International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Brain Tumor Detection and MRI Images Classification Using Advance Deep Learning: Advancements, Difficulties, and Clinical Deployment

Vikas Kumar Singh

[1]KPGU University, Krishna School of Diploma Studies, Varnama, Vadodara, Gujarat, India
Email: *vikas.s0905[at]gmail.com*

Abstract: ***A brain tumor is one of the most cancerous diseases in the world, which can be identified on time and significantly treated using magnetic resonance imaging. In the last few years, deep learning models have provided the expected performance in automatically detecting and classifying brain tumors from MRI scans, repeatedly attaining accuracies more than 95% on measurable point datasets. But so many of these strategies fail when utilized in real-time clinical environments because of deviations in the quality of the image, differences in scanners, and limited numbers of resources in several healthcare environments. This review aims to give substantial enhancements in deep learning for brain tumor detection and classification during 2025 and 2026. It includes a number of methods covering traditional hybrid classifiers, convolutional neural networks (CNNs), and explainable AI, which utilizes techniques like GRAD-CAM to make more transparent decisions. Special concentration is provided to difficulties like low performance on various datasets and low-resource data (e.g., BraTS-Africa), the requirement of effective models appropriate for deployments, and establishing clinical trust via enhanced interpretability. After carefully analyzing the latest findings, this paper discovers alternative drawbacks in generalization, cost of computation, and its practical application. This paper advises a promising way, such as adaptive learning and integrated explainable technologies, to assist in more reliable and acquirable brain tumor identification in real-world scenarios.***

Keywords: Brain tumor detection, Magnetic Resonance Imaging (MRI), Deep Learning, Explainable AI, Convolutional Neural Network (CNN), Domain generalization, GRAD-CAM, clinical translation, BrsTS-Africa.

1. Introduction

Brain tumors are the most serious and death-causing diseases impacting the central nervous system [1]. These are caused by the unrestrained expansion of abnormal cells within the brain and nearby tissues, which can be classified as benign or malignant. The most usual types discovered in medical practice are represented by gliomas, meningiomas, and pituitary tumors [2]. In spite of advancements in treatments, huge mortality rates persist, especially in aggressive forms like glioblastoma. Therefore, early and correct detection is mandatory for effective diagnosis, treatment, surgery planning, and patient survival [3].

Because of its soft tissue contrast capability to provide anatomical data, Magnetic Resonance Imaging (MRI) remains the most widely used imaging technology [4]. Traditionally, a radiologist was manually analyzing the MRI scans to identify the disease and classify tumors. This process was time-consuming, subjective, and heavily dependent on radiologists' experience, and it became challenging when dealing with a large number of medical images [5].

Deep learning techniques have appeared as the most powerful tools for medical image analysis. The increasing cases of brain tumors and increasing workload of healthcare professionals have created an urgent need for computer-aided automated brain tumor detection. Deep learning models, especially convolution neural networks (CNNs), transformers, and object detection frameworks such as YOLO, have shown exceptional performance in tumor detection, classification, and segmentation tasks [6], [7]. Classification accuracies exceeded 95-99% on publicly available datasets like Figshare and BraTS [8], [9]. Modern techniques, including explainable AI, which influenced Grad-CAM for model refinement [10], hybrid ensemble classifiers [11], self-attention mechanisms [12], and lightweight architectures [13], have achieved outstanding performance and also attempt to address efficiency and interpretability concerns. Even with these advancements, various challenges remain before deep learning-based brain tumor detection systems are generally accepted in medical practice. Many models show performance degradation when tested on data from different scanners, institutions, or populations [14]. It continues to be difficult for real-world deployment because of issues such as limited annotated datasets, class imbalance, model interpretability, lack of generalization across diverse patients, and high computational power [15]. This review paper presents modern advances and studies of deep learning techniques for MRI-based brain tumor detection and classification. This paper evaluates a wide variety of deep learning techniques, starting from traditional machine learning hybrids to modern deep learning architectures and explainable AI systems. A special focus is given to practical limitations like domain generalization, model efficiency, and medical trustworthiness.

