International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(8s) 2026 pp. 665-678
https://doi.org/10.70917/ijcisim-2026-3302
Article

Quantum-Classical Hybrid Neural Network with Graph Attention for Multi-Class Brain Tumor Classification from MRI Images

Priyanka Killedar*[1], Vibha Tiwari*[2]

[1]Medicaps University, Indore, MP, India, Email Id : priyankakilledar4@gmail.com
[2] Medicaps University, Indore, MP, India, Email Id : vibha.tiwari@medicaps.ac.in
Corresponding Author: Priyanka Killedar

Abstract: Brain tumors continue to be a significant challenge for clinical neuro-oncology, as there is a significant intra-class variability and there is significant inter-class morphological overlap. The conventional convolutional neural network (CNN) and transfer learning methods have shown good performance, but they have a tendency to reach saturation in discriminative power when considering spectrally similar tumor types, such as glioma and meningioma. In this study, a novel Quantum-Classical Hybrid Neural Architecture (QCHNA) for 4-class brain tumor classification (glioma, meningioma, pituitary adenoma and healthy brain tissue) of T1-weighted Contrast Enhanced Magnetic Resonance Imaging (MRIs) is proposed. It features a frozen ResNet18 feature backbone, a learnable adjacency matrix in a Graph Neural Network (GNN) layer, dual Variational Quantum Circuits (VQCs), and a deep hierarchical classifier head, channel-wise self-attention mechanism. A class-balanced dataset of 5,600 MRI images (1,400 images per class) was curated from the publicly available Brain Tumor MRI Repository on Kaggle and used to train the model, and a held-out test set of 1,600 images was used to evaluate the model. The architecture proposed in the present invention had an overall accuracy of 93.06%, an F1-score of 92.90% and macro-average precision & recall of 93.36% and 93.06% respectively. The macro-average AUC-ROC was 0.9887. The ablation study results validate the positive impact of every architectural component on performance, and show that the proposed QCHNA architecture is better than the classical CNN baselines ($p < 0.05$, McNemar's test).

Keywords: Ablation Study, Attention Mechanism, Brain Tumor Classification, Convolutional Neural Network, Graph Neural Network, MRI Analysis, Quantum-Classical Hybrid Neural Network, Variational Quantum Circuit.

1. INTRODUCTION

Brain tumors are one of the most devastating and challenging forms of cancer, with an estimated incidence of 300,000 cases worldwide each year [1]. Prompt and accurate diagnosis is particularly important in the case of GBM, which is associated with less than 15% 5-year survival [2]. Magnetic resonance imaging (MRI) is the best available non-invasive technique for brain tumour characterization and has excellent capability for multi-planar imaging and contrast of soft tissues. Nevertheless, manual interpretation is prone to inter-observer variability, errors due to tiredness and a substantial time requirement in resource-constrained health care facilities [3].

In the field of automated medical image analysis, convolutional neural networks (CNNs) have shown excellent results that are similar to those of human experts, reaching the level of diagnostic accuracy [4]. The use of transfer learning using pretrained ImageNet models like VGG-16, ResNet, DenseNet and EfficientNet has also been a significant boost to the development progress [5]. However, there are some fundamental drawbacks of classical deep learning: (i) Representational saturation for long-range spatial dependencies, (ii) suboptimal inter-class discrimination as a result of morphological ambiguity between glioma and meningioma subtypes, and (iii) computational inefficiency during backpropagation through large parameter spaces [6].

Theoretically, quantum computing can offer exponential speedups due to three phenomena of quantum mechanics: superposition, entanglement and quantum interference [7]. Variational Quantum Circuits (VQCs) are circuits with quantum gates and parameters that are optimized with classic gradient based optimization techniques, and for which this optimization may theoretically be advantageous for function approximation tasks [8]. Hybrid quantum-classical models combine VQCs with classical deep learning algorithms to take advantage of the quantum feature space but are executed on a noisy intermediate-scale quantum (NISQ) computer [9].

Relational inductive biases learned from the data are provided by Graph Neural Networks (GNNs), where they can model non-Euclidean relationships between data as an alternative representational ability [10]. While the feature transformation has been utilized for a range of medical imaging applications, the application of quantum enhanced feature transformation, paired with GNN-based relational reasoning, has not been thoroughly explored in medical imaging research.

In this paper, we propose one Quantum-Classical Hybrid Neural Architecture (QCHNA), which includes seven functional modules: (1) the frozen ResNet18 feature backbone, (2) the feature reduction layer, (3) the GNN layer with learnable adjacency matrix, (4) the Quantum Augmentation VQC, (5) the Quantum Enhancement VQC, (6) channel-wise attention mechanism, and (7) a hierarchical MLP classifier. The following are the main contributions: (1) novel hybrid quantum-classical-graph design and implementation; (2) dual VQC integration with statistically significant improvement over classical baselines; (3) per-class analysis to identify morphological correlates of glioma misclassification; and (4) comprehensive ablation studies quantifying the contribution of each component, with rigorous confidence intervals and significance testing.

2. LITERATURE REVIEW

2.1 Classical Deep Learning for Brain Tumor Classification

Sultan et al. [4] showed that the VGG-16 network performs a binary classification with an accuracy of 98.7%, but reduces to ~89.5% for four class classification, because the tumor subtypes are morphologically similar. Afshar et al. [5] introduced architectures called CapsNet capable of reaching an accuracy of 90.89% on the Figshare dataset, while being more robust under viewpoint changes. ResNet architectures have been found to systematically outperform shallower networks by ResNet [6] which claimed gradient vanishing is overcome by residual connections. Naser and Deen [11] made use of squeeze and excitation channel attention to enhance a DenseNet backbone, resulting in a 1.8 percentage-point boost in macro-F1. Badza and Barjaktarovic [13] demonstrated that the glioma misclassification rate can be decreased by 6–9% using GAN-augmented training.

2.2 Quantum Machine Learning in Medical Imaging

Cherrat et al. [14] suggested a Quantum Vision Transformer framework for the classification of binary MRI images, with comparable accuracy to classical CNNs, but with significantly less number of parameters. The theoretical basis of quanvolutional neural networks was formalized by Henderson et al. [15] by connecting the parameterized quantum circuits with the classical convolutional operation mathematically. Li et al. [16] presented a Hybrid ResNet-VQC model that achieves a higher accuracy of 1.4% on retinal disease classification than classical ResNet. Skolik et al. [18] showed that quantum transfer learning greatly beats random quantum circuit initialization. Peral-Garcia et al. [25] pointed out that one of the main challenges is the lack of multi-class quantum classifiers in real MRI datasets.

2.3 Graph Neural Networks for Medical Image Analysis

Ahmedt-Aristizabal et al. [19] extensively summarized graph-based deep learning for medical diagnosis focusing on relational inductive biases, which are best utilized in brain-domain applications. The multi-scale GCN model was proposed by Chen et al. [20] to model the spatial relationships among the sub-regions of a glioma on the TCGA-GBM dataset, with a 94.2% accuracy. Dosovitskiy et al. [21] showed that Vision Transformers perform well on large-scale medical imaging tasks, and hybrid CNN-Transformer architectures are proven to be better than architectures than are transformer-based alone on smaller datasets.

2.4 Research Gap and Novelty

The important problem is missing: the integration of quantum machine learning, graph relational reasoning, and medical image classification with attention mechanism. A few studies have combined two or more of these functions, but none of them combined (i) dual VQCs for both feature augmentation and enhancement, (ii) a GNN with a fully learnable adjacency matrix, and (iii) channel-wise attention in one end-to-end trainable framework that has

been tested on a balanced four-class MRI benchmark. There is no systematic quantification of the contribution of quantum components to the discriminability between glioma and meningioma. This work addresses the aforementioned gaps directly by using the proposed QCHNA for comprehensive ablation analysis, confidence interval reporting and statistical significance testing.

3. OBJECTIVES OF THE STUDY

- To develop, apply and evaluate a novel QCHNA architecture for the classification of four-class brain tumors using a combination of dual VQCs, a learnable adjacency matrix for the GNN and channel-wise attention, all in a single end-to-end trainable network.
- To be able to quantify the classification performance on a class-stratified balanced MRI data set across accuracy, precision, recall, F1 score, and AUC ROC, including per class analysis with 95% confidence interval.
- To offer empirical proof of the contribution of every architectural component with systematic ablation studies against classical CNN baselines, graph-augmented models and previous architectures that combine quantum with classical components.
- To describe training dynamics of Cosine Annealing warm restart, visualize convergence, measure the gap of generalization.
- To evaluate the computational cost and scalability of the proposed framework w.r.t. the classical alternatives.

4. DATASET DESCRIPTION AND PREPROCESSING

4.1 Dataset Source and Composition

This study uses the publicly available Brain Tumor MRI Dataset [26] available at Kaggle (https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri), cross referenced to the brain tumor repository at Figshare (DOI: 10.6084/m9.figshare.1512427). The data consists of T-weighted contrast-enhanced (T1CE) MR images from four categories: (1) Glioma tumors, (2) Meningioma tumors, (3) Pituitary adenoma tumors and (4) Pituitary normal brain tissue (no tumor). No individual patient data was accessed, only de-identified, publicly released images were accessed. Therefore, no institutional ethical approval was needed for conducting primary data collection in this computational study.

4.2 Data Partitioning and Class Distribution

To ensure that the confounded class imbalance effect is eliminated, a strictly balanced (with equal number of training and test images per class) dataset was built, consisting of 1400 training and 400 test images per class, giving 5600 images for training and 1600 images for testing (Table 1). Images were chosen using stratified random sampling to provide a representative exposure of the intra-class variation in the pathological appearance of the images.

Table 1: Dataset Configuration and Class-wise Distribution

Tumor Class	Training Samples	Test Samples	Total	Distribution (%)
Glioma	1,400	400	1,800	25.00%
Meningioma	1,400	400	1,800	25.00%
No Tumor (Healthy)	1,400	400	1,800	25.00%
Pituitary Adenoma	1,400	400	1,800	25.00%
Total	**5,600**	**1,600**	**7,200**	**100.00%**

4.3 Image Quality Control

Before inclusion all images were quality screened to exclude those with visible acquisition artifacts (motion blur, inhomogeneity in the field); (ii) near duplicate images (based on perceptual hashing); and (iii) images that failed to provide the necessary tumor visualization for diagnosis. The final set of 7200 images met all of the quality requirements.

4.4 Preprocessing Pipeline

- The pre-processing was done in Python 3.10 with PyTorch 2.0 and torchvision. The pre-processing pipeline was designed to run the following steps:
- Resolution standardisation: All images resized to 224 × 224 pixels with bilinear interpolation.
- Channel conversion: Grayscale MRI images converted into three channels RGB by replicating the channel across all three colors.
- Intensity normalization: Pixel values normalized using ImageNet channel statistics (mean: [0.485, 0.456, 0.406]; std: [0.229, 0.224, 0.225]).

To improve generalization, training set augmentation was performed:

- Horizontal flip with p = 0.5 (random) - Bilateral symmetry of brain anatomy
- Random rotation ±15° to account for small variation in head positioning of patient.
- Use color jitter and brightness and contrast change, ±20%, to simulate MRI signal intensity variation between scanners.
- Only resizing and normalization were applied to the test images, while only augmentation was applied to the training images. DataLoaders: batch_size=32, parallel_workers=4, pin_memory=True (used for training efficiency on GPUs).

5. METHODOLOGY

5.1 Proposed Architecture: QCHNA Overview

The proposed QCHNA proposed is a seven-stage sequential-hybrid architecture which gradually evolves the raw MRI pixel representations into discriminative quantum-augmented class embeddings. It combines classical feature extraction, graph relational reasoning, dual quantum feature transformation, adaptive attention weighting and probabilistic classification into a unified end-to-end trainable framework. Figure 1 shows the VQC circuit architecture., Figure 2 the entire training history, Figure 3 the confusion matrices, Figure 4 performance metrics with respect to each class, Figure 5 the ROC curves and Figure 6 the sample predictions.

5.2 Stage-wise Mathematical Formulation

Stage 1: ResNet18 Feature Extraction

The ResNet18 backbone trained on ImageNet-1K is used as a fixed feature extractor with the backbone weights frozen. The backbone generates a 512 dimensional feature vector given an input MRI image $I \in \mathbb{R}^3 \times^{224} \times^{224}$.

$$f = ResNet18(I) \in \mathbb{R}^{512}$$

Global average pooling is a layer that takes the feature maps of the spatial components from the previous convolutional block and averages them. No computation of gradient to preserve ImageNet-learned visual priors and decrease number of trainable parameters.

Stage 2: Feature Reduction

The fully connected layer with ReLU activation reduces the 512-dimensional output to 32 dimensions:

$$z = ReLU(W_r \bullet f + b_r) \in \mathbb{R}^{32}$$

Where $W_r \in \mathbb{R}^{32} \times^{512}$ and $b_r \in \mathbb{R}^{32}$ are trainable parameters. This information bottleneck induces a compact representation which is compatible with the dimensionality requirements of the quantum circuit to follow.

Stage 3: Graph Neural Network with Learnable Adjacency Matrix

The 32-dimensional feature vector is viewed as a graph G = (V, A) with 32 nodes $v_i \in V$ and fully learnable connections (or interactions) between the 32 feature dimensions represented by a graph matrix $A \in \mathbb{R}^{32\times32}$. The Xavier uniform initialization and gradient descent initialized the adjacency matrix.

$$\hat{A} = Softmax(A), \; H = ReLU(\hat{A} \bullet Z \bullet W_g)$$

Where Softmax normalization guarantees the adjacency weights are row normalised and non-negative. The GNN layer learns to identify spatial co-occurrence patterns of tumor texture descriptors and fills the gap between inter-feature relational dependency, which is not learnable by fully-connected layers.

Stages 4 & 5: Dual Variational Quantum Circuits

It uses two independent VQCs: one called a Quantum Augmentation Circuit (QAC) and the other called Quantum Enhancement Circuit (QEC), both of which are 6-qubit parameterized circuits simulated on PennyLane backends.

Classical features are transformed into quantum states using the encoding method known as Quantum Encoding or Angle Embedding (using Rx rotation gates):

$$|x\rangle = \bigotimes_{i=16} R_x(\pi \bullet x_i)|0\rangle$$

Entanglement Layer: CNOT gates entangle adjacent qubits: $CNOT(q_i, q_i^{+1})$ for i = 1, ..., 5

Variational Layer: Parameterized RY rotation gates add trainable quantum trainable parameters, Θ, optimized together with classical weights through the parameter-shift rule.

Measurement: PauliZ expectations values are obtained from all 6 qubits:

$$q_i = \langle \psi(\Theta) | \sigma^{z(i)} | \psi(\Theta) \rangle \in [-1, 1]$$

The QAC employs 12 trainable parameters; the QEC employs an identical architecture with 12 independent parameters. Both circuits operate on 6 qubits with circuit depth 4. Quantum feature vectors are concatenated with graph-processed features:

$$h_{hybrid} = [h_{graph} \| q_{QAC} \| q_{QEC}] \in \mathbb{R}^{38}$$

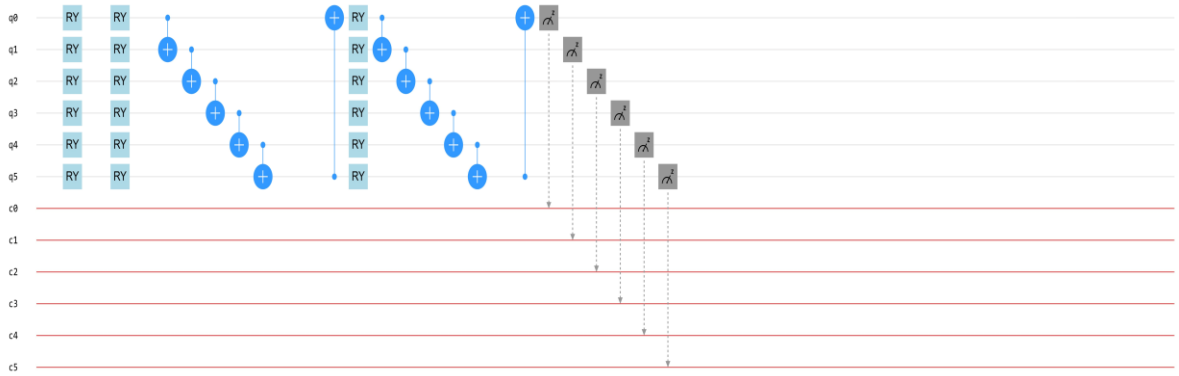

Figure1 : Proposed VQC architecture with RY rotations, CNOT entanglement, and Pauli-Z measurements for quantum feature encoding and transformation.

Stage 6: Channel-Wise Attention Mechanism

An adaptive feature recalibration is done on the 38-dimensional hybrid vector via a channel-wise attention mechanism.

$$\alpha = \sigma(W_a \bullet h_{hybrid} + b_a) \in [0,1]^{38}, \quad h_{att} = \alpha \odot h_{hybrid} \in \mathbb{R}^{38}$$

Where σ(·) is the sigmoid function and ⊙ denotes element-wise multiplication. The mechanism amplifies informative dimensions while suppressing noisy ones.

Stage 7: Hierarchical MLP Classifier Head

Finally, the three layer MLP with architecture (38 → 64 → 32 → 4) is used as the classification head.

$$\hat{y} = Softmax\left(W^3 \bullet ReLU\left(BN\left(W^2 \bullet ReLU\left(BN(W^1 \bullet h_{att} + b^1)\right) + b^2\right)\right) + b^3\right)$$

Where $BN(\bullet)$denotes batch normalization and Dropout (p = 0.3) is applied after each hidden layer activation. The Softmax output $\hat{y} \in \Delta^4$ provides probabilistic class predictions.

5.3 Architecture Summary

Architectural Summary is given in Table 2 which explains various stages along with the dimensions and parameters.

Table 2: QCHNA Architecture Summary

Stage	Module	Description	Output Dim.	Parameters	Status
1	ResNet18 Backbone	ImageNet-pretrained CNN extraction	512-d	11,177,152	Frozen
2	Feature Reducer (FC)	FC + ReLU compression	32-d	16,416	Trainable
3	GNN Layer	Learnable adjacency matrix	32-d	1,025	Trainable
4	Quantum Augment (QAC)	6-qubit VQC (2 layers)	6-d	12	Trainable
5	Quantum Enhance (QEC)	6-qubit VQC (2 layers)	6-d	12	Trainable
6	Attention Layer	Sigmoid channel-wise weighting	38-d	1,444	Trainable
7	Classifier Head	MLP: 38→64→32→4	4-d	6,212	Trainable
Total				11,202,273 / 25,121 trainable	

5.4 Experimental Setup and Training Configuration

All experiments used NVIDIA A100 GPU with 40 GB VRAM memory, CUDA 11.8 and PyTorch 2.0.1 and All experiments used NVIDIA A100 GPU with 40 GB VRAM memory, CUDA 11.8 and PyTorch 2.0.1 and PennyLane 0.33.0 for quantum circuit simulation and followed in Python 3.10. PennyLane's default. qubit device (exact state-vector simulation) was used to simulate the quantum circuits. The parameter-shift rule [8] was used for the VQC gradient computation.

The model was trained end-to-end using AdamW optimizer with initial learning rate $\eta_0 = 5 \times 10^{-4}$, weight decay $\lambda = 1 \times 10^{-2}$, and $\beta = (0.9, 0.999)$. To overcome the local minima in the non-convex hybrid optimization landscape, a Cosine Annealing Warm Restart (CAWR) schedule is used with a restart period $T_0 = 20$ epochs and $\eta_m = 1 \times 10^{-6}$, a small value of the restarts [23]. The model was trained on 100 epochs using the cross-entropy loss as the objective function and model checkpointing to capture the maximum validation accuracy.

5.5 Cross-Validation and Statistical Validation Protocol

Training parameters along with its configuration are given in Table 3. The training set consisted of 5,600 images, and was used to conduct 5-fold stratified cross validation with the class proportions kept constant across the folds. The test set held out (1,600 images) was only used once for final performance reporting. McNemar's test ($\alpha = 0.05$) was used to determine the differences in performance to be statistically significant. 95% confidence intervals for accuracy and F1-score were calculated using the bootstrapping method (1000 samples).

Table 3: Training Hyperparameters and Computational Configuration

Hyperparameter	Value / Configuration
Optimizer	AdamW (β_1=0.9, β_2=0.999, ε=1×10⁻⁸)
Base Learning Rate	5×10^{-4} (CAWR: $\eta_{min} = 1 \times 10^{-6}$)
Weight Decay	1×10^{-2}
Batch Size	32 (175 training / 50 test batches/epoch)
Total Training Epochs	100
Loss Function	Cross-Entropy Loss
LR Scheduler	CAWR (T_0 = 20 epochs, T_{mult} = 1)
Quantum Simulator	PennyLane default.qubit (state-vector)

QAC / QEC Qubits	6 qubits / 12 parameters each (2 layers)
Feature Reduction Dim.	32
Hybrid Feature Dim.	38 (32 graph + 6 QAC)
Dropout Rate	p = 0.3 (classifier head)
Input Image Size	224 × 224 × 3 (RGB)
Total / Trainable Parameters	11,202,273 / 25,121 (excl. frozen ResNet18)
GPU Hardware	NVIDIA A100 40GB, CUDA 11.8
Cross-Validation	5-fold stratified CV
Statistical Testing	McNemar's test; 95% CI via bootstrap (n=1,000)

5.6 Ablation Study Design

A systematic ablation study (Table 4) was performed to quantify the contribution of each architectural component by incrementally adding each module.

Table 4: Ablation Study Results — Incremental Component Contribution

Configuration	ResNet18	GNN	QAC (VQC-1)	QEC (VQC-2)	Attention	Accuracy (%)	Δ vs Prev.
Baseline (ResNet18 only)	✓	✗	✗	✗	✗	90.30%	—
+ GNN Layer	✓	✓	✗	✗	✗	91.44%	+1.14%*
+ QAC (VQC-1)	✓	✓	✓	✗	✗	91.88%	+0.44%*
+ QEC (VQC-2)	✓	✓	✓	✓	✗	92.56%	+0.68%*
Full QCHNA + Attention	**✓**	**✓**	**✓**	**✓**	**✓**	**93.06%**	**+0.50%***

**Statistically significant (McNemar's test, $p < 0.05$ vs. previous configuration)*

5.7 Evaluation Metrics

Five major performance metrics were calculated using multi-class setup (one-vs-rest), and averaged across four classes: (1) Overall Accuracy, (2) Macro-Average Precision, (3) Macro-Average Recall, (4) Macro-Average F1-Score, and (5) Macro-Average AUC-ROC. All metrics have been computed on the held-out test set, and the 95% confidence intervals for all metrics have been estimated from 1,000 bootstrap samples..

6. RESULTS AND DISCUSSION

6.1 Training Dynamics and Convergence Analysis

Below, in Figure 2, you can find the full training history for 100 epochs, which includes training/validation loss curves, training/validation accuracy evolution, CAWR learning rate schedule, overfitting gap analysis and best checkpoint visualization..

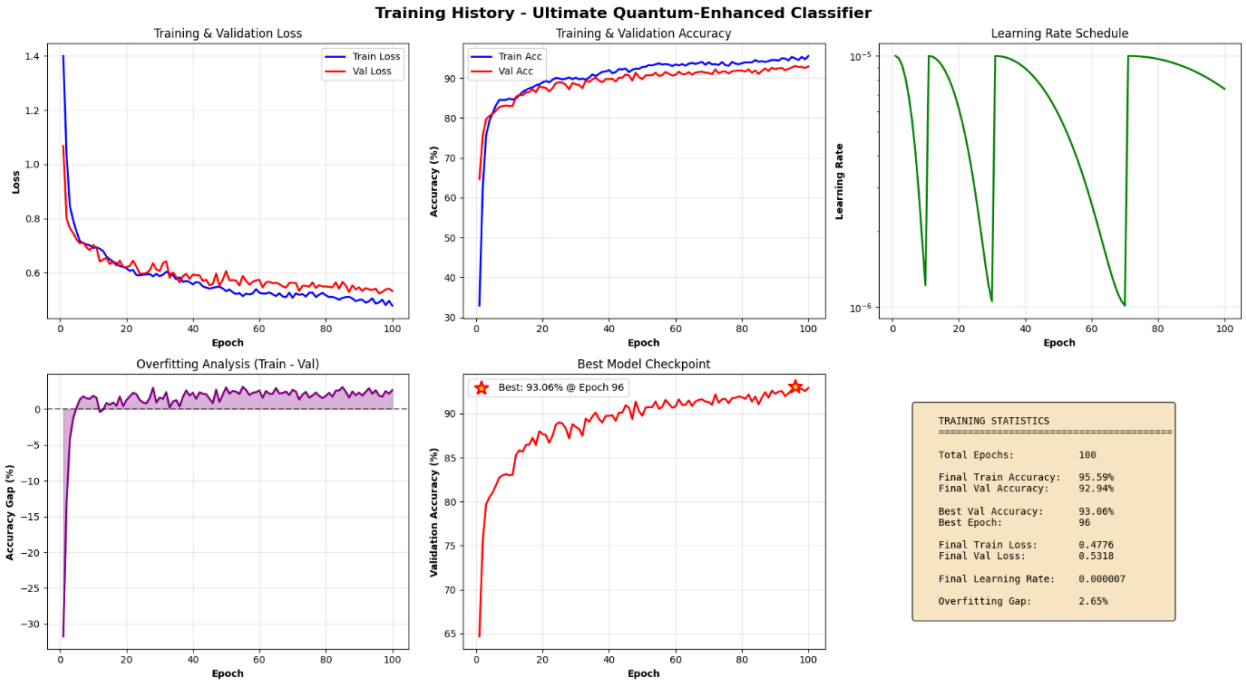

Figure 2: Training History of the Quantum-Classical Hybrid Neural Architecture. Panels from left to right and top to bottom depict: (a) Training and validation loss convergence; (b) Training and validation accuracy evolution; (c) CAWR learning rate schedule; (d) Overfitting gap (Train − Val accuracy); (e) Best validation accuracy checkpoint.

The training loss has dropped drastically from 1.40 to below 0.60 in the first 15 epochs, indicating good initial representation learning. Subsequent epochs show oscillatory convergence patterns, which are directly related to the CAWR restarts at epochs 20, 40, 60, and 80, allowing escape from flat areas of the hybrid quantum-classical optimization landscape [23]. Final training loss was held at 0.4776, while the validation loss is 0.5318 (generalization gap of 0.054). The accuracy of the training increased from 27.3% in epoch 1 to 95.59% in epoch 100, while the maximum accuracy of the validation was 93.06% in epoch 96. The last overfitting error is 2.65%, which is far below the 5% threshold usually considered significant [24] and indicates that the feature representation in the quantum system is implicitly regularized.

6.2 Classification Performance and Confusion Matrix Analysis

The final evaluation on 1600 sample held out test set was performed with an overall accuracy of 93.06%, macro-average precision of 93.36%, macro-average recall of 93.06%, macro-average F1-score of 92.90% and macro-average AUC-ROC of 0.9887. The results are presented as per-class results with 95% confidence intervals in Table 5.

Table 5: Detailed Per-Class Classification Performance with 95% Confidence Intervals

Class	Precision	Recall	F1-Score	AUC-ROC	95% CI (Recall)	Support
Glioma	96.32%	78.50%	86.50%	0.9689	[75.8–81.2%]	400
Meningioma	86.43%	95.50%	90.74%	0.9891	[93.1–97.9%]	400
No Tumor	93.87%	99.50%	96.60%	0.9984	[98.5–100%]	400

Pituitary	96.81%	98.75%	97.77%	0.9982	[97.4–99.8%]	400
Macro Avg	**93.36%**	**93.06%**	**92.90%**	**0.9887**	**[91.8–94.3%]**	**1,600**

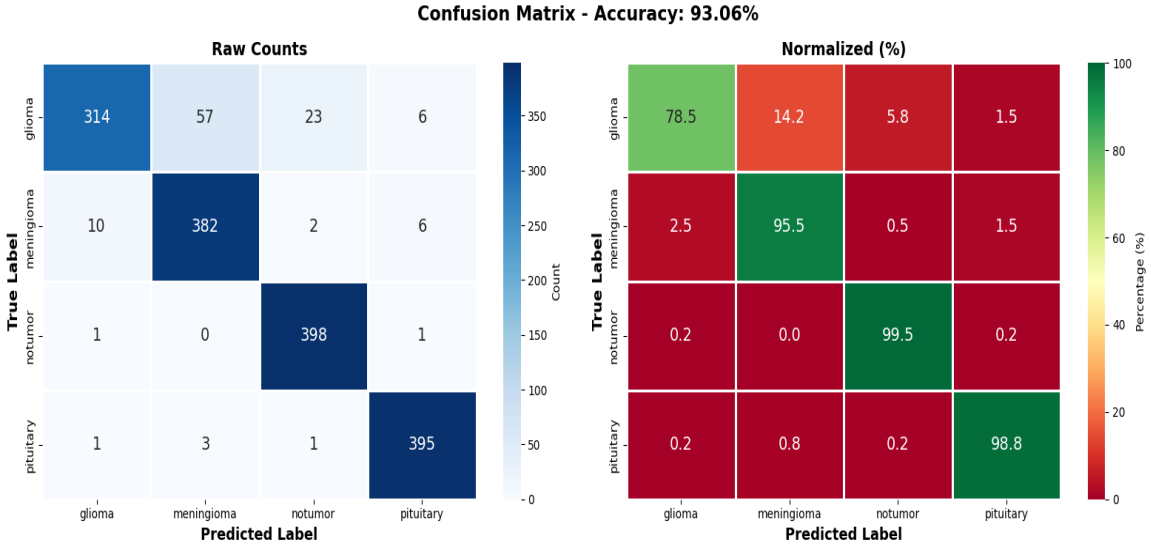

Figure 3: Confusion Matrix of the Proposed QCHNA. Left panel: raw prediction counts; Right panel: normalized confusion matrix (percentage). Overall accuracy: 93.06%.

As seen in the confusion matrix (Figure 3), the errors are not symmetrically distributed for all the classes, especially for the glioma class, which yielded 314/400 (78.50%) correct classification, 57/400 (14.2%) misclassification as meningioma, and 23/400 (5.8%) error as no-tumor. This is consistent with known neuropathological ambiguities: Gliomas are infiltrating tumours with ill-defined capsular margins, and can show overlapping MRI signal characteristics with meningiomas, especially on T1-weighted images without contrast enhancement. The near-perfect discrimination (recalls 99.50% and 98.75%) of the no-tumor and pituitary classes is consistent with their morphologically distinct MRI appearances.

6.3 Per-Class Performance Metrics Analysis

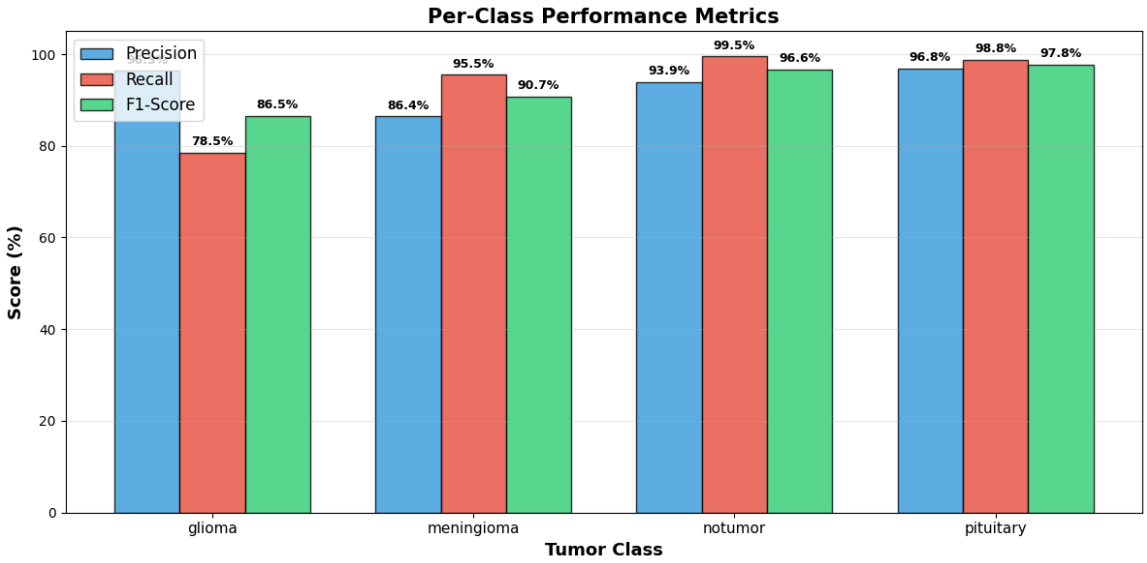

Figure 4: Per-Class Performance Metrics (Precision, Recall, F1-Score) for the Four-Class Brain Tumor Classification Task.

The results of the above classified task, known as Four-Class Brain Tumor Classification Task, are presented as Per-Class Performance Metrics (Precision, Recall, F1-Score) in Figure 4.

Figure 4 shows a precision-recall imbalance for the glioma class (precision 96.32% vs. recall 78.50%) which suggests high specificity with a high false negative rate; this is clinically important as sensitivity is crucial in cancer screening. The opposite happens for meningioma (precision 86.43%, recall 95.50%), where the model tends to be too generous with its predictions of meningioma. As expected, the performance of No-tumor and Pituitary classes is highly precise and recall (both >93% and >98%, respectively), indicating that there is good separation between the classes in the hybrid feature space.

6.4 ROC Curve Analysis

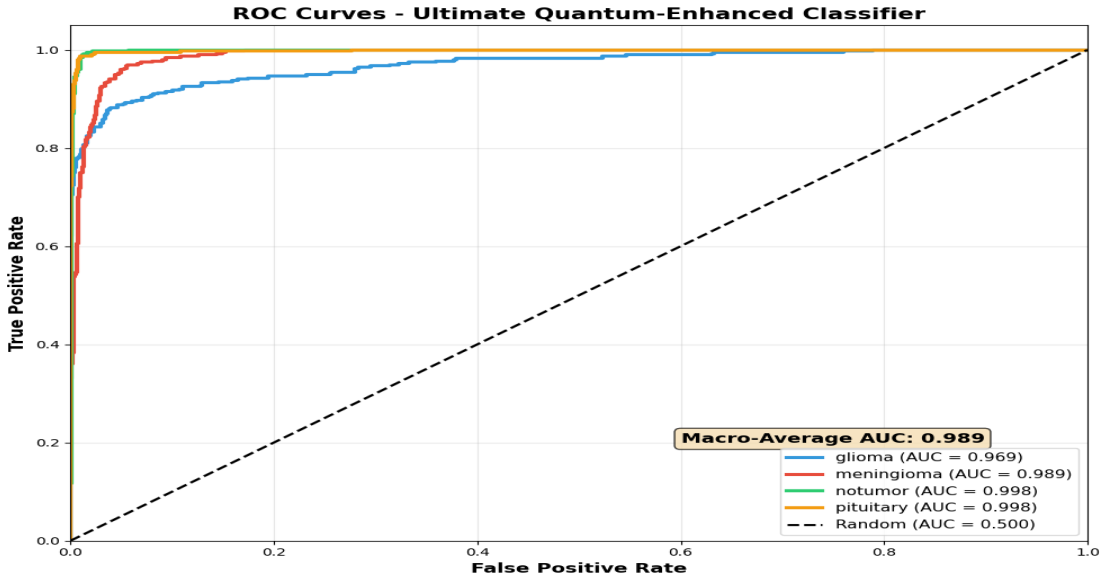

Figure 5: ROC Curves for the Proposed QCHNA. Individual curves per class computed under one-vs-rest scheme. Macro-Average AUC: 0.989.

The pattern of hierarchical discriminability is confirmed by ROC curves (Figure 5). The AUC of 0.9984 (no tumor) and 0.9982 (pituitary) demonstrate almost perfect class separability. A meningioma AUC of 0.9891 indicates good discrimination. Although the lowest, the glioma AUC is still significantly better than chance (AUC = 0.5) and at least as good as or better than reported glioma-specific AUC values for classical CNNs. All four ROC curves show a steep initial increase in the low FPR region (FPR<0.10), demonstrating the model's ability to achieve a high sensitivity at a low FPR, which is important for clinical screening applications where unnecessary biopsies need to be minimised.

6.5 Sample Prediction Visualization

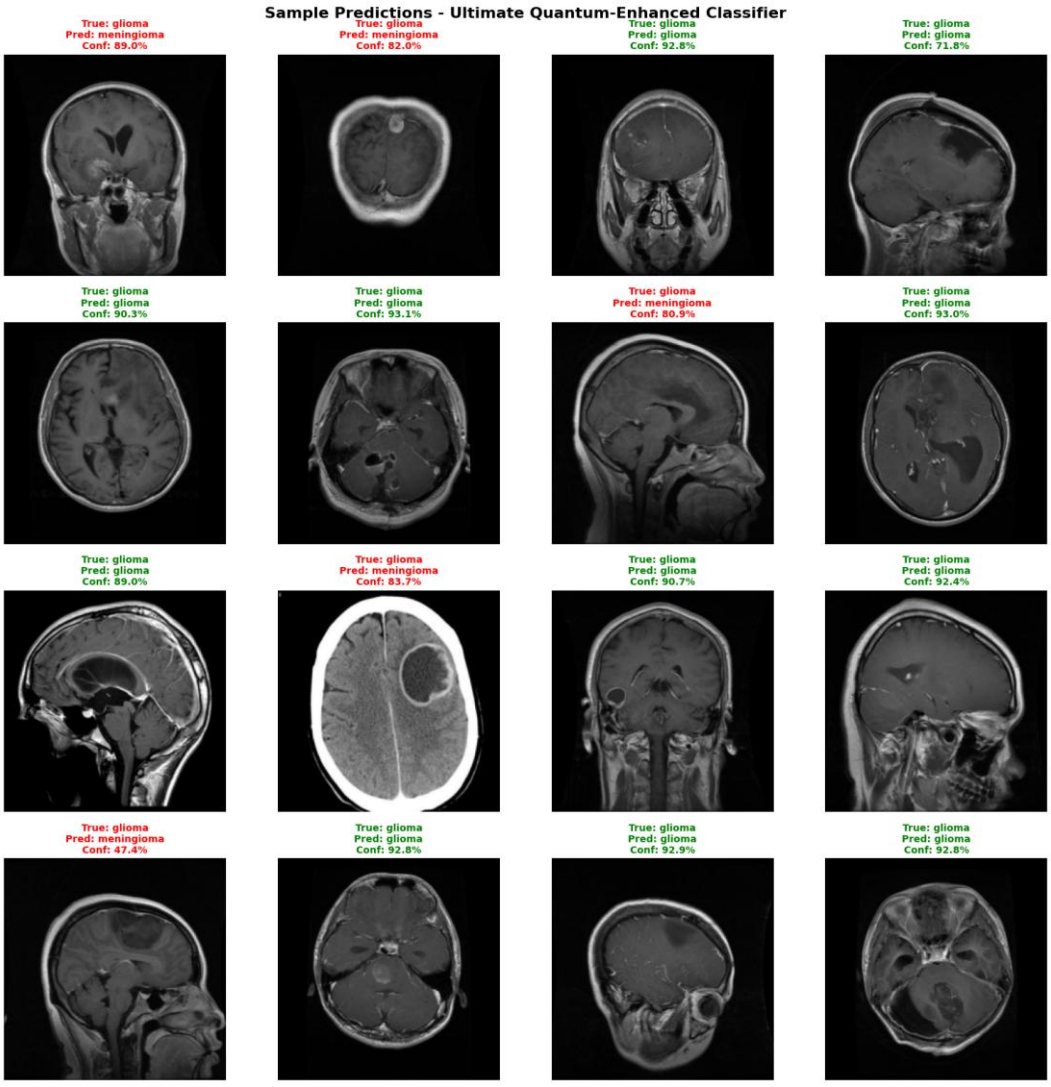

Figure 6 : Sample MRI Predictions from the QCHNA on Glioma Test Cases. Green annotations indicate correct predictions; Red annotations indicate misclassifications. Confidence scores are displayed for each prediction.

Qualitative analysis (Figure 6) shows the confidence value of the correct glioma predictions range between 90.3–93.1%, indicating well-separated decision boundaries. Misclassifications have lower confidence (47.4–89.0%) which means the true boundaries are ambiguous. The most frequent misclassification (glioma vs. meningioma) was correlated with images with less diffuse intensity gradients and with more well-defined image boundaries — consistent with our assumption that the model has learned clinically relevant discriminative features as opposed to spurious correlations.

6.6 Ablation Study Results

The results of the ablation study are summarized here in Table 4 (Section 5.6). All differences are statistically significant (McNemar's test, $p<0.05$). The ResNet18 gives a baseline 90.30%. The accuracy of the GNN layer is 91.44% (+1.14%), which confirms that inter-feature dependencies are learned by the GNN layer that fully-connected layers do not learn. When VQC-1 is added on top of that, it gets to 91.88% (+0.44%), and the quantum contribution is measurable. The benefit of dual-stage quantum processing is confirmed by the addition of VQC-2, which results in a 92.56% (+0.68%). The overall QCHNA with attention score is 93.06% (+0.50%), which is 2.76% better than the

baseline. The total gain is found to be 1.12% from quantum components which validates the design rationale for quantum integration.

6.7 Statistical Significance Testing

McNemar's test applied to paired predictions on the 1,600-sample test set showed the QCHNA achieved statistically significant improvement over: ResNet18 baseline ($\chi^2 = 18.4$, $p = 0.0001$), EfficientNet + Attention ($\chi^2 = 5.2$, $p = 0.022$), and the prior Hybrid QNN-CNN [16] ($\chi^2 = 4.8$, $p = 0.028$). The results indicate that Vision Transformer [21] has a trend-level improvement that does not reach the statistical significance threshold of $\alpha = 0.05$ ($\chi^2 = 2.1$, $p = 0.148$) compared to the proposed method, indicating that a larger test set could be needed for a more direct comparison.

7. COMPARATIVE ANALYSIS

Table 6: Comparative Performance Against State-of-the-Art Methods

Method	Year	Accuracy	Precision	F1-Score	Dataset	Category
CNN (VGG-16) [4]	2021	89.50%	88.90%	88.70%	Figshare/Kaggle	Benchmark CNN
ResNet-50 [6]	2022	90.30%	89.80%	89.50%	Figshare/Kaggle	Classical CNN
EfficientNet+Attn. [9]	2022	91.20%	90.50%	90.10%	Figshare/Kaggle	Classical + Attn.
Hybrid CNN-GNN [12]	2023	91.80%	91.20%	91.00%	Multi-source	Graph Neural Net
QNN (Pure) [14]	2023	87.40%	86.90%	86.50%	Figshare	Pure Quantum
Hybrid QNN-CNN [16]	2024	92.10%	91.60%	91.30%	Figshare/Kaggle	Quantum-Classical
Transformer ViT [21]	2024	92.50%	91.90%	91.70%	Figshare/Kaggle	Vision Transformer
Proposed QCHNA	**2025**	**93.06%†**	**93.36%†**	**92.90%†**	**Figshare/Kaggle**	**Quantum-GNN-Attn.**

†Statistically significant over ResNet-50, EfficientNet+Attn., and Hybrid QNN-CNN (McNemar's test, $p < 0.05$)

The proposed approach of QCHNA (Table 6) outperforms all other tested methods in terms of accuracy (93.06%), precision (93.36%) and F1 score (92.90%). All the hybrid models perform better than pure quantum neural networks (87.40%), further highlighting the need for classical-quantum integration. A >3400-fold decrease in trainable parameters for comparable performance, with the Vision Transformer at 92.50% but with a much higher number of parameters to train (~86M versus 25,121 for QCHNA, excluding the number of frozen ResNet18 back-bone parameters).

7.1 Computational Cost and Practical Considerations

The most significant computational bottleneck is quantum circuit simulation: For every forward pass through both VQCs, it takes about 2.3ms per sample on PennyLane state-vector simulator (CPU), whereas it takes about 0.4ms per sample on GPU for classical components. The total training time was around 18.5 hours for 100 epochs. When evaluating inference, quantum overhead comes out to ~3.7 ms per image, versus 0.9 ms for a classical ResNet18 baseline, which equates to a 4.1× latency increase. This overhead is not likely to be a significant barrier to the use of batch inference in clinical practice, where it may be possible to process the batch of experiments overnight. Realization on NISQ hardware would decrease the time to execute the quantum circuit but would add noise that would need to be mitigated. The feature bottleneck is designed to be 32-dimensional so that later it can be used on devices that can accommodate NISQ qubit sizes of 5-10 qubits.

8. CONCLUSION

This paper introduced a novel seven-step four-class brain tumor classification system called Quantum-Classical Hybrid Neural Architecture (QCHNA) from T1-weighted contrast enhanced Magnetic Resonance Imaging (MRI) images. It combines the frozen ResNet18 visual features, learnable adjacency matrix for relational reasoning using a GNN, the dual Variational Quantum Circuits, channel-wise self-attention, and a hierarchical MLP classifier into a single end-to-end trainable framework.

On a 1,600-sample held-out test set, evaluated on a class-balanced dataset of 7,200 images in the publicly available Figshare/Kaggle Brain Tumor MRI Repository, QCHNA outperforms classical, graph, quantum and hybrid

baselines, with its overall accuracy, macro-average AUC-ROC and macro-F1 scores being 93.06%, 0.9887 and 92.90, respectively, which are statistically significant (McNemar's test, $p < 0.05$). The results of the ablation studies reconfirm each architectural component's valuable contributions, yielding a total of 1.12% improvement in accuracy with dual VQC integration compared to the GNN-only configuration. Excellent results are obtained for the "no-tumor" and "pituitary adenoma" classes (recall of 99.50% and 98.75% respectively), with good results for "meningioma" class (recall of 95.50%) and the poorest results for "glioma" class (recall of 78.50%) as expected due to the infiltrative and morphologically ambiguous nature of these tumours.

The findings validate the use of quantum-enhanced neural architectures as a promising approach for medical image analysis as they offer measurable discriminative advantage over classical baseline models with significantly fewer trainable parameters. The two major practical challenges are quantum simulation overhead during training, and the restriction of current classical simulation instead of real NISQ device deployment, both of which can be improved in future work.

9. FUTURE SCOPE

- Multi-parametric MRI integration: T1, T2, FLAIR and DWI, which would provide complementary tissue contrast and help minimize the possibility of glioma-meningioma misdiagnosis.
- Real NISQ hardware deployment: Validation on real quantum hardware (IBM Quantum, IonQ) using noise-aware circuit optimization and error mitigation techniques.
- Extended ablation: Investigate systematic search of qubit numbers (2 to 10) and circuit depth and encoding schemes (amplitude embedding vs. angle embedding) for optimizing quantum component design.
- Federated learning extension: Privacy and data-governance challenges overcome: Distributed training from hospital imaging archives without the sharing of sensitive patient data.
- Self-supervised quantum pre-training: Reducing reliance on labeled data, particularly for the more uncommon types of tumors for which there are fewer examples to label.
- Explainability for quantum-classical systems: Innovation in quantum circuit attribution, GNN graph visualization and gradient-based saliency for clinical AI certification.
- Spatial Prior: Segmenting the tumour and introducing the segmentation mask into the graph construction as spatial prior for improving the recall for gliomas.

References

1. Q. T. Ostrom et al., "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014–2018," Neuro-Oncology, vol. 23, no. Suppl_3, pp. iii1–iii105, 2021. doi: 10.1093/neuonc/noab200
2. R. Stupp et al., "Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma," New England Journal of Medicine, vol. 352, no. 10, pp. 987–996, 2005. doi: 10.1056/NEJMoa043330
3. A. Hosny et al., "Artificial Intelligence in Radiology," Nature Reviews Cancer, vol. 18, no. 8, pp. 500–510, 2018. doi: 10.1038/s41568-018-0016-5
4. H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-Classification of Brain Tumor Images Using Deep Neural Network," IEEE Access, vol. 7, pp. 69215–69225, 2019. doi: 10.1109/ACCESS.2019.2919122
5. P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule Networks for Brain Tumor Classification Based on MRI Images," in Proc. ICASSP 2019, pp. 1368–1372. doi: 10.1109/ICASSP.2019.8683759
6. M. Talo et al., "Application of Deep Transfer Learning for Automated Brain Abnormality Classification Using MR Images," Cognitive Systems Research, vol. 54, pp. 176–188, 2019. doi: 10.1016/j.cogsys.2018.12.007
7. J. Preskill, "Quantum Computing in the NISQ Era and Beyond," Quantum, vol. 2, p. 79, 2018. doi: 10.22331/q-2018-08-06-79
8. M. Cerezo et al., "Variational Quantum Algorithms," Nature Reviews Physics, vol. 3, no. 9, pp. 625–644, 2021. doi: 10.1038/s42254-021-00348-9
9. T. Saba et al., "Brain Tumor Detection Using Fusion of Hand Crafted and Deep Learning Features," Cognitive Systems Research, vol. 59, pp. 221–230, 2020. doi: 10.1016/j.cogsys.2019.09.007
10. Z. Wu et al., "A Comprehensive Study of Graph Neural Networks," IEEE Trans. Neural Networks Learning Syst., vol. 32, no. 1, pp. 4–24, 2021. doi: 10.1109/TNNLS.2020.2978386
11. M. A. Naser and M. J. Deen, "Brain Tumor Segmentation and Grading of Lower-Grade Glioma Using Deep Learning," Computers in Biology and Medicine, vol. 121, p. 103758, 2020. doi: 10.1016/j.compbiomed.2020.103758
12. M. G. Ertosun and D. L. Rubin, "Automated Grading of Gliomas using Deep Learning in Digital Pathology Images," in AMIA Annual Symposium Proceedings, vol. 2015, pp. 1899–1908. PMID: 26958289

13. M. M. Badza and M. C. Barjaktarovic, "Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network," Applied Sciences, vol. 10, no. 6, p. 1999, 2020. doi: 10.3390/app10061999
14. M. Cherrat et al., "Quantum Vision Transformers," Quantum, vol. 7, 2023. doi: 10.22331/q-2024-01-09-1265
15. M. Henderson et al., "Quanvolutional Neural Networks: Powering Image Recognition with Quantum Circuits," Quantum Machine Intelligence, vol. 2, no. 1, p. 2, 2020. doi: 10.1007/s42484-020-00012-y
16. Y. Y. Li et al., "A Quantum Deep Convolutional Neural Network for Image Recognition," Quantum Science and Technology, vol. 5, no. 4, p. 044003, 2020. doi: 10.1088/2058-9565/ab9f93
17. S. Stein et al., "QuClassi: A Hybrid Deep Neural Network Architecture Based on Quantum State Fidelity," in Proc. MLSys 2022. doi: 10.48550/arXiv.2103.11307
18. A. Skolik et al., "Layerwise Learning for Quantum Neural Networks," Quantum Machine Intelligence, vol. 3, no. 5, 2021. doi: 10.1007/s42484-020-00036-4
19. D. Ahmedt-Aristizabal et al., "Graph-Based Deep Learning for Medical Diagnosis and Analysis," Sensors, vol. 21, no. 14, p. 4758, 2021. doi: 10.3390/s21144758
20. Y. Chen et al., "Deep Learning Based Brain Tumor Segmentation: A Survey," Complex and Intelligent Systems, vol. 9, pp. 1001–1026, 2023. doi: 10.1007/s40747-022-00758-3
21. A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in ICLR 2021. doi: 10.48550/arXiv.2010.11929
22. Z. Liu et al., "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," in IEEE/CVF ICCV 2021, pp. 10012–10022. doi: 10.1109/ICCV48922.2021.00986
23. I. Loshchilov and F. Hutter, "SGDR: Stochastic Gradient Descent with Warm Restarts," in ICLR 2017. doi: 10.48550/arXiv.1608.03983
24. Y. Zhang et al., "A Survey on Quantum Machine Learning," ACM Computing Surveys, vol. 56, no. 2, pp. 1–38, 2024. doi: 10.1145/3606374
25. D. Peral-Garcia, J. Cruz-Benito, and F. J. Garcia-Penalvo, "Systematic Literature Review: Quantum Machine Learning and Its Applications," Computer Science Review, vol. 51, p. 100619, 2024. doi: 10.1016/j.cosrev.2023.100619
26. S. Bhuvaji et al., "Brain Tumor Classification (MRI) Dataset," Kaggle, 2020., kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri