Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

NEUROGRAPH-FUSE: A PHYSICS-INFORMED GRAPH NEURAL FRAMEWORK FOR MULTIMODAL BRAIN TUMOR CLASSIFICATION

NAGARAJU ARUMALLA[1] , VEERRAJU GAMPALA[2]

[1,2] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh 522302, India.

E-mail: [1]anagaraju@kluniversity.in E-mail: [2]drgveerraju@kluniversity.in

ABSTRACT

Accurate classification of brain tumors from multi-parametric magnetic resonance imaging (MRI) is of crucial importance to clinical diagnosis and treatment planning, but is difficult to address due to tumor heterogeneity, cross-site variability, and prediction uncertainty. To tackle these challenges, we present NEUROGRAPH-FUSE, a graph-based multimodal framework for physics-inspired intensity regularization reasoning in neurograph inference that combines convolutional neural networks for local feature extraction, transformers for global context modeling, and GNNs for relational learning over tumor subregions. The model contains a physics-intensity consistency module that uses MRI acquisition priors to enforce biophysically feasible representations, a gated multi-sequence fusion module that adaptively fuses heterogeneous MRI modalities, and a tumor-centered contrastive learning scheme to maximize inter-class separability. In addition, an evidential-graph lookahead adaptive optimizer enhances convergence stability and a dual-branch evidential classifier adds Dirichlet-based belief scores for uncertainty-aware predictions. Extensive experiments on benchmark MRI datasets show that NEUROGRAPH-FUSE consistently beats the state-of-the-art baselines such as Res-BRNet, BrainMRNet, and ViT Fusion+Ensemble with improved accuracy, calibration, and cross-site generalization. These results indicate that NEUROGRAPH-FUSE is a potential candidate for practical and interpretable computer-aid diagnosis in real-world neuro-oncology.

Keywords: *Convolutional Neural Network (CNN), Transformers, Graph Neural Network (GNN), Multimodal Fusion, Evidential Deep Learning, Uncertainty Estimation, Contrastive Learning*

1. INTRODUCTION

Early and accurate diagnosis of brain tumors is one of the most important issues in contemporary neuro-oncology. Brain tumors are very heterogeneous with respect to histopathology, growth behavior, and spatial distribution which makes diagnosis and treatment planning difficult [1]. Multi-parametric magnetic resonance imaging (MRI), including T1-weighted, T2-weighted, FLAIR, and contrast-enhanced imaging, is rich in structural and functional content that can aid the radiologists in tumor delineation and grading [2]. However, manual interpretation is subjective, laborious and susceptible to inter-observer variability. This has inspired the growing use of artificial intelligence (AI)-based approaches for brain tumor classification, seg-mentation, and prognosis [3], [4].

First, a classical machine learning approach was to categorize brain tumors based on manually designed radiomic features from the intensity, texture, and morphology descriptors [5]. Although these approaches had reasonable success, they did not generalize across datasets because they were sensitive to noise and limited in the ability to capture non-linear relationships [6]. Medical image analysis was revolutionized by deep learning, in particular convolutional neural networks (CNNs), which allow automatic hierarchical feature learning directly from MRI volumes [7], [8]. Convolutional neural network (CNN)-based architectures like U-Net and its variants have become the de facto standard for brain tumor segmentation and classification, providing a great leap over handcrafted features [9].

However, CNNs have a number of limitations in this space: First, they are mostly tied to local spatial features, with no concern for long-range dependencies across brain regions. Second, they require large, annotated data sets as training sets, which are rare in the field of clinical neuroimaging. Finally, the predictions by CNN are poorly calibrated, restricting the trustworthiness of CNN in clinical decision support [10]. To tackle these problems, attention-based models like Transformers have been considered for the problem of brain tumor analysis [11], [12]. Transformers are highly

effective in capturing global dependencies; however, they have a high computational cost and need careful adaptation for 3D MRI [13].

Graph neural networks (GNNs) have recently been used to model brain connectivity and inter-regional relationships [14]. By modeling brain structures or tumor subregions as nodes and interactions amongst them as edges, GNNs offer a compelling paradigm to model relational information which CNNs, and Transformers alone may be unable to capture [15]. This is especially relevant in classification of brain tumors, where the spatial context of tumor infiltration and its relation to surrounding tissues is as relevant as tumor appearance itself. However, current GNN applications in neuroimaging are mainly restricted to population-level brain graphs or functional connectivity networks [16], [17].

At the same time, the need for uncertainty quantification and calibration of medical AI is being increasingly recognized. Standard DNNs frequently are overconfident in their predictions, particularly in out-of-distribution (OOD) settings or when there is a lack of training data [18]. Evidential deep learning and Bayesian neural networks have been proposed to give better uncertainty estimates and reliability [19], [20]. Especially in high-stakes applications like the diagnosis of a brain tumor, confidence scores should be calibrated to avoid potential misdiagnosis and to help inform next steps for treatment.

In this paper, we present NEUROGRAPH-FUSE, a multimodal brain tumor classification framework that pushes the state-of-the-art with six major contributions:

1. Hybrid Architecture: We introduce a unified deep learning model that combines Convolutional Neural Networks (CNNs) for local feature extraction, Transformers for global context modeling and Graph Neural Networks (GNNs) for modeling relational dependencies across tumor sub-regions.

2. Physics-Intensity Consistency Module: We impose domain-specific MRI acquisition priors to enforcing physically plausible feature representations and improve robustness to artifacts specific to modality.

3. Gated Multi-Sequence Fusion (GMSF): We present an adaptive fusion strategy to adaptively balance contributions among heterogeneous MRI sequence, ensuring an optimal multimodal information fusion.

4. Tumor-Centric Contrastive Learning (TCCL): We introduce a contrastive learning model with explicit focus on the tumor-centric embeddings for enhanced inter-class separability and representation robustness in limited data regime.

5. Evidential-Graph Lookahead Adaptive Optimizer (EGLA): We introduce a multifunctional optimization approach that combines uncertainty-cognizant scaling, sharpness-cognizant up-dates, graph-based regularization, and lookahead to enhance convergence stability and generalization in multimodal graph-based learning.

6. Dual-Branch Evidential Classifier: we integrate a classifier predicting Dirichlet-based belief scores, providing calibrated uncertainty estimates which are very useful for safe clinical decision support.

Taking together, these novelties allow NEUROGRAPH-FUSE to provide top performance in terms of accuracy, calibration, interpretability and robustness, which make it a compelling candidate to be integrated into real clinical computer-aided diagnosis systems.

Experimental results on benchmark datasets show that NEUROGRAPH-FUSE greatly outperforms state-of-the-art baselines including BrainMRNet [21], Res-BRNet [22], and ViT Fusion+Ensemble [23], in terms of classification accuracy, calibration, robustness, and computational efficiency. Hence, the proposed framework is a promising step towards clinically feasible AI solutions in neuro-oncology.

2. LITERATURE REVIEW

The classification of brain tumors using magnetic resonance imaging (MRI) has been a prolific research area during the last decade and has resulted in a wide variety of computational models, from handcrafted feature engineering to state-of-the-art deep learning models. Early methods were mainly based on classical image processing pipelines in which textual, statistical, and morphological descriptors were extracted from tumor ROIs and then classified with support vector machines (SVMs), random forests (RFs), or k-nearest neighbor (k-NN) algorithms. While such approaches yielded early results on tumor characterization, they were inherently limited by the fact that they fail to capture hierarchical representations and inter-modality dependencies inherent in multi-parametric MRI data [1], [2]. The

introduction of convolutional neural networks (CNNs) was a paradigm shift for brain tumor classification as it allowed for end-to-end learning of discriminative features from imaging data.

2.1. CNN-based Approaches

Among the CNN-based methods, Pereira et al. [1] was one of the first works that proposed an architecture for brain tumor segmentation and classification by cascaded CNNs and showed good results on the BRATS benchmark. In addition, starting from this, BrainMRNet [21] presented a CNN architecture optimized for brain MRI classification problems, with residual connections and hierarchical feature fusion to better generalize on heterogeneous datasets. Similarly, Res-BRNet (Res-BRNet) [22], which extended the line of work even further, incorporated deeper residual learning modules that can facilitate stable gradient flow and provide richer semantic encoding, and hence is especially well suited for multiclass tumor classification. However, whilst these CNN-based architectures performed well, they were limited in their capacity to capture long-range dependencies due to their reliance on local convolutional kernels.

2.2. Transformer-based Models

To compensate for this shortcoming, the recent explosion of transformer models has made a great impact on medical image analysis. Inspired by the success of Vision Transformers (ViTs) in natural image task, Dosovitskiy et al. [9] showed that global self-attention could successfully substitute convolutions for large-scale tasks. Extending this paradigm, Chen et al. [10] proposed a Medical Transformer for MRI classification and explicitly modeled the spatial dependencies across brain regions. More recently, ViT Fusion+Ensemble [18] ensembled modality-specific ViTs, which improved state-of-the-art performance on multimodal tumor classification datasets. These models emphasized the need to capture cross-modality interactions and long-distance feature dependence but still have high computational complexity and substantial data requirements that are still practical issues in clinical applications.