2. Literature Study

Qasem et al. [1] developed a learning-based system with the help of watershed segmentation for tumor boundary detection, followed by feature extraction and K-Nearest Neighbours (KNN) classification. This method achieved acceptable accuracy while tested on large MRI datasets, demonstrating the viability of non-deep learning techniques in resource-constrained environments.

Garg and Garg [2] implemented an ensemble classifier (KNN-RF-DT), which is based on majority voting. For segmentation and extracted features, they utilized Otsu's thresholding using Stationary Wavelet Transform (SWT), PCA, and GLCM. This model scored an accuracy of 97.30% when tested on datasets of 2556 images, whereas it involved smaller training data contrasted to deep learning approaches. This approach provides the benefits of low computational cost and interpretability but frequently faces difficulty with complex tumor boundaries and multi-class scenarios.

Das and Biswas [3] developed the two novel architectures: SAETCN (Self-Attention Enhancement Tumor Classification Network) and SAS-Net (Self-Attentive Segmentation Network). Validation accuracy achieved by SAETCN is 99.38% for classifying glioma, meningioma, pituitary tumor, and non-tumor cases, while SAS-Net has achieved 99.23% pixel accuracy for the segmentation. It can capture long-range dependencies effectively with their self-attention mechanisms.

Swin UNETR, proposed by Hatamizadeh et al. [4], combines Swin Transformers with U-Net architecture for 3D brain tumor segmentation. This model had performed the challenges in BraTS 2021 by effectively modeling multi-scale features via hierarchical shifted windows.

Shen et al. [5] proposed a lightweight MBDRes-U-Net multiscale 3D variant that incorporates a multibranch residual block and fused attention. This model scored a strong performance on the BraTS 2018/2019 datasets, which significantly reduced the computational cost.

BGF-YOLO was proposed by Kang et al. [6], which enhances YOLOv8 with bi-level routing attention and a generalized feature pyramid network, achieving a 4.7% mAP improvement on the Br35H dataset. Approaches for object detection have also obtained traction.

Paul et al. [7] utilized YOLOv5 for brain cancer instance segmentation, demonstrating real-time capabilities.

Das Gupta et al. [8] introduced an explainability-driven CNN refinement framework. Rather than utilizing Grad-CAM only for post-hoc visualization, they have measured layer-wise relevance to prune low-contribution layers, lowering model depth and parameters while maintaining high accuracy of 98.21% on the primary dataset and 95.74% accuracy on unseen data. Moreover, they evaluated the decisions using SHAP and LIME, managing the black-box nature of deep models.

Mathivanan et al. [9] proposed a BTDN (Brain-Tumor Detection Network), which is combined with Secure-Net for data protection. This model scores an accuracy up to 99.68% across multiple datasets like Br35H, Kaggle, and BraTS while integrating the security features for safe transmission.

Ghosh et al. [10] have developed the dataset UCSF-PDGM-VQA with QA pairs of 2387 to benchmark vision-language models (VLMs) for interactive MRI integration that highlights the limitations such as modality collapse in current VLMs.

Isett et al. [11] proposed a lightweight multi-cancer tumor localization that demonstrates the cross-cancer generalization.

Balcerak et al. [12] examined the physics-informed approaches (GliODIL) for personalizing the glioma radiotherapy planning by enhancing data and physics-informed discrete loss, moving beyond uniform margins.

3. Methodology

The methodology sued in this review follows a Preferred reporting Items for Systematic Reviews and Meta-Analyses (PRISM) guidelines [18] to confirm transparency, reproducibility, and severity in literature selection process.

