



Machine Learning and Deep Learning for Brain Tumor Prediction and Segmentation in MRI: A Systematic Review of Methods, Datasets and Open Challenges

Namrata Vijayvargiya¹, Nirupma Singh²

¹Research Scholar, School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

²Assistant Professor, School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

Abstract- Magnetic resonance imaging (MRI) is the reference modality for brain tumor diagnosis, yet manual delineation remains slow, expertise-dependent and subject to inter-observer variability. Over the last four years the literature has shifted decisively from handcrafted feature pipelines to convolutional, encoder-decoder, transformer and hybrid learning architectures. This paper presents a systematic review of forty-one primary studies published between 2023 and 2026, selected from 412 records through a PRISMA-style protocol. The reviewed work is organised into a four-branch taxonomy - classical machine learning, CNN backbones, encoder-decoder segmentation networks and transformer/hybrid ensembles - and is compared on a common footing of accuracy, precision, recall, F1-score, Dice similarity coefficient and intersection over union. Reported classification accuracy ranges from 93.30% to 99.70% and Dice from 0.650 to 0.961, but the spread is shown to be driven as much by dataset choice and evaluation protocol as by architecture. A reference hybrid detection-segmentation pipeline that couples classical preprocessing and region-based localisation with CNN feature learning and a margin-based or meta-learned decision layer is formalised as two algorithms, and priorities for trustworthy, federated and explainable neuro-oncology AI are set out.

Keywords- brain tumor, MRI, systematic review, deep learning, U-Net, transformer, hybrid model, segmentation, explainable AI.

I. INTRODUCTION

Brain tumors are among the most lethal neurological disorders; outcome depends heavily on how early and how precisely the lesion is identified and delineated. Magnetic resonance imaging (MRI) is the



modality of choice because it delivers high soft-tissue contrast, multi-parametric structural information and no ionising radiation [1], [2]. Its interpretation, however, is a specialist task: manual reading is time consuming, is sensitive to tumor size, shape, intensity and location, and shows measurable inter-observer variability. This has made computer-aided detection and segmentation an active area of neuro-oncology research [3], [4].

The methodological centre of gravity of that research has moved several times, as summarised in Fig. Early automation relied on intensity thresholding, atlas registration and rule-based heuristics; these methods were transparent but brittle against the natural variability of tumor appearance. Classical machine learning then paired handcrafted descriptors - grey-level co-occurrence matrix (GLCM) texture, shape and intensity histograms - with support vector machines (SVM), random forests and boosting [5], [6]. From 2015 onward convolutional neural networks (CNN) and encoder-decoder architectures such as U-Net displaced feature engineering altogether, and since 2023 the dominant direction has been hybridisation: attention modules, vision and shifted-window transformers, ensembles and stacked meta-learners, increasingly accompanied by explainability layers [7]-[9].

That rapid progress has produced a difficult secondary problem. Reported figures are extremely high and extremely close together - many recent studies claim above 98% accuracy or above 0.90 Dice - yet they are obtained on different datasets, with different splits, different class definitions (binary tumor/non-tumor versus three- or four-class typing) and different post-processing. Headline numbers therefore cannot be compared directly, and the field lacks a consolidated, protocol-driven picture of what is actually established.

This paper addresses that need. Its contributions are: (i) a PRISMA-style review protocol, stated as an explicit algorithm, applied to 412 records and yielding 41 primary studies from 2023-2026; (ii) a four-branch taxonomy of methods with the design assumptions and failure modes of each branch made explicit; (iii) a normalised quantitative comparison across classification and segmentation metrics, with the confounding role of dataset and protocol separated from architectural merit; (iv) a structured gap analysis mapping each observed limitation to concrete research directions; and (v) a reference hybrid pipeline that consolidates the practices most consistently associated with strong and reproducible results.

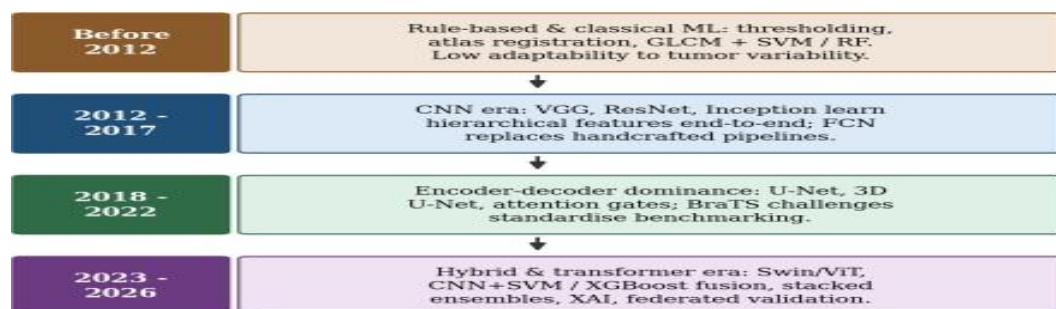


Fig. 1. Methodological eras in MRI-based brain tumor analysis, from rule-based processing to hybrid and transformer architectures.



