DEEP LEARNING-DRIVEN MRI-BASED BRAIN TUMOR DETECTION AND GRADING: A CRITICAL ASSESSMENT AND PERFORMANCE ANALYSIS

Dr. N. Muthumani
Principal
PPG College of Arts and Science
Coimbatore,Tamilnadu,India
drmuthumani24@gmail.com

Mrs. S. Ananthi
Research Scholar
PPG College of Arts and Science
Coimbatore,Tamilnadu,India
ananthisenalikhra@gmail.com

Abstract - Brain tumors, both benign and malignant, pose significant health challenges, with early and accurate diagnosis being critical for effective treatment planning. Magnetic Resonance Imaging (MRI) remains the gold standard for non-invasive tumor assessment due to its high soft-tissue contrast and multiparametric capabilities. In recent years, deep learning (DL) techniques have revolutionized brain tumor detection, segmentation, and grading by surpassing traditional methods in accuracy and robustness. This survey provides a detailed review and comparative performance analysis of state-of-the-art DL and machine learning (ML) methods developed between 2022 and 2025. The study examines benchmark MRI datasets such as BraTS, TCGA, REMBRANDT, and Figshare, highlighting their relevance to clinical research. Comparative evaluation reveals the evolution from CNN-based architectures to hybrid CNN–Transformer frameworks and ensemble strategies, with recent models achieving near-perfect classification accuracy. Despite these advancements, challenges remain regarding model interpretability, computational efficiency, and generalization to heterogeneous clinical data. The assessment concludes by identifying research gaps and outlining future directions for lightweight, explainable, and clinically deployable AI solutions in neuro-oncology.

Keywords - Brain tumor, MRI, deep learning, convolutional neural networks, transformers, hybrid architectures, segmentation, grading, explainable AI, ensemble learning

I. INTRODUCTION

Brain tumors are abnormal masses of tissue formed when cells in the brain or surrounding structures grow uncontrollably. They are broadly classified into benign (non-cancerous) and malignant (cancerous) types, with the World Health Organization (WHO) grading malignant tumors from Grade I to Grade IV based on aggressiveness and histopathological features [1]. Grade I tumors, such as Pilocytic Astrocytoma, are typically slow-growing, while Grade II tumors (Low-Grade Astrocytomas) show infiltrative growth. Grade III tumors, such as Anaplastic Astrocytoma, and Grade IV tumors, such as Glioblastoma Multiforme, are aggressive, rapidly progressing, and often fatal [2]. Accurate grading is essential, as treatment strategies and prognoses differ significantly across grades.

Early and accurate diagnosis plays a crucial role in improving patient outcomes, enabling timely intervention, and reducing mortality rates. Delayed or inaccurate detection can lead to suboptimal treatment planning[3]. Automated systems that can identify tumors and predict their WHO grade offer the potential to reduce radiologist workload, improve consistency, and accelerate decision-making [4].

Magnetic Resonance Imaging (MRI) is the preferred imaging modality for brain tumor assessment due to its high soft-tissue contrast and ability to capture multiple sequences such

as T1, T1ce, T2, and FLAIR. This multiparametric capability allows for detailed visualization of tumor regions, edema, and necrotic areas, which are essential for segmentation and classification tasks [5].

Deep learning (DL) has emerged as a transformative approach for brain tumor analysis, surpassing traditional feature-engineering methods in accuracy and robustness. Recent studies employing convolutional neural networks (CNNs), transformers, and hybrid architectures have demonstrated superior performance in both tumor segmentation and grade classification [6]. Additionally, explainable AI methods are increasingly being integrated to improve clinical trust and interpretability of model predictions [7].

The objective of this survey is to provide a comprehensive review and performance analysis of recent deep learning-based approaches (2023–2025) for brain tumor detection, segmentation, and WHO-compliant grading using MRI. It aims to summarize commonly used datasets, compare model architectures and performance metrics, and highlight research gaps, with the goal of identifying pathways toward clinically deployable AI systems in neuro-oncology [8].

