RGNN3D: A Hybrid Radiomic Graph Neural Network for 3D MRI Glioma Grading

Md. Aiyub Ali[a,b], Md Shakhawat Hossain[a,b], Taslima Ferdaus Shuva[b], Muhammad Ali Abdullah Almoyad[c], Nabil Anan Orka[a], Risala Tasin Khan[e], M. Shamim Kaiser[e], Md. Tanvir Rahman[a,d,*], Mohammad Ali Moni[a,b,f,g,h,**]

[a]*School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia, Brisbane, 4072, Queensland, Australia*
[b]*Department of Computer Science and Engineering, Daffodil International University, Daffodil Smart City, Savar, 1341, Dhaka, Bangladesh*
[c]*Department of Basic Medical Sciences, College of Applied Medical Sciences, King Khalid University, Guraiger, 62521, Abha, Saudi Arabia*
[d]*Department of Information and Communication Technology, Mawlana Bhashani Science and Technology University, Santosh, Tangail, 1902, Dhaka, Bangladesh*
[e]*Institute of Information Technology, Jahangirnagar University, Savar, 1342, Dhaka, Bangladesh*
[f]*School of Information Technology, Washington University of Science and Technology, Alexandria, 22314, Virginia, USA*
[g]*AI and Cyber Futures Institute, Charles Sturt University, Bathurst, 2795, New South Wales, Australia*
[h]*Rural Health Research Institute, Charles Sturt University, Orange, 2800, New South Wales, Australia*

Abstract

The diagnosis of glioma, a complex and often deadly brain tumor, involves extensive medical examinations. Still, accurately grading and classifying gliomas is difficult, as different areas within the same tumor can exhibit varying characteristics. The integration of radiomics, a clinically relevant feature extraction method, with machine learning (ML) is becoming increasingly popular in addressing this issue, but several research gaps persist. To this end, this study proposes a novel deep neural network, RGNN3D, that combines Graph Neural Networks with LSTM layers to precisely grade gliomas in 3D magnetic resonance imaging (MRI) data. To train our proposed model, we meticulously extracted 112 radiomic biomarkers. Utilizing the biomarkers, RGNN3D constructs a graph, channels essential information within layers, and preserves only pertinent information through its integrated memory cells. The proposed framework attained an accuracy of 98.58%, aligning with the performance of previous state-of-the-art architectures and surpassing prior radiomic-based ML models. We further employed an explainable AI approach (LIME) to highlight the most significant features, assisting radiologists in making more informed decisions. In short, RGNN3D offers a reliable and robust computer-aided solution for potential clinical application in the automated identification of gliomas.

Keywords: Radiomics, Biomarkers, Graph Neural Networks, Glioma Grading.

1. Introduction

Gliomas, which originate from glial cells, are the most prevalent primary intracranial tumors in adults [1]. Gliomas account for approximately 74.6% of all malignant brain tumors [2]. Patients with low-grade

*Corresponding author
**Corresponding author
Email addresses: aiyubali15-13456@diu.edu.bd (Md. Aiyub Ali), shakhawat15-14283@diu.edu.bd (Md Shakhawat Hossain), shuva.cse@diu.edu.bd (Taslima Ferdaus Shuva), maabdulllah@kku.edu.sa (Muhammad Ali Abdullah Almoyad), n.orka@uq.edu.au (Nabil Anan Orka), risala@juniv.edu (Risala Tasin Khan), mskaiser@juniv.edu (M. Shamim Kaiser), tanvirrahman@mbstu.ac.bd (Md. Tanvir Rahman), m.moni@uq.edu.au (Mohammad Ali Moni)

gliomas (LGG, grades I and II) generally have a survival rate averaging seven years [3]. In stark contrast, only 3–5% of patients with glioblastoma (GBM, grade IV) survive beyond five years, with a median survival time of approximately 12 months [4]. Intraoperative grading of gliomas, especially distinguishing between GBM and LGG, is crucial for making informed diagnostic decisions in clinical practice. To this end, magnetic resonance imaging (MRI) is an essential technique for screening, treatment planning, and assessing tumor response to therapy because it assists in analyzing the phenotypic and structural differences of gliomas [5]. In the same vein, radiomic features have shown a lot of promise for distinguishing between different grades of glioma. Radiomics is an emerging field that extracts high-throughput quantitative features from medical images, particularly MRI [6]. Radiomic analysis involves separating the tumor area from the rest of the image and extracting clinically useful information about its shape, appearance, size, intensity, location, and texture [7]. For example, high heterogeneity in intensity and texture features can sometimes indicate GBM [8, 9]. Although the combination of MRI and radiomic analysis offers vital clinical information regarding gliomas, identifying patterns among the hundreds of variables and understanding how each one affects glioma grading is challenging. To this end, machine learning (ML) algorithms emerge as a viable solution. For instance, earlier studies used various ML classifiers such as logistic regression (LR) [10], support vector machines (SVM) [11, 12], random forest (RF) [2, 13, 14], LASSO [15], and multi-layer perceptron [12, 16]. These models performed comparably to state-of-the-art convolutional neural networks, showing the promise of radiomic-based automated glioma grading. Han Li et al. [17] introduced a transfer learning-based optimizer (MSAS-DMOA) to improve adaptability in dynamic tasks. Peishu Wu et al. [18] proposed GLA-TD, a CNN-transformer model using attention and tensor decomposition for efficient medical image analysis. These works align with our RGNN3D approach in enhancing learning and interpretability. However, notable research gaps still persist. First of all, despite the availability of numerous types of radiomic features, prior research has focused on distinct radiomic feature types. For example, some studies used only texture-based features [19, 20, 21], while others relied on wavelet-based features [22, 11]. To the best of our knowledge, no studies holistically explored radiomics, i.e., modeling multiple classes of radiomic features simultaneously. Second, existing studies lack interpretability. While often accurate, black-box models do not foster trust between clinicians and these models because the variables critical to the decision-making process remain unclear. Finally, no framework has yet outperformed current state-of-the-art classifiers regardless of earlier studies' comparable efficacy. Given the aforementioned research gaps, we propose a novel framework, RGNN3D, which comprises graph convolutional networks (GCNs) and long short-term memory (LSTM) cells to form a hybrid classifier. The advantages of the proposed hybrid model are twofold. The primary benefit of GCNs, or graph neural networks in general, is the ability to embed complex relational data into a graph and pass only the relevant information to the subsequent layers [23]. In addition, LSTM layers use memory cells to retain pertinent input data and avoid the vanishing gradient problem [24]. We train our proposed model on 112 clinically significant radiomic biomarkers extracted from 3D MRI scans and later employ local interpretable model-agnostic explanations (LIME) [25]. LIME quantitatively explains which radiomic features have the most significant impact on grading LGG and GBM, ensuring utmost reliability. The key contributions of this study are summarized as follows:

- Introduced RGNN3D, a novel hybrid method that combines LSTMs and GCNs, leading to highly accurate glioma grading from complex, multimodal 3D MRI scans.
- Demonstrated that GCNs captured 112 significant radiomic biomarkers, while the LSTMs serve long-range dependencies and alleviate gradient vanishing by gated memory mechanisms.
- To enhance model transparency, we incorporate LIME-based interpretability analysis, which identifies and ranks the most influential features contributing to the grading outcomes.
- Identified the top 10 radiomic biomarkers that play a significant role in terms of glioma grading, which enriches the clinical trust.

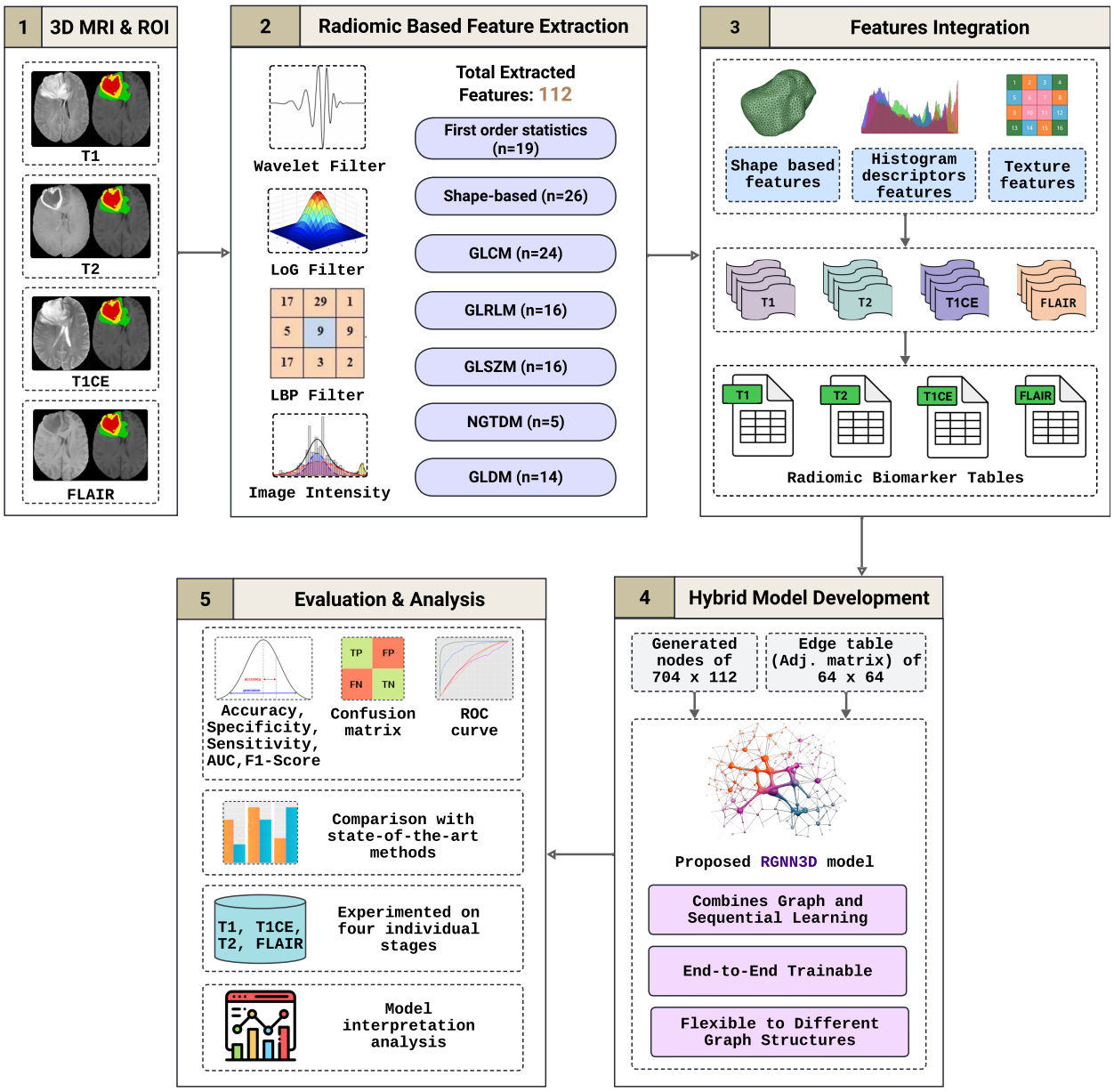

Figure 1: Comprehensive workflow of the proposed radiomic-based hybrid RGNN3D Model for glioma grading.

2. Materials and Methods

The overall methodology adopted in this study has five key stages as shown in Fig. 1: (i) 3D MRI & ROI involves acquiring 3D MRI scans and identifying regions of interest (ROI) for GBM and LGG, (ii) Radiomic Based Feature Extraction employs the PyRadiomics toolbox to extract features from the ROIs using various filters, (iii) Features Integration incorporates shape-based features, histogram descriptors, and texture features to create a comprehensive feature set, (iv) Hybrid Model Development proposes the RGNN3D model that combines graph neural networks and LSTM-based sequential learning, ensuring end-to-end trainability and interpretability, and (v) Evaluation & Analysis assesses the performance of the proposed framework with interpretability. In the following subsections, we delve into the specifics of every

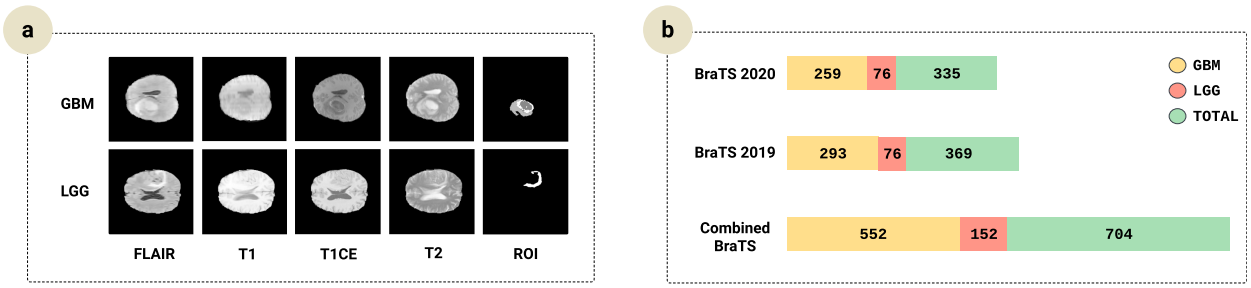

Figure 2: Dataset used in this study: (a) Sample images from each category; (b) Integration of BRATS 2019 and BRATS 2020

step.

2.1. Dataset Description

This study used the Brain Tumor Segmentation (BraTS) datasets, where multimodal 3D MRI scans of GBM and LGG are available. The BraTS'2019 [26] contains data of 369 subjects with GBM (n = 293) and LGG (n = 76). Besides, BraTS'2020 [27] comprises 335 subjects with GBM (n = 259) and LGG (76). We merged these two datasets, and the compilation incorporated a total of 704 subjects: GBM (n = 552) and LGG (n = 152), as shown in Fig. 2 (b). The merged dataset includes neuroimaging files with multiparametric 3D MRI scans, encompassing (a) T2 Fluid Attenuated Inversion Recovery (T2-FLAIR): suppresses cerebrospinal fluid signals, enhancing lesion visibility, (b) Native (T1): captures the brain's baseline anatomy, highlighting tissue density differences, (c) Post-contrast T1-weighted (T1CE): acquires T1-weighted images with a contrast agent, highlighting areas of increased vascularization and blood-brain barrier disruption to identify abnormal tissue, such as glioma lesions, (d) T2-weighted (T2): provides detailed information on edema, cysts, and tissue abnormalities, aiding in understanding glioma characteristics, and (e) ROI: allows feature extractions specifically from the tumor region or areas of interest that are relevant, precise, and clinically meaningful. Sample MRI scans from each category are also visualized in Fig. 2 (a).

2.2. Feature Extraction and Integration

Radiomics, a transformative approach in medical imaging, involves the extraction and analysis of clinically significant information embedded within medical images, transcending what is perceptible to the human eye [28]. To address this issue, we used the PyRadiomics [29] tool to identify relevant radiomic features in our study. Here, for each subject, we get 112 radiomic features corresponding to seven distinct types as described in Table 1: First Order, Shape-Based, Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM). The extracted features encompass a range of statistical metrics, shape-based attributes, and matrix-based analyses, offering insights into voxel intensity distribution, geometric properties, spatial relationships, and textural patterns within ROI. However, each type of these radiomic features contributes to unique and valuable information, and their integration ensures a holistic understanding of the clinically significant attributes. To achieve this, for each MRI stage (T1, T2, T1CE, FLAIR), we organize these extracted features into tables (Radiomic Biomarker Table: 704 x 112) where each row corresponds to a specific subject, and each column represents different radiomic features.

2.3. RGNN3D

We introduce a Radiomic Graph Neural Network (RGNN3D) architecture that integrates LSTM cells within its message-passing framework to enhance the learning of node embeddings as depicted in Fig. 3. This approach allows the model to capture both short-term and long-term dependencies inherent in graph structures. The design of the RGNN3D model consists of several layers, including an LSTM layer, GCN layers with appropriate message passing among each block, and a dense output layer. The input layer receives 112 features and constructs a graph. Here, the node feature vector and the graph's adjacency matrix are

Table 1: Types of Radiomic Features Extracted in Our Study.

Feature Type	Description
First Order	Captures fundamental statistical metrics, offering voxel intensity distribution in ROIs.
Shape-based	Encompasses shape-related attributes in 3D and 2D, which understand the brain's geometric properties.
GLCM	Reveals voxel intensity spatial relationship patterns.
GLRLM	Quantifies consecutive voxel lengths with identical intensity values.
GLSZM	Offers size and spatial distribution of homogeneous intensity regions.
NGTDM	Highlights tone differences between neighboring voxels and textural.
GLDM	Analyzes voxel pair dependence.

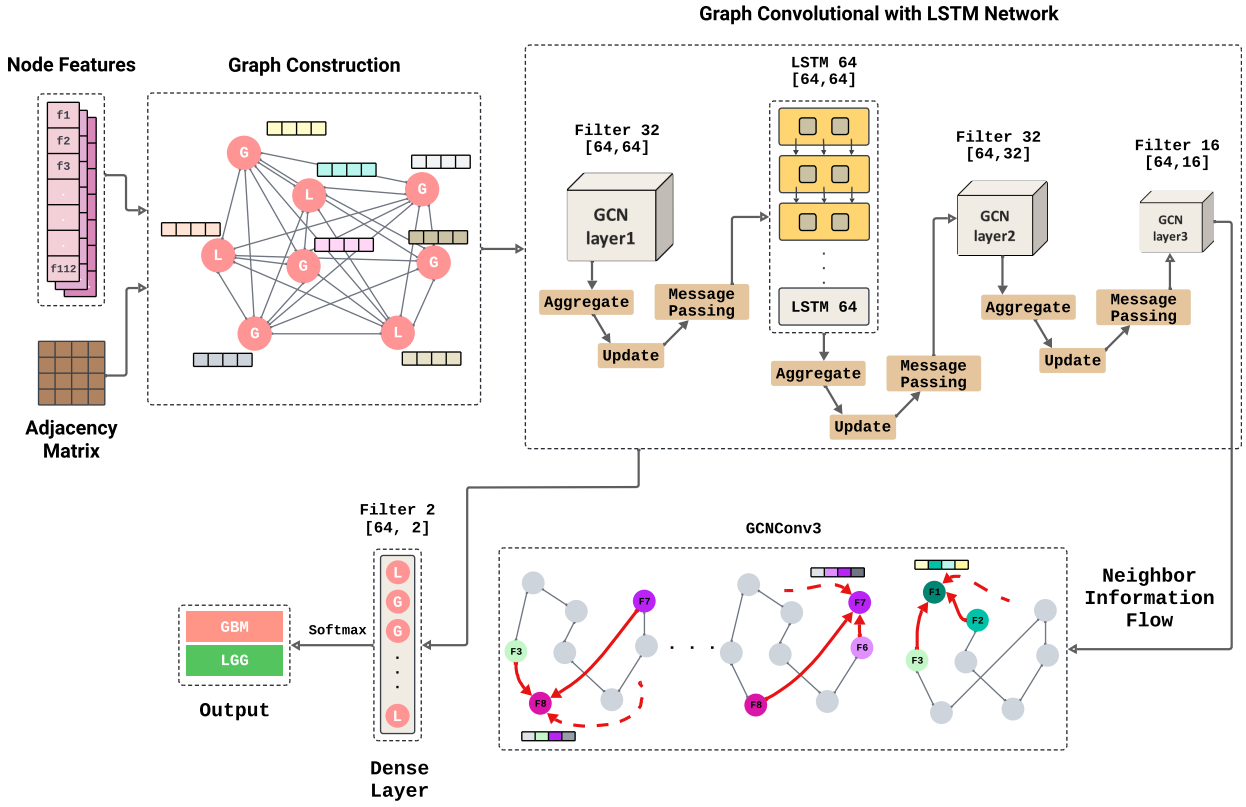

Figure 3: Proposed RGNN3D model architecture.

initially fed into the Feed Forward Network (FFN) [30] blocks, followed by passing these to the GCN layers after processing. This architecture includes three GCN layers, each employing the ReLU activation function within a message-passing framework. The embeddings updated by the LSTM cell and GCN layers are then channeled back into the FFN blocks, producing the final node embeddings as logit values. Finally, a Softmax activation function is applied to generate a probability distribution for the potential node labels. The inclusion of LSTM enables gradient flow across time steps due to its gating mechanisms (input, forget, and output gates), which retain essential information and reduce vanishing gradient effects. This makes it suitable for preserving long-range dependencies and refining spatial-temporal feature representations learned from the graph structure.

2.4. Graph Representation and Initial Node Embeddings

We generate a graph with 704 rows and 112 columns using the radiomic biomarker table, representing image features and the target class. Each patient is considered a single node in the graph and the corresponding radiomic feature vector serves as its initial embedding. In this regard, $G = (V, E)$ is a graph, where V is the set of patient nodes and E represents the set of edges. The node feature matrix $\mathbf{X} \in \mathbb{R}^{704\times112}$, that encodes the radiomic features for all patients.

To avoid an artificial dependency from different patients, we employ an identity adjacency matrix $\mathbf{A} = \mathbf{I}_{704}$. Because we decided to connect by self-loops. In this matter, each node maintains independence during message passing, where there is no possibillity of information leakage from other patients. This clinical practice especially relevant in the healthcare domain, where patient to patient connections are not always meaningful. We ensure a simple graph structure but clinically efficient. This approach confirms that node embeddings are rely on radiomic features of each patient's record.

Theoretically, initial representation of a node $v \in V$ is given by:

$$h_v^{(0)} = \mathbf{x}_v \in \mathbb{R}^d,$$

where $d = 112$ is the dimensionality of the feature vector. After that, these embeddings pass through a series of graph convolution and recurrent (LSTM gates) operations. That enables the model to capture higher level abstractions from radiomic features.

For better computational efficiency and scalability, we generate a mini batch (each batch is 64 training sizes). In this regard, each batch feeds to the model during iteration. For a batch size B, we construct the corresponding subgraph with a feature matrix $\mathbf{X}_b \in \mathbb{R}^{B\times112}$ and adjacency matrix $\mathbf{A}_b = \mathbf{I}_B$. This approach ensures the message passing within a batch. It is also prevents corss batch dependencies during training.

While we use an identity adjacency matrix in this study, the framework is flexible and robust. That can extend to more complex graphs such as kNN based graphs or fully connected graphs but they did not offer reliable performance. By using a simple identity graph with LSTM cells, the approach is truly robust to the dataset. It is also adaptable to broader clinical applications. Overall, it supports the clinical generalisability.

2.4.1. Graph Convolutions with LSTM Network

One of the core innovations of the RGNN3D model lies in its use of LSTM cells for aggregating and updating node embeddings inside GCN blocks. The underlying message-passing process involves aggregating node features from its neighbors and updating the node's embedding using LSTM cells, followed by normalization and the final nonlinear embeddings. These steps are described as follows:

(a) *Aggregation of Neighboring Node Features:* For a node u at iteration i, the aggregated message $\mathbf{m}_u^{(i)}$ from its neighbors $\mathcal{N}(u)$ is computed as:

$$\mathbf{m}_u^{(i)} = \sum_{v\in\mathcal{N}(u)} \mathbf{W}_{\text{Nei}}^{(i)} \mathbf{h}_v^{(i-1)} \tag{1}$$

where $\mathbf{W}_{\text{Nei}}^{(i)}$ is a learnable weight matrix.

(b) *LSTM-based Node Updating:* The updated embedding $\mathbf{h}_u^{(i)}$ for node u is obtained using an LSTM cell [31] that takes the node's previous embedding $\mathbf{h}_u^{(i-1)}$ and the aggregated message $\mathbf{m}_u^{(i)}$ as inputs:

$$\mathbf{h}_u^{(i)} = \text{LSTM}(\mathbf{h}_u^{(i-1)}, \mathbf{m}_u^{(i)}) \tag{2}$$

The internal operations of the LSTM cell are defined by the following equations:

$$\mathbf{i}_u^{(i)} = \sigma(\mathbf{W}_i^{(i)}\mathbf{h}_u^{(i-1)} + \mathbf{U}_i^{(i)}\mathbf{m}_u^{(i)} + \mathbf{b}_i^{(i)}) \tag{3}$$

$$\mathbf{f}_u^{(i)} = \sigma(\mathbf{W}_f^{(i)}\mathbf{h}_u^{(i-1)} + \mathbf{U}_f^{(i)}\mathbf{m}_u^{(i)} + \mathbf{b}_f^{(i)}) \tag{4}$$

$$\mathbf{o}_u^{(i)} = \sigma(\mathbf{W}_o^{(i)}\mathbf{h}_u^{(i-1)} + \mathbf{U}_o^{(i)}\mathbf{m}_u^{(i)} + \mathbf{b}_o^{(i)}) \tag{5}$$

$$\begin{aligned}\mathbf{c}_u^{(i)} = \mathbf{f}_u^{(i)} \odot \mathbf{c}_u^{(i-1)} + \mathbf{i}_u^{(i)} \odot \tanh(\mathbf{W}_c^{(i)}\mathbf{h}_u^{(i-1)} + \\ \mathbf{U}_c^{(i)}\mathbf{m}_u^{(i)} + \mathbf{b}_c^{(i)})\end{aligned} \tag{6}$$

$$\mathbf{h}_u^{(i)} = \mathbf{o}_u^{(i)} \odot \tanh(\mathbf{c}_u^{(i)}) \tag{7}$$

Here, $\mathbf{i}_u^{(i)}$, $\mathbf{f}_u^{(i)}$, $\mathbf{o}_u^{(i)}$ are the input, forget, and output gates, respectively, and $\mathbf{c}_u^{(i)}$ is the cell state. The parameters $\mathbf{W}_i^{(i)}$, $\mathbf{U}_i^{(i)}$, $\mathbf{b}_i^{(i)}$, etc., are the learnable weights and biases of the LSTM cell at iteration i.

(c) *Normalization of Node Embeddings:* After several iterations of message passing, the node embeddings are normalized to improve their representational power:

$$\mathbf{h}_u^{\text{norm}} = \frac{\mathbf{h}_u^{(i)}}{\|\mathbf{h}_u^{(i)}\|} \tag{8}$$

where $\|\mathbf{h}_u^{(i)}\|$ denotes the Euclidean norm of the vector $\mathbf{h}_u^{(i)}$.

(d) *Final Node Embeddings via GCN Layer:* The normalized embeddings are then passed through a Graph Convolutional Network (GCN) [32] layer to obtain the final nonlinear node embeddings:

$$\mathbf{H}^{(i+1)} = \sigma(\mathbf{A}\mathbf{H}^{(i)}\mathbf{W}_{\text{GCN}}) \tag{9}$$

where $\mathbf{A}$ is the adjacency matrix, where ones are on the diagonal and zeros elsewhere. This approach ensures that each node is only connected to itself, focusing on the node-specific features without considering inter-node dependencies. Here, $\mathbf{H}^{(i)}$ is the matrix of node embeddings at iteration i, $\mathbf{W}_{\text{GCN}}$ is the weight matrix of the GCN layer, and σ is a nonlinear activation function.

2.4.2. Output Layer and Grading

The final embeddings $\mathbf{H}^{(i+1)}$ from the GCN layer are passed through a dense output layer to predict the node labels. The dense layer applies a linear transformation followed by a nonlinear activation function of Softmax to generate a probability distribution for each node's label. The output for a node u is computed as follows:

$$\mathbf{y}_u = \sigma(\mathbf{W}_{\text{out}}\mathbf{h}_u^{\text{norm}} + \mathbf{b}_{\text{out}}) \tag{10}$$

Here, $\mathbf{W}_{\text{out}}$ is the weight matrix of the output layer, $\mathbf{b}_{\text{out}}$ is the bias term, and σ denotes the Softmax activation function, which maps the output to each probability value representing the grading of glioma.

2.5. Implementation

The implementation of the RGNN3D model commenced with the combined BraTS dataset comprising 704 records, which was further divided into training and testing sections in an 80:20 ratio. The model was optimized using the Adam optimizer, with a batch size of 64, and trained over 50 epochs. The Glorot Uniform initializer was employed for the kernel initialization. Sparse Categorical Cross-Entropy was utilized for the loss function. The implementation was executed on a system equipped with an RTX 3060 GPU and 32GB of RAM, running the NVIDIA driver version 535.104.05 with CUDA version 12.2.

3. Results and Discussion

First, we start with the performance analysis across all four 3D MRI stages (T1, T1CE, T2, and FLAIR) using standard classification performance metrics [33] i.e., precision, recall, F1-score, and accuracy. From the confusion matrices as shown in Fig. 4 we observe: (i) for classifying LGG, the proposed model acquired identical performance across all stages with 28 accurate cases and only two misclassifications, (ii) for GBM, we get no false negative result for both T1 and T2 stages. However, there were only 2 and 6 misclassifications out of 141 cases with T1CE and FLAIR stages, respectively. from the confusion matrices, it is clear that GBM is more consistently detected compared to LGG, which demonstrated strong power for aggressive gliomas.

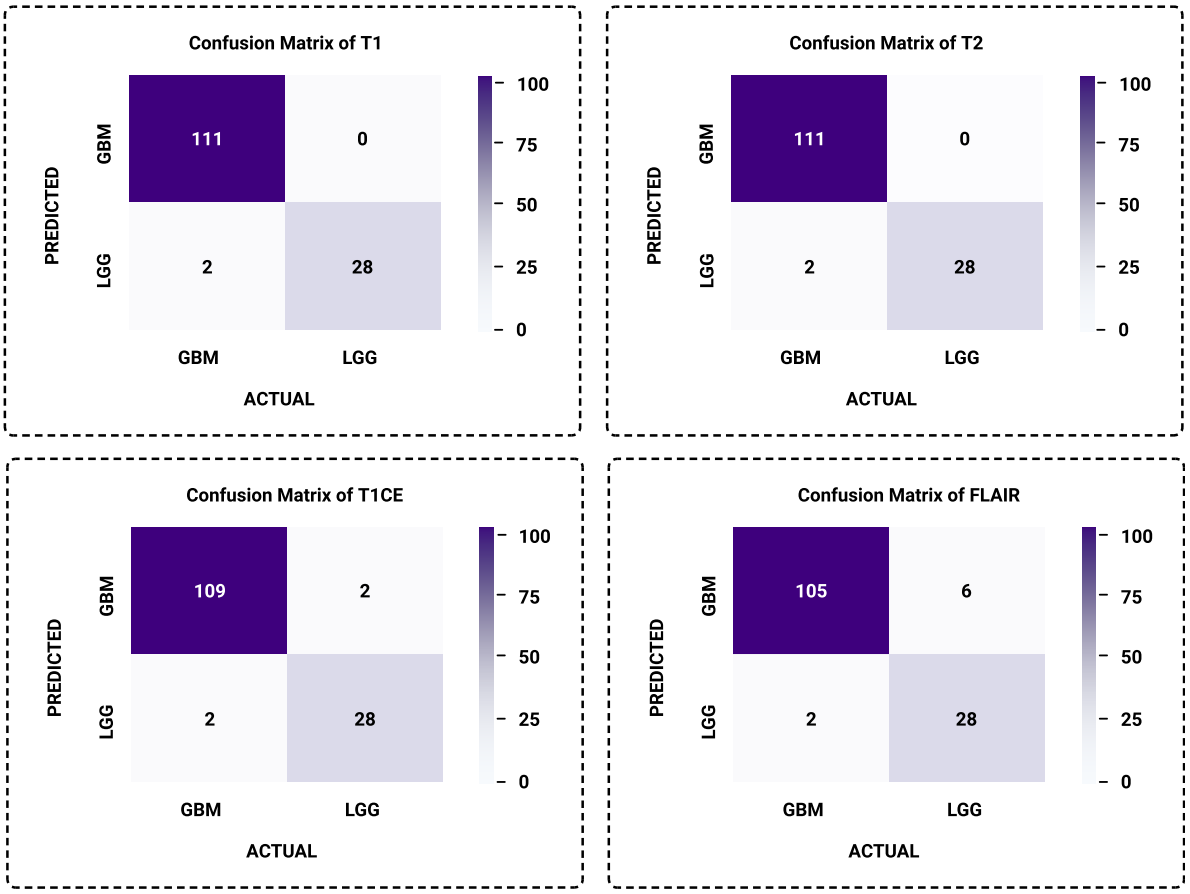

Figure 4: Confusion Matrices of RGNN3D Model for MRI Stages T1, T2, T1CE, and FLAIR

Using these matrices, we calculate different scores and discover a similar trend for each evaluation criterion as depicted in Fig. 5. The proposed RGNN3D system achieved outstanding accuracy scores of 98.58% for T1 and T2, 97.16% for T1CE, and 94.33% for FLAIR. The F1-scores also reinforce the system's robustness, maintaining values of 99.10% for T1 and T2, 98.20% for T1CE, and 96.33% for FLAIR. Furthermore, the precision values reached 100% for F1 and F2, which are able to completely avoid false-positive decisions for glioma grading. Even for T1CE and FLAIR modalities, precision scores remain high at 98.20% and 94.59% respectively. Moreover, our system shows strong recall: 98.23% (T1), 98.23% (T2), 98.20% (T1CE), and 98.13% (FLAIR). Overall, we notice the best and identical performance of the system with both T1 and T2 (accuracy: 98.58%, precision: 100%, f1-score: 99.10%, and recall: 98.23%) for each category.

Besides, the accuracy and loss curves (outlined in Fig. 6) illustrate the performance of the proposed model over 50 epochs. Here, we notice that the training accuracy reaches 100.00%, and we get the validation accuracy to be 98.58%. We noted that the training loss decreases smoothly for 4 different MRI stages, but

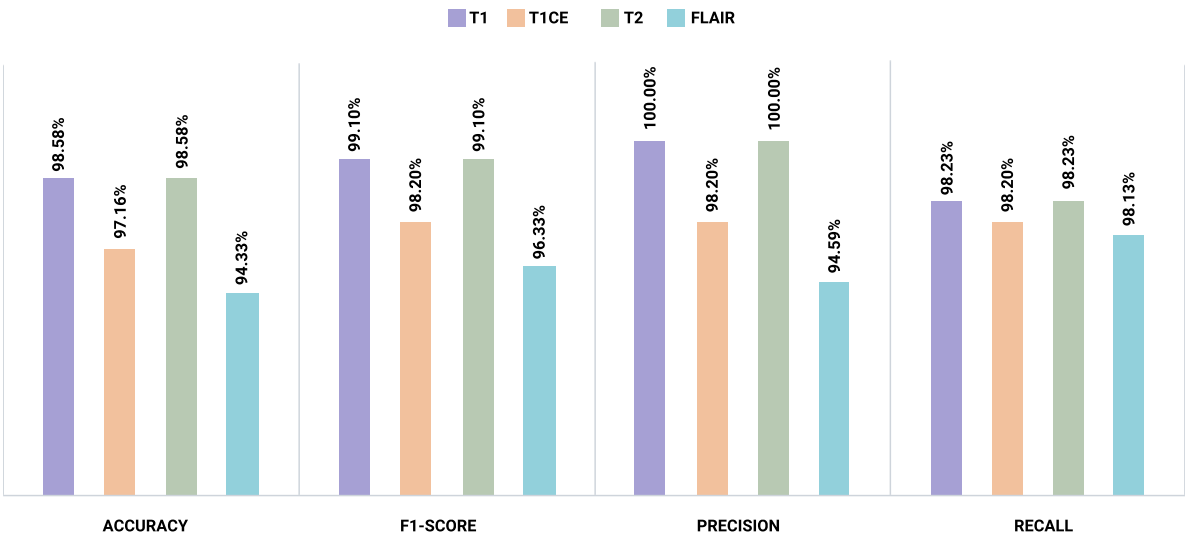

Figure 5: Comparison of Scores of RGNN3D Model for 4 MRI Stages

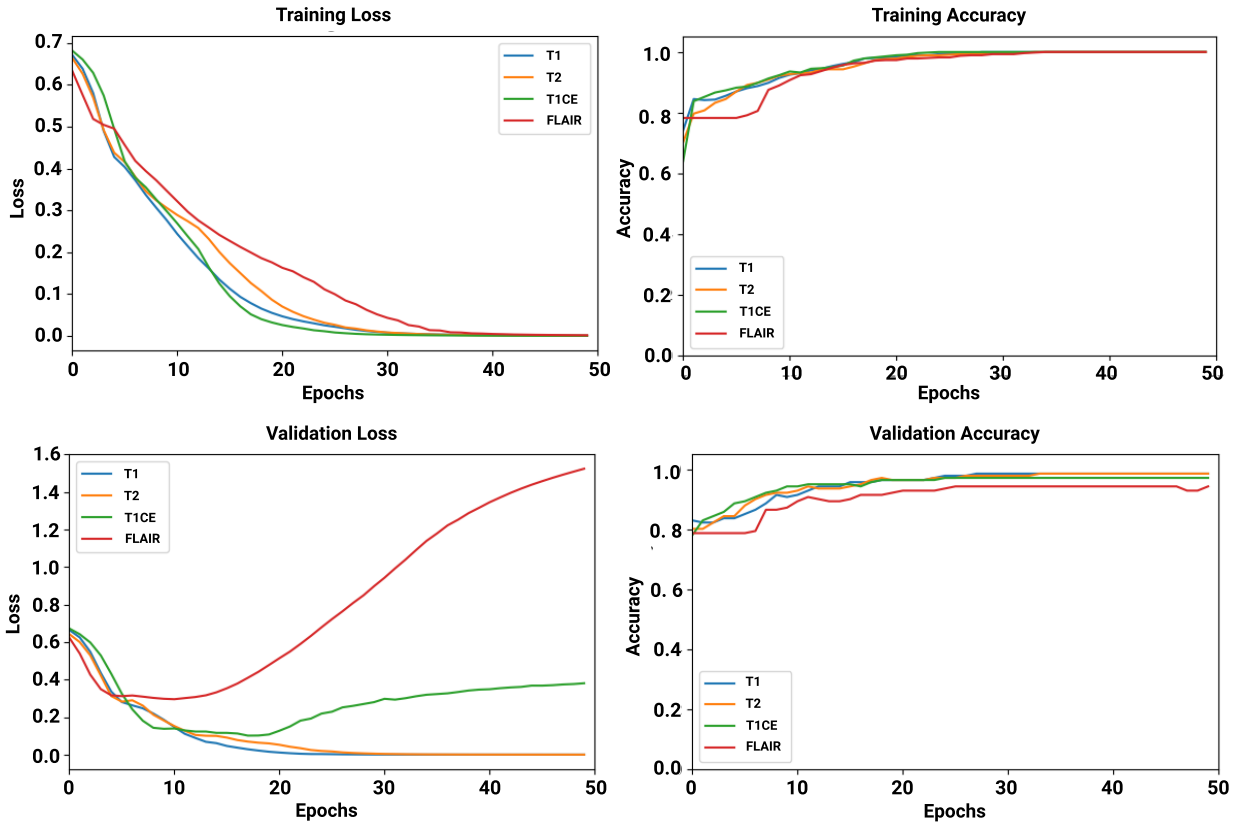

Figure 6: Accuracy and Loss Curves of RGNN3D Model for 4 MRI Stages

the validation losss shows differently. Here, for T1 and T2, the validation loss follows training loss closely, which indicates stable learning in our proposed RGNN3D method. In T1CE, the validation loss shows

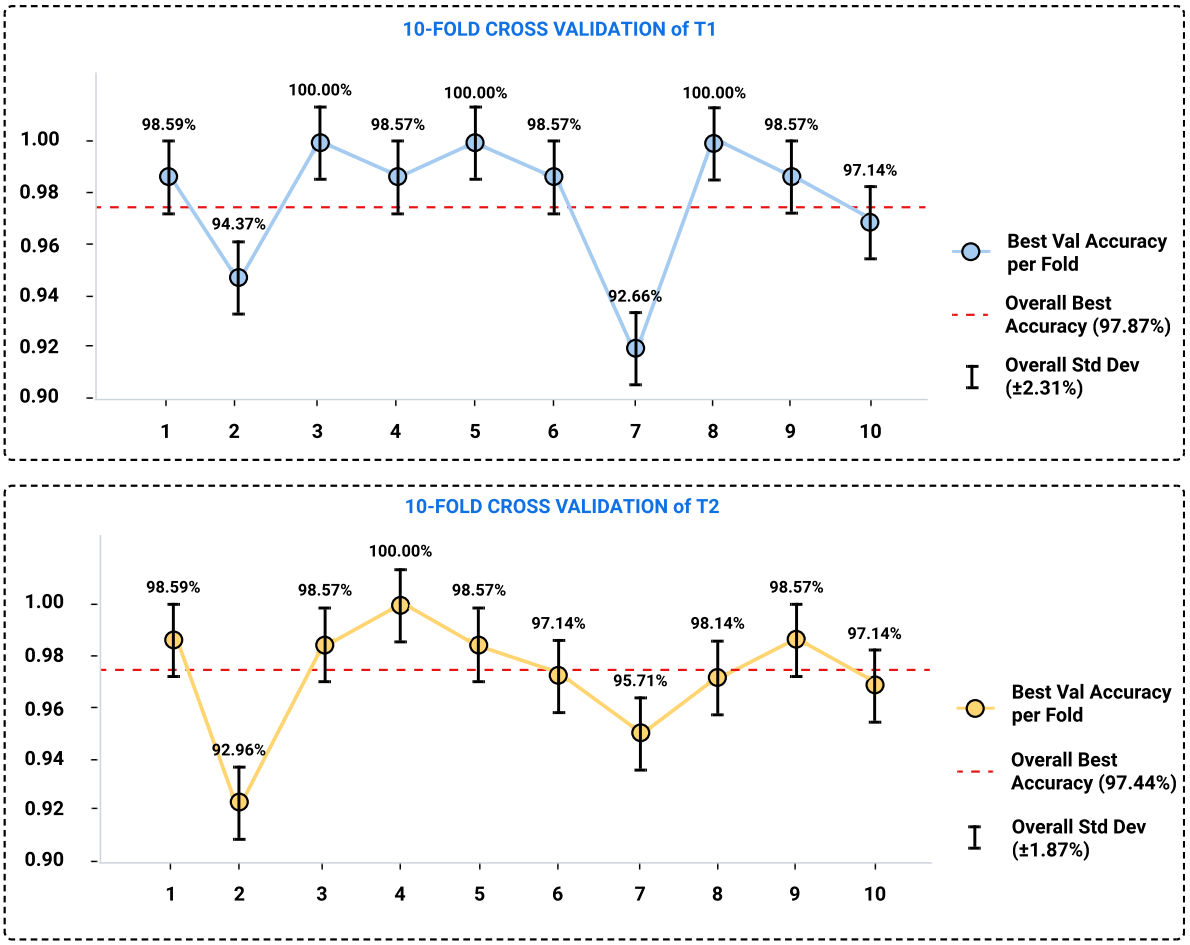

Figure 7: Cross-Validation Performance of Glioma Grading with T1 and T2 MRI Modalities

a bit higher, but it has stability. For the FLAIR stage, it significantly rises. This is why, the FLAIR modality achieved the lowest accuracy compared to other modalities. Overall, the training loss curve shows a consistent decrease compared to the validation loss trend, where we only observe consistency with T1 and T2. This phenomenon also supports its effectiveness with our architecture.

We extend our investigation into contemporary glioma grading systems that are not explicitly focused on radiomics features for 10-fold cross-validation. Among all the configurations tested, T1 and T2 were chosen since they attained the highest overall accuracies. Figure 7 shows T1's (upper part) validation accuracies (x-axis: fold numbers 1-10, y-axis: accuracy %), ranging from 92.66% (fold 7) to 100% (folds 3, 5, 8). The overall average accuracy across all folds was 97.87% with a standard deviation of ± 2.31%, indicating strong performance with some variation across folds. Figure 2 (lower part) shows T2's accuracies, which ranged from 92.96% (fold 2) to 100% (fold 3), with an average accuracy of 97.44% ± 1.87%, reflecting slightly lower but more consistent results compared to T1. Furthermore, the ROC curves concerning MRI modality are displayed in Fig. 8 to confirm the proposed RGNN3D model's strength in identifying glioma grades. Remarkably, both GBM and LGG perform exceptionally prominent; their ROC curves for T1 and T2 approach the optimal top-left corner, which is consistent with their high accuracy and the prior described F1-scores. Conversely, FLAIR demonstrates a marginally lower AUC, confirming its comparatively poorer performance. Taken together, our proposed architecture demonstrates better performance with T1 and T2 stages compared to T1CE and FLAIR while grading glioma tumors.

Subsequently, to illustrate the interpretability of our proposed architecture, we implement LIME [34].