Section II states the review methodology, Section III develops the taxonomy, Section IV covers datasets and metrics, Section V the comparative analysis, Section VI the gap analysis, Section VII the reference pipeline, and Section VIII concludes.

II. REVIEW METHODOLOGY

The review follows a reproducible protocol, given as Algorithm 1 and illustrated in Fig. 2. Six sources were queried

- IEEE Xplore, SpringerLink, ScienceDirect, MDPI, PubMed and arXiv - for the window January 2023 to March 2026, using the Boolean expression ("brain tumor" OR glioma OR meningioma OR pituitary) AND (MRI OR "magnetic resonance") AND ("deep learning" OR "machine learning" OR segmentation OR classification). The search returned 412 records; 144 duplicates were removed.

Screening criteria are listed in Table I. A study was retained only if it evaluated an ML or DL method on brain MRI, reported at least one quantitative metric on a named dataset, and described its architecture in enough detail to be reproduced in principle. Editorials, non-MRI modalities, purely clinical papers without a computational contribution and works reporting only qualitative results were excluded. Title and abstract screening removed 172 records; full-text assessment removed a further 55, leaving 41 primary studies. Of these, 27 report accuracy or Dice on a recognised benchmark and enter the quantitative comparison of Section V.

For each retained study the architecture family, backbone, preprocessing chain, dataset and modality, task definition, train-test protocol and every reported metric were extracted. Where several configurations were reported, the best configuration on the primary dataset was recorded together with its protocol, so that like is compared with like in Section V.

Table 1. Inclusion and exclusion criteria applied during screening

Inclusion criterion	Exclusion criterion
Peer-reviewed article, conference paper or indexed preprint, 2023-2026	Published before 2023 or not retrievable in full text
Brain MRI as the primary imaging modality	CT, PET, histopathology or non-brain imaging
ML or DL method for detection, classification or segmentation	Purely clinical, epidemiological or hardware-oriented study
At least one quantitative metric on a named dataset	Qualitative results only, or dataset not identified
Architecture and training protocol described	Method described only at block-diagram level

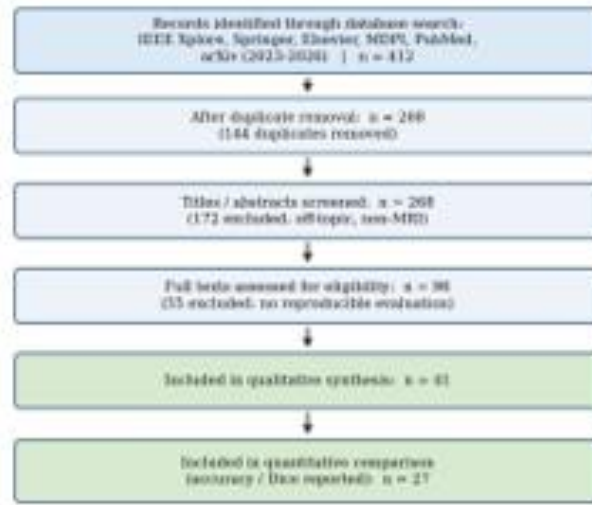


Fig. 2. PRISMA-style selection flow: 412 records screened down to 41 primary studies, of which 27 are quantitatively comparable.

Algorithm 1: Systematic Review and Evidence Synthesis Protocol

Input: source list S , query Q , window $W = [2023, 2026]$ **Output:** synthesised evidence table E , gap set G

```
1:  $R \leftarrow \emptyset$ 
2: for each source  $s \in S$  do
3:    $R \leftarrow R \cup \text{Query}(s, Q, W)$ 
4:  $R \leftarrow \text{DeduplicateByDOIAndTitle}(R)$ 
5:  $R_1 \leftarrow \{ r \in R : \text{ScreenTitleAbstract}(r) = \text{accept} \}$ 
6:  $R_2 \leftarrow \{ r \in R_1 : \text{FullTextEligible}(r, \text{criteria of Table I}) \}$ 
7: for each  $r \in R_2$  do
8:    $e \leftarrow \text{Extract}(r)$ : family, backbone, preprocessing,
dataset, task, split, metrics
9:    $e.\text{metrics} \leftarrow \text{NormaliseMetrics}(e)$  D unify Acc, Dice, IoU
10:  $E \leftarrow E \cup \{e\}$ 
11:  $T \leftarrow \text{GroupBy}(E, \text{family})$  D taxonomy of Section III
```



12: $C \leftarrow \text{CompareWithinDatasetStrata}(T)$ D removes dataset bias
13: $G \leftarrow \text{IdentifyRecurringLimitations}(E, C)$
14: return E, G

III. TAXONOMY OF EXISTING APPROACHES

Fig. 3 summarises the four branches into which the 41 primary studies fall. The branches are ordered by how much of the representation is learned rather than specified.

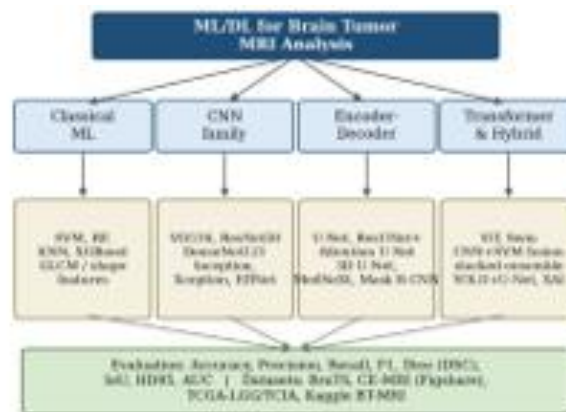


Fig. 3. Four-branch taxonomy of MRI brain tumor analysis methods, with representative models, datasets and evaluation metrics.

1. Classical machine learning and feature-engineered pipelines

This branch separates description from decision. Preprocessing removes acquisition artefacts (median and Wiener filtering, N4ITK bias-field correction, z-score normalisation); a segmentation stage isolates candidate regions using Otsu thresholding, seeded region growing, marker-controlled watershed or fuzzy C-means clustering; GLCM texture and contour-based shape descriptors are then classified by SVM, k-nearest neighbours, random forest or gradient boosting. Shemanto et al. reached 96.46% accuracy with high-order filtering followed by edge and morphological segmentation [9], while Tseng and Tang combined CLAHE enhancement, k-means segmentation and particle-swarm feature selection with XGBoost for 97% accuracy and 98% precision [6]. Alqhtani et al. coupled fuzzy C-means with an SVM classifier on the CE-MRI corpus and obtained 98.2% accuracy with a Dice of 0.961 and markedly lower runtime than deep baselines [19]. These pipelines remain competitive, are cheap to train, and are inherently auditable - each stage can be inspected - but they degrade when tumor appearance departs from the assumptions encoded in the descriptors.

2. CNN backbones for detection and typing

Transfer-learned CNN backbones dominate classification. Ahmed et al. fine-tuned VGG-16 to 97.33% accuracy and added layer-wise relevance propagation to expose the evidence behind each decision [8];



Saeedi et al. compared a 2D CNN (96.47%) and an auto-encoder (95.63%) against several classical learners on a four-class task [10]; Mahmud et al. reported 93.3% accuracy and 98.43% AUC for a purpose-built CNN evaluated against ResNet-50, VGG-16 and Inception-V3 on 3,264 images [16]. Ariful Islam et al. benchmarked DenseNet-121 with Grad-CAM++ against seventeen architectures including ViT, ConvNeXt and MobileNetV3, reaching up to 99.3% accuracy across two datasets and performing particularly well on non-enhancing gliomas and small metastatic lesions [27]. Abdusalomov et al. instead treated the problem as detection, extending YOLOv7 with CBAM attention and a bidirectional feature pyramid [12]. The recurring weakness of this branch is that a single backbone produces correlated errors and is sensitive to class imbalance on small medical corpora.

Table 2. Representative classification-oriented studies

Study	Method	Dataset	Best result
Ahmed [8]	VGG-16 + LRP	Kaggle MRI	Acc 97.33%
Saeedi [10]	2D CNN, auto-encoder, KNN	4-class MRI	Acc 96.47%
Mahmud [16]	Custom CNN vs. ResNet-50 / VGG-16 / Inception-V3	3,264 MRI	Acc 93.30%, AUC 98.43%
Tseng [6]	CLAHE + k-means + PSO + XGBoost	MRI (public)	Acc 97%, Prec 98%
Alqhtani [19]	CLAHE + FCM + SVM	CE-MRI	Acc 98.2%, Dice 0.961
Islam [27]	DenseNet-121 + Grad-CAM++	2 MRI corpora	Acc 99.3%
Asif [35]	Ensemble of deep models	MRI scans	Acc 98.3%
Vijayvargiya [40]	Xception + ConvNeXt + Swin, meta-learner stacking	Kaggle 4-class (7,212 images)	Acc 98.92%

3. Encoder-decoder segmentation networks

Voxel- and pixel-wise delineation is dominated by U-Net and its descendants. Asiri et al. fine-tuned a CNN with a ResNet-50 encoder and a U-Net decoder on TCGA-LGG and TCIA, reporting IoU 0.91 and Dice 0.95 [11]; Metlek and Cetiner introduced ResUNet+ with convolution and attention blocks for Dice up to 0.931 [13]; Rajagopal et al. fused GLCM features with U-Net and a 3D CNN for whole-tumor accuracy of 99.4% [15]. Tiwary et al. replaced the U-Net encoder with EfficientNet to improve long-range context modelling, reporting 99.25% accuracy on the 3,064-slice CE-MRI corpus [28]. Rahman et al. embedded topological data analysis into a SegUNet to make the network shape-aware [37], and Alanazi et al. combined densely connected convolutions with bidirectional ConvLSTM layers [42]. The BraTS challenge family supplies the standardised annotations and lesion-wise metrics on which most of this branch is benchmarked [20], [26].