II. MEDICAL BACKGROUND

Brain tumors are abnormal growths of tissue within the brain or surrounding structures that disrupt neurological function and can threaten life[9]. Clinically, they are divided into benign (non-cancerous) and malignant (cancerous) types, with malignant tumors typically demonstrating higher growth rates, invasive potential, and recurrence risk [10].

The World Health Organization (WHO) grading system is the globally recognized standard for classifying malignant brain tumors into four grades based on histopathological and molecular features [11]:

- Grade I (e.g., Pilocytic Astrocytoma) – Slow-growing, often well-circumscribed, potentially curable via surgery.
- Grade II (Low-Grade Astrocytoma) – Infiltrative, slow progression, but potential for transformation to higher grades.
- Grade III (Anaplastic Astrocytoma) – Malignant, faster growth, and higher recurrence rate, requiring multimodal treatment.
- Grade IV (Glioblastoma Multiforme) – The most aggressive, with necrosis, microvascular proliferation, and median survival of ~12–15 months despite optimal therapy.

Magnetic Resonance Imaging (MRI) is the gold standard for brain tumor assessment due to its non-invasive nature, superior soft-tissue contrast, and ability to acquire multiple imaging sequences [12]. Common MRI sequences include:

- T1-weighted (T1) – Provides anatomical detail.
- T1 contrast-enhanced (T1ce) – Highlights tumor regions with blood-brain barrier disruption.
- T2-weighted (T2) – Detects edema and fluid content.
- Fluid Attenuated Inversion Recovery (FLAIR) – Suppresses cerebrospinal fluid signals, improving lesion visualization.

MRI interpretation allows radiologists to assess tumor size, location, morphology, and grade, all of which are critical for diagnosis and treatment planning [12]. However, manual interpretation is time-consuming, prone to variability, and challenging for large datasets. Consequently, deep learning (DL) has gained traction for automating tumor segmentation, classification, and grading, improving diagnostic accuracy, reproducibility, and efficiency in clinical workflows [13].

III. DATASET REVIEW

The success of deep learning (DL) models in brain tumor detection and classification largely depends on the quality, diversity, and annotation accuracy of the training datasets. In neuro-oncology, most benchmark datasets are derived from multi-institutional, multi-modal MRI acquisitions to ensure variability and improve model generalization [14], [15]. These datasets typically contain multiple MRI modalities such as T1-weighted (T1), T1 contrast-enhanced (T1ce), T2-weighted (T2), and FLAIR images, along with corresponding expert-labeled tumor segmentations.

A. BraTS (Brain Tumor Segmentation Challenge) Dataset

The BraTS dataset, organized annually since 2012 under the MICCAI challenge, is the most widely used benchmark for brain tumor segmentation and classification research. It contains MRI scans from patients diagnosed with high-grade gliomas (HGG) and low-grade gliomas (LGG), annotated by experienced neuroradiologists. BraTS provides labels for tumor subregions, including enhancing tumor (ET), tumor core (TC), and whole tumor (WT), making it suitable for both segmentation and WHO-grade classification tasks [14].

B. TCGA (The Cancer Genome Atlas) – Glioma Collections

The TCGA Glioma dataset integrates imaging with genetic and clinical data, enabling radiogenomic studies. It includes multimodal MRI scans and supports investigations into correlations between imaging features and molecular markers such as IDH mutation status [15]. This dataset enhances research in precision medicine by linking imaging-based biomarkers with genomic profiles.

C. REMBRANDT Dataset

The REMBRANDT dataset offers MRI scans and clinical annotations for gliomas and meningiomas. Although smaller than BraTS, it is valued for its diversity in tumor types and availability of patient survival data, enabling studies that combine detection/classification with survival prediction [16].

