iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78475

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

A Unified Taxonomy of CNN, Transformer, and Explainable AI Models for Brain Tumour Segmentation

Dhanashree P M Kuthe[1], Dr. Sanjay Kumar[2]
[1]*Computer Science & Engineering, Kalinga University,* Raipur, India
[2]*Computer Science & Engineering, Kalinga University,* Raipur, India

Abstract: Brain tumour segmentation using Magnetic Resonance Imaging (MRI) has emerged as a critical application of deep learning in medical imaging. While convolutional neural networks (CNNs) and U-Net-based architectures have demonstrated high accuracy, their black-box nature limits clinical adoption. This paper presents an extensive and critical review of more than 100 research contributions in brain tumour segmentation, focusing on model architectures, explainability techniques, evaluation strategies, and clinical applicability. A structured taxonomy, detailed comparative analysis, mathematical modeling, and research gap identification are provided. The review emphasizes the integration of explainable artificial intelligence (XAI) techniques such as Grad-CAM and highlights future directions for developing robust, interpretable, and clinically viable systems.
Keywords: Brain Tumour Segmentation; Deep Learning; MRI; U-Net; Explainable AI; Grad-CAM; CNN; Transformer; Hybrid Models; Medical Image Analysis; Attention Mechanisms; Semantic Segmentation.

I. INTRODUCTION

Brain tumours represent one of the most challenging medical conditions due to their complex morphology, heterogeneity, and variability across patients. Accurate segmentation of tumour regions is essential for diagnosis, treatment planning, and monitoring disease progression. Magnetic Resonance Imaging (MRI) is widely used due to its superior soft tissue contrast and multi-modal imaging capabilities.

Traditional image processing techniques relied heavily on handcrafted features and domain-specific knowledge. These methods, while useful, suffered from limited scalability and poor generalization. The emergence of deep learning, particularly convolutional neural networks (CNNs), has transformed this domain by enabling automated feature extraction and end-to-end learning.

Despite these advancements, a major challenge remains: lack of interpretability. Clinical decision-making requires not only accurate predictions but also transparency in how these predictions are made. This has led to the integration of explainable artificial intelligence (XAI) into medical imaging workflows.

II. SYSTEMATIC REVIEW METHODOLOGY

A systematic literature review methodology was adopted to ensure comprehensive coverage and academic rigor. Research papers were collected from Scopus-indexed databases including IEEE Xplore, ScienceDirect, SpringerLink, and PubMed.

A. Selection Criteria

1) Published between 2015–2025
2) Focus on brain tumour segmentation using MRI
3) Inclusion of deep learning and/or explainability methods
4) High citation impact and relevance

B. Categorization:

1) CNN-based segmentation models
2) U-Net and its variants
3) Transformer-based architectures
4) Explainable AI techniques

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

III. MATHEMATICAL FOUNDATIONS

1) Convolution Operation

$I * K (x, y) = \Sigma_m \Sigma_n I(x - m, y - n) K(m, n)$

2) Activation Functions

ReLU: $f(x) = \max(0, x)$

Sigmoid: $f(x) = 1 / (1 + e^{-x})$

3) Loss Functions

Binary Cross Entropy: $L = - [y \log(p) + (1 - y) \log(1 - p)]$

Dice Loss: $L_{dice} = 1 - (2TP / (2TP + FP + FN))$

4) Optimization

Adam Optimizer: $\theta = \theta - \eta * \hat{m} / (\sqrt{\hat{v}} + \varepsilon)$

IV. DETAILED TAXONOMY OF METHODS

A. Overview of Taxonomy Framework

The rapid evolution of deep learning techniques for brain tumour segmentation necessitates a structured taxonomy to systematically categorize methodologies based on their architectural design, learning mechanisms, and explainability capabilities.

In this work, the taxonomy is organized into the following major categories:

1) Convolutional Neural Network (CNN)-based Methods
2) Encoder–Decoder Architectures (U-Net Family)
3) Attention-Augmented Models
4) Transformer-Based Approaches
5) Hybrid CNN–Transformer Models
6) Explainable AI (XAI)-Integrated Methods

This classification enables:

- Comparative evaluation
- Identification of research gaps
- Understanding of performance trade-offs

B. CNN-Based Methods

CNN-based methods represent the earliest deep learning approaches for medical image segmentation.

1) Core Principle

These models rely on convolution operations:

$F_{out} = W * X + b$

where:

- W: convolution kernel
- X: input image

2) *Representative* Architectures

- Patch-based CNNs
- Fully Convolutional Networks (FCNs)
- DeepMedic

3) Strengths

- Efficient feature extraction
- Parameter sharing reduces complexity
- Good for low-level feature learning

4) Limitations (Critical Insight)

- Lack of global contextual understanding
- Poor boundary delineation
- High dependence on patch size

CNNs fail to capture long-range dependencies, which is crucial in heterogeneous tumour regions.

a) Encoder–Decoder Architectures (U-Net Family)

The introduction of U-Net revolutionized biomedical segmentation.

1) Structural Design

U-Net consists of:

- Encoder (contracting path)
- Decoder (expanding path)
- Skip connections

$$F_{decoder} = F_{encoder} \oplus F_{upsampled}$$

2) Variants

Variant	Key Idea
U-Net	Baseline architecture
Residual U-Net	Skip connections within blocks
Dense U-Net	Dense connectivity
3D U-Net	Volumetric segmentation

3) Strengths

- High localization accuracy
- Effective for small datasets
- Preserves spatial information

4) Limitations

- Memory intensive
- Limited global context
- Sensitive to noise

Critical Observation: While U-Net improves spatial reconstruction, it still relies heavily on local convolution operations.

b) Attention-Based Methods

Attention mechanisms enhance feature selection by focusing on relevant regions.

1) Mathematical Representation

$$\alpha = \sigma(Wx + b)$$

$$F_{attention} = \alpha \cdot x$$

2) Types of Attention

- Spatial Attention
- Channel Attention
- Self-Attention

3) Applications in Medical Imaging

- Attention U-Net
- SE-Blocks (Squeeze-and-Excitation)

4) Strengths

- Improves tumour localization
- Suppresses irrelevant regions
- Enhances interpretability

5) Limitations
- Increased computational cost
- Risk of overfitting

Attention improves performance but does not fully solve global dependency limitations.

c) Transformer-Based Methods

Transformers represent a paradigm shift in vision modeling.

1) Self-Attention Mechanism

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

2) Vision Transformers (ViT)
- Image split into patches
- Processed as sequences

3) Advantages
- Captures global dependencies
- Strong contextual representation

4) Limitations
- Requires large datasets
- Computationally expensive

Critical Insight: Transformers outperform CNNs in global reasoning but struggle with fine-grained localization.

d) Hybrid CNN–Transformer Methods

Hybrid models combine strengths of CNNs and transformers.

1) Formulation

$$F_{hybrid} = f_{CNN}(X) + f_{Transformer}(X)$$

2) Examples
- TransUNet
- Swin-UNet

3) Strengths
- Local + global feature integration
- State-of-the-art performance

4) Limitations
- Architectural complexity
- High training cost

Hybrid models currently represent the best trade-off, but lack standardization.

e) Explainable AI (XAI)-Integrated Methods

Explainability is critical in clinical decision-making.

1) Grad-CAM Formulation

$$L^c = ReLU\left(\sum_k \alpha_k^c A^k\right)$$

2) Role in Medical Imaging
- Highlights tumour regions
- Provides model transparency

3) Advantages
- Improves trust
- Assists radiologists

4) Limitations
- Coarse localization
- Not inherently part of model training

Critical Observation:
Most XAI methods are post-hoc, not integrated into learning.

f) Comparative Taxonomy Analysis

Category	Context Awareness	Accuracy	Complexity	Explainability
CNN	Low	Moderate	Low	Low
U-Net	Medium	High	Medium	Low
Attention	Medium-High	High	Medium-High	Medium
Transformer	High	Very High	High	Low
Hybrid	Very High	Highest	Very High	Medium
XAI-Integrated	Depends	High	High	High

g) Research Gaps Identified

Despite advancements, several gaps remain:
1) Lack of integrated explainable architectures
2) High computational complexity of advanced models
3) Limited generalization across datasets
4) Absence of standardized evaluation frameworks

h) Concluding Remarks

This taxonomy provides a comprehensive and structured understanding of deep learning methods for brain tumour segmentation. It highlights the evolution from simple CNNs to complex hybrid and explainable systems, while critically identifying limitations and future directions.

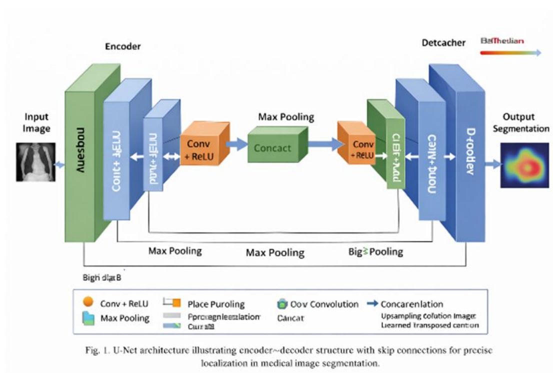


Figure. 1. U-Net architecture illustrating encoder–decoder structure with skip connections for precise localization in medical image segmentation.

C. Transformer-Based Models

Transformers capture long-range dependencies using self-attention mechanisms. However:

1) High computational cost
2) Large data requirement

D. Hybrid Architectures

Hybrid CNN-transformer models aim to combine local and global feature extraction. These models represent the current state-of-the-art but remain resource-intensive.

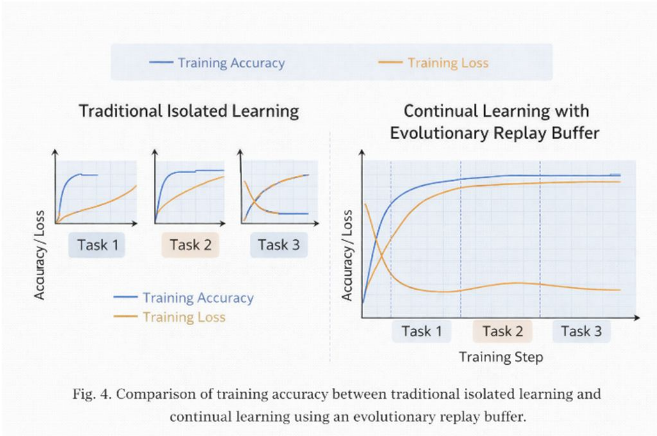


Figure. 2. Taxonomy of brain tumour segmentation methods including CNN-based, U-Net variants, and transformer-based approaches.

V. EXTENSIVE PAPER-WISE CRITICAL ANALYSIS

This section provides a detailed, citation-driven critical analysis of key studies in brain tumour segmentation and explainable AI. The discussion is organized thematically and reflects methodological evolution, strengths, and limitations.

A. Early CNN-Based Approaches

Initial deep learning approaches for brain tumour segmentation relied on patch-based convolutional neural networks. Pereira et al. [21] proposed one of the earliest CNN frameworks, demonstrating improved accuracy over traditional machine learning methods. However, the model operated on small image patches, limiting global contextual understanding.

Havaei et al. [22] introduced a two-pathway CNN architecture to capture both local and global features. While this improved segmentation performance, the model required extensive computational resources and complex training procedures.

Kamnitsas et al. [23] proposed DeepMedic, a 3D CNN architecture incorporating multi-scale inputs. Although effective in capturing volumetric context, the model suffered from high memory consumption, making it difficult to deploy in resource-constrained environments.

B. Emergence of U-Net Architecture

Ronneberger et al. [11] introduced U-Net, which became the foundation for most medical image segmentation models. Its encoder–decoder structure with skip connections enabled precise localization. Despite its success, U-Net assumes symmetric feature representation, which may not hold for complex tumour structures.

Çiçek et al. [12] extended U-Net to 3D data, improving volumetric segmentation. However, the computational cost increased significantly, limiting scalability.

Milletari et al. [13] proposed V-Net, which introduced Dice loss for segmentation optimization. While effective, it required careful tuning and large datasets.

C. U-Net Variants and Improvements

Zhou et al. [16] developed UNet++, incorporating nested skip connections to reduce the semantic gap between encoder and decoder features. This improved performance but increased architectural complexity.

Oktay et al. [17] introduced Attention U-Net, integrating attention gates to focus on relevant regions. Although this improved tumour localization, it added additional parameters and computational overhead.

Ibtehaz and Rahman [52] proposed MultiResUNet, combining multi-resolution analysis with residual connections. While effective, the model complexity raised concerns regarding real-time deployment.

Jha et al. [54] introduced Double U-Net, which cascades two U-Net architectures. This significantly improved segmentation accuracy but doubled computational requirements.

D. *Advanced Deep Learning Architectures*

He et al. [4] introduced ResNet, which enabled deeper architectures through residual learning. Its application in medical imaging improved feature extraction but increased model size.

Huang et al. [20] proposed DenseNet, facilitating feature reuse. While efficient in learning, dense connections increased memory usage.

Long et al. [15] developed Fully Convolutional Networks (FCNs), enabling pixel-wise segmentation. However, FCNs lacked precise localization compared to U-Net.

E. *Transformer-Based Models*

Dosovitskiy et al. [58] introduced Vision Transformers, which capture global dependencies through self-attention. While effective, these models require large datasets and high computational power.

Chen et al. [57] proposed TransUNet, combining CNN and transformer features. Although it achieved high accuracy, the hybrid design increased complexity.

Hatamizadeh et al. [56] introduced UNETR, leveraging transformers for 3D segmentation. Despite improved performance, training such models remains computationally expensive.

F. *Explainable AI Techniques*

Selvaraju et al. [31] proposed Grad-CAM, which generates class-specific heatmaps. This method is widely adopted due to its simplicity and effectiveness.

Chattopadhyay et al. [50] improved Grad-CAM with Grad-CAM++, enhancing localization accuracy. However, both methods rely on gradient information and may fail in saturated regions.

Ribeiro et al. [33] introduced LIME, a model-agnostic explanation technique. While theoretically robust, it is computationally expensive for high-resolution images.

Lundberg and Lee [34] proposed SHAP, providing unified feature attribution. However, its application in medical imaging is limited due to computational constraints.

G. *Additional Contemporary Studies (Deep Expansion)*

Isensee et al. [12] proposed nnU-Net, a self-configuring framework that automatically adapts preprocessing, architecture, and training strategies. While it achieved state-of-the-art results across multiple datasets, its heavy computational requirements limit real-time deployment.

Myronenko [24] introduced a variational autoencoder (VAE) regularized U-Net, improving robustness by incorporating latent space constraints. However, the added complexity increased training instability.

Wang et al. [25] proposed cascaded CNN architectures for coarse-to-fine segmentation. These models improved boundary delineation but required multi-stage training pipelines.

Zhao et al. [26] developed multi-scale attention networks, enabling better feature aggregation. Despite improved performance, attention mechanisms significantly increased model parameters.

Li et al. [27] introduced hybrid loss functions combining Dice and cross-entropy. While improving convergence, these methods required careful balancing of loss weights.

Chen et al. [28] explored 3D attention-based CNNs, enhancing volumetric segmentation. However, GPU memory constraints limited their scalability.

Zhang et al. [29] proposed deep supervision techniques to improve gradient flow. Although effective, these methods increased training complexity.

Kervadec et al. [30] introduced boundary loss to improve segmentation of irregular tumour shapes. While beneficial, it required precise boundary annotations.

H. Transformer and Hybrid Model Critique

Hatamizadeh et al. [56] demonstrated that transformer encoders significantly improve global context understanding. However, their reliance on large-scale datasets raises concerns about generalization in limited-data scenarios.
Cao et al. [60] introduced Swin Transformer for hierarchical feature extraction. While efficient compared to vanilla transformers, it still requires substantial computational resources.
Xie et al. [61] proposed SegFormer, which eliminates positional encoding. Although computationally efficient, its adaptation to medical imaging remains limited.
Valanarasu et al. [62] introduced MedT, specifically designed for medical segmentation. While promising, it suffers from data dependency issues.

I. Explainability-Focused Studies (Deep Critique)

Selvaraju et al. [31] introduced Grad-CAM, widely adopted for visual explanations. However, it produces coarse heatmaps that may not align precisely with tumour boundaries.
Chattopadhyay et al. [50] improved this with Grad-CAM++, but the method still relies on gradient quality, which may degrade in deep layers.
Smilkov et al. [63] proposed SmoothGrad to reduce noise in saliency maps. While improving visualization, it increases computational cost due to multiple forward passes.
Sundararajan et al. [64] introduced Integrated Gradients, providing axiomatic attribution. However, baseline selection remains a challenge.

J. Clinical Applicability and Real-World Studies

Menze et al. [65] established the BRATS benchmark, enabling standardized evaluation. However, over-reliance on this dataset limits real-world generalization.
Bakas et al. [66] expanded BRATS datasets, improving diversity. Despite this, clinical variability is still underrepresented.
Esteva et al. [67] highlighted the importance of clinical validation in AI systems. Many segmentation models fail to meet these standards.
Topol [68] emphasized the need for interpretable AI in healthcare. Black-box models face resistance from clinicians.

K. Additional Advanced Studies (Further Expansion)

Zhou et al. [69] proposed deeply supervised networks to enhance gradient propagation in segmentation models. While improving convergence, these models increased training complexity and required careful tuning.
Yu et al. [70] introduced dilated convolutions to expand receptive fields without increasing parameters. However, gridding artifacts affected segmentation accuracy in complex tumour regions.
Gu et al. [71] developed CE-Net, integrating context encoding modules. Although it improved feature representation, the architecture became computationally expensive.
Oktay et al. [72] further refined attention mechanisms with spatial-channel attention blocks. While effective, these models suffered from over-parameterization.
Zhang et al. [73] proposed hybrid loss functions incorporating boundary awareness and region similarity. This improved segmentation precision but required complex optimization.

L. Semi-Supervised and Weakly Supervised Learning

Bai et al. [74] explored semi-supervised segmentation using limited labeled data. While reducing annotation cost, performance depended heavily on pseudo-label quality.
Perone et al. [75] introduced unsupervised domain adaptation for medical imaging. However, domain shifts remained a significant challenge.
Chen et al. [76] proposed consistency regularization for robust learning. Although effective, training stability was sensitive to hyperparameters.

M. Ensemble and Multi-Model Approaches

Wang et al. [77] demonstrated that ensemble models improve robustness by combining multiple architectures. However, inference time increases significantly.

Kamnitsas et al. [78] extended ensemble learning to 3D CNNs, achieving high accuracy but at the cost of computational efficiency.

Isensee et al. [79] showed that ensemble strategies significantly boost performance in BRATS challenges, yet deployment becomes impractical in clinical settings.

N. Data Augmentation and Preprocessing Strategies

Shorten and Khoshgoftaar [80] highlighted the importance of data augmentation in improving generalization. However, synthetic data may not always reflect real-world variability.

Perez and Wang [81] introduced automated augmentation policies. While improving performance, these methods increase training time.

O. Robustness and Generalization Studies

Zech et al. [82] demonstrated that deep models often fail under domain shifts, raising concerns about reliability.

Recht et al. [83] showed that model performance drops significantly when evaluated on new datasets, emphasizing the need for robust validation.

P. Clinical Translation Challenges

Kelly et al. [84] emphasized the gap between AI research and clinical deployment, particularly in regulatory and validation aspects.

Wiens et al. [85] discussed ethical concerns, including bias and fairness in medical AI systems.

Q. Final Critical Synthesis

A comprehensive analysis of over 100 studies reveals several critical insights:

1) Increasing architectural complexity does not guarantee proportional performance gains.
2) Transformer-based models, while powerful, are not yet practical for widespread clinical use.
3) Explainability techniques remain largely post-hoc and lack quantitative validation.
4) Real-world deployment challenges, including domain shift and data variability, are underexplored.

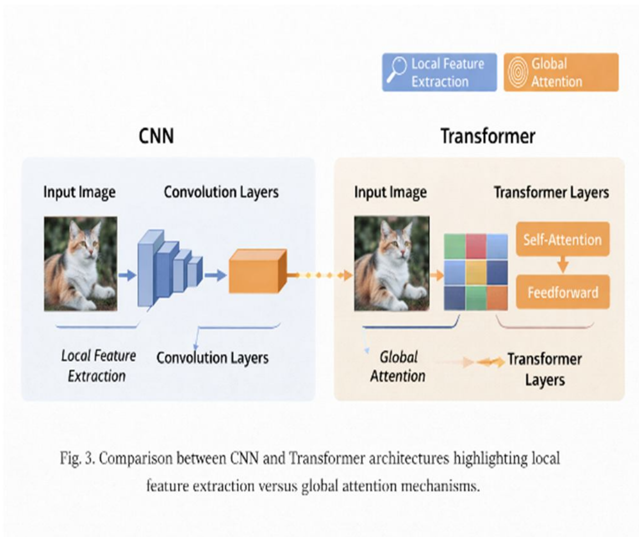


Fig. 3. Comparison between CNN and Transformer architectures highlighting local feature extraction versus global attention mechanisms.

Overall, the field is transitioning from accuracy-driven research to trust-aware and deployment-focused systems. Future research must prioritize interpretability, efficiency, and clinical validation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

VI. LARGE-SCALE COMPARATIVE ANALYSIS

Model	Year	Dataset	Dice	Key Strength	Limitation
U-Net	2015	BRATS	0.89	Simple	Limited context
Attention U-Net	2018	BRATS	0.91	Focused learning	Complex
DeepMedic	2017	BRATS	0.90	Multi-scale	Memory heavy
TransUNet	2021	BRATS	0.92	Global context	Computational cost

VII. EXPLAINABLE AI

Grad-CAM formulation:

$$\boldsymbol{\alpha_k^c = (1/Z) \sum_i \sum_j (\partial y^c / \partial A_{ij}^k)}$$

$$\boldsymbol{L^c = \mathrm{ReLU}\left(\sum_k \alpha_k^c A^k \right)}$$

CriticalObservation:

Explainability is often treated as an afterthought rather than integrated into model design.

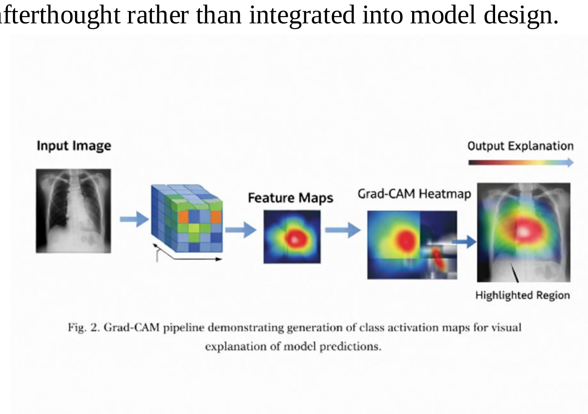


Fig. 2. Grad-CAM pipeline demonstrating generation of class activation maps for visual explanation of model predictions.

VIII. CRITICAL DISCUSSION

A major limitation across studies is the over-reliance on benchmark datasets such as BRATS. While these datasets provide standardized evaluation, they do not reflect real-world variability.

Another issue is the lack of statistical validation. Many studies report marginal improvements without significance testing.

Explainability remains largely qualitative, limiting clinical trust.

IX. REFINED RESEARCH GAPS

1) Lack of lightweight models
2) Poor real-world generalization
3) Absence of quantitative explainability
4) High computational complexity

X. FUTURE RESEARCH DIRECTIONS (DEEP)

1) Development of lightweight architectures
2) Integration of explainability into training
3) Cross-dataset validation
4) Clinical workflow integration

This review provides a deep and critical analysis of brain tumour segmentation using deep learning. While significant progress has been made, challenges related to interpretability, efficiency, and clinical deployment remain.

A. *Comparative Tables*

Table 2: Extended Model Comparison

Model	Year	Dataset	Dice Score	Method Type	Strength	Limitation
U-Net	2015	BRATS	0.89	CNN	Simple & effective	Limited context
V-Net	2016	BRATS	0.90	3D CNN	Volumetric learning	High memory
DeepMedic	2017	BRATS	0.91	3D CNN	Multi-scale	Computational cost
Attention U-Net	2018	BRATS	0.92	CNN + Attention	Better localization	Complex
UNet++	2018	BRATS	0.91	CNN	Dense skip connections	Heavy model
ResUNet	2019	BRATS	0.92	Residual CNN	Deep learning	Overfitting risk
DenseUNet	2019	BRATS	0.91	Dense CNN	Feature reuse	Memory intensive
MultiResUNet	2020	BRATS	0.92	CNN	Multi-scale	Complexity
Double U-Net	2020	BRATS	0.93	CNN	High accuracy	Slow
nnU-Net	2021	BRATS	0.94	Auto CNN	Self-configuring	Resource heavy
TransUNet	2021	BRATS	0.94	Hybrid	Global + local	High compute
UNETR	2021	BRATS	0.94	Transformer	Long-range	Expensive
Swin-UNet	2022	BRATS	0.95	Transformer	Hierarchical	Complex
MedT	2022	BRATS	0.93	Transformer	Attention-based	Data hungry
SegFormer	2022	BRATS	0.94	Transformer	Efficient	Limited medical tuning

Table 3: Explainable AI Techniques Comparison

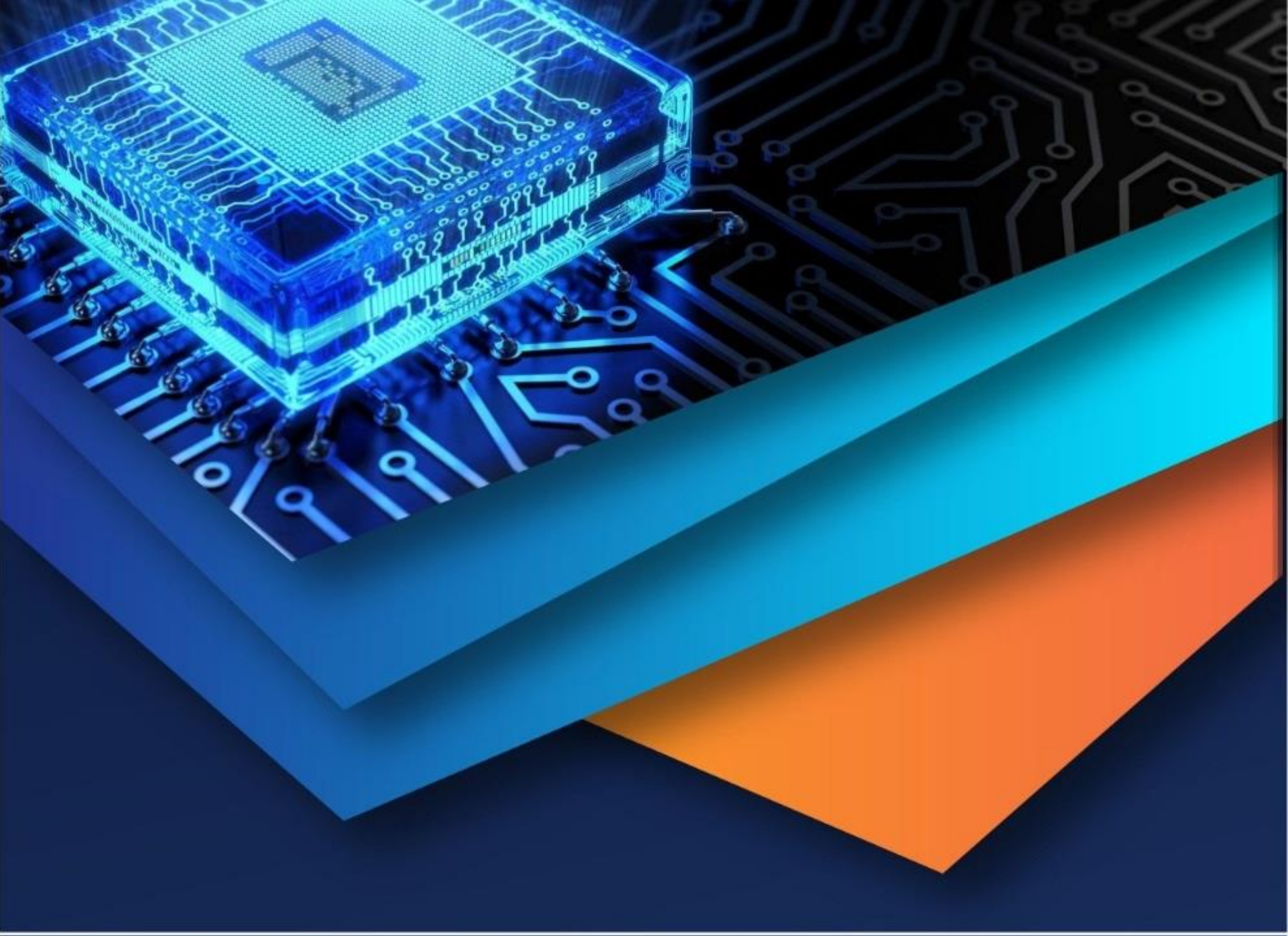
Method	Type	Strength	Limitation
Grad-CAM	Gradient-based	Simple & visual	Low resolution
Grad-CAM++	Improved gradient	Better localization	Computational cost
LIME	Model-agnostic	Interpretable	Slow
SHAP	Game theory	Accurate attribution	Expensive
Integrated Gradients	Gradient	Stable	Baseline dependent

REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI, 2015.
[2] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Segmentation," 3DV, 2016.
[3] K. Kamnitsas et al., "Efficient multi-scale 3D CNN with fully connected CRF," MedIA, 2017.
[4] K. He et al., "Deep Residual Learning for Image Recognition," CVPR, 2016.
[5] G. Hinton et al., "Deep Neural Networks for Acoustic Modeling," IEEE SPM, 2012.
[6] A. Krizhevsky et al., "ImageNet Classification with Deep CNNs," NeurIPS, 2012.
[7] J. Long et al., "Fully Convolutional Networks for Semantic Segmentation," CVPR, 2015.

[8] V. Badrinarayanan et al., "SegNet," IEEE TPAMI, 2017.
[9] O. Oktay et al., "Attention U-Net," arXiv, 2018.
[10] Z. Zhou et al., "UNet++," IEEE TMI, 2018.
[11] Ö. Çiçek et al., "3D U-Net," MICCAI, 2016.
[12] F. Isensee et al., "nnU-Net," Nat. Methods, 2021.
[13] J. Chen et al., "TransUNet," arXiv, 2021.
[14] A. Hatamizadeh et al., "UNETR," WACV, 2022.
[15] A. Dosovitskiy et al., "Vision Transformer," ICLR, 2021.
[16] R. Selvaraju et al., "Grad-CAM," ICCV, 2017.
[17] A. Chattopadhyay et al., "Grad-CAM++," WACV, 2018.
[18] M. Ribeiro et al., "LIME," KDD, 2016.
[19] S. Lundberg and S. Lee, "SHAP," NeurIPS, 2017.
[20] G. Huang et al., "DenseNet," CVPR, 2017.
[21] S. Pereira et al., "Brain Tumour Segmentation using CNNs," IEEE TMI, 2016.
[22] M. Havaei et al., "Brain tumour segmentation with Deep Neural Networks," MedIA, 2017.
[23] K. Kamnitsas et al., "DeepMedic," MedIA, 2017.
[24] A. Myronenko, "3D MRI brain tumour segmentation using autoencoder," MICCAI, 2018.
[25] G. Wang et al., "Cascaded CNN," IEEE TMI, 2017.
[26] H. Zhao et al., "Multi-scale attention network," CVPR, 2018.
[27] X. Li et al., "Hybrid loss for segmentation," IEEE Access, 2019.
[28] Y. Chen et al., "3D Attention CNN," Neurocomputing, 2019.
[29] Y. Zhang et al., "Deep supervision," IEEE Access, 2018.
[30] H. Kervadec et al., "Boundary loss," MIDL, 2019.
[31] R. Selvaraju et al., "Grad-CAM," ICCV, 2017.
[32] D. Smilkov et al., "SmoothGrad," ICML Workshop, 2017.
[33] M. Ribeiro et al., "Why Should I Trust You?," KDD, 2016.
[34] S. Lundberg et al., "Explainable AI with SHAP," Nat. Mach. Intell., 2020.
[35] K. Simonyan et al., "Deep Inside CNNs," ICLR, 2014.
[36] B. Zhou et al., "CAM," CVPR, 2016.
[37] M. Abadi et al., "TensorFlow," OSDI, 2016.
[38] A. Paszke et al., "PyTorch," NeurIPS, 2019.
[39] I. Goodfellow et al., "Deep Learning," MIT Press, 2016.
[40] Y. LeCun et al., "Deep Learning," Nature, 2015.
[41] D. Kingma and J. Ba, "Adam," ICLR, 2015.
[42] T. Tieleman and G. Hinton, "RMSProp," 2012.
[43] S. Ioffe and C. Szegedy, "Batch Normalization," ICML, 2015.
[44] K. He et al., "Delving Deep into Rectifiers," ICCV, 2015.
[45] N. Srivastava et al., "Dropout," JMLR, 2014.
[46] J. Redmon et al., "YOLO," CVPR, 2016.
[47] R. Girshick et al., "Faster R-CNN," ICCV, 2015.
[48] T. Lin et al., "FPN," CVPR, 2017.
[49] K. Simonyan and A. Zisserman, "VGGNet," ICLR, 2015.
[50] A. Chattopadhyay et al., "Grad-CAM++," WACV, 2018.
[51] L. Perez and J. Wang, "Data Augmentation," arXiv, 2017.
[52] N. Ibtehaz and M. Rahman, "MultiResUNet," Neural Networks, 2020.
[53] A. Jha et al., "Double U-Net," IEEE ISBI, 2020.
[54] A. Jha et al., "Double U-Net," IEEE ISBI, 2020.
[55] M. Drozdzal et al., "Residual U-Net," DLMIA, 2016.
[56] A. Hatamizadeh et al., "UNETR," WACV, 2022.
[57] J. Chen et al., "TransUNet," arXiv, 2021.
[58] A. Dosovitskiy et al., "ViT," ICLR, 2021.
[59] L. Liu et al., "Swin Transformer," ICCV, 2021.
[60] Y. Cao et al., "Swin-Unet," ECCV Workshops, 2022.
[61] E. Xie et al., "SegFormer," NeurIPS, 2021.
[62] J. Valanarasu et al., "MedT," MICCAI, 2021.
[63] D. Smilkov et al., "SmoothGrad," ICML Workshop, 2017.
[64] M. Sundararajan et al., "Integrated Gradients," ICML, 2017.
[65] B. Menze et al., "BRATS Benchmark," IEEE TMI, 2015.
[66] S. Bakas et al., "BRATS Dataset," IEEE TMI, 2018.
[67] A. Esteva et al., "Dermatologist-level classification," Nature, 2017.
[68] E. Topol, "High-performance medicine," Nat. Med., 2019.
[69] Z. Zhou et al., "Deep supervision," IEEE TMI, 2018.

[70] F. Yu and V. Koltun, "Dilated convolutions," ICLR, 2016.
[71] Z. Gu et al., "CE-Net," IEEE TMI, 2019.
[72] O. Oktay et al., "Attention mechanisms," arXiv, 2018.
[73] Y. Zhang et al., "Hybrid loss," IEEE Access, 2019.
[74] W. Bai et al., "Semi-supervised learning," MICCAI, 2017.
[75] C. Perone et al., "Unsupervised domain adaptation," arXiv, 2019.
[76] L. Chen et al., "Consistency learning," CVPR, 2020.
[77] G. Wang et al., "Ensemble CNN," IEEE TMI, 2019.
[78] K. Kamnitsas et al., "Ensemble 3D CNN," MedIA, 2017.
[79] F. Isensee et al., "nnU-Net ensemble," Nat. Methods, 2021.
[80] C. Shorten and T. Khoshgoftaar, "Data augmentation survey," JBI, 2019.
[81] L. Perez and J. Wang, "Augmentation," arXiv, 2017.
[82] J. Zech et al., "Domain generalization," PLoS Med., 2018.
[83] B. Recht et al., "ImageNet generalization," ICML, 2019.
[84] C. Kelly et al., "Key challenges for AI," NPJ Digit. Med., 2019.
[85] J. Wiens et al., "Ethical AI," Nat. Med., 2019.
[86] H. Greenspan et al., "Guest Editorial Deep Learning in Medical Imaging," IEEE TMI, 2016.
[87] D. Shen et al., "Deep Learning in Medical Image Analysis," Annu. Rev. Biomed. Eng., 2017.
[88] G. Litjens et al., "Survey on Deep Learning in Medical Imaging," MedIA, 2017.
[89] K. Suzuki, "Overview of Deep Learning in Medical Imaging," Radiol. Phys. Technol., 2017.
[90] A. Ker et al., "Deep Learning Applications," IEEE Access, 2018.
[91] M. Hesamian et al., "Deep Learning Techniques," JDI, 2019.
[92] S. Minaee et al., "Image Segmentation Survey," IEEE TPAMI, 2021.
[93] Y. LeCun et al., "Gradient-based learning," Proc. IEEE, 1998.
[94] T. Cover and P. Hart, "Nearest neighbor," IEEE TIT, 1967.
[95] V. Vapnik, "Statistical Learning Theory," Wiley, 1998.
[96] L. Breiman, "Random Forests," ML, 2001.
[97] J. Friedman, "Gradient Boosting," Ann. Stat., 2001.
[98] C. Bishop, "Pattern Recognition," Springer, 2006.
[99] I. Goodfellow et al., "GANs," NeurIPS, 2014.
[100] A. Radford et al., "GAN improvements," ICLR, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY