, F., Ahmed, S., Mehmood, A., Chuhan, I. S., Manan, M. A., & Pathan, M. S. (2023). DeepLabV3, IBCO-based ALCResNet: A fully automated classification, and grading system for brain tumor. Alexandria Engineering Journal, 76, 609-627. DOI:10.1016/j.aej.2023.06.062
77. Jun, W., & Liyuan, Z. (2022). Brain tumor classification based on attention guided deep learning model. International Journal of Computational Intelligence Systems, 15(1), 35. DOI:10.1007/s44196-022-00090-9
78. Oladimeji, O. O., & Ibitoye, A. O. J. (2023). Brain tumor classification using ResNet50-convolutional block attention module. Applied Computing and Informatics. DOI:10.1108/ACI-09-2023-0022
79. Pacal, I., Celik, O., Bayram, B., & Cunha, A. (2024). Enhancing EfficientNetv2 with global and efficient channel attention mechanisms for accurate MRI-Based brain tumor classification. Cluster Computing, 27(8), 11187-11212. DOI:10.1007/s10586-024-04532-1

80. Kar, S., & Singh, P. K. (2025). MBTC-Net: Multimodal brain tumor classification from CT and MRI scans using deep neural network with multi-head attention mechanism. Medicine in Novel Technology and Devices, 27, 100382. DOI:10.1016/j.medntd.2025.100382
81. Panigrahi, S., Adhikary, D. R. D., & Pattanayak, B. K. (2025). Hybrid transfer learning and self-attention framework for robust MRI-based brain tumor classification. Scientific reports, 15(1), 21343. https://doi.org/10.1038/s41598-025-09311-5
82. Yazdan, S. A., Ahmad, R., Iqbal, N., Rizwan, A., Khan, A. N., & Kim, D. H. (2022). An Efficient Multi-Scale Convolutional Neural Network Based Multi-Class Brain MRI Classification for SaMD. Tomography (Ann Arbor, Mich.), 8(4), 1905–1927. https://doi.org/10.3390/tomography8040161
83. Zhou, Y., Yang, X., Yin, J., & Liu, S. (2024). Research on multi-scale feature fusion network algorithm based on brain tumor medical image classification. Computers, Materials, & Continua, 79(3), 5313. DOI:10.32604/cmc.2024.052060
84. Li, B., Mao, J., Sun, Z., Li, C., Zhao, Q., & Tanaka, T. (2024). Maprotonet: A multi-scale attentive interpretable prototypical part network for 3d magnetic resonance imaging brain tumor classification. arXiv preprint arXiv:2404.08917. https://doi.org/10.48550/arXiv.2404.08917
85. Batool, A., & Byun, Y. C. (2025). A lightweight multi-path convolutional neural network architecture using optimal features selection for multiclass classification of brain tumor using magnetic resonance images. Results in Engineering, 25, 104327. DOI:10.1016/j.rineng.2025.104327
86. Zahoor, M. M., Khan, S. H., Alahmadi, T. J., Alsahfi, T., Mazroa, A. S. A., Sakr, H. A., Alqahtani, S., Albanyan, A., & Alshemaimri, B. K. (2024). Brain Tumor MRI Classification Using a Novel Deep Residual and Regional CNN. Biomedicines, 12(7), 1395. https://doi.org/10.3390/biomedicines12071395
87. Tummala, S., Kadry, S., Bukhari, S. A. C., & Rauf, H. T. (2022). Classification of Brain Tumor from Magnetic Resonance Imaging Using Vision Transformers Ensembling. Current oncology (Toronto, Ont.), 29(10), 7498–7511. https://doi.org/10.3390/curroncol29100590
88. Aloraini, M., Khan, A., Aladhadh, S., Habib, S., Alsharekh, M. F., & Islam, M. (2023). Combining the transformer and convolution for effective brain tumor classification using MRI images. Applied Sciences, 13(6), 3680. DOI:10.3390/app13063680
89. Reddy, C. K. K., Reddy, P. A., Janapati, H., Assiri, B., Shuaib, M., Alam, S., & Sheneamer, A. (2024). A fine-tuned vision transformer based enhanced multi-class brain tumor classification using MRI scan imagery. Frontiers in oncology, 14, 1400341. https://doi.org/10.3389/fonc.2024.1400341
90. Ferdous, G. J., Sathi, K. A., Hossain, M. A., Hoque, M. M., & Dewan, M. A. A. (2023). LCDEiT: A linear complexity data-efficient image transformer for MRI brain tumor classification. IEEe Access, 11, 20337-20350. DOI:10.1109/ACCESS.2023.3244228
91. Anaya-Isaza, A., Mera-Jiménez, L., Verdugo-Alejo, L., & Sarasti, L. (2023). Optimizing MRI-based brain tumor classification and detection using AI: A comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network. European journal of radiology open, 10, 100484. https://doi.org/10.1016/j.ejro.2023.100484
92. Satushe, V., Vyas, V., Metkar, S., & Paul Singh, D. (2025). Advanced CNN Architecture for Brain Tumor Segmentation and Classification using BraTS-GOAT 2024 Dataset. Current medical imaging, 21, e15734056344235. https://doi.org/10.2174/0115734056344235241217155930
93. Asiri, A. A., Shaf, A., Ali, T., Aamir, M., Irfan, M., Alqahtani, S., ... & Alqhtani, S. M. (2023). Brain tumor detection and classification using fine-tuned CNN with ResNet50 and U-Net model: A study on TCGA-LGG and TCIA dataset for MRI applications. Life, 13(7), 1449. https://doi.org/10.3390/life13071449
94. Thahiruddin, M., & Wulandari, A. (2025). CNNs vs. Hybrid Transformers for Brain Tumor Classification on the BRISC Dataset. Jurnal Aplikasi Teknologi Informasi dan Manajemen (JATIM), 6(1), 24-33. DOI:10.31102/jatim.v6i1.3545
95. Müller, D., Soto-Rey, I., & Kramer, F. (2022). Towards a guideline for evaluation metrics in medical image segmentation. BMC research notes, 15(1), 210. https://doi.org/10.1186/s13104-022-06096-y