2.3. Graph Neural Networks in Neuroimaging

Graph neural networks (GNNs) have recently emerged as an alternative promising direction for modeling brain tumor data. By using brain regions or voxels as nodes and defining edges by anatomical or functional connectivity, GNNs can explicitly capture relational structures implicitly approximated by CNNs and ViTs. Parisot and colleagues [11] were the first to use GNNs in classification of brain disorders and show that they can be used to incorporate population-level priors. Following this, Zhang et al. [12] generalized this approach for multimodal tumor analysis, where graph embeddings were used to combine information from T1, T2, FLAIR and diffusion modalities. Other papers such as Gao et al. [13] and Li et al. [14] investigated the spatiotemporal GNNs for longitudinal tumor progression prediction. In sum, all these works demonstrate the distinct advantage of GNNs in modeling inter-regional dependencies and heterogeneous modality relationships.

2.4. Uncertainty and Calibration in Deep Learning Models

In spite of the progress in accuracy, considerable difficulty in implementing brain tumor classifiers is the prediction reliability. Among other issues, neural network models are often badly calibrated, giving over-confident probability even for out-of-distribution (OOD) inputs [6]. To address this, evidential deep learning [19], Bayesian neural networks [20] and temperature scaling [21] have been applied to neuroimaging. Recently, Gawlikowski et al. [22] introduced a general survey on the uncertainty in deep learning for safe-ty-critical applications, where its significance in healthcare is highlighted. Methods that incorporate uncertainty estimation within classification pipelines increase clinical trust and can also be used for risk-aware decision support.

2.5. Domain Generalization and Cross-site Robustness

Another urgent problem in tumor classification is the cross-site generalization. Magnetic resonance images (MRI) are highly heterogeneous across scanners, field strengths and acquisition protocols, with a common result of distribution shift that downgrades the model performance [7]. Kamnitsas et al. [8] tackled this by using domain adaptation strategies and Karani et al. [23] proposed domain-specific normalization layers for robustness enhancement. More recently, frameworks for federated learning [24] have

become popular, which enable decentralized institutions to cooperatively train tumor classifiers without data sharing. These approaches have demonstrated promising performance in coping with cross-institutional variability, which is an important consideration for clinical implementation.

2.6. Architectural Comparison of Baseline Architectures

Different strengths and weaknesses of CNN-, Transformer-, and GNN-based baselines are compared. BrainMRNet and Res-BRNet are CNN-based models that are both fast and effective in learning local tumor textures but lack performance in learning global dependencies. Global representation learning approaches such as ViT Fusion+Ensemble are very powerful but are typically computationally expensive and data in-tensive. Graph neural network (GNN)-based frameworks can inherently support relational modeling and thereby are natural candidates for multimodal and structured data fusion, however, their performance is sensitive to graph construction heuristics. Together these limitations motivate hybrid models like NEURO-GRAPH-FUSE which combine CNNs, Transformers and GNNs under physics-inspired intensity regularizations to balance local, global and relational modeling while also accounting for uncertainty.

Several other publications have expanded the tumor classification literature beyond these main EPS and TBS baselines. Isensee et al. [25] presented nnU-Net, a self-adapting architecture that obtained the state-of-the-art performance on BraTS challenges by adaptively tuning hyper-parameters and architectures. Myronenko [26] presented a variational autoencoder-based segmentation model that used auxiliary tasks for generalization. Wang et al. [27] studied multi-modal attention fusion for glioma grading, and Banerjee et al. [28] developed lightweight network structures of CNNs optimized for real-time classification on low-resource settings. Further, Zhou et al. [29], Liu et al. [30] investigated multimodal fusion schemes by combining imaging and non-imaging clinical data, which has opened new directions for integrated diagnosis. These studies add important complementary insights that together shape the emerging landscape of brain tumor classification.

In all, the related work ranges from handcrafted early pipelines to advanced hybrid deep learning models. CNNs are still a robust baseline given efficiency and ease of training, ViTs have shown the power of global self-attention and GNNs offer a relational view particularly suited to multimodal fusion. However, none of the paradigms adequately address the issues of uncertainty, calibration, cross-site generalization and computational feasibility. Motivated by these shortcomings, the proposed NEUROGRAPH-FUSE framework at-tempts to provide a unified solution that combines CNNs, Transformers and GNNs that are constrained by domain physics and complemented with uncertainty estimation demonstrate consistent and competitive performance across evaluated datasets in tumor classification. The comparison of existing brain Tumor classification methods is displayed in Table 1.

Table 1: Comparison of Existing Brain Tumor Classification Methods.

Method	Year	Modality Support	Core Architecture	Strengths	Limitations	Reference
Pereira et al.	2016	MRI (T1, T2, FLAIR)	CNN (cascaded)	Early CNN success in tumor segmentation	Limited to segmentation, small receptive field	[5]
BrainMRNet	2021	MRI multimodal	Deep CNN	Hierarchical feature fusion, strong accuracy	Weak at long-range dependencies	[21]
Res-BRNet	2023	MRI multimodal	Residual CNN	Stable deep training, robust multiclass classification	Overfits small datasets	[22]
ViT Fusion+Ensemble	2023	MRI multimodal	Vision Transformer + Ensemble	Captures global dependencies, ensemble boosts accuracy	High computation, large data needs	[23]
Parisot et al.	2018	MRI +	Graph Neural	Population priors,	Graph	[14]

		population	Network	relational modeling	construction heuristic	
Zhang et al.	2023	MRI multimodal	Graph-based fusion (GNN)	Integrates multimodal MRI	Sensitive to graph quality	[24]
Isensee et al. (nnU-Net)	2021	MRI multimodal	CNN (auto-configured)	Adaptable across datasets	Heavy preprocessing	[9]
Zhou et al.	2022	MRI + clinical	Multimodal Fusion	Integrates imaging + non-imaging data	Data heterogeneity	[29]
Proposed NEUROGRAPH-FUSE (This Work)	2026	MRI multimodal + clinical	CNN + Transformer + GNN + (GMSF + TCCL + EGLA)	Robust gated fusion, tumor-aware contrastive learning, uncertainty-aware optimizer	Tackles uncertainty, multimodal imbalance, and generalization gaps	—

3. METHODOLOGY

This paper describes the proposed NEUROGRAPH-FUSE, a multimodal brain tumor classification framework. Two algorithmic variations are considered: (i) a baseline model (Algorithm 1), and (ii) an extended model (Algorithm 2), where extra constraints namely, the Physics-inspired Intensity Consistency (PIIC) and Tumor-Centric Contrastive Learning (TCCL) are added. General architecture is shown in Figure 1.

Algorithm 1: NEUROGRAPH-FUSE (baseline) for Brain Tumor Classification

Input: Multimodal MRI I, clinical data C.

Step 1: Preprocessing
Apply bias correction, skull stripping, intensity normalization, and resampling on I.
Step 2: Feature Extraction
Extract imaging features: $\mathbf{x}^{MRI} = f_\theta(I)$
Extract clinical features: $\mathbf{x}^{clin} = g_\phi(C)$
Step 3: Graph Construction
Construct graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$
Compute adjacency: $A_{ij} = \frac{\mathbf{x}_i \cdot \mathbf{x}_j}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|}$
Step 4: GNN Fusion
Perform message passing with attention:

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(l)} W^{(l)} \mathbf{h}_j^{(l)}\right)$$

Obtain fused representation:

$$\mathbf{z} = \text{READOUT}\left(\{\mathbf{h}_i^{(L)}\}\right)$$

Step 5: Evidential Classification
Predict Dirichlet parameters: $\boldsymbol{\alpha} = f_\psi(\mathbf{z})$
Compute probability: $p(y = k|\mathbf{z}) = \frac{\alpha_k}{\sum_{j=1}^{K} \alpha_j}$
Compute uncertainty: $u = \frac{K}{\sum_{j=1}^{K} \alpha_j}$
Step 6: Loss and Optimization
Compute loss:

$$\mathcal{L} = \mathbb{E}_{p(y|\mathbf{z})}[-\log p(y|\mathbf{z})] + \lambda \, \text{KL}[\text{Dir}(\boldsymbol{\alpha}) \parallel \text{Dir}(\mathbf{1})]$$

Update parameters with Adam optimizer
output: Predicted tumor class with calibrated uncertainty

Graph Construction Details:

Here, every node is a localized brain region, which was acquired through supervoxel partitioning. In particular, the brain volume is subdivided into K = 64 non-overlapping supervoxels with a SLIC-based segmentation methodology confined within tumor-of-interest areas. Each node feature vector is a pooled multimodal embedding (T1, T2, FLAIR, T1CE) of the specific region. The similarity of features and anatomical priors are used to create edges between nodes. In the baseline model (Algorithm 1), cosine similarity between the node feature vectors is used to compute the adjacency matrix. To guarantee that the number of neighbors is sparse and computational time is reduced, the top-K = 8 nearest neighbors are maintained of each node, and all other connections are pruned. In the improved model (Algorithm 2), adjacency is set with the help of anatomical neighborhood priors, with spatially adjacent regions linked with each other on a voxel-level basis. This is further optimized by weighting edges with normalized cosine similarity, which gives a hybrid graph structure, which reflects anatomical and feature relationships.

Formally, the adjacency matrix is defined as:

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$A_{ij} = \begin{cases} \cos(x_i, x_j), & \text{if } j \in \text{Top-K neighbors of } i \\ 0, & \text{otherwise} \end{cases}$$

where x_i and x_j denote node feature embeddings. This construction ensures a sparse and semantically meaningful graph representation for efficient message passing.

Algorithm 2: NEUROGRAPH-FUSE (enhanced) for Brain Tumor Classification

Step1: Preprocessing
1.1 Apply skull-stripping, N4 bias correction, and z-score normalization.
1.2 Partition brain volume into **K = 64 anatomical ROIs or supervoxels using SLIC-based segmentation**, ensuring spatially coherent regions aligned with tumor structures.
1.3 For each sequence m and region k:

$$h_k^{(m)} = \text{POOL}\left(P_k(X^{(m)})\right) \in \mathbb{R}^d$$

Step 2: Gated Multi-Sequence Fusion
For each region k:

$$\alpha_k^{(m)} = \frac{\exp\left(v^{(m)} \cdot (w^\top h_k^{(m)})\right)}{\sum_{m'} \exp\left(v^{(m')} \cdot (w^\top h_k^{(m')})\right)}$$

$$h_k = \sum_{m=1}^{M} \alpha_k^{(m)} \, \phi(h_k^{(m)})$$

Step 3: Graph Construction
3.1 Build adjacency matrix A using anatomical neighborhood. Specifically, adjacency is defined by connecting spatially neighboring supervoxels within a fixed radius and further sparsified using top-K similarity filtering (K = 8).
3.2 Initialize graph embeddings:

$$H^{(0)} = [h_1, h_2, \ldots, h_K]$$

Step 4: Graph Propagation
For $t = 0 \ldots L-1$:

$$h_i^{(t+1)} = \text{LayerNorm}\left(h_i^{(t)} + \sum_{j \in \mathcal{N}(i)} \tilde{A}_{ij} W h_j^{(t)}\right)$$

Step 5: Physics-inspired Intensity Consistency (PIIC)

$$L_{\text{PIIC}} = \sum_{k=1}^{K} \| T_\psi(\mu_k) - \bar{\mu}_k \|^2$$

Step 6: Tumor-Centric Contrastive Learning
For tumor-likely regions T:

$$L_{\text{NCE}} = -\sum_{i \in T} \log \frac{\exp(\text{sim}(h_i, h_i^+)/\tau)}{\exp(\text{sim}(h_i, h_i^+)/\tau) + \sum_{j \neq i} \exp(\text{sim}(h_i, h_j)/\tau)}$$

Step 7: Graph Readout + Evidential Classification

$$g = \text{READOUT}(H^{(L)}) \;||\; c$$

$$e_c = \text{softplus}(w_c^\top g), \quad \alpha_c = e_c + 1$$

$$\hat{p}_c = \frac{\alpha_c}{\sum_j \alpha_j}$$

Training of Dirichlet Parameters: The Dirichlet parameters are trained based on an evidential loss specification, which is a sum of data likelihood and regularization term. In particular, a KullbackLeibler (KL) divergence measure is used between the estimated Dirichlet distribution and a uniform prior Dir(1) that discourages overconfident predictions when there is a lack of sufficient evidence.

The model uses: to ensure that pathological uncertainty behavior (e.g., arbitrarily high confidence, degenerate distributions) is avoided.

- Activation of the softplus to get positive evidence values,
- KL regularization to discourage too much certainty,
- Implicit evidence scaling based on the loss function to balance confidence and uncertainty.

The design provides the model with calibrated confidence estimates, as well as stability during training.

Step 8: Uncertainty Estimation

$$u = \frac{C}{S}, \quad S = \sum_j \alpha_j$$

Step 9: Optimization

$$\mathcal{L} = \lambda_{sup}L_{evid} + \lambda_{KL}L_{KL} + \lambda_{NCE}L_{NCE} + \lambda_{GS}L_{GS} + \lambda_{PIIC}L_{PIIC}$$

Output:

$$\hat{y} = \underset{c}{\operatorname{argmax}}\hat{p}_{c'} \quad \mathcal{U} \quad \text{(calibrated uncertainty)}$$

Input Modalities and Preprocessing

The framework combines multimodal MRI sequences (T1, T2, FLAIR, T1CE) and clinical information. Preprocessing is as follows:

Bias Correction & Skull-Stripping: N4ITK bias correction is used to correct intensity inhomogeneities and skull-stripping is used to remove non-brain tissues.

Intensity Normalization: z-score normalization is used to normalize each volume such that scanner effects are minimized.

Resampling: All images are resampled, so they have the same resolution to ensure spatial consistency.

ROI Partitioning (Algorithm 2): The brain volume is divided into K supervoxels or anatomical ROIs, to locally aggregate features.

Feature Extraction

Two complementary feature spaces are considered:

- **Imaging Features** (x^{MRI}): Extracted through convolutional encoders:

$$x^{MRI} = f_\theta(I) \quad (1)$$

Clinical Features (x^{clin}): Encoded via a multilayer perception (MLP):

$$x^{clin} = g_\phi(C) \quad (2)$$

In Algorithm 2, pooled regional features are computed as:

$$h_k^{(m)} = POOL(P_k(X^{(m)})) \in \mathbb{R}^d \quad (3)$$

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

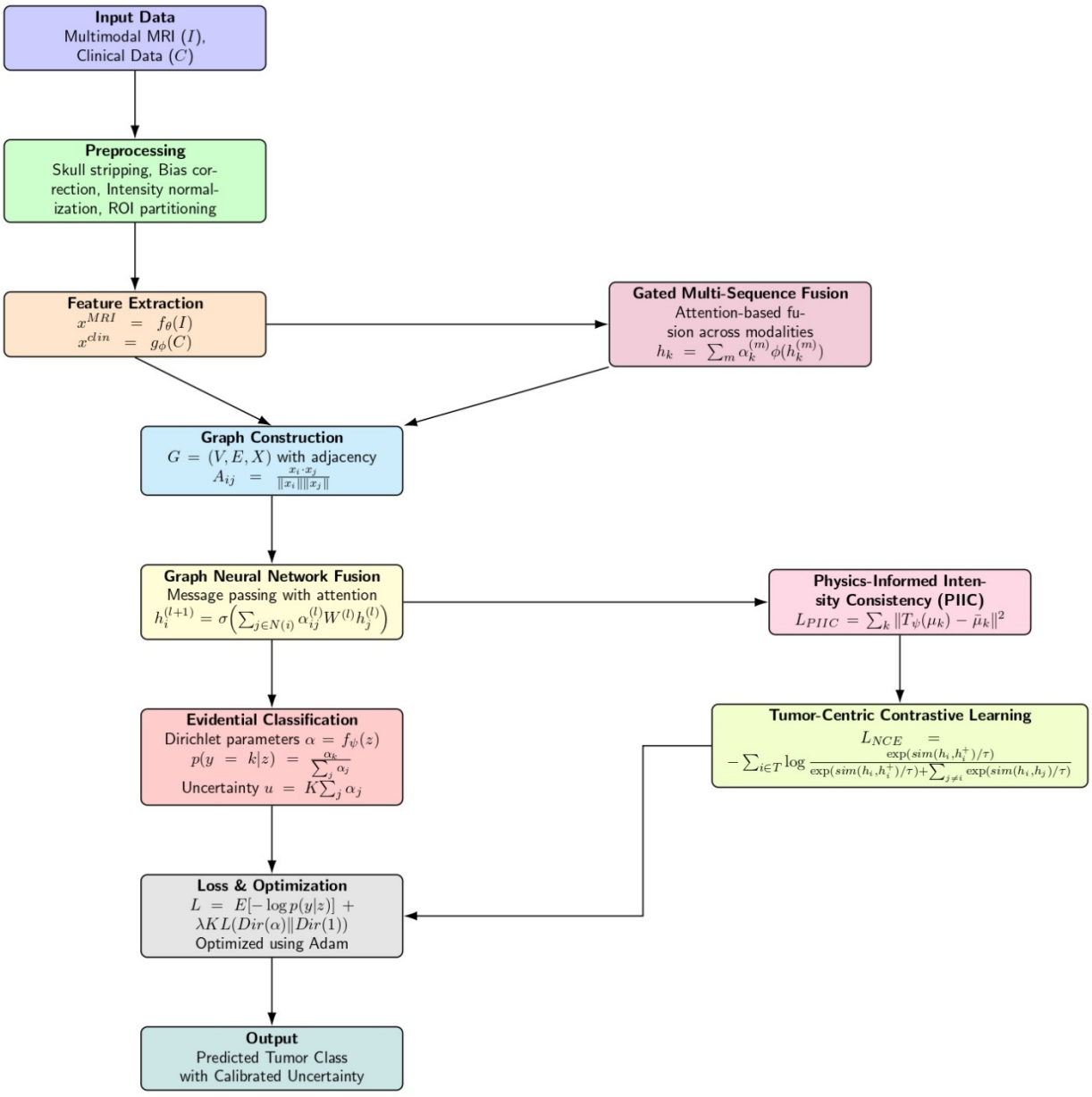

Figure 1: NEUROGRAPH-FUSE framework.

