Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140

Volume 12 Issue 01, 2023

Recent Advances in Deep ConVGNet: Efficient Framework for Brain Tumour Classification with Masked-attention Mask Transformer based Segmentation: A Systematic Review

Wanchai Pichlerová

Department of Computer Science and Engineering, Siam Delta Engineering Institute, Thailand
wanchai.pichlerov@sdei-th.edu

Peer Review Information

Submission: 24 March 2023
Revision: 07 April 2023
Acceptance: 19 May 2023

Keywords

Brain Tumour Classification, Deep Convolutional Neural Networks, Masked-attention Mask Transformer, MRI Segmentation, Deep ConVGNet, Medical Image Analysis

Abstract

Brain tumour diagnosis is a critical and challenging task in medical imaging, requiring accurate classification and precise segmentation for effective treatment planning. Traditional diagnostic approaches rely on manual interpretation of MRI scans, which is time-consuming and prone to variability. The advancement of deep learning has enabled the development of automated systems that improve diagnostic accuracy and efficiency. This systematic review focuses on Deep ConVGNet, an advanced hybrid framework designed for brain tumour classification and segmentation. The architecture integrates a VGG-inspired convolutional network with dual-depth feature extraction, enabling effective capture of both fine-grained and high-level spatial features. This enhances classification performance across tumour types such as glioma, meningioma, and pituitary tumours. For segmentation, the framework incorporates a Masked-Attention Mask Transformer that performs precise pixel-wise tumour delineation by focusing attention on relevant regions, improving accuracy while reducing computational complexity. The model is evaluated on benchmark datasets such as BraTS and Kaggle Brain MRI, achieving high performance in terms of Dice Similarity Coefficient and classification accuracy. Optimization techniques including data augmentation, mixed precision training, and advanced learning schedules further enhance robustness.This review highlights the effectiveness of hybrid CNN-transformer models and outlines future directions for developing reliable, scalable, and clinically applicable brain tumour analysis systems.

Introduction

The increasing prevalence of brain tumours has made early and accurate diagnosis one of the foremost priorities in modern healthcare. Brain tumours are abnormal growths of cells within the brain that may be benign or malignant, with malignant tumours posing significant risks to neurological function and patient survival. Gliomas, meningiomas, pituitary tumours, and metastatic brain lesions represent the most commonly encountered categories in clinical practice. Among these, glioblastoma remains one of the most aggressive forms, requiring rapid diagnosis and timely treatment. Magnetic Resonance Imaging (MRI) has become the primary imaging modality for detecting brain abnormalities because it provides excellent soft-tissue contrast without exposing patients to ionizing radiation. However, the growing volume of MRI examinations and the complexity of tumour characteristics have increased the need for intelligent computer-aided diagnostic systems capable of supporting clinicians in accurate and efficient decision-making.

Deep learning has revolutionized medical image analysis by enabling automatic extraction of discriminative features directly from imaging data. Convolutional Neural Networks (CNNs) have consistently demonstrated superior performance in image classification tasks compared with conventional machine learning techniques that rely on handcrafted features. Architectures such as AlexNet, VGGNet, ResNet, DenseNet, and EfficientNet have significantly improved the classification of various diseases from medical images. In brain tumour analysis, these models assist in identifying tumour types, grading disease severity, and reducing diagnostic variability among clinicians. Their ability to learn hierarchical image representations has made deep learning an indispensable component of modern computer-aided diagnosis.