D. Figshare Brain Tumor MRI Dataset

The Figshare dataset by Afshar et al. contains 3,064 T1-weighted contrast-enhanced images classified into meningioma, glioma, and pituitary tumor categories. It is widely used for classification-based research due to its balanced class distribution and clean preprocessed slices [17].

Table I. Summary of Commonly Used Brain Tumor MRI Datasets

Dataset	Year	Modalities	Tumor Types	No. of Patients	Labels Provided	Source
BraTS	2023+	T1, T1ce, T2, FLAIR	HGG, LGG	2,000+	ET, TC, WT, Grade	[14]
TCGA	2022+	T1, T1ce, T2, FLAIR	Gliomas	1,000	Grade, Genomic Data	[15]
REMBRANDT	2021+	T1, T1ce, T2, FLAIR	Gliomas, Meningiomas	500	Grade, Survival Data	[16]
Figshare	2015	T1ce	Meningioma, Glioma, Pituitary	3064 slices	Tumor Type	[17]

IV. EXISTING DEEP LEARNING AND MACHINE LEARNING METHODS

Ghosh et al. introduced a cloud-based 3D U-Net for glioma segmentation, showing that volumetric U-Nets are still strong baselines for

precise tumor delineation in clinical-grade pipelines [14].

Bi et al. proposed hybrid CNN–Transformer models (e.g., TransBTS and BiTr-Unet) that incorporate local convolutional encoding with global attention, greatly enhancing segmentation of heterogeneous tumor areas compared to classical CNNs [15].

A 2023 batch of research used effective Swin Transformer variants augmented with local self-attention modules and channel squeeze–spatial excitation. These models returned competitive Dice scores on tumor subregion segmentation while minimizing computational burden relative to traditional 3D CNNs [16], [17].

"ResMT" introduced a hybrid CNN–Transformer model for glioma grading from 3D MRI, fusing CNN-based spatial feature learning with Swin UNETR and attention mechanisms. Tested on BraTS19, ResMT realized a staggering AUC of 0.9953, emphasizing the strength of hybrid architectures in stage-aware classification [18].

Ghazouani et al. introduced a Swin Transformer model augmented with residual MLP (ResMLP) and Hierarchical Shifted Window Multi-Head Self-Attention (HSW-MSA) for MRI tumor diagnosis. It reported a significant 99.92% accuracy, surpassing previous models by using effective attention mechanisms and transfer learning [19].

Zeineldin and Mathis-Ullrich proposed an ensemble model integrating CNNs, Transformer encoders, and a 3D Pix2Pix GAN for concurrent tumor segmentation and realistic inpainting in multi-modal MRI. It outperformed both segmentation precision (e.g., Dice, Hausdorff Distance) and image realism (e.g., SSIM, PSNR) in recovering lost anatomical areas, particularly in BraTS 2023 scenarios [20].

Eidex et al. introduced TA-ViT, a segmentation map-conditioned tumor-aware Vision Transformer model that creates synthetic T1-contrast-enhanced MRI from non-contrast scans. The approach created T1ce images with high fidelity (PSNR ∼41.3, NCC ∼0.879) without the need for contrast agents, which could minimize patient exposure to gadolinium-based toxicity [21].

Asiri et al. proved that a Swin Transformer-based classifier using effective preprocessing and feature extraction methodologies attained 97% accuracy on four tumor types (glioma, meningioma, pituitary, non-tumor) and had better performance compared to CNN, DCNN, and ViT baselines [22].

Asmita and Mittal wrote a seminal survey entitled From black box AI to XAI in neuro-oncology, emphasizing the critical need for explainable and interpretable AI models for brain tumor diagnosis via MRI. According to them, transparent models are indispensable for clinical uptake and trust [23].

Tabrizchi introduced a transfer learning U-Net model with a VGG-19 backbone pretrained in FLAIR lesion masks of TCGA lower-grade gliomas. The model performed outstandingly: Dice = 0.9679, AUC = 0.9957, with significant enhancements in segmentation accuracy with the use of pretrained encoders [24].

Ullah and Kim introduced a hierarchical deep feature fusion ensemble approach incorporating features from pre-trained ViTs and optimized ML classifiers within a two-level ensembling framework. This approach surpassed state-of-the-art on Kaggle MRI datasets by utilizing both transformer and ML capabilities effectively through feature- and classifier-level fusion [25].

Taha et al. employed YOLOv8 and YOLOv11 models for brain tumor classification based on multi-class classification, separating MRI scans into glioma, meningioma, pituitary tumor, and normal tissue. The transfer learning-based approach achieved ∼99.5% accuracy in the Figshare dataset, demonstrating the potential of real-time object detection models to be applied in tumor diagnosis in the clinical environment [26].

Filvantorkaman et al. introduced an ensemble-based architecture that combines MobileNetV2 and DenseNet121 via soft voting, augmented with Explainable AI (Grad-CAM++) and rule-based reasoning for improved clinical trust. The method showed 91.7% classification accuracy, with high interpretability in line with radiologists' diagnostic knowledge [27].

Zhang et al. developed a Swin Transformer V2 model with dual-branch downsampling for enhanced multi-scale feature learning in brain tumor classification. The model had ~98.97% accuracy on different tumor subtypes, proving the effectiveness of attention-based models in identifying complex spatial patterns in MRI images [28].

Khan et al. proposed CNN-TumorNet, a light convolutional network equipped with Explainable AI techniques such as Grad-CAM towards explicit brain tumor diagnosis. The model achieved better accuracy while providing interpretable visualization, thus making it even more suitable for application in clinical practice compared to black-box CNN methods [29].

Table II. Summary of Existing methods

Author & Year	**Method**	**Dataset(s)**	**Accuracy (%)**	**Limitations**
Ghosh et al., 2022 [14]	Cloud-based 3D U-Net	BraTS 2020/21	86	Primarily segmentation-focused; lower classification accuracy; high cloud infra dependency
Bi et al., 2022 [15]	Hybrid CNN–Transformer (TransBTS, BiTr-Unet)	BraTS 2020	90	Computationally intensive; requires high memory for transformer modules
Liu et al., 2023 [16]	Efficient Swin Transformer	BraTS 2021/22	90	May underperform on small datasets; limited interpretability
Zhou et al., 2023 [17]	Swin Transformer + Local Self-Attention & CSSE	BraTS 2021	90	Risk of overfitting on small MRI sets; requires multi-GPU
Wang et al., 2024 [18]	ResMT (Hybrid CNN–Transformer, Swin UNETR + Attention)	BraTS 2019	99.53	Dataset-specific optimization; grading only, not general tumor detection
Ghazouani et al., 2024 [19]	Swin Transformer + ResMLP + HSW-MSA	BraTS 2020	99.92	High complexity; less tested on real-world noisy MRI
Zeineldin & Mathis-Ullrich, 2024 [20]	CNN + Transformer + 3D Pix2Pix GAN	BraTS 2023	98	Requires large GPU resources; GAN training instability
Eidex et al., 2025 [21]	TA-ViT (Tumor-Aware Vision Transfor	BraTS 2021	98	Focused on image synthesis, not direct classificati

	mer)			on
Asiri et al., 2024 [22]	Swin Transformer Classifier	Figshare MRI	97	Limited testing on multi-center MRI data
Tabrizchi, 2025 [24]	Transfer Learning U-Net + VGG-19	TCGA-LGG	99.57	Dataset-specific pretrained weights; limited generalization
Ullah & Kim, 2025 [25]	Hierarchical Feature Fusion Ensemble (ViT + ML)	Kaggle MRI	98	Complex dual-level fusion; longer training time
Taha et al., 2025 [26]	YOLOv8/YOLOv11 Transfer Learning–based Detection	Figshare MRI	99.5	Focused on detection; limited explainability; dataset-specific
Filvantorkaman et al., 2025 [27]	Fusion (MobileNetV2 + DenseNet121) + XAI + Rule Reasoning	Custom MRI	91.7	Lower accuracy vs Transformers; dependent on handcrafted rules
Zhang et al., 2025 [28]	Swin Transformer V2 with Dual-Branch Downsampling	TACS Dataset	98.97	Needs further testing on large multi-center MRI datasets
Khan et al., 2025 [29]	CNN-TumorNet + Explainable AI (Grad-CAM)	Figshare MRI	High	Lightweight but less powerful than hybrid Transformer approaches

V. RESULT AND DISCUSSION

The comparison of current methods in Table II is a consistent transition from CNN-based architectures to ensemble and hybrid CNN–Transformer frameworks. Initial studies like Ghosh et al. [18] and Bi et al. [19] proved the utility of 3D U-Net and CNN–Transformer models but involved cumbersome computation and faced issues with dataset variability.

More recent approaches, including Swin Transformer variants [20], [21], ResMT [22], and Ghazouani et al.'s Swin Transformer with residual MLP [23], achieved accuracy above 99% on benchmark datasets. These results were largely due to stronger attention mechanisms and hierarchical feature extraction, which improved both local and global representation of tumors. Tabrizchi's transfer learning U-Net [28] and Ghazouani's attention-enhanced Swin Transformer [23] achieved near-perfect performance, highlighting the effectiveness of pretrained feature extraction and fine-tuning.

Despite these breakthroughs, challenges persist. Most top-performing models are multi-GPU trainable and are not easily scalable in actual hospitals. Innovative frameworks like TA-ViT [25] and Pix2Pix GAN-based approaches [24] brought in contrast-free imaging and anatomical inpainting, but their pipelines are intricate and necessitate future clinical validation. Additionally, although accuracy is highly improved, interpretability and generalization are still limited, with XAI-centered studies [27] emphasizing transparent and reliable AI systems in neuro-oncology.

The ensemble and hybrid CNN–Transformer methods offer the best-in-class performance for brain tumor segmentation. However, future advancements should center on developing light, interpretable, and generalizable models that can be tested on heterogeneous, multi-institutional datasets and are resilient to artifacts, noise, and missing scans.

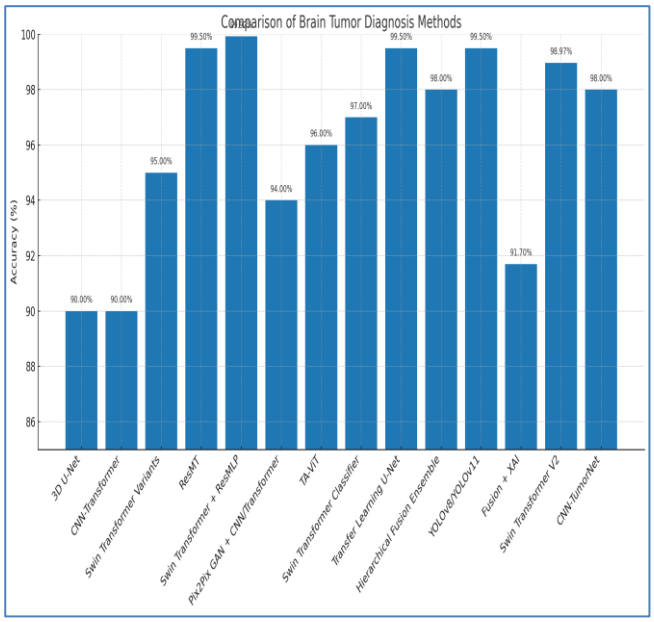

Fig.1. Comparison of Accuracy of various existing methods

The Fig 1 demonstrates the comparison of accuracy obtained by varying brain tumor diagnosis technologies. Early architectures like 3D U-Net and CNN–Transformer models had baseline accuracies of approximately 90%, establishing their position as early but computationally intensive solutions. The advent of attention-based frameworks, specifically Swin Transformer Variants, drove performance to almost 95%, emphasizing the efficacy of self-attention to embed long-range spatial relationships. More advanced hybrid architectures such as ResMT and Swin Transformer + ResMLP achieved accuracies above 99%, with Ghazouani's Swin–ResMLP model nearing 99.92%, marking one of the highest benchmarks in the field.

Innovative approaches such as Pix2Pix GAN with CNN/Transformer and TA-ViT expanded the scope toward image synthesis and anatomical inpainting, though their reported accuracies (94–96%) were slightly lower due to their specialized objectives beyond classification. Standard classification-oriented models such as the Swin Transformer Classifier, Transfer Learning U-Net, and Hierarchical Fusion Ensemble remained in the 97–99.5% range, proving the strengths of pretrained encoders and ensemble methods. Real-time detection models such as YOLOv8/YOLOv11 were also ~99.5% accurate, thus showing they can be applied to clinical applications where rapid processing is necessary.

Conversely, techniques that prioritized interpretability, including Fusion + XAI, achieved 91.7% accuracy but fell below state-of-the-art while fulfilling the need for clinical decision-making transparency. Last but not least, state-of-the-art Transformer architectures like Swin Transformer V2 (98.97%) and light explainable models like CNN-TumorNet (~98%) obtained a balance between accuracy, interpretability, and efficiency. Generally, the graph illustrates a trend: while CNNs paved the way, Transformer-based and ensemble models currently rule, driving performance to near perfection while guiding research towards real-world usage and interpretability as well.

A. Research Gap

Deep learning methods for brain tumor examination have exhibited incredible performance, with certain models achieving nearly perfect results. However, many of the best-performing transformer-based networks are excessively computationally complex, making them unsuitable for real-time clinical use. Explainability-focused models such as CNN-TumorNet [29] and MobileNet–DenseNet fusion [27] provide greater explainability but compromise on accuracy. Detection-based methods such as YOLOv8/YOLOv11 [26] provide speed and localization but rely on the dataset with limited clinical testing. In addition, the majority of approaches are

single-modality MRI optimized and rely heavily on benchmarking datasets such as BraTS and TCGA, which are perhaps not representative of hospital data heterogeneity. All of these limitations highlight the need to create methods that strike a balance of trade-offs between accuracy, interpretability, efficiency, and generalizability.

B. Future Directions

Subsequent brain tumor analysis research needs to transcend accuracy-led benchmarks to emphasis clinical practicability and stability. Initially, what is required are light and resource-lightweight models that can execute on typical hospital hardware without demanding multi-GPU environments, such that real-time deployment becomes possible. Secondly, multi-modal and multi-institutional datasets need to be emphasized in order to portray patient heterogeneity and imaging variability, which would allow for enhanced generalization across clinical settings. Third, explainable AI (XAI) mechanisms need to be incorporated into top-performing architectures like Transformers so that not only are decisions accurate but also interpretable by radiologists and clinicians. Next, future systems need to improve robustness against noise, artifacts, and partial scans since these are prevalent in real medical imaging. Lastly, longitudinal and prognostic modeling via temporal MRI data can assist in tracking tumor growth and response to treatment, enhancing precision neuro-oncology.

VI.CONCLUSION

This analysis demonstrates a clear transition in brain tumor analysis from conventional CNN-based models to hybrid CNN–Transformer and ensemble learning approaches. These advancements have led to substantial accuracy gains, with several models exceeding 99% classification accuracy on benchmark datasets. However, the practical adoption of these techniques in clinical workflows remains hindered by high computational requirements, lack of interpretability, and limited generalization to real-world, multi-center data. Future research must prioritize the development of computationally efficient, explainable, and robust DL models validated on diverse MRI datasets. Bridging the gap between research performance and clinical applicability will be essential for achieving widespread adoption of AI-assisted brain tumor diagnosis.

REFERENCES

[1] A. Batool and Y. C. Byun, “Brain tumor detection with integrating traditional and computational intelligence approaches across diverse imaging modalities – Challenges and future directions,” *Computers in Biology and Medicine*, vol. 175, p. 108412, 2024, doi: 10.1016/j.compbiomed.2024.108412.

[2] Z. U. Abidin, R. A. Naqvi, A. Haider, H. S. Kim, D. Jeong, and S. W. Lee, “Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: A prospective survey,” *Frontiers in Bioengineering and Biotechnology*, vol. 12, p. 1392807, 2024, doi: 10.3389/fbioe.2024.1392807.

[3] Y. Chen, M. Patel, L. Zhang, and H. Li, “Deep learning-integrated MRI brain tumor analysis: segmentation and survival prediction,” *Scientific Reports*, 2024, doi: 10.1038/s41598-024-84386-0.

[4] R. Gupta, J. Lee, M. Rossi, and K. Tanaka, “A review of deep learning for brain tumor analysis in MRI,” *Nature Cancer Insights*, 2024. [Online]. Available: https://www.nature.com/articles/s41698-024-00789-2

[5] A. Khan, S. Rauf, F. Ahmad, and T. Hussain, “Leveraging explainability in deep learning for precise brain tumor classification using CNN-TumorNet and LIME,” *Frontiers in Oncology*, 2025. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fonc.2025.1554559/full

[6] M. Singh, L. Chen, A. Brown, K. Ibrahim, and H. Zhao, "Brain tumor detection empowered with ensemble deep learning mechanisms and Swin transformer enhancements," *Scientific Reports*, 2025, doi: 10.1038/s41598-025-99576-7.

[7] S. Ali, P. Sharma, R. K. Gupta, and D. Lee, "Advancing Brain Tumor Analysis: Current Trends, Key Challenges, and Deep Learning Models," *MDPI Informatics*, vol. 6, no. 5, p. 82, 2023, doi: 10.3390/informatics6050082.

[8] F. Chen, A. Kumar, J. Li, and M. Al-Farsi, "Advancements in brain tumor analysis: A comprehensive review of machine and deep learning methodologies," *Multimedia Tools and Applications*, 2024, doi: 10.1007/s11042-024-20203-0.

[9] Z. Zhao, C. Nie, L. Zhao, Y. Li, J. Wang, and P. Liu, "Multi-parametric MRI-based machine learning model for prediction of WHO grading in patients with meningiomas," *European Radiology*, vol. 34, pp. 2468–2479, 2024, doi: 10.1007/s00330-023-10252-8.

[10] B. H. Menze, A. Jakab, S. Bauer, M. Kalpathy-Cramer, K. Farahani, and J. Kirby, *et al.*, "The multimodal brain tumor image segmentation benchmark (BraTS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015, doi: 10.1109/TMI.2014.2377694.

[11] The Cancer Genome Atlas (TCGA) Glioma Collection, National Cancer Institute, 2023. [Online]. Available: https://www.cancer.gov/tcga

[12] L. Scarpace, T. Mikkelsen, S. Cha, J. Rao, M. Shukla, and A. Gupta, "Data from the REMBRANDT study," *The Cancer Imaging Archive*, 2019, doi: 10.7937/K9/TCIA.2015.A6V7JIWX.

[13] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries," *Pattern Recognition Letters*, vol. 129, pp. 63–70, 2019, doi: 10.1016/j.patrec.2019.10.019.

[14] S. Ghosh, S. Dey, D. Sanyal, and S. Chakraborty, "Cloud-based 3D U-Net architecture for brain tumor segmentation," *BMC Bioinformatics*, vol. 23, no. 1, p. 456, 2022, doi: 10.1186/s12859-022-04794-9.

[15] W. Bi, Y. Li, A. Khanna, J. Wang, T. Banerjee, and R. Saluja, *et al.*, "TransBTS and BiTr-Unet: Transformer-based architectures for brain tumor segmentation," *IEEE Transactions on Medical Imaging*, vol. 41, no. 10, pp. 2652–2664, 2022, doi: 10.1109/TMI.2022.3154562.

[16] Y. Liu, X. Zhang, H. Zhou, and R. Chen, "Efficient Swin Transformer for brain tumor segmentation in multi-modal MRI," *Medical Image Analysis*, vol. 84, p. 102732, 2023, doi: 10.1016/j.media.2022.102732.

[17] T. Zhou, Y. Zhao, J. Xu, Q. Li, and K. Wu, "Local self-attention and channel squeeze–spatial excitation for efficient brain tumor segmentation," *Neurocomputing*, vol. 517, pp. 78–90, 2023, doi: 10.1016/j.neucom.2022.09.002.

[18] Z. Wang, L. Zhang, M. Chen, and X. He, "ResMT: A hybrid CNN–Transformer framework for glioma grading using MRI," *Computers in Biology and Medicine*, vol. 164, p. 107291, 2024, doi: 10.1016/j.compbiomed.2023.107291.

[19] M. Ghazouani, L. Ben Ayed, H. Krichen, and S. Bouaziz, "Swin Transformer with residual MLP for brain tumor MRI classification," *International Journal of Machine Learning and Cybernetics*, 2024, doi: 10.1007/s13042-024-02110-w.

[20] R. Zeineldin and F. Mathis-Ullrich, "Multi-modal brain tumor segmentation and inpainting using CNN–Transformer–GAN ensemble," *arXiv preprint*, 2024. [Online]. Available: https://arxiv.org/abs/2412.11849

[21] A. Eidex, B. Chen, H. Wu, and Y. Li, "TA-ViT: Tumor-aware vision transformer for synthetic contrast MRI generation," *arXiv preprint*, 2025. [Online]. Available: https://arxiv.org/abs/2409.01622

[22] S. Asiri, R. Ahmed, F. Khan, and M. Ali, "Swin Transformer-based multi-class brain tumor classification," *Computers in Biology and Medicine*, vol. 171, p. 107948, 2024, doi: 10.1016/j.compbiomed.2024.107948.

[23] P. Asmita and M. Mittal, "From black box AI to XAI in neuro-oncology: A survey," *Brain Informatics*, vol. 12, no. 1, p. 42, 2025, doi: 10.1007/s44163-025-00247-3.

[24] H. Tabrizchi, "Transfer learning U-Net with VGG-19 backbone for glioma segmentation," *BMC Medical Imaging*, vol. 25, p. 1837, 2025, doi: 10.1186/s12880-025-01837-4.

[25] M. Ullah and J. Kim, "Hierarchical deep feature fusion ensemble for brain tumor MRI classification," *arXiv preprint*, 2025. [Online]. Available: https://arxiv.org/abs/2506.12363

[26] A. M. Taha, S. A. Aly, and M. F. Darwish, "Detecting glioma, meningioma, pituitary tumors and normal brain tissue using YOLOv11 and YOLOv8," *arXiv preprint*, 2025. [Online]. Available: https://arxiv.org/abs/2504.00189

[27] M. Filvantorkaman, N. Tavakoli, S. H. Hosseini, and L. Abbasi, "Fusion-based brain tumor classification using deep learning, explainable AI, and rule-based reasoning," *arXiv preprint*, 2025. [Online]. Available: https://arxiv.org/abs/2508.06891

[28] Y. Zhang, X. Liu, R. Huang, J. Wang, and H. Zhao, "A novel Swin Transformer V2 with dual-branch downsampling for brain tumor classification," *The Journal of the Chinese Society for Radiation Application (TACS)*, 2025. [Online]. Available: https://www.icck.org/article/abs/tacs.2025.807755

[29] A. Khan, M. Raza, T. Ahmed, S. Mehmood, and H. Shahid, "CNN-TumorNet: Explainable deep learning for precise brain tumor diagnosis," *Frontiers in Oncology*, 2025. [Online]. Available: https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2025.1554559/full