Table 3. Representative segmentation-oriented studies

Study	Method	Dataset	Best result
Asiri [11]	ResNet-50 encoder + U-Net	TCGA-LGG, TCIA	Dice 0.95, IoU 0.91
Metlek [13]	ResUNet+ (conv. + attention)	Multiple MRI	Dice 0.931
Rajagopal [15]	GLCM + U-Net + 3D CNN	BraTS	Acc 99.4%
Hashmi [22]	MedNeXt + ensembling	BraTS-SSA / pediatric	Dice 0.896 / 0.830
Moawad [26]	Challenge winner, lesion-wise scoring	BraTS-METS	Dice 0.65 $\pm$ 0.25
Ahsan [24]	YOLOv5 detector + 2D U-Net	MRI	mAP 89.5%, Dice 0.881
Tiary [28]	EfficientNet-enhanced UNet	CE-MRI (3,064 slices)	Acc 99.25%

4. Transformer and hybrid ensemble architectures

The newest branch treats architectural diversity itself as the design variable. Self-attention and shifted-window attention supply the global context that convolution cannot, and are combined with convolutional streams either by feature fusion or by probabilistic stacking. Vijayvargiya and Singh trained Xception, ConvNeXt and a Swin Transformer independently and fused their class-probability vectors through a meta-learner fitted on out-of-fold predictions, reaching 98.92% test accuracy on a balanced four-class task with a train-test gap of only 0.1% [40]. Hussain et al. and Musthafa et al. integrate segmentation and classification into single multi-task pipelines [36], [34];

Ahsan et al. cascade a YOLOv5 detector into a 2D U-Net and outperform both standalone U-Net and Mask R-CNN [24]; Hassan and Boulila combine fuzzy thresholding with deep classifiers [39]. Recent surveys agree that hybrid CNN-transformer designs give the best trade-off between local texture sensitivity and global context, at the cost of parameter count and training complexity [31], [38].

IV. DATASETS AND EVALUATION METRICS

Four corpora account for most reported results (Table IV). The Figshare CE-MRI collection contains 3,064 T1-weighted contrast-enhanced slices from 233 patients across glioma, meningioma and pituitary tumor, each 512x512 with a manually delineated border and binary mask [41]. The BraTS family supplies multi-parametric, expertly annotated volumes with standardised lesion-wise evaluation and has expanded to metastases, meningioma radiotherapy planning, sub-Saharan African and pediatric cohorts [20], [22], [26]. TCGA-LGG/TCIA links imaging to genomic annotation, and the balanced Kaggle four-class corpus is widely used for classification benchmarking.



Table 4. Benchmark MRI datasets used across the reviewed studies

Dataset	Content	Task	Used in
CE-MRI (Figshare)	3,064 T1-CE slices, 233 patients, 3 tumor types, borders + binary masks	Typing, segmentation	[19], [28], [41]
BraTS (2012-2025)	Multi-parametric volumes (T1, T1-CE, T2, FLAIR), expert annotation	Voxel-wise segmentation	[20], [22], [26], [29]
TCGA-LGG / TCIA	Imaging linked to genomic and clinical annotation	Segmentation, grading	[11], [15]
Kaggle BT-MRI	~7,000 images, balanced glioma / meningioma / pituitary / no-tumor	4-class classification	[8], [16], [35], [40]

Classification studies report accuracy, precision, recall and F1-score computed from the confusion-matrix counts TP, TN, FP and FN, while segmentation studies report the Dice similarity coefficient (DSC), the intersection over union (IoU) and, in the BraTS protocol, the 95th-percentile Hausdorff distance:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

$$Precision = TP / (TP + FP) \quad (2)$$

$$Recall = TP / (TP + FN) \quad (3) F1 = 2 \cdot Precision \cdot Recall / (Precision + Recall) \quad (4) DSC(P, G) = 2 |P \cap G| / (|P| + |G|) \quad (5)$$

$$IoU(P, G) = |P \cap G| / |P \cup G| \quad (6)$$

where P is the predicted tumor region and G the ground-truth annotation. Two protocol details determine comparability far more than the choice of metric: whether Dice is computed per case or per lesion - lesion-wise scoring penalises missed small metastases heavily, which is why the BraTS-METS winner records 0.65 while whole-tumor studies exceed 0.90 - and whether accuracy refers to a binary or a four-class decision.

V. COMPARATIVE PERFORMANCE ANALYSIS

Fig. 4 orders the reviewed classification studies by reported accuracy and colours each by architectural family. The distribution is compressed into a narrow band between 93.3% and 99.7%, and no family is uniformly superior: the highest single figure comes from a hybrid DCNN-LuNet pipeline (99.70% [5]), the second from a U-Net/3D-CNN combination (99.40% [15]), and a CNN backbone with an explainability layer is third (99.30% [27]). Differences of a few tenths of a percent between studies on different corpora are not evidence of architectural superiority.

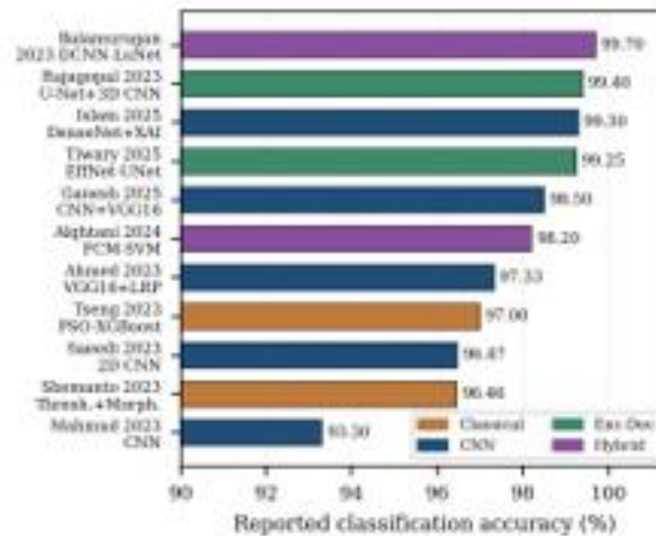


Fig. 4. Reported classification accuracy across the reviewed studies, grouped by architectural family.

Fig. 5 shows the same effect more sharply for segmentation. Studies evaluated on whole-tumor tasks with case-level Dice cluster above 0.93, whereas the BraTS-METS metastasis task under lesion-wise scoring falls to 0.65 and the pediatric BraTS-SSA cohort to 0.830. The variance here is attributable to lesion size, cohort and scoring rule, not to the networks, all of which are modern encoder-decoders.

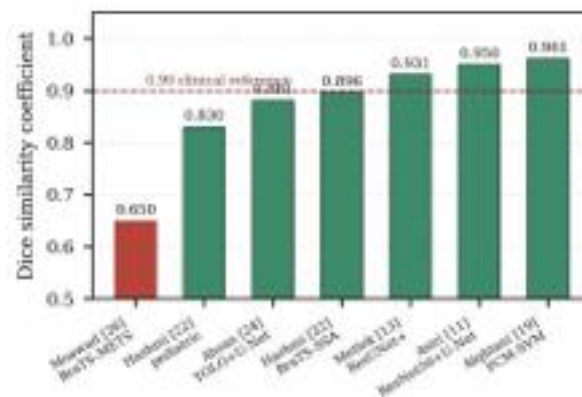


Fig. 5. Reported Dice similarity coefficients. Variation tracks task difficulty and scoring protocol rather than architecture.

Fig. 6 shows how the corpus is distributed over publication year and family. Classical feature-engineered work contracts from four studies in 2023 to none in 2026, while transformer and hybrid contributions grow from one to three, confirming the trajectory of Fig. 1. Table V consolidates the four families.

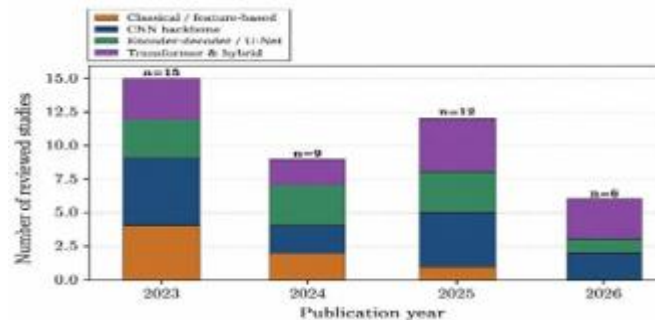


Fig. 6. Distribution of the 41 reviewed studies by publication year and architectural family.

Table 5. Consolidated comparison of the four architectural families

Family	Typical Acc.	Typical Dice	Strength	Principal limitation
Classical ML	96-98%	0.90-0.96	Low cost, fully auditable stages	Descriptors fail on atypical appearance
CNN backbone	93-99%	n/a	Strong local texture, transfer learning	Correlated errors, imbalance sensitivity
Encoder-decoder	98-99%	0.65-0.95	Precise boundaries, BraTS-standardised	Weak on small and metastatic lesions
Transformer / hybrid	98-99%	0.88-0.96	Global context + local detail, low overfitting gap	Parameter count, training complexity, data hunger

VI. RESEARCH GAPS AND OPEN CHALLENGES

Six limitations recur across the corpus and are set out in Table VI. First, benchmarking is not comparable: datasets, splits, class definitions and scoring rules differ, so the near-saturated accuracies of Fig. 4 cannot be ranked. Second, performance collapses on small and metastatic lesions - the BraTS-METS analysis attributes the fall to false negatives and spatial misregistration below 100 mm³ [26]. Third, generalisation is demographically and technically narrow; the BraTS-SSA and pediatric results show measurable degradation under distribution shift [22]. Fourth, almost all pipelines are unimodal, discarding the clinical metadata - age, presenting symptoms, prior histology - that a radiologist uses routinely. Fifth, the highest-performing backbones remain opaque; layer-wise relevance propagation and Grad-CAM++ are applied post hoc in only a minority of studies [8], [27]. Sixth, evaluation is retrospective throughout: no reviewed study reports prospective clinical deployment, and regulatory and workflow-integration barriers are acknowledged but unaddressed [33], [38].

Table 6. Identified gaps and corresponding research directions

Gap (evidence)	Research direction
Non-comparable benchmarking [2], [26], [31]	Common protocol: fixed splits, lesion-wise and case-wise Dice reported together, released inference containers



Small and metastatic lesion failure [26]	Size-stratified loss weighting, high-resolution decoders, detection-then-segmentation cascades
Demographic and scanner bias [22], [24]	Augmentation-aware and federated training across sites; explicit subgroup reporting
Unimodal input [17], [23]	Joint fusion of imaging with clinical and genomic metadata in a shared latent space
Opacity of deep backbones [8], [27]	Intrinsically interpretable modules plus calibrated uncertainty, not post-hoc saliency alone
Retrospective evaluation only [33], [38]	Prospective reader studies, workflow integration and regulatory-grade validation

VII. A REFERENCE HYBRID PIPELINE

The practices most consistently associated with strong and stable results across the reviewed corpus can be assembled into the single reference pipeline of Fig. 7, formalised in Algorithm

2. It retains the classical preprocessing and region-based localisation that make the early stages auditable, learns the representation with a CNN or Mask R-CNN encoder augmented by GLCM and shape descriptors, and delegates the decision either to a margin-based classifier or to a meta-learner fitted on out-of-fold base-model probabilities. Post-processing suppresses false positives and an explainability overlay accompanies every prediction. increments on saturated retrospective benchmarks, will determine whether MRI-based tumor analysis reaches routine neuro-oncology practice.

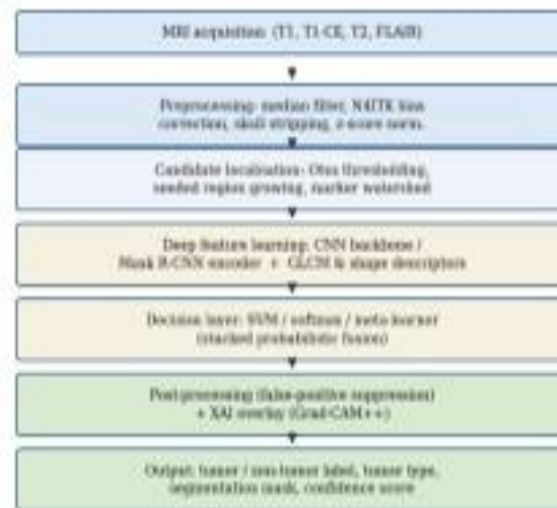


Fig. 7. Reference hybrid detection-segmentation pipeline consolidating the practices observed across the reviewed literature.



Algorithm 2: Reference Hybrid Detection and Segmentation Pipeline

Input: MRI volume x , trained encoder f_θ , decision head g_ϕ **Output:** label $\hat{y}$, tumor mask M , confidence c , saliency map S

```
1:  $x \leftarrow \text{MedianFilter}(x)$ ;  $x \leftarrow \text{N4ITKBiasCorrect}(x)$ 
2:  $x \leftarrow \text{SkullStrip}(x)$ ;  $x \leftarrow \text{ZScoreNormalise}(x)$ 
3:  $R \leftarrow \text{OtsuThreshold}(x) \cup \text{RegionGrow}(x, \text{seeds})$ 
4:  $R \leftarrow \text{MarkerWatershed}(x, R)$   $D$  candidate regions 5:  $F_{\text{deep}} \leftarrow f_\theta(x)$   $D$  learned representation 6:  $F_{\text{hand}} \leftarrow \text{GLCM}(R) \oplus \text{ShapeDescriptors}(R)$ 
7:  $z \leftarrow \text{Concat}(F_{\text{deep}}, F_{\text{hand}})$  8: if ensemble mode then
9:   for each base learner  $k$  do  $p^k \leftarrow M_k(x)$ 
10:  $z \leftarrow [p^1 \parallel p^2 \parallel \dots \parallel p^k]$   $D$  out-of-fold stacking
11:  $\hat{y} \leftarrow \arg \max_c g_\phi(z)$   $D$  SVM / softmax / meta-learner 12:  $M \leftarrow \text{Decoder}(F_{\text{deep}})$ ;  $M \leftarrow \text{RemoveSmallComponents}(M)$  13:  $c \leftarrow \max_c g_\phi(z)$ ;  $S \leftarrow \text{GradCAMpp}(f_\theta, x, \hat{y})$ 
14: return  $\hat{y}, M, c, S$ 
```

Two design choices in this pipeline are directly supported by the evidence. Fitting the decision layer on out-of-fold predictions rather than on training-set outputs is what keeps the train-test gap of the best stacked ensemble at 0.1% [40]. Retaining handcrafted descriptors alongside learned features is what allows the fuzzy-clustering plus SVM design to match deep baselines at a fraction of the runtime [19], and preserves stage-level auditability that a monolithic network cannot offer.

VIII. CONCLUSION AND FUTURE DIRECTIONS

This review synthesised 41 primary studies published between 2023 and 2026 on machine learning for brain tumor detection and segmentation in MRI. Four architectural families were distinguished and compared on normalised metrics. Reported accuracy occupies a narrow 93.30-99.70% band and Dice a wide 0.650-0.961 band, and the analysis shows that the variance in both is governed principally by dataset composition, lesion size distribution and scoring protocol rather than by architecture. Hybrid designs that combine convolutional local sensitivity with transformer global context and fuse decisions through a learned meta-classifier currently offer the best accuracy-stability trade-off, while classical feature-engineered pipelines remain competitive where interpretability and runtime matter.

The field's binding constraint is no longer model capacity but evidence quality. Priorities are therefore protocol standardisation with size-stratified and lesion-wise reporting; federated and augmentation-aware training to address demographic and scanner shift; genuine multimodal fusion of imaging with clinical and genomic metadata; intrinsically interpretable architectures with calibrated uncertainty rather than post-hoc saliency; and prospective, reader-study-based clinical validation. Progress on these fronts, rather than further



REFERENCES

1. B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
2. Z. Liu et al., "Deep learning based brain tumor segmentation: A survey," *Complex & Intelligent Systems*, vol. 9, no. 1, pp. 1001–1026, 2023.
3. V. Gupta and V. Bibhu, "Deep residual network-based brain tumor segmentation and detection with MRI using improved invasive bat algorithm," *Multimedia Tools and Applications*, vol. 82, no. 8, pp. 12445–12467, 2023.
4. A. Kumar, "Study and analysis of different segmentation methods for brain tumor MRI application," *Multimedia Tools and Applications*, vol. 82, no. 5, pp. 7117–7139, 2023.
5. T. Balamurugan and E. Gnanamanoharan, "Brain tumor segmentation and classification using hybrid deep CNN with LuNetClassifier," *Neural Computing and Applications*, vol. 35, no. 6, pp. 4739–4753, 2023, doi: 10.1007/s00521-022-07934-7.
6. C.-J. Tseng and C. Tang, "An optimized XGBoost technique for accurate brain tumor detection using feature selection and image segmentation," *Healthcare Analytics*, vol. 4, art. 100217, 2023.
7. M. Ahmadi, A. Sharifi, M. Jafarian Fard, and N. Soleimani, "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA," *International Journal of Neuroscience*, vol. 133, no. 1, pp. 55–66, 2023.
8. F. Ahmed, M. Asif, M. Saleem, U. F. Mushtaq, and M. Imran, "Identification and prediction of brain tumor using VGG-16 empowered with explainable artificial intelligence," *International Journal of Computational and Innovative Sciences*, vol. 2, no. 2, pp. 24–33, 2023.
9. T. H. Shemanto, L. B. Billah, and M. A. Ibtesham, "A novel method of thresholding for brain tumor segmentation and detection," in *Proc. ICICTD 2022*, Singapore: Springer Nature, 2023, pp. 277–289.
10. S. Saeedi, S. Rezayi, H. Keshavarz, and S. R. Niakan Kalhori, "MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, art. 16, 2023.
11. A. A. Asiri et al., "Brain tumor detection and classification using fine-tuned CNN with ResNet50 and U-Net model: A study on TCGA-LGG and TCIA dataset for MRI applications," *Life*, vol. 13, no. 7, art. 1449, 2023.
12. A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain tumor detection based on deep learning approaches and magnetic resonance imaging," *Cancers*, vol. 15, no. 16, art. 4172, 2023.
13. S. Metlek and H. Çetiner, "ResUNet+: A new convolutional and attention block-based approach for brain tumor segmentation," *IEEE Access*, vol. 11, 2023.
14. Y. M. Mohammed, S. El Garouani, and I. Jellouli, "A survey of methods for brain tumor segmentation based on MRI images," *Journal of Computational Design and Engineering*, vol. 10, no. 1, pp. 266–293, 2023.
15. S. Rajagopal, T. Thanarajan, Y. Alotaibi, and S. Alghamdi, "Brain tumor: Hybrid feature extraction based on UNet and 3DCNN," *Computer Systems Science and Engineering*, vol. 45, no. 2, 2023.
16. M. I. Mahmud, M. Mamun, and A. Abdelgawad, "A deep analysis of brain tumor detection from MRI images using deep learning networks," *Algorithms*, vol. 16, no. 4, art. 176, 2023.



17. S. Bouhafra and H. El Bahi, "Deep learning approaches for brain tumor detection and classification using MRI images (2020 to 2024): A systematic review," *Journal of Imaging Informatics in Medicine*, vol. 38, no. 3, pp. 1403–1433, 2025.
18. Y. Zakeri, B. Karasfi, and A. Jalalian, "A review of brain tumor segmentation using MRIs from 2019 to 2023," *Journal of Medical and Biological Engineering*, vol. 44, no. 2, pp. 155–180, 2024.
19. S. M. Alqhtani et al., "Improved brain tumor segmentation and classification in brain MRI with FCM-SVM: A diagnostic approach," *IEEE Access*, vol. 12, pp. 61312–61335, 2024.
20. D. LaBella et al., "Brain tumor segmentation (BraTS) challenge 2024: Meningioma radiotherapy planning automated segmentation," *arXiv:2405.18383*, 2024.
21. U. Humayun, M. T. Yaseen, A. Shahwaiz, and A. Iftikhar, "Deep learning approaches for brain tumor detection and segmentation in MRI imaging," *Journal of Computing and Biomedical Informatics*, vol. 8, no. 1, 2024.
22. S. Hashmi et al., "Optimizing brain tumor segmentation with MedNeXt: BraTS 2024 SSA and pediatrics," *arXiv:2411.15872*, 2024.
23. C. M. Umarani, S. G. Gollagi, S. Allagi, K. Sambrekar, and S. B. Ankali, "Advancements in deep learning techniques for brain tumor segmentation: A survey," *Informatics in Medicine Unlocked*, art. 101576, 2024.
24. R. Ahsan, I. Shahzadi, F. Najeeb, and H. Omer, "Brain tumor detection and segmentation using deep learning," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 38, no. 1, pp. 13–22, 2025.
25. S. Ganesh, R. Gomathi, and S. Kannadhasan, "Brain tumor segmentation and detection in MRI using convolutional neural networks and VGG16," *Cancer Biomarkers*, vol. 42, no. 3, 2025.
26. A. W. Moawad et al., "The brain tumor segmentation-metastases (BraTS-METS) challenge 2023: Brain metastasis segmentation on pre-treatment MRI," *arXiv:2306.00838*, 2024.
27. M. Ariful Islam, M. F. Mridha, M. Safran, S. Alfarhood, and M. Mohsin Kabir, "Revolutionizing brain tumor detection using explainable AI in MRI images," *NMR in Biomedicine*, vol. 38, no. 3, art. e70001, 2025.
28. P. K. Tiwary, P. Johri, A. Katiyar, and M. K. Chhipa, "Deep learning-based MRI brain tumor segmentation with EfficientNet-enhanced UNet," *IEEE Access*, 2025.
29. N. Gökmen, "AI techniques for brain tumor segmentation in MRI: A review (2019–2024)," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 14, no. 1, art. 168, 2025.
30. S. Kaur, U. Mittal, and A. Wadhawan, "A systematic review of deep learning approaches for brain tumor segmentation in MRI: Trends, challenges and future directions," *Archives of Computational Methods in Engineering*, 2025.
31. A. Hamza and R. Damaševičius, "Deep learning for brain tumor segmentation and classification: A systematic review of methods and trends," *Computers, Materials and Continua*, vol. 86, no. 1, pp. 1–41, 2026.
32. N. Rasool and J. I. Bhat, "A critical review on segmentation of glioma brain tumor and prediction of overall survival," *Archives of Computational Methods in Engineering*, vol. 32, no. 3, pp. 1525–1569, 2025.



33. N. Musthafa, Q. A. Memon, and M. M. Masud, "Advancing brain tumor analysis: Current trends, key challenges and perspectives in deep learning-based brain MRI tumor diagnosis," *Eng*, vol. 6, no. 5, art. 82, 2025.
34. N. Musthafa, M. M. Masud, Q. Memon, and M. Farhad, "VISION: View-specific integrated segmentation-classification framework for accurate brain tumor detection in MRI scans," *PLoS ONE*, vol. 20, no. 10, art. e0332395, 2025.
35. R. N. Asif et al., "Brain tumor detection empowered with ensemble deep learning approaches from MRI scan images," *Scientific Reports*, vol. 15, no. 1, art. 15002, 2025.
36. S. S. Hussain, N. A. Wani, J. Kaur, N. Ahmad, and S. Ahmad, "Next-generation automation in neuro-oncology: Advanced neural networks for MRI-based brain tumor segmentation and classification," *IEEE Access*, vol. 13, pp. 41141–41158, 2025.
37. A. Rahman, A. Satti, A. R. Shahid, Q. M. Shafi, K. Farooq, and A. A. Safi, "TDA SegUNet: Topological data analysis-based shape-aware brain tumor segmentation," *IEEE Access*, vol. 13, pp. 36190–36200, 2025.
38. M. Aamir, Z. Rahman, N. Choudhry, J. A. Bhutto, W. A. Abro, and Z. Zhu, "From CNNs to transformers: A review of evolving deep learning architectures for brain tumor classification," *IEEE Access*, vol. 13, pp. 184918–184936, 2025.
39. N. M. H. Hassan and W. Boulila, "Efficient approach for brain tumor detection and classification using fuzzy thresholding and deep learning algorithms," *IEEE Access*, vol. 13, pp. 78808–78832, 2025.
40. N. Vijayvargiya and N. Singh, "A meta-learner driven CNN–transformer stacked ensemble for high-accuracy brain tumor classification," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 18, no. 3s, pp. 1272–1291, 2026, doi: 10.70917/ijcism-2026-2460.
41. J. Cheng et al., "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS ONE*, vol. 10, no. 10, art. e0140381, 2015 (brain tumor dataset, figshare, doi: 10.6084/m9.figshare.1512427).
42. M. D. Alanazi et al., "Magnetic resonance imaging brain segmentation using bi-directional convolutional long short-term memory U-Net with densely connected convolutions," *IEEE Access*, vol. 13, pp. 75847–75860, 2025.