(i) Research Questions:

The reviews was directed by the given research questions:
1) Which deep learning techniques are mainly used for brain tumor detection and classification from MRI scans?
2) How accurate, efficient and interpretable these models are?
3) Which key challenges these model faces in generalizing to real-world clinical environment especially in low-resources environment?
4) How have lightweight architectures and explainable AI been utilized to enhance trustworthiness and deployability?

(ii) Search Methods.

A detailed literature review was carried out across multiple datasets such as PubMed, IEE Xplore, Scopus, Google Scholar, and arXiv. This review includes publications from 2020 to June 2026.

(iii) Inclusion and Exclusion Criteria.

Inclusion Criteria:
The study proposes deep learning, hybrid, or explainable AI methods for tasks like brain tumor detection, classification and segmentation. All the papers have been published in conferences, peer-reviewed journals, or reputable re-print platforms (arXiv). Works that has been addressed are clinical adaptability, generalization and lightweight models.

Exclusion criteria:
The main primary focus is studies using non-MRI modalities. Papers without methodological details and clear performance. Non-English, editorial and review articles.

(iv) Study Selection Process

Author provided total 18 papers and thoroughly examined in detail. All the papers were carefully analysed and thematically synthesized. Additionally to support broader studies supportive papers were taken into consideration to provide context and strengthen the analysis. Traditional methods, CNNs, YOLO variants, transformers, explainable AI, lightweight models and emerging technology are covered by all the chosen papers.

(v) Data Extraction and synthesis

Information related to developed method, datasets utilized, performance metrics, strengths, limitation and contribution towards real-world problems such as efficiency, generalization and explainability was derived. Instead of purely quantitative meta-analysis, a thematic synthesis was executed.

4. Comparative Analysis and Findings

The comparative analysis of all the reviewed paper have been presented in Table 1. This table summarizes all the key findings regarding performance, progress, limitations, strengths and real-word applicability.

Deep learning models particularly SAETCN, BTDN, and YOLO consistently achieves an accuracy of 95-99% or Dice scores on public datasets like Figshare, Kaggle, and BraTS. This represents a substantial technical advancement. The work carried out by Das Gupta et al. [8] is remarkable as it uses Grad-CAM for active model development, not only for visualization. This method maintains the competitive and improving transparency while reducing the parameters. Architectures like BGF-YOLO, MBDRes-U-Net, and MuCTal demonstrates the growing concentration on minimizing the computational cost, which makes it more suitable for real-world deployments in limited-resource enviroments. Although the accuracy is excellent on training distributions, various studies involving cross-dataset testing in Das Gupta et al. illustrates the noticeable declines on unseen data. Few studies deeply examined on challenging datasets like BraTS-Africa or low-quality scans. Although these fields are in their early stages of developments, integration of security (Secure-Net), vision-languages models (UCSF-PDGM-VQA0, and physics-informed techniques (GliODIL) indicates a change toward therapeutically applicable solutions.

Table 1: Comparative Analysis of Selected Studies

Method	Task	Key Strengths	Major Limitations
Watershed + K-NN	Detection	Less computation, generic	Finite multi-class performance
KNN-RF-DT Ensemble	Detection + Classification	Small amount of data needed	Conventional features, fewer robust
SAETCN / SAS-Net	Classification + Segmentation	Self-focus, Great accuracy	Controlled generalization testing
BGF-YOLO (YOLOv8 variant)	Detection	Lightweight, real-time	Concentrated on one dataset
XAI-driven CNN Refinement	Classification	Transparency + pruning	Reasonable generalization
BTDN + Secure-Net	Classification	High security + accuracy	Limited external validation
MBDRes-U-Net	Segmentation	Lightweight 3D model	Computational compromise
Swin UNETR	Segmentation	Transformer-based long-range	Expensive computational cost

5. Result Analysis

Several significant trends, accomplishment, and constraints in the field of brain tumor detection and classification form MRI images are identified by the comparative analysis of studied research. In terms of accuracy, deep learning-based techniques reliably surpass traditional method. The highest classification accuracies were achieved by the modern architectures including SAETCN, BTDN, and YOLO variations, which ranged from 98.21% to 99.68% on standard benchmark datasets. With pixel-level accuracy and Dice scores repeatedly exceeding 99% and 0.85-0.90 respectively on BraTS datasets, Segmentation models such as SAS-Ne6t, MBDRes-U-Net, and Swin UNETR also reported powerful results.

Explainable AI techniques especially Das Gupta et al.'s XAI-driven CNN refinement methods, demonstrated that Grad-CAM directed pruning can drastically decrease the model complexity while achieving the high accuracy of 98.21%.

Various deep learning models outperforms when examined on clean, publicly available datasets. Promising minimization in computational cost and inference time, shown by lightweight models such as BGF-YOLO, MBDRes-U-Net, and pruned CNNs, enabling them more suitable for real-time and edge-device applications. By offering visual explanations of model decisions the integration of Grad-CAM, SHAP, and LIME in explainable frameworks supports building clinical trust. Works such as detection, classification, and segmentation combined successfully in unified structures.

6. Research Gaps & Open Challenges

The Even though advancement in deep learning for brain tumor identification and classification has made significant remarks, there are still a number of crucial gaps that prevent these technologies from being applied in real clinical settings.

Gap-1: Domain generalization and open challenges.

Poor Generalization is one of the most significant gaps of models across different datasets and clinical enviroments. The majority of research, such as SAETCN [3], BTDN [9], and BGF-YOLO [4], exhibit significant performance loss on unseen data, poor-quality scans, or data from several institutions, still represent outstanding performance on clean public benchmarks (Figshare, Kaggle,BraTS). In low resource enviroments such as Sub-Saharan Africa, where imaging equipment and protocols vary greatly, this domain shift problem in especially important. Datasets such as BraTS-Africa or pediatric cohorts (BraTS-PEDs) fully assess model.

Gap-2: Limited Clinical Translation and Real-World Validation.

Clinical applicability and laboratory results are clearly totally different. Most models does not have patient level evaluation instead of slice-level accuracy, external validation across multi-centres, integration with real-world clinical processes, and testing in real world conditions such as class imbalance, noisy images, motion artifacts. Additionally problems like late-stage tumor and different tumor size in real hospital are rarely addressed.

Gap-3: Computational Efficiency and Deployment Barriers.
Most of high-performing architectures, like Swin UNETR, are nevertheless computationally costly, despite the proposal of some lightweight models (MBDRes-U-Net, BGF-YOLO). This restricts their use in hospitals, particularly in underdeveloped nations, where GPU resources are scarce. More effective models that are appropriate for mobile platforms or edge devices and maintain high accuracy are desperately needed.

Gap-4: Explainability and Trustworthiness.
Most of explainable AI techniques are currently only used for visualizing pot-hoc rather than actively improving the model, although Das Gupta et al [8] made significant progress by applying Grad-CAM for model refinement. More thorough evaluation is needed to determine the usefulness of explanations (Grad-CAM, SHAP, LIME). For critical decisions, clinicals still hesitate to trust black-box or semi-transparent system.

Gap-5: Data-Related Challenges.
Real-world tumor cases are very less as compared to normal scans, although most studies use balanced datasets. Secure data transmission is addressed in very few studies (e.g., Mathivanan et al. [9]). Most of the models do not incorporate longitudinal or multi-modal (MRI + PET/CT + genomics) data: instead they depends on sinlg-timepoint MRI.

7. Conclusion

The rapid advancements in deep learning techniques for brain tumor detection and classification from MRI scans over the last six years have been systematically examined. The discipline has made significant strides in accuracy, from conventional hybrid techniques to contemporary architectures like SAETCN, BGF-YOLO, Swin UNETR, and explainable models. Several research have reported performance exceeding 95–99% on benchmark datasets. The performance and transparency of contemporary models have been significantly improved by combining interpretable AI frameworks, simplified architectures, and attention-based components. However, despite these advancements, there is still a significant gap between what can be successfully used in clinical situations and what is shown in controlled research settings. The wider implementation of these systems in actual healthcare is still hampered by a number of issues, including poor adaptation across various data domains, a lack of thorough external testing, stringent hardware requirements, and limited acceptance among medical professionals.

In order to create reliable, trustworthy, and egalitarian AI systems that can operate successfully in a variety of healthcare contexts, including low-resource settings, it is imperative that these gaps be filled. Domain generalization, patient-level assessment, edge-device compatibility, and human-AI cooperation via vision-language models and federated learning should be the top priorities for future research.

In conclusion, deep learning has demonstrated enormous potential to transform the detection of brain tumors; yet, achieving this potential necessitates a change from aiming for greater accuracy on public benchmarks to developing dependable, comprehensible, and clinically applicable solutions. In addition to advancing medical imaging research, closing this gap will significantly enhance patient outcomes and promote global healthcare equity.

References

[1] S. N. Qasem, A. Nazar, A. Qamar, S. Shamshirband and A. Karim, "A Learning Based Brain Tumor Detection System," CMC-Computers, Materials & Continua, vol. 59, no. 3, pp. 713–727, 2019.
[2] G. Garg and R. Garg, "Brain Tumor Detection and Classification based on Hybrid Ensemble Classifier," 2023.
[3] S. Das and A. Biswas, "Novel Deep Learning Architectures for Classification and Segmentation of Brain Tumors from MRI Images," arXiv preprint arXiv:2512.06531, 2025.
[4] M. Kang, C.-M. Ting, F. F. Ting and R. C.-W. Phan, "BGF-YOLO: Enhanced YOLOv8 with Multiscale Attentional Feature Fusion for Brain Tumor Detection," in MICCAI 2024, 2024.
[5] S. Paul, M. T. Ahad and M. M. Hasan, "Brain Cancer Segmentation Using YOLOv5 Deep Neural Network," 2023.
[6] L. Shen, Y. Hou, J. Chen, L. Diao and Y. Duan, "MBDRes-U-Net: Multi-Scale Lightweight Brain Tumor Segmentation Network," 2024.
[7] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth and D. Xu, "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images," arXiv preprint arXiv:2201.01266, 2022.
[8] R. Das Gupta, M. I. H. Showmick, L. Wei, M. R. Abir, S. Akter, M. Y. Rahat and M. J. Hossen, "From Explanations to Architecture: Explainability-Driven CNN Refinement for Brain Tumor Classification in MRI," arXiv preprint arXiv:2506.09161, 2025.
[9] S. K. Mathivanan, S. Srinivasan, M. S. Koti, V. S. Kushwah, R. B. Joseph and M. A. Shah, "A secure hybrid deep learning framework for brain tumor detection and classification," Journal of Big Data, vol. 12, no. 72, 2025.
[10] B. Isett, R. Dadey, A. Li, R. C. Augustin, K. Smith, A. D. Singhi, Q. Gu and R. Bao, "A Lightweight Multi-Cancer Tumor Localization Framework for Deployable Digital Pathology," arXiv preprint arXiv:2603.08844, 2025.
[11] S. Ghosh, J. Lateef, C.-H. Liu, Y. Yu, A. M. Rauschecker and M. Sushil, "UCSF-PDGM-VQA: Visual Question Answering dataset for brain tumor MRI interpretation," arXiv preprint arXiv:2605.17140, 2026.
[12] M. Balcerak, J. Weidner, P. Karnakov, I. Ezhov, S. Litvinov, P. Koumoutsakos, T. Amiranashvili, R. Z. Zhang, J. S. Lowengrub, I. Yakushev, B. Wiestler and B. Menze, "Individualizing Glioma Radiotherapy Planning by Optimization of a Data and Physics-Informed Discrete Loss," arXiv preprint arXiv:2312.05063, 2025.