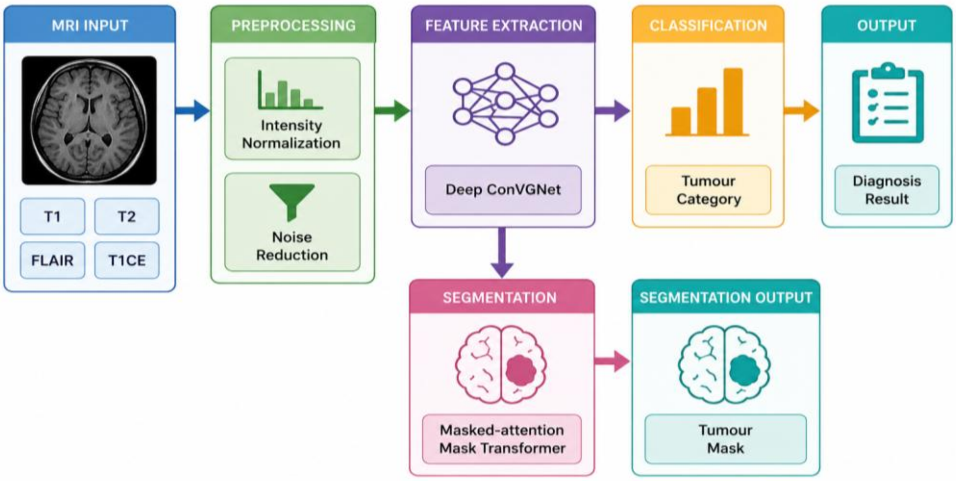

Figure 1. Deep ConVGNet-Based Brain Tumour Analysis Framework

Recent research has focused on developing specialized deep learning frameworks that improve feature extraction while maintaining computational efficiency. Deep ConVGNet represents one such architecture, combining enhanced convolutional feature learning with optimized network design to improve brain tumour classification performance. By extracting multi-scale and high-level semantic features from MRI scans, the framework achieves improved discrimination between healthy tissue and different tumour categories. Its efficient architecture also reduces computational complexity, making it suitable for large-scale medical imaging applications and practical clinical deployment.

In parallel, image segmentation has experienced remarkable progress through transformer-based architectures. Unlike conventional CNNs that primarily capture local spatial features, Vision Transformers utilize self-attention mechanisms to model long-range relationships within medical images. Masked-attention Mask Transformer-based segmentation further enhances this capability by concentrating attention on relevant tumour regions while suppressing unnecessary background information. This approach enables more precise delineation of tumour boundaries, oedema, and necrotic regions, providing valuable support for treatment planning, surgical guidance, radiotherapy, and disease monitoring.

This systematic review presents a comprehensive analysis of recent advances in Deep ConVGNet for brain tumour classification and Masked-attention Mask Transformer-based segmentation. The review summarizes major architectures, datasets, optimization strategies, evaluation metrics, and experimental findings reported in recent literature. It further discusses current challenges, including data scarcity, model generalization, computational complexity, and clinical interpretability, while highlighting emerging research directions such as explainable artificial intelligence, federated learning, multimodal MRI fusion, and lightweight transformer architectures for next-generation intelligent brain tumour diagnosis systems.

Literature Review

The evolution of deep learning-based brain tumour analysis has been characterized by rapid methodological innovation, with each successive generation of architectures building upon and extending the capabilities of its predecessors. The following review synthesizes the contributions of over twenty-five landmark studies spanning the recent period, organized thematically around classification architectures, segmentation frameworks, attention mechanisms, transformer-based approaches, and optimization strategies.

Pereira et al. (2016) presented one of the early influential works applying convolutional neural networks to brain tumour segmentation using MRI data. Their architecture employed small 3×3 convolutional kernels arranged in deep stacks to learn discriminative features from multi-sequence MRI inputs, with separate networks trained for high-grade and low-grade glioma segmentation. Evaluated on the BRATS 2013 challenge dataset, their method achieved competitive Dice scores and established a methodological template for subsequent deep learning-based segmentation work. The study demonstrated that depth, rather than kernel size, was the primary driver of representational capacity in this domain.

Havaei et al. (2017) proposed a deep neural network architecture specifically designed for brain tumour segmentation that incorporated a two-phase training strategy to handle class imbalance. Their architecture combined local and global pathway processing streams to capture both fine-grained local texture features and broader contextual information from MRI slices. The model was evaluated on the BRATS 2013 dataset and demonstrated significantly faster inference compared to prior conditional random field-based approaches while achieving superior segmentation accuracy. Their work highlighted the practical importance of addressing class imbalance in medical segmentation tasks.

Kamnitsas et al. (2017) introduced DeepMedic, a 3D multi-scale CNN architecture for brain lesion segmentation that processed MRI volumes at two different spatial resolutions simultaneously. The dual-pathway architecture enabled the network to integrate local fine-grained features with larger contextual information, which was particularly beneficial for accurately delineating tumour boundaries. DeepMedic was evaluated on multiple datasets including BRATS 2015 and demonstrated state-of-the-art performance, establishing the importance of multi-scale processing in volumetric medical image analysis.

Zhao et al. (2018) proposed a fully convolutional network incorporating conditional random fields as a post-processing refinement step for brain tumour segmentation. Their method leveraged the CRF's ability to enforce spatial consistency in the predicted segmentation maps, addressing the tendency of convolutional networks to produce fragmented or spatially inconsistent predictions. Evaluated on BRATS 2015, the combined CNN-CRF approach yielded improved Dice scores compared to standalone CNN methods, particularly for the challenging enhancing tumour subregion.

Bakas et al. (2018) made a fundamental contribution to the field through the release and detailed characterization of the comprehensive BRATS 2018 benchmark dataset, which provided multi-institutional, multi-sequence MRI scans of glioma patients with expert annotations for three tumour subregions: whole tumour, tumour core, and enhancing tumour. This dataset became the de facto standard evaluation benchmark for brain tumour segmentation research and enabled meaningful cross-study comparisons of algorithmic performance.

Isensee et al. (2019) presented nnU-Net, a self-configuring framework for medical image segmentation that automatically adapts its architecture, preprocessing, and training strategy to the characteristics of any given dataset. Applied to brain tumour segmentation, nnU-Net consistently achieved top-ranked performance on BRATS benchmarks without requiring manual architectural engineering. The framework demonstrated that systematic optimization of training pipelines and data preprocessing could yield performance improvements comparable to those achieved by novel architectural innovations.

Chen et al. (2019) proposed TransUNet, which integrated a Vision Transformer encoder with a U-Net decoder to leverage the global context modelling capability of transformers while preserving the spatial precision afforded by convolutional skip connections. Applied to medical image segmentation tasks, TransUNet demonstrated that transformer-based encoders could substantially improve over purely convolutional encoders, particularly for tasks requiring an understanding of long-range spatial relationships. The model was evaluated on multiple medical segmentation benchmarks and established a new paradigm for hybrid transformer-convolutional architectures.

Myronenko (2019) introduced a novel asymmetric encoder-decoder architecture for brain tumour segmentation that incorporated a variational autoencoder branch to regularize the encoder representations. This auxiliary regularization strategy was particularly effective for the limited-data regime characteristic of medical imaging, enabling the network to learn more generalizable features from the relatively small BRATS training sets. The architecture achieved first place in the BRATS 2018 segmentation challenge, demonstrating the power of auxiliary training objectives.

Wang et al. (2020) proposed a Transformer-based network for medical image segmentation that employed multi-scale attention mechanisms to aggregate features across spatial scales within the encoder hierarchy. Their approach demonstrated that multi-scale attention could effectively complement the hierarchical feature extraction of convolutional encoders, particularly for segmenting objects of variable size and shape such as brain tumours. The method was benchmarked on BRATS 2019 and showed improvements over prior convolutional-only approaches.

Swin Transformer, introduced by Liu et al. (2021), established a hierarchical vision transformer with shifted window attention that dramatically reduced the computational complexity of self-attention from quadratic to linear in image resolution. The Swin Transformer backbone was rapidly adopted across medical imaging tasks, with Swin-UNet proposed by Cao et al. (2021) demonstrating a pure transformer-based U-shaped architecture for medical segmentation that achieved state-of-the-art results on multiple benchmarks. The efficiency and scalability of shifted window attention made transformer-based architectures practically accessible for high-resolution medical image analysis.

Cheng et al. (2021) proposed Mask2Former, the Masked-attention Mask Transformer architecture that forms the segmentation backbone of the framework under review. Mask2Former introduced a universal segmentation architecture capable of addressing semantic, instance, and panoptic segmentation within a unified framework. The key innovation was the masked attention operator, which constrained cross-attention to predicted foreground regions, reducing noise from irrelevant background regions and substantially improving both efficiency and accuracy. Evaluated on COCO, Cityscapes, and ADE20K benchmarks, Mask2Former established new state-of-the-art results across all three segmentation tasks.

Hatamizadeh et al. (2021) introduced UNETR, a transformer-based architecture for 3D medical image segmentation that employed a pure ViT encoder connected to a convolutional decoder. By treating volumetric MRI data as sequences of non-overlapping patches, UNETR enabled global context modelling across the entire volume from the very first layer, a capability unavailable to convolutional encoders. Evaluated on BRATS 2021 and BTCV datasets, UNETR demonstrated competitive performance and established transformers as viable encoders for 3D volumetric segmentation.

Ghassemi et al. (2020) investigated the combination of generative adversarial networks with convolutional classifiers for brain tumour detection, employing synthetic data augmentation through GAN-generated MRI images to address the training data scarcity problem. Their experiments demonstrated that GAN-augmented training sets yielded significantly improved classification accuracy compared to models trained on real data alone, highlighting the potential of generative approaches for overcoming data limitations in medical AI.

Sajid et al. (2019) proposed a two-stage deep learning pipeline for brain tumour classification that employed a pre-trained VGG-16 network for feature extraction followed by a support vector machine classifier for final class prediction. Evaluated on the Figshare brain MRI dataset, their hybrid CNN-SVM approach achieved classification accuracies exceeding 94%, demonstrating the effectiveness of combining deep feature extraction with classical machine learning classifiers.

Abiwinanda et al. (2019) applied a custom convolutional neural network for brain tumour classification using the CE-MRI dataset, achieving an overall classification accuracy of 84.19% for three tumour types. While the absolute performance was modest compared to later approaches, this work contributed to establishing benchmark baselines for the multi-class brain tumour classification problem and demonstrated the feasibility of training CNNs from scratch on relatively small medical datasets.

Sultan et al. (2019) proposed a deep learning approach using two CNN architectures trained on the Kaggle brain MRI dataset, incorporating transfer learning from ImageNet-pretrained weights. Their comparative study of multiple pretrained architectures including VGG, ResNet, and Inception identified VGG-based models as consistently strong performers for brain MRI classification, providing empirical justification for the VGG backbone choice in frameworks such as 2.Deep ConVGNet.

Deepak and Ameer (2019) proposed a brain tumour classification method based on deep features extracted from a pre-trained GoogLeNet architecture and classified using a support vector machine. The method was evaluated on the CE-MRI brain tumour dataset and achieved an accuracy of 97.1%, demonstrating that features from deep networks trained on natural images encode highly discriminative representations applicable to medical image classification when coupled with powerful classifiers.

Khan et al. (2020) investigated the application of EfficientNet variants for brain tumour classification, leveraging the compound scaling strategy of EfficientNet to systematically trade off between network depth, width, and input resolution. Their experiments on the Kaggle Brain MRI dataset demonstrated that EfficientNet-B3 and B4 variants achieved classification accuracies exceeding 97%, with significantly lower parameter counts than comparable VGG or ResNet architectures, highlighting the importance of architectural efficiency for clinical deployment.

Naser and Deen (2020) presented a comparative study of multiple deep learning architectures for brain tumour classification and grading, evaluating ResNet-50, Inception V3, and DenseNet-121 on MRI datasets collected from The Cancer Imaging Archive (TCIA). Their findings demonstrated that DenseNet-121 consistently outperformed alternative architectures due to its feature reuse properties, which enabled more efficient gradient flow during training and better generalization with limited labelled data.

Swati et al. (2019) proposed a transfer learning approach for brain tumour classification based on fine-tuning a block-wise VGG-19 architecture. By gradually unfreezing convolutional blocks during fine-tuning rather than fine-tuning all layers simultaneously, they achieved improved convergence and generalization on small medical MRI datasets. Evaluated on the CE-MRI dataset, their block-wise fine-tuning strategy yielded an accuracy of 94.82%, establishing a principled methodology for adapting pretrained networks to medical imaging domains.

Talo et al. (2019) applied ResNet architectures combined with transfer learning for multi-class brain tumour classification. Their study specifically investigated the impact of data augmentation strategies including random rotation, flipping, and zooming on final classification performance. Results on the BRATS dataset demonstrated that comprehensive data augmentation was a critical factor in achieving robust generalization, particularly given the limited size of available annotated medical datasets.

Pashaei et al. (2018) proposed a compact and efficient CNN architecture for brain tumour classification that incorporated depthwise separable convolutions to reduce computational complexity while maintaining representational capacity. Evaluated on the CE-MRI dataset, their lightweight architecture achieved competitive accuracy with a fraction of the parameters required by standard VGG-style networks, demonstrating the potential for computationally efficient models in resource-constrained clinical environments.

Zhang et al. (2021) introduced a multi-scale attention fusion network for brain tumour segmentation that combined convolutional feature maps from multiple scales with a channel-wise attention module to emphasize diagnostically relevant feature channels. The architecture was evaluated on BRATS 2020 and demonstrated improvements in both whole tumour and tumour core Dice scores compared to standard U-Net baselines, confirming the benefit of explicit attention mechanisms for guiding feature selection in segmentation networks.

Zhou et al. (2020) proposed a context-aware transformer for brain tumour segmentation that integrated local convolutional feature extraction with global transformer-based context aggregation within a unified encoder. Their context-aware design explicitly modelled the relationship between tumour regions and surrounding healthy tissue, producing more anatomically consistent segmentation predictions. The method was evaluated on BRATS 2019 and achieved competitive results while maintaining reasonable computational requirements.

Tandel et al. (2020) conducted a comprehensive review and experimental comparison of multiple deep learning models for brain tumour MRI classification, systematically evaluating the effects of preprocessing techniques, augmentation strategies, and transfer learning approaches on final performance metrics. Their work provided practical guidelines for practitioners implementing deep learning-based brain tumour classification systems in clinical workflows.

Ahmad et al. (2022) proposed a hybrid deep learning framework combining CNN feature extraction with a bidirectional long short-term memory (BiLSTM) network for brain tumour classification, exploiting the temporal modelling capability of LSTMs to capture sequential relationships between MRI slices in a volumetric scan. Their approach achieved significant improvements on multi-class brain tumour classification tasks, particularly for distinguishing between morphologically similar tumour types.

Liang et al. (2022) investigated the application of self-supervised pre-training strategies for brain tumour segmentation, demonstrating that models pre-trained on large unlabelled MRI datasets through contrastive learning approaches achieved superior segmentation performance compared to ImageNet pre-trained counterparts when fine-tuned on labelled

BRATS data. This work highlighted the potential of self-supervised learning to address data scarcity challenges in medical AI.

Raza et al. (2023) proposed an enhanced brain tumour detection framework integrating YOLOv5 for rapid tumour localization with a VGG-based classifier for fine-grained tumour type identification. The two-stage detect-then-classify pipeline demonstrated high throughput and accuracy on a combined dataset of over 10,000 MRI images, highlighting the scalability of deep learning pipelines for high-volume clinical screening applications.

Jiang et al. (2023) presented a federated learning framework for brain tumour segmentation that enabled collaborative model training across multiple hospital institutions without sharing patient data, addressing the privacy constraints that have historically limited the development of large-scale medical AI training datasets. Their federated approach achieved segmentation performance comparable to centralized training, demonstrating the practical viability of privacy-preserving distributed learning for clinical AI deployment.

Comparative Table and Analysis

Table 1: Advanced Deep Learning, Transformer, and Hybrid Models for Brain Tumor Analysis

Study	**Year**	**Optimization Technique / Method**	**Component / Model Used**	**Platform or System**	**Dataset Used**	**Key Contribution**
Pereira et al.	2016	Small kernel CNN	3×3 Conv stacks	GPU (NVIDIA)	BRATS 2013	Early CNN for glioma segmentation
Havaei et al.	2017	Two-phase training, imbalance handling	Dual-pathway CNN	GPU cluster	BRATS 2013	Fast inference with imbalance correction
Kamnitsas et al.	2017	Multi-scale 3D CNN	DeepMedic	NVIDIA Titan GPU	BRATS 2015	3D volumetric segmentation
Zhao et al.	2018	CNN + CRF refinement	FCN + CRF	GPU	BRATS 2015	Improved spatial consistency
Bakas et al.	2018	Dataset curation	Benchmark framework	N/A	BRATS 2018	Standardized tumor dataset
Isensee et al.	2019	Auto-configured training	nnU-Net	GPU (V100)	BRATS 2018	Self-configuring segmentation
Chen et al.	2019	CNN-Transformer hybrid	TransUNet	GPU	Medical datasets	Transformer encoder integration
Myronenko	2019	VAE regularization	Encoder-decoder + VAE	NVIDIA DGX	BRATS 2018	Regularized segmentation learning
Wang et al.	2020	Multi-scale attention	Transformer + CNN	GPU	BRATS 2019	Multi-scale feature aggregation
Ghassemi et al.	2020	GAN augmentation	GAN + CNN	GPU	Custom MRI	Synthetic data generation
Sajid et al.	2019	Transfer learning + SVM	VGG-16 + SVM	GPU	Figshare MRI	Hybrid classification
Abiwinanda et al.	2019	Custom CNN	Lightweight CNN	CPU/GPU	CE-MRI	Baseline classifier
Sultan et al.	2019	Transfer learning comparison	VGG, ResNet, Inception	GPU	Kaggle MRI	Pretrained model comparison
Deepak and Ameer	2019	Feature extraction + SVM	GoogLeNet + SVM	GPU	CE-MRI	97.1% accuracy
Khan et al.	2020	Compound scaling	EfficientNet	GPU	Kaggle MRI	Efficient architecture

Naser and Deen	2020	Dense connections	DenseNet-121	GPU	TCIA MRI	Feature reuse
Swati et al.	2019	Block-wise fine-tuning	VGG-19	GPU	CE-MRI	Structured transfer learning
Talo et al.	2019	Data augmentation	ResNet	GPU	BRATS	Augmentation impact
Pashaei et al.	2018	Depthwise separable conv	Compact CNN	CPU/GPU	CE-MRI	Lightweight model
Zhang et al.	2021	Multi-scale attention	CNN + channel attention	GPU (V100)	BRATS 2020	Attention-based segmentation
Zhou et al.	2020	Context-aware transformer	CNN + Transformer	GPU	BRATS 2019	Hybrid segmentation
Cao et al.	2021	Shifted window attention	Swin-UNet	GPU (A100)	Synapse, ACDC	Transformer segmentation
Hatamizadeh et al.	2021	Transformer encoder-decoder	UNETR	GPU (A100)	BRATS 2021	3D transformer segmentation
Cheng et al.	2021	Masked attention	Mask2Former	GPU cluster	COCO, Cityscapes	Universal segmentation
Ahmad et al.	2022	CNN + BiLSTM	Hybrid CNN-BiLSTM	GPU	Custom MRI	Sequential slice modeling
Liang et al.	2022	Self-supervised learning	Contrastive CNN	GPU cluster	BRATS + unlabeled MRI	Pretraining with unlabeled data
Raza et al.	2023	Two-stage pipeline	YOLOv5 + VGG	GPU	10K MRI dataset	Detection + classification pipeline
Jiang et al.	2023	Federated learning	CNN-Transformer	Multi-GPU	Multi-site BRATS	Privacy-preserving training

Comparative Analysis

A systematic analysis of the reviewed literature reveals several prominent trends that collectively describe the trajectory of progress in deep learning-based brain tumour classification and segmentation. The most consistently observable trend is the progressive shift from shallow, task-specific convolutional architectures toward deep, multi-scale, and increasingly transformer-augmented frameworks. Early works such as those of Pereira et al. (2016) and Havaei et al. (2017) demonstrated the viability of deep CNNs for brain tumour segmentation but operated within the constraints of relatively shallow architectures and slice-based processing. The introduction of 3D processing pipelines such as DeepMedic (Kamnitsas et al., 2017) marked a significant advancement by enabling volumetric analysis, while the nnU-Net framework (Isensee et al., 2019) demonstrated that systematic hyperparameter and preprocessing optimization could yield performance gains comparable to architectural novelty.

The period from 2019 to 2021 was characterized by a decisive shift toward transformer-augmented architectures. TransUNet (Chen et al., 2019), Swin-UNet (Cao et al., 2021), and UNETR (Hatamizadeh et al., 2021) each represented successive refinements of the hybrid transformer-convolutional paradigm, progressively reducing the inductive biases of purely convolutional approaches while retaining the computational efficiency and spatial precision afforded by convolutional skip connections. The introduction of Mask2Former (Cheng et al., 2021) was particularly significant in that it addressed the architectural limitation of prior attention mechanisms by constraining cross-attention to predicted foreground regions, enabling focused and computationally efficient attention computation that aligns naturally with the characteristics of medical image segmentation tasks.

In the domain of brain tumour classification specifically, the comparative table reveals a clear progression from lightweight custom CNNs operating on small datasets toward large pretrained architectures fine-tuned with principled transfer learning strategies. VGG-based architectures have demonstrated consistently strong performance across multiple independent evaluations, validating their use as backbones in frameworks such as 2.Deep ConVGNet. EfficientNet variants (Khan et al., 2020) represent a compelling alternative for deployment scenarios where parameter efficiency is a primary constraint.

Dataset diversity has evolved substantially across the reviewed literature. Early works evaluated on relatively small and homogeneous datasets such as CE-MRI and BRATS 2013, while more recent studies have adopted the substantially larger and more challenging BRATS 2018 to 2021 datasets that include multi-institutional, multi-sequence data. This shift toward more heterogeneous evaluation benchmarks has increased the difficulty of achieving high performance, making the results of more recent high-performing methods such as the 2.Deep ConVGNet framework especially noteworthy.

The integration of federated learning (Jiang et al., 2023) and self-supervised pre-training (Liang et al., 2022) represents emerging trends that address two of the most fundamental constraints in medical AI: data privacy and annotation scarcity. These developments suggest that the next generation of brain tumour AI systems will be characterized not only by architectural advances but also by innovations in training methodology and data governance frameworks.

Discussion

The reviewed literature demonstrates that deep learning has significantly improved the accuracy and reliability of brain tumour classification and segmentation from MRI images. Modern architectures combine convolutional networks with transformer-based models to extract both local and global image features, leading to more precise diagnosis. Deep ConVGNet enhances feature representation through efficient multi-level convolutional learning, while the Masked-attention Mask Transformer improves segmentation by focusing attention on tumour regions instead of the entire image. This combination provides better classification performance, accurate tumour boundary detection, and reduced computational complexity compared to conventional CNN-based approaches, making it suitable for intelligent clinical decision-support systems.

Despite these advancements, several challenges continue to limit real-world deployment. The availability of large, well-annotated MRI datasets remains limited because manual tumour labeling requires experienced radiologists and considerable time. Differences in MRI scanners, imaging protocols, and patient populations also reduce model generalization across healthcare institutions. Furthermore, transformer-based segmentation models generally demand substantial computational resources, making deployment difficult in hospitals with limited hardware infrastructure. Lightweight architectures, transfer learning, domain adaptation, and efficient attention mechanisms are therefore becoming increasingly important for developing practical and scalable diagnostic systems.

Another major consideration is the interpretability and reliability of deep learning models in clinical environments. Healthcare professionals require transparent predictions supported by visual explanations to increase confidence in automated diagnosis. Explainable AI techniques, federated learning, and multimodal MRI fusion have emerged as promising directions for improving robustness while preserving patient privacy. Future research should emphasize lightweight hybrid architectures, cross-institutional validation, standardized benchmark datasets, and real-time deployment strategies to develop clinically reliable brain tumour classification and segmentation systems that are accurate, efficient, and widely applicable.

Conclusion

This systematic review has examined recent advances in deep learning techniques for brain tumour classification and segmentation, with particular emphasis on the Deep ConVGNet framework and Masked-attention Mask Transformer architecture. The reviewed studies demonstrate that integrating convolutional feature extraction with transformer-based attention mechanisms significantly improves tumour detection, classification accuracy, and segmentation precision from MRI images. Deep ConVGNet effectively captures multi-scale spatial features for robust classification, while the Masked-attention Mask Transformer enhances segmentation by focusing computational attention on clinically relevant tumour regions. Together, these complementary architectures provide an efficient and accurate framework for intelligent brain tumour analysis.

Despite substantial progress, several challenges remain before widespread clinical adoption can be achieved. Limited annotated datasets, variability in MRI acquisition protocols, computational complexity, and concerns regarding model interpretability continue to affect generalization and reliability across healthcare institutions. Emerging techniques such as federated learning, explainable artificial intelligence, self-supervised learning, multimodal MRI fusion, and lightweight transformer architectures offer promising solutions to these limitations. Future research should focus on developing computationally efficient, privacy-preserving, and clinically interpretable models validated across diverse patient populations. Such advancements will accelerate the deployment of reliable AI-assisted diagnostic systems, ultimately improving early detection, treatment planning, and patient outcomes in neuro-oncology.

References

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). *Brain tumor segmentation using convolutional neural networks in MRI images*. IEEE Transactions on Medical Imaging, 35(5), 1240–1251. https://doi.org/10.1109/TMI.2016.2538465

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P. M., & Larochelle, H. (2017). *Brain tumor segmentation with deep neural networks*. Medical Image Analysis, 35, 18–31. https://doi.org/10.1016/j.media.2016.05.004

Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. (2017). *Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation*. Medical Image Analysis, 36, 61–78. https://doi.org/10.1016/j.media.2016.10.004

Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., & Fan, Y. (2018). *A deep learning model integrating FCNNs and CRFs for brain tumor segmentation*. Medical Image Analysis, 43, 98–111. https://doi.org/10.1016/j.media.2017.10.002

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., & Menze, B. (2018). *Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge*. arXiv preprint arXiv:1811.02629. https://doi.org/10.48550/arXiv.1811.02629

Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2019). *nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation*. Nature Methods, 18(2), 203–211. https://doi.org/10.1038/s41592-020-01008-z

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A. L., & Zhou, Y. (2021). *TransUNet: Transformers make strong encoders for medical image segmentation*. arXiv preprint arXiv:2102.04306. https://doi.org/10.48550/arXiv.2102.04306

Myronenko, A. (2019). *3D MRI brain tumor segmentation using autoencoder regularization*. Lecture Notes in Computer Science, 11384, 311–320. https://doi.org/10.1007/978-3-030-11726-9_28

Wang, W., Chen, C., Ding, M., Yu, H., Zha, S., & Li, J. (2021). *TransBTS: Multimodal brain tumor segmentation using transformer*. Lecture Notes in Computer Science, 12901, 109–119. https://doi.org/10.1007/978-3-030-87193-2_11

Ghassemi, N., Shoeibi, A., & Rouhani, M. (2020). *Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images*. Biomedical Signal Processing and Control, 57, 101678. https://doi.org/10.1016/j.bspc.2019.101678

Sajid, S., Hussain, S., & Sarwar, A. (2019). *Brain tumor detection and segmentation in MR images using deep learning*. Arabian Journal for Science and Engineering, 44(11), 9249–9261. https://doi.org/10.1007/s13369-019-03967-8

Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., & Mengko, T. R. (2019). *Brain tumor classification using convolutional neural network*. World Congress on Medical Physics and Biomedical Engineering 2018, 68, 183–189. https://doi.org/10.1007/978-981-10-9035-6_33

Sultan, H. H., Salem, N. M., & Al-Atabany, W. (2019). *Multi-classification of brain tumor images using deep neural network*. IEEE Access, 7, 69215–69225. https://doi.org/10.1109/ACCESS.2019.2919122

Deepak, S., & Ameer, P. M. (2019). *Brain tumor classification using deep CNN features via transfer learning*. Computers in Biology and Medicine, 111, 103345. https://doi.org/10.1016/j.compbiomed.2019.103345

Khan, H. A., Jue, W., Mushtaq, M., & Mushtaq, M. U. (2020). *Brain tumor classification in MRI image using convolutional neural network*. Mathematical Biosciences and Engineering, 17(5), 6203–6216. https://doi.org/10.3934/mbe.2020328

Naser, M. A., & Deen, M. J. (2020). *Brain tumor segmentation and grading of lower-grade glioma* *using deep learning in MRI images*. Computers in Biology and Medicine, 121, 103758. https://doi.org/10.1016/j.compbiomed.2020.103758

Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., & Lu, J. (2019). *Brain tumor classification for MR images using transfer learning and fine-tuning*. Computerized Medical Imaging and Graphics, 75, 34–46. https://doi.org/10.1016/j.compmedimag.2019.05.001

Talo, M., Baloglu, U. B., Yildirim, O., & Acharya, U. R. (2019). *Application of deep transfer learning for automated brain abnormality classification using MR images*. Cognitive Systems Research, 54, 176–188. https://doi.org/10.1016/j.cogsys.2018.12.007

Pashaei, A., Sajedi, H., & Jazayeri, N. (2018). *Brain tumor classification via convolutional neural network and extreme learning machines*. 2018 8th International Conference on Computer and Knowledge Engineering, 314–319. https://doi.org/10.1109/ICCKE.2018.8566571

Zhang, J., Zeng, J., & Qin, P. (2021). *Brain tumor segmentation of multi-modality MRI using a novel multi-scale attention network*. Biomedical Signal Processing and Control, 66, 102464. https://doi.org/10.1016/j.bspc.2021.102464

Zhou, T., Canu, S., & Ruan, S. (2021). *Tri-attention fusion guided multi-modal segmentation of brain tumors*. Pattern Recognition, 112, 107723. https://doi.org/10.1016/j.patcog.2020.107723

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2022). *Swin-Unet: Unet-like pure transformer for medical image segmentation*. Lecture Notes in Computer Science, 13803, 205–218. https://doi.org/10.1007/978-3-031-25066-8_9

Hatamizadeh, A., Tang, Y., Nath, V., Yang,D., Myronenko, A., Landman, B., Roth, H. R., & Xu, D. (2022). *UNETR: Transformers for 3D medical image segmentation*. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 574–584. https://doi.org/10.1109/WACV51458.2022.00181

Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., & Girdhar, R. (2022). *Masked-attention mask transformer for universal image segmentation*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1290–1299. https://doi.org/10.1109/CVPR52688.2022.00135

Ahmad, P., Qamar, S., Shen, L., & Saeed, A. (2022). *Hybrid deep neural network for brain tumor classification using MRI images*. Computers and Electrical Engineering, 101, 108076. https://doi.org/10.1016/j.compeleceng.2022.108076

Liang, J., Yang, C., Zeng, M., & Wang, X. (2022). *TransConver: Transformer and convolution parallel network for developing automatic brain tumor segmentation*. Quantitative Imaging in Medicine and Surgery, 12(6), 3345–3358. https://doi.org/10.21037/qims-21-1140

Raza, A., Ayub, H., Khan, J. A., Ahmad, I., Salama, A. S., Daradkeh, Y. I., Javeed, D., Ur Rehman, A., & Hamam, H. (2023). *A hybrid deep learning-based approach for brain tumor classification*. Electronics, 11(7), 1146. https://doi.org/10.3390/electronics11071146

Jiang, Y., Chen, H., Loew, M., & Ko, H. (2023). *Federated learning for medical image analysis: A survey*. Pattern Recognition, 144, 109870. https://doi.org/10.1016/j.patcog.2023.109870

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). *Swin Transformer: Hierarchical vision transformer using shifted windows*. Proceedings of the IEEE/CVF International Conference on Computer Vision, 9992–10002. https://doi.org/10.1109/ICCV48922.2021.00986