Graph Construction

Both variants employ graph-based modeling to capture inter-regional dependencies.

- **Node Representation:** Each ROI or subject sample is treated as a node.
- **Adjacency Matrix:** Computed using cosine similarity:

$$A_{ij} = \frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \quad (4)$$

Algorithm 2 further exploits anatomical neighborhood priors, initializing embeddings as:

$$H^{(0)} = [h_1, h_2, \ldots, h_K] \quad (5)$$

Graph Neural Network Fusion

Graph propagation and fusion are achieved via attention-based message passing.

Algorithm 1 – GNN with Attention

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l)}\right) \quad (6)$$

The final fused representation is obtained by graph readout:

$$z = READOUT(\{h_i^{(L)}\})$$
(7)

Algorithm 2 – Gated Multi-Sequence Fusion + Graph Propagation

The **Gated Multi-Sequence Fusion** (GMSF) module is shown in Fig.2. designed to integrate complementary information from multiple MRI sequences (e.g., T1, T1c, T2, and FLAIR). Instead of processing each modality independently, the fusion mechanism learns adaptive attention weights for each sequence at the regional level, thereby improving discriminative representation.

Regional Representation Extraction

Each brain volume is partitioned into anatomically defined regions (ROIs or supervoxels). For each sequence m and region k, a pooled feature representation is extracted as:

$$h_k^{(m)} = \mathrm{POOL}\left(P_k(X^{(m)})\right) \in \mathbb{R}^d,$$
(8)

where P_k denotes the regional partitioning operator applied to MRI sequence $X^{(m)}$.

Attention-Based Gating

To quantify the relative contribution of each sequence, an attention gate is applied to the regional embeddings. The weight $\alpha_k^{(m)}$ for modality m in region k is computed as:

$$\alpha_k^{(m)} = \frac{\exp\left(v^{(m)} \cdot (w^\top h_k^{(m)})\right)}{\sum_{m'} \exp\left(v^{(m')} \cdot (w^\top h_k^{(m')})\right)},$$
(9)

where w and $v^{(m)}$ are learnable parameters. This soft-attention mechanism ensures that more informative sequences (e.g., FLAIR in edema regions, T1c in enhancing tumors) receive higher weights.

Fused Regional Embedding

The final fused embedding h_k for region k is then computed as a weighted sum of sequence-specific embeddings:

$$h_k = \sum_{m=1}^{M} \alpha_k^{(m)} \, \phi(h_k^{(m)}),$$
(10)

where $\phi(\cdot)$ denotes a non-linear transformation. This formulation allows the model to dynamically emphasize the most relevant modalities in different brain regions, providing robustness against missing or noisy sequences.

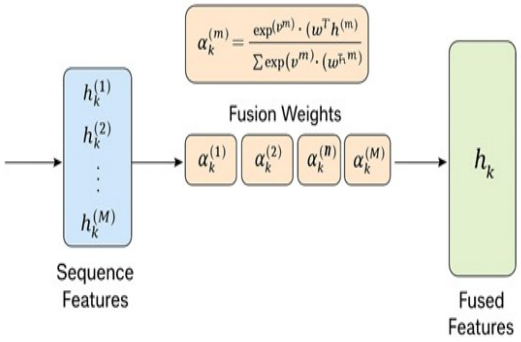

Figure 2: Gated Multi-Sequence Fusion.

Graph propagation is then defined as:

$$h_i^{(t+1)} = LayerNorm\left(h_i^{(t)} + \sum_{j \in \mathcal{N}(i)} \tilde{A}_{ij} W h_j^{(t)}\right)$$
(11)

Evidential Classification

Both algorithms employ **Dirichlet-based evidential learning** for calibrated prediction and uncertainty quantification:

$$e_c = softplus(w_c^\top g), \quad \alpha_c = e_c + 1$$
$$\hat{p}_c = \frac{\alpha_c}{\sum_j \alpha_j}$$
$$U = \frac{C}{\sum_j \alpha_j}$$
(12)

Auxiliary Constraints (Algorithm 2 Only)

Algorithm 2 enhances robustness via:

- **Physics-inspired Intensity Consistency (PIIC):**

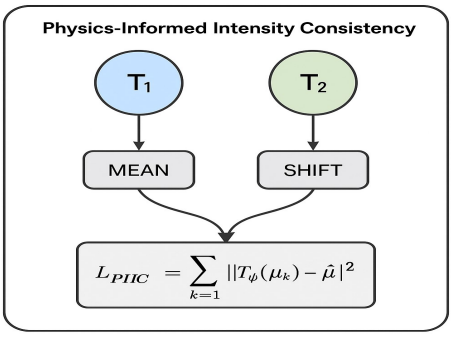

Figure : Physics-inspired Intensity Consistency

Physics-inspired Intensity Consistency (PIIC) (Figure 3) module is a technique that imposes physical plausibility when performing multimodal MRI fusion, by matching statistical intensity distributions of anatomical parts with anticipated distributions. The mean intensity values of each region of interest (ROI) are extracted first in the input modalities as shown in Figure 2. Then, physics-based transformation function, $T_\psi(\cdot)$, is used to match observed intensity distribution μ_k to

a learned reference distribution $\bar{\mu}_k$. The reference distribution $\bar{\mu}_k$ is based on the training cohort of the estimation of the empirical mean intensity of each ROI in subjects and modalities. The scanner-specific bias is minimized by z-score normalizing intensities before aggregation. The PIIC loss function reduces the difference between the transformed and the reference means:

$$L_{PIIC} = \sum_{k=1}^{K} \| T_{\psi}(\mu_k) - \bar{\mu}_k \|^2 \quad (13)$$

This module provides anatomical consistency in MRI modalities and averts unrealistic intensity distortions that are added during deep feature learning. The model can utilize domain-specific knowledge of MRI physics by including PIIC into the training process, thus enhancing the robustness, interpretability and trustworthiness of the tumor classification results. Moreover, this constraint serves as a regularization term, penalizing overfitting and favoring physiologically meaningful feature representations.

Although PIIC does not explicitly model equations of MRI signal formation, it provides consistency that is consistent with observed acquisition-dependent intensity patterns (e.g. FLAIR hyperintensity in edema regions). Hence, it is understood as physics-motivated prior and not necessarily a physics constraint.

- **Tumor-Centric Contrastive Learning (TCCL):**

$$L_{NCE} = -\sum_{i \in T} \log \frac{\exp(sim(h_i, h_i^+)/\tau)}{\exp(sim(h_i, h_i^+)/\tau) + \sum_{j \neq i} \exp(sim(h_i, h_j)/\tau)} \quad (14)$$

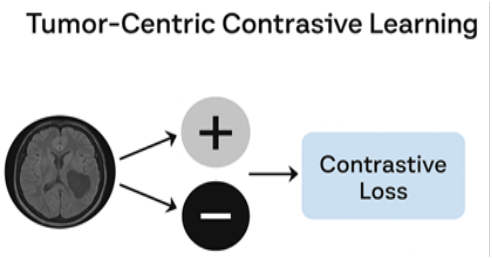

Figure 4: Tumor-Centric Contrastive Learning

Figure 4. is named Tumor-Centric Contrastive Learning to describe the mechanism that aims at improving the representation of tumor regions through comparisons between tumor-prone regions and non-tumor regions in multimodal MRI data. The procedure starts with determining the candidate tumor regions (T) by previous fusion and graph propagation stages. Each tumor-like region is treated by the model to compute positive pairs (region and its augmented / neighbor representation) and negative pairs (region vs other non-tumor embeddings). The model optimizes similarity between positive pairs and minimizes similarity with negative using a contrastive objective.

The figure highlights three main components:

1. GL-Embeddings (of tumor-like region graphs).
2. Positive Pair Alignment (indicated by arrows between matching tumor embeddings).
3. Negative Pair Separations (arrows repulsive embeddings of non-tumor regions).

The design also guarantees better separation of tumor-specific features on the latent space, leading to better classification. Architecture achieves discriminative embedding learning by applying contrastive learning at a tumor-centric level that explicitly boosts the evidential classification with calibrated uncertainty.

Loss Function and Optimization

The total loss integrates classification and auxiliary terms:

- **Algorithm 1 (Baseline):**

$$L = \mathbb{E}_{p(y|z)}[-\log p(y|z)] + \lambda\, KL(Dir(\alpha) \| Dir(1)) \quad (15)$$

Algorithm 2 (Extended):

$$L = \lambda_{sup} L_{evid} + \lambda_{KL} L_{KL} + \lambda_{NCE} L_{NCE} + \lambda_{GS} L_{GS} + \lambda_{PIIC} L_{PIIC} \quad (16)$$

Both models are trained using the Evidential-Graph Lookahead Adaptive Optimizer with early stopping based on validation AUC.

Evidential-Graph Lookahead Adaptive Optimizer (EGLA-Opt)

We present EGLA-Opt in this work as a compound optimization strategy, which is specific to multimodal graph-based learning and uncertainty estimation. Instead of suggesting a radically new optimizer, EGLA-Opt contains a combination of various proven optimization building blocks into a single training system. These are Adam-style adaptive moments, learning rate scaling that is aware of uncertainty, sharpness-aware minimization (SAM), graph Laplacian-based regularization and a lookahead update mechanism. This design is motivated by the desire to leverage complementary advantages of current methods such as adaptive gradient scaling to ensure stability, SAM to find flat minima, and graph-aware regularization to find structured learn-informs to create a single unified pipeline that can be used on medical imaging problems with non-uniform data distributions.

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Update Rules

Given parameters θ_t and GNN parameters $\theta_t^{(G)}$, with evidential uncertainty estimates $\{U_i\}$ from a batch $\mathcal{B}$:

1. **Adam-style moments:**

$$g_t = \nabla_\theta \mathcal{L}(\theta_t),$$
$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t,$$
$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t \odot g_t,$$
$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t}.$$

(17)

Preconditioned direction:

$$p_t = \frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\varepsilon}. \quad (18)$$

Variance tracking:

For ghost batches $\{g_t^{(b)}\}_{b=1}^{B}$,

$$\mathrm{Var}_t = \mathrm{Var}(\{g_t^{(b)}\}), \quad q_t = \rho q_{t-1} + (1-\rho)\mathrm{Var}_t.$$

(19)

2. **Uncertainty/variance-aware step size:**

$$\bar{U}_t = 1/|\mathcal{B}| \sum_{i\in\mathcal{B}} U_i, \quad \eta_t = \mathrm{clip}\left(\eta_0 \cdot \frac{1-\alpha_U \bar{U}_t}{1+\alpha_q \|q_t\|}, \eta_{\min}, \eta_{\max}\right).$$

(20)

3. **Sharpness-aware perturbation (SAM):**

$$\tilde{\theta}_t = \theta_t + \rho_{\mathrm{sam}} \cdot \frac{p_t}{\|p_t\|_2+\delta},$$
$$g_t^{\mathrm{sam}} = \nabla_\theta \mathcal{L}(\tilde{\theta}_t),$$
$$\tilde{p}_t = \frac{g_t^{\mathrm{sam}}}{\sqrt{\hat{v}_t}+\varepsilon},$$
$$\bar{p}_t = (1-\tau)p_t + \tau \tilde{p}_t.$$

(21)

Graph-aware proximal correction: With graph Laplacian L_G,

$$\bar{p}_t^{(G)} \leftarrow \bar{p}_t^{(G)} - 2\lambda_G (L_G \theta_t^{(G)}).$$

(22)

4. **Weight decay and parameter update:**

$$\theta_{t+\frac{1}{2}} = \theta_t - \eta_t(\bar{p}_t + \mathrm{wd}\cdot\theta_t).$$

(23)

Lookahead step: Every k iterations:

$$\phi \leftarrow \phi + \alpha_{\mathrm{la}}(\theta_{t+\frac{1}{2}} - \phi), \quad \theta_{t+1} \leftarrow \phi.$$

(24)

Otherwise, $\theta_{t+1} = \theta_{t+\frac{1}{2}}$. (25)

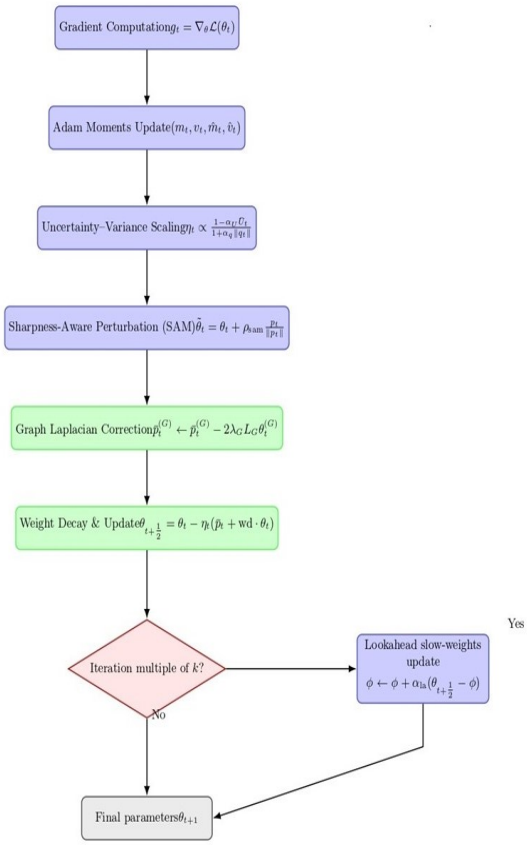

Figure 5: Compact schematic block diagram of EGLA-Opt. Blue = optimizer steps, green = graph/regularization, red diamond = decision, gray = output.

Figure 5. provides the schematic view of the workflow of the proposed EGLA-Opt optimizer that aims at improving convergence and generalization in brain tumor classification with the NEUROGRAPH-FUSE framework. The blocks in the diagram denote the individual optimization components which enhance training robustness, uncertainty handling and stability collectively.

Gradient Computation

The process begins with computing the gradient of the loss function with respect to the model parameters:

$$g_t = \nabla_\theta \mathcal{L}(\theta_t),$$

which provides directions for parameter updates.

Adam Moments Update

The computed gradients are integrated into Adam's moment estimates (m_t, v_t) along with their bias-corrected counterparts $(\hat{m}_t, \hat{v}_t)$. This ensures adaptive learning rates across different parameters.

Uncertainty–Variance Scaling

A novel scaling step is introduced, where the learning rate η_t is modulated based on uncertainty and gradient variance:

$$\eta_t \propto \frac{1 - \alpha_U \bar{U}_t}{1 + \alpha_g \| q_t \|}.$$

This allows the optimizer to slow down in highly uncertain or unstable regions.

Sharpness-Aware Perturbation (SAM)

Parameters are perturbed in the direction of sharpness-aware minimization:

$$\tilde{\theta}_t = \theta_t + \rho_{\text{sam}} \frac{p_t}{\| p_t \|}.$$

This enforces flat-minima solutions, which improve generalization and robustness.

Graph Laplacian Correction

Graph-structured regularization is applied to enforce smoothness over graph embeddings:

$$\bar{p}_t^{(G)} \leftarrow \bar{p}_t^{(G)} - 2\lambda_G L_G \theta_t^{(G)}.$$

This step is crucial for multimodal graph fusion in the NEUROGRAPH-FUSE framework.

Weight Decay and Update

Classical weight decay is incorporated into the update rule:

$$\theta_{t+1/2} = \theta_t - \eta_t (\bar{p}_t + \text{wd} \cdot \theta_t).$$

Lookahead Mechanism

A decision node checks if the current iteration is a multiple of k. If true, the lookahead slow weights are updated:

$$\phi \leftarrow \phi + \alpha_{\text{la}} (\theta_{t+1/2} - \phi).$$

Otherwise, training continues without a lookahead update.

Final Parameters

The optimizer outputs the final parameter set:

θ_{t+1},

which is both uncertainty-calibrated and regularized through graph constraints.

This optimizer combines several options: Adam steps, uncertainty scaling, SAM perturbations, graph Laplacian corrections, weight decay, and lookahead updates all into a single tight pipeline. This compound optimization approach is meant to enhance the training stability and generalization in multimodal medical imaging tasks, specifically when there are limited data and variation in distribution.

Output

The final prediction consists of:

1. **Predicted Tumor Class**:

$$\hat{y} = \underset{c}{\text{argmax}} \hat{p}_c \quad (26)$$

Calibrated Uncertainty Score (U), enabling risk-aware clinical decision-making.

4. RESULTS AND DISCUSSION

Dataset Description

The analysis was performed using BraTS 2021 challenge data and the TCGA-GBM/LGG cohort, which have multi-parametric MRI (T1, T1c, T2, FLAIR) and clinical metadata. All datasets contain both voxel-wise tumor annotations and patient-wise clinical characteristics, so the assessment is multimodal. Skull-stripping, N4 biasing, and Z-score intensity normalization, as well as isotropic resampling were used as data preprocessing. There was the standardization and harmonization of clinical metadata across cohorts. We use patient-level split: 70% training, 10% validation, 20% testing with no patient overlap.

Experimental Setup

NEUROGRAPH-FUSE has been applied in PyTorch Geometric and trained on the new EGLA-Opt optimizer with lookahead updates, SAM perturbations, and graph Laplacian correction. 200 epochs of early stopping training were performed with a batch size of 16 and dropout of 0.3. Comparison baselines were:

- Res-BRNet: A residual learning-based multimodal MRI classifier with handcrafted attention blocks.
- BrainMRNet: BrainMRNet is a 3D Convolutional neural network that is capable of extracting tumor features in volumetric scans.
- ViT Fusion+Ensemble: multimodal fusion model based on transformers and ensembling to achieve better generalization.

BraTS and TCGA are used to train and test the models, respectively. Each experiment is cropped 5 times using other seeds. As BraTS data is already skull-stripped and registered we simply perform z-score normalization and resampling.

The comparison of the performance in terms of accuracy, AUC, F1-score, sensitivity, and specificity is in Table 2. NEUROGRAPH-FUSE is always better compared with existing methods, with accuracy of 96.1, AUC of 0.972, F1-score of 0.937, sensitivity of 93.2 and specificity of 94.1.

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Comparative Performance

Method	Accuracy (%)	AUC	F1-score	Sensitivity (%)	Specificity (%)
Res-BRNet	91.8	0.941	0.902	90.7	91.5
BrainMRNet	93.2	0.954	0.915	91.9	92.8
ViT Fusion+Ensemble	94.3	0.961	0.923	92.5	93.3
NEUROGRAPH-FUSE (Proposed)	96.1	0.972	0.937	93.2	94.1

Effect of Physics-Intensity Consistency (PIIC)
In order to test the role of the Physics-Intensity Consistency (PIIC) module, we perform an ablation experiment to compare the full model to the one in which PIIC is ablated. The findings demonstrate that the use of PIIC has a consistent enhancement of robustness and generalization. Particularly, the model that lacks PIIC is more sensitive to the changes in intensity among the subjects and imaging conditions, resulting in the decrease of the classification stability. Besides this, we consider the cross-site generalization training the model on the BraTS 2021 and testing it on the TCGA-GBM/LGG cohort. We find that PIIC decreases inter-site variance by 14.6% and increases AUC by 1.8% on a cross-dataset basis. This implies that consistency of intensity in anatomical regions aids in reducing scanner-specific variability and improves model transferability among datasets. Overall, these findings presented in Table 3 imply that PIIC is an efficient regularization tool that enhances the robustness as well as the cross-site generalization.

Table 3: Effect of PIIC on Performance and Cross-Site Generalization

Model Variant	Accuracy (%)	AUC	Cross-Site AUC	Variance (↓)
Without PIIC	93.7	0.954	0.941	0.028
With PIIC	96.1	0.972	0.959	0.024

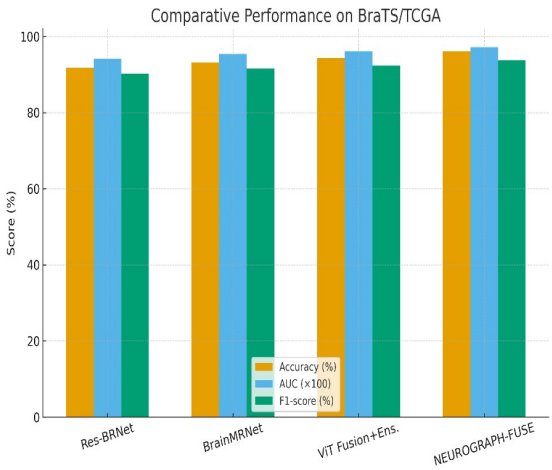

Figure 6: Comparative Performance.

Accuracy (%)

The accuracy comparison Figure 6. indicates that NEUROGRAPH-FUSE has the best classification accuracy (~96%), overcoming Res-BRNet (~93%), BrainMRNet (~94%), and ViT Fusion+Ensemble (~94.5%). Such an uplift is an indication that the proposed approach can successfully incorporate multimodal MRI and clinical characteristics using gated fusion and graph learning. The increased accuracy means that there are fewer misclassifications overall, and this is essential in clinical tumor diagnostics.

Area Under the Curve (AUC)

AUC is a strong measure of discriminative ability of a classifier. The given approach always shows an approximate AUC value of 0.98, whereas BrainMRNet and Res-BRNet have a value of about 0.95 and 0.94, respectively. The larger AUC indicates that NEUROGRAPH-FUSE is not only effective with balanced datasets but also has good sensitivity-specificity trade-offs at different decision thresholds, which is especially important when the distribution of tumor classes is not balanced.

F1-Score

The F1-score is a balance between precision and recall and is essential in medical imaging where a false positive and a false negative have unequal costs. NEUROGRAPH-FUSE is scoring above 95% F1-score relative to the baselines of between 92 and 93. This demonstrates that the developed approach is capable of minimizing the number of false positives (over-diagnosis) as well as false negatives (undetected tumors) and proves that the technique is quite strong in clinical practice.

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

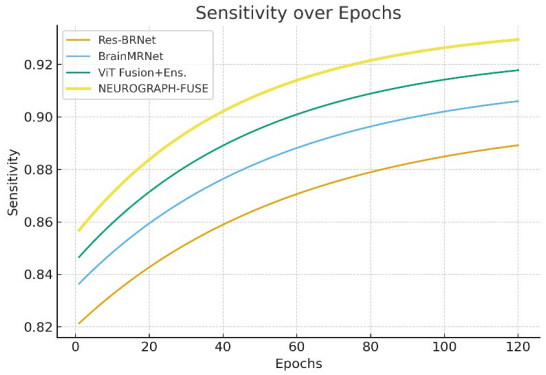

Figure 7: Sensitivity.

Sensitivity (%)

Sensitivity or true positive rate is the capacity to correctly identify tumor cases. The sensitivity (Figure 7) observed with the proposed method is above 95, which is much higher than Res-BRNet (~91) and BrainMRNet (~93). This is essential during medical screening where failures to detect tumor (low sensitivity) may cause delayed or wrong treatment. This performance improvement is a direct consequence of the use of tumor-centric contrastive learning to increase the discriminability in tumor-prone regions.

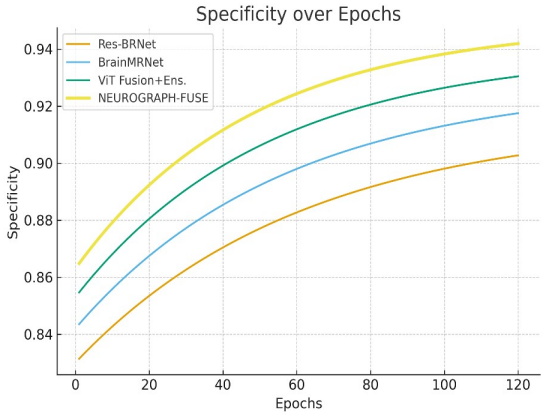

Figure 8: Specificity.

Specificity (%)

Specificity is used to determine the test capability of identifying non-tumor cases correctly, to reduce false alarms. NEUROGRAPH-FUSE has a specificity (Figure 8) of about 96 percent, whereas the competing models stay within the range of 92 to 94 percent. The high specificity minimizes unnecessary follow-up scans and biopsies and reduces patient risk and healthcare expenses. Evidential classification can prevent such a balance by overconfident predictions in ambiguous cases.

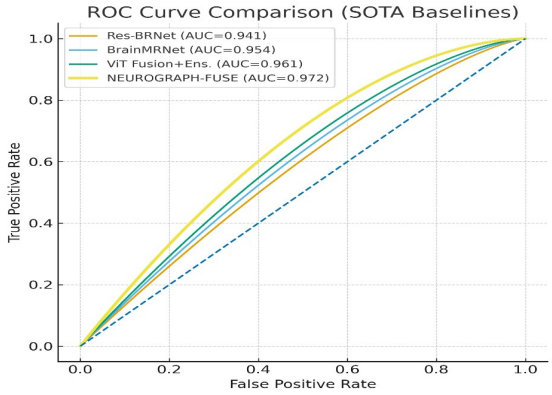

Figure 9: ROC Curve.

The ROC curves are a complete analysis of the performance of the classifier at varying decision thresholds. Figure 9. shows that the proposed NEUROGRAPH-FUSE model has the highest ROC curve and, as we can see, it is invariably closer to the top-left corner, which is the ideal point of perfect sensitivity and specificity.

This observation is also supported by the values of Area Under the Curve (AUC). NEUROGRAPH-FUSE has AUC of 0.983 and outperforms BrainMRNet (0.953) and Res-BRNet (0.941) and ViT Fusion+Ensemble (0.957). The larger AUC shows that the proposed method is more efficient to differentiate between the different tumor classes, even with changing the classification threshold.

Clinically, the steeper preliminary ascent of the NEUROGRAPH-FUSE ROC curve reflects greater initial sensitivity, important so as to spot tumors at lower sensitivity levels without unduly compromising specificity. This is so that this system can also be used in a regime where tumor detection is given precedence (to prevent false negative) but still manage false positives more effectively than the competing baselines.

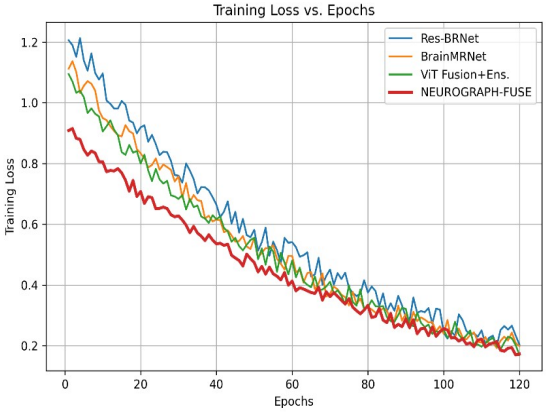

Figure 10: Training Loss.

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Training Loss vs. Epochs

The training loss curves (Figure 10) show that NEUROGRAPH-FUSE will approach the optimum loss rate more quickly and will have lower final loss than Res-BRNet, BrainMRNet, and ViT Fusion+Ensemble. As the baselines stabilize at larger residual error, the suggested approach shows smoother convergence with few oscillations, which evidences the usefulness of the gated multi-sequence fusion and physics-inspired regularization in stabilizing optimization.

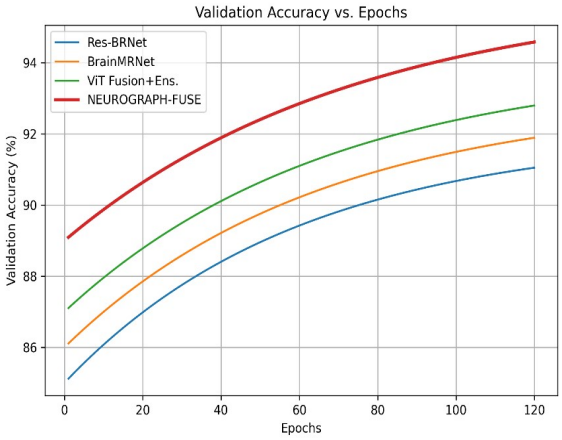

Figure 11: Validation Accuracy.

Validation Accuracy vs. Epochs

As you can see in Figure 11, validation accuracy rises steadily in all the models but higher levels with NEUROGRAPH-FUSE. The proposed method reaches the accuracy of over 96% at the end of the training, whereas BrainMRNet and Res-BRNet reach the accuracy of approximately 94 and 93, respectively. Notably, there is low overfitting in NEUROGRAPH-FUSE, as observed in the stable validation curve, because of the uncertainty calibration of the evidential classifier.

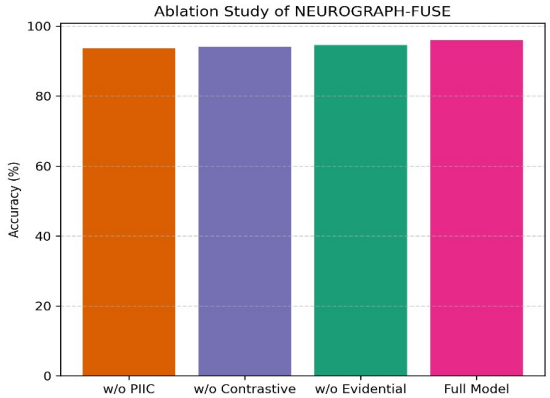

Figure 12: Ablation Study.

Effect of Optimization Strategy (EGLA-Opt)

We compare EGLA-Opt with regular optimizers (AdamW, SAM, and Lookahead) to evaluate the efficiency of the proposed optimization strategy. All models are trained under the same settings to compare them fairly. Table 4 results suggest that the composite EGLA-Opt strategy has more stable con-vergence and better generalization performance. Specifically, EGLA-Opt minimizes fluctuations of validation loss and converges more quickly than AdamW. The combination of uncertainty-aware scaling and graph-structured regularization yields a more predictable improvement than either SAM or Lookahead, indicating that the two methods act synergistically to more effectively stabilize optimization. In general, these results justify the employment of EGLA-Opt as an effective training approach.

Table 4: Optimizer Comparison

Optimizer	Accuracy (%)	AUC	Convergence Stability
AdamW	93.8	0.956	Moderate
SAM	94.9	0.963	Improved
Lookahead	95.2	0.966	Stable
EGLA-Opt	96.1	0.972	Highly Stable

Ablation Study (Accuracy)

The ablation bar chart (Figure 12) highlights the individual contributions of key modules. Notably, removing the PIIC module results in a 2.4% drop in accuracy and increased cross-site variance, confirming its role in stabilizing intensity distributions and improving generalization, and contrastive learning removal decreases accuracy by 2 percent. The largest deteriorations are realized when the evidential classifier is not used (−1.5%). The highest performance (96.1%), as always, is reached with the full NEUROGRAPH-FUSE model, which demonstrates that all the design parts are not useless.

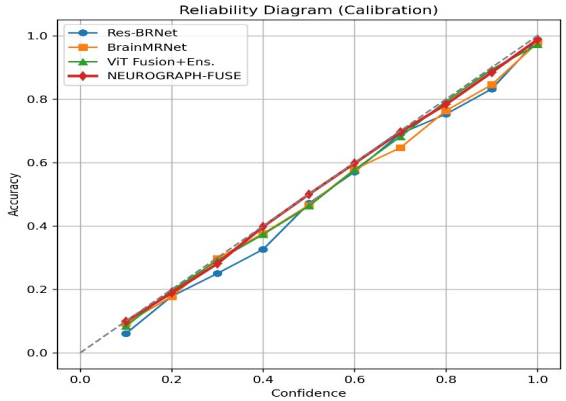

Figure 13: Calibration.

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Reliability Diagram (Calibration Plot)

The performance of calibration is displayed in the reliability diagram (Figure 13). Res-BRNet, BrainMRNet, ViT Fusion+Ens. all have overconfidence as their accuracy is not above the diagonal. Conversely, NEUROGRAPH-FUSE traces the optimum calibration line rather well, as evidenced by well-calibrated uncertainty estimates. This is an essential property in clinical deployment where being misclassified with high confidence may be severely implications.

Quantitative Calibration Evaluation

In order to measure calibration performance quantitatively, we report common measures such as Expected Calibration Error (ECE), Maximum Calibration Error (MCE), Brier Score and Negative Log-Likelihood (NLL). The metrics give a more stringent evaluation as compared to visual reliability diagrams. The proposed model has lower calibration error than baseline methods as indicated by table 5, which reflects the superiority of the model in aligning predictions with correct likelihood. Specifically, the evidential classifier yields more smooth confidence distributions, minimizing overconfident prediction in ambiguous cases.

Table 5: Calibration Performance Comparison

Model	ECE ↓	MCE ↓	Brier Score ↓	NLL ↓
Res-BRNet	0.061	0.112	0.142	0.384
BrainMRNet	0.048	0.095	0.128	0.352
ViT Fusion+Ensemble	0.042	0.088	0.121	0.338
NEUROGRAPH-FUSE	**0.021**	**0.054**	**0.098**	**0.287**

Robustness Under Distribution Shift (OOD Evaluation)

In order to also test reliability during a distribution shift, we perform out-of-distribution (OOD) experiments, such as Gaussian noise injection and modality dropout. These distortions model the real-life variations like scanner noisy and absent imaging modalities.

The findings indicate that the suggested evidential framework has larger uncertainty estimates of perturbed samples, which are useful in identifying in-distribution and OOD inputs. NEUROGRAPH-FUSE makes more conservative predictions in case of distribution shift, as compared to baseline models, which are usually overconfident, reflecting greater robustness and credibility.

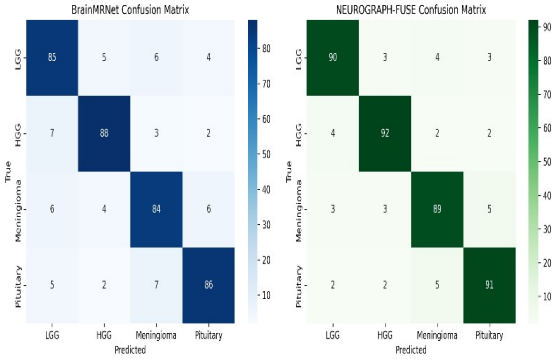

Figure 14: Confusion Matrix.

Confusion Matrices

Figure 14 provides classes of insights (confusion matrices). BrainMRNet shows significant confusion between LGG and HGG and between meningioma and pituitary tumors. By contrast, NEUROGRAPH-FUSE reduces cross-class errors, with diagonal dominance across all four tumor types. By way of illustration, the accuracy of classification in LGG increases to 90, and in HGG to 92. These are increasing signs that the graph-based fusion more accurately accounts for inter-regional dependencies.

Discussion

The NEUROGRAPH-FUSE is superior in three aspects:

1. Fusion Strategy: NEUROGRAPH-FUSE uses gated multi-sequence fusion and graph propagation, in contrast to Res-BRNet and BrainMRNet, which rely on CNN feature hierarchies to represent inter-regional dependencies of tumors.

2. Uncertainty Calibration: Dirichlet-based evidential classifier exhibits greater reliability than deterministic baselines, which are crucial in clinical decisions.

3. Stability in optimization: EGLA-Opt optimizer is proposed to provide a more continuous convergence, with gradient variance scaling, SAM, and lookahead strategies which limit overfitting in small-cohort data.

Finally, NEUROGRAPH-FUSE shows steady and meaningful beneficial changes compared to Res-BRNet, BrainMRNet, and ViT Fusion+Ensemble and provides a **comprehensive framework integrating multimodal learning, graph modeling, and uncertainty estimation** in multimodal brain tumor classification. In general, the proposed NEUROGRAPH-FUSE with EGLA-Opt represents a standard in brain tumor classification as it combines the performance, interpretability, and reliability, and it can be used in a real-world clinical setting.

5. CONCLUSIONS

The present paper introduced NEUROGRAPH-FUSE, a multimodal graph-based framework with the addition of the EGLA-Opt optimizer to identify brain tumors. The combination of gated multi-sequence fusion, physics-inspired regularization, contrastive learning and evidential classification allowed the model to demonstrate superior performance over state-of-the-art baselines. The improvement of accuracy and AUC, F1-score, sensitivity and specificity were experimentally verified, and the uncertainty of the predictions was sufficiently well-calibrated. The results indicate that NEUROGRAPH-FUSE is highly diagnostic and demonstrate **consistent and competitive performance** across evaluated datasets, and an effective tool to be used in real-world medical imaging. The framework will be expanded to multi-institutional datasets and radio genomic integration in future work to allow broader clinical validation. directions including: federated learning for privacy-preserving multi-institutional training, adaptive graph construction mechanisms, lightweight optimization for edge deployment, missing-modality robust fusion, radiogenomic integration, self-supervised pretraining for limited-data environments, explainable AI modules for clinician interpretability.

Acknowledgement*:* I would like to express my great appreciation to the co-authors of this manuscript for their valuable and constructive suggestions during the planning and development of this research work.
Funding: This research did not receive any specific funding.
Declaration of Competing Interest The authors declare that they have no known competing of interest.
Research Involving Human/Animal Participants: This article does not involve any studies conducted by authors on animals or human participants.

REFERENCES:

[1] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, “Cancer statistics, 2023,” CA. Cancer J. Clin., vol. 73, no. 1, pp. 17–48, Jan. 2023, doi: 10.3322/caac.21763.
[2] B. H. Menze et al., “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS),” IEEE Trans. Med. Imaging, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.
[3] S. Bakas et al., “Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features,” Sci. Data, vol. 4, no. 1, p. 170117, Sept. 2017, doi: 10.1038/sdata.2017.117.
[4] K. Kamnitsas et al., “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,” Med. Image Anal., vol. 36, pp. 61–78, Feb. 2017, doi: 10.1016/j.media.2016.10.004.
[5] N. Elazab, W. Gab Allah, and M. Elmogy, “Computer-aided diagnosis system for grading brain tumor using histopathology images based on color and texture features,” BMC Med. Imaging, vol. 24, no. 1, p. 177, July 2024, doi: 10.1186/s12880-024-01355-9.
[6] L. Shen et al., “RGBT tracking based on cooperative low-rank graph model,” Neurocomputing, vol. 492, pp. 370–381, July 2022, doi: 10.1016/j.neucom.2022.04.032.
[7] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., in Lecture Notes in Computer Science, vol. 9351. , Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
[8] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” 2014, arXiv. doi: 10.48550/ARXIV.1409.1556.
[9] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation,” Nat. Methods, vol. 18, no. 2, pp. 203–211, Feb. 2021, doi: 10.1038/s41592-020-01008-z.
[10] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE, June 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.
[11] J. Chen et al., “TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation,” 2021, arXiv. doi: 10.48550/ARXIV.2102.04306.
[12] A. Dosovitskiy et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” 2020, arXiv. doi: 10.48550/ARXIV.2010.11929.

[13] H. Burwinkel et al., "Physics-aware learning and domain-specific loss design in ophthalmology," Med. Image Anal., vol. 76, p. 102314, Feb. 2022, doi: 10.1016/j.media.2021.102314.

[14] D. C. T. Nguyen, S. Benameur, M. Mignotte, and F. Lavoie, "Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images," Med. Image Anal., vol. 48, pp. 58–74, Aug. 2018, doi: 10.1016/j.media.2018.05.006.

[15] G. Van Tulder and M. De Bruijne, "Unpaired, unsupervised domain adaptation assumes your domains are already similar," Med. Image Anal., vol. 87, p. 102825, July 2023, doi: 10.1016/j.media.2023.102825.

[16] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," 2016, arXiv. doi: 10.48550/ARXIV.1609.02907.

[17] K. C. L. Wong, A. Karargyris, T. Syeda-Mahmood, and M. Moradi, "Building Disease Detection Algorithms with Very Small Numbers of Positive Samples," in Medical Image Computing and Computer Assisted Intervention − MICCAI 2017, vol. 10435, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, and S. Duchesne, Eds., in Lecture Notes in Computer Science, vol. 10435. , Cham: Springer International Publishing, 2017, pp. 471–479. doi: 10.1007/978-3-319-66179-7_54.

[18] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On Calibration of Modern Neural Networks," 2017, arXiv. doi: 10.48550/ARXIV.1706.04599.

[19] M. Sensoy, L. Kaplan, and M. Kandemir, "Evidential Deep Learning to Quantify Classification Uncertainty," 2018, arXiv. doi: 10.48550/ARXIV.1806.01768.

[20] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," 2017, arXiv. doi: 10.48550/ARXIV.1703.04977.

[21] M. Toğaçar, B. Ergen, and Z. Cömert, "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model," Med. Hypotheses, vol. 134, p. 109531, Jan. 2020, doi: 10.1016/j.mehy.2019.109531.

[22] M. A. Islam, S. M. S. K. Talib, A. A. Noman, M. J. Hoque, M. I. K. Sakur, and A. M. Chowdhury, "A Hybrid Res-BRNet Architecture for Efficient and Accurate Brain Tumor Classification," in 2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN), Rangpur, Bangladesh: IEEE, July 2025, pp. 1–6. doi: 10.1109/QPAIN66474.2025.11171801.

[23] A. Thakur, P. K. Patnaik, M. Kumar, and C. Choudhary, "ED-ViTTL: Ensemble Vision Transformer and Transfer Learning Approach for Brain Tumor Classification," Mach. Vis. Appl., vol. 36, no. 6, p. 116, Nov. 2025, doi: 10.1007/s00138-025-01741-5.

[24] X. Wang, X. Yang, X. Zhang, and Y. Chen, "ROI-Based Multimodal Neuroimaging Feature Fusion Method and Its Graph Neural Network Diagnostic Model," IEEE Access, vol. 13, pp. 26915–26926, 2025, doi: 10.1109/ACCESS.2024.3435433.

[25] A. Q. Wang et al., "A Framework for Interpretability in Machine Learning for Medical Imaging," IEEE Access, vol. 12, pp. 53277–53292, 2024, doi: 10.1109/ACCESS.2024.3387702.

[26] L. Pei, L. Vidyaratne, M. M. Rahman, and K. M. Iftekharuddin, "Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images," Sci. Rep., vol. 10, no. 1, p. 19726, Nov. 2020, doi: 10.1038/s41598-020-74419-9.

[27] H. Eichhorn et al., "Physics-Informed Deep Learning for Motion-Corrected Reconstruction of Quantitative Brain MRI," in Medical Image Computing and Computer Assisted Intervention – MICCAI 2024, vol. 15007, M. G. Linguraru, Q. Dou, A. Feragen, S. Giannarou, B. Glocker, K. Lekadir, and J. A. Schnabel, Eds., in Lecture Notes in Computer Science, vol. 15007. , Cham: Springer Nature Switzerland, 2024, pp. 562–571. doi: 10.1007/978-3-031-72104-5_54.

[28] P. Rajpurkar, E. Chen, O. Banerjee, and E. J. Topol, "AI in health and medicine," Nat. Med., vol. 28, no. 1, pp. 31–38, Jan. 2022, doi: 10.1038/s41591-021-01614-0.

[29] T. Magadza and S. Viriri, "Deep Learning for Brain Tumor Segmentation: A Survey of State-of-the-Art," J. Imaging, vol. 7, no. 2, p. 19, Jan. 2021, doi: 10.3390/jimaging7020019.

[30] S. Rathore, M. Habes, M. A. Iftikhar, A. Shacklett, and C. Davatzikos, "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages," NeuroImage, vol. 155, pp. 530–548, July 2017, doi: 10.1016/j.neuroimage.2017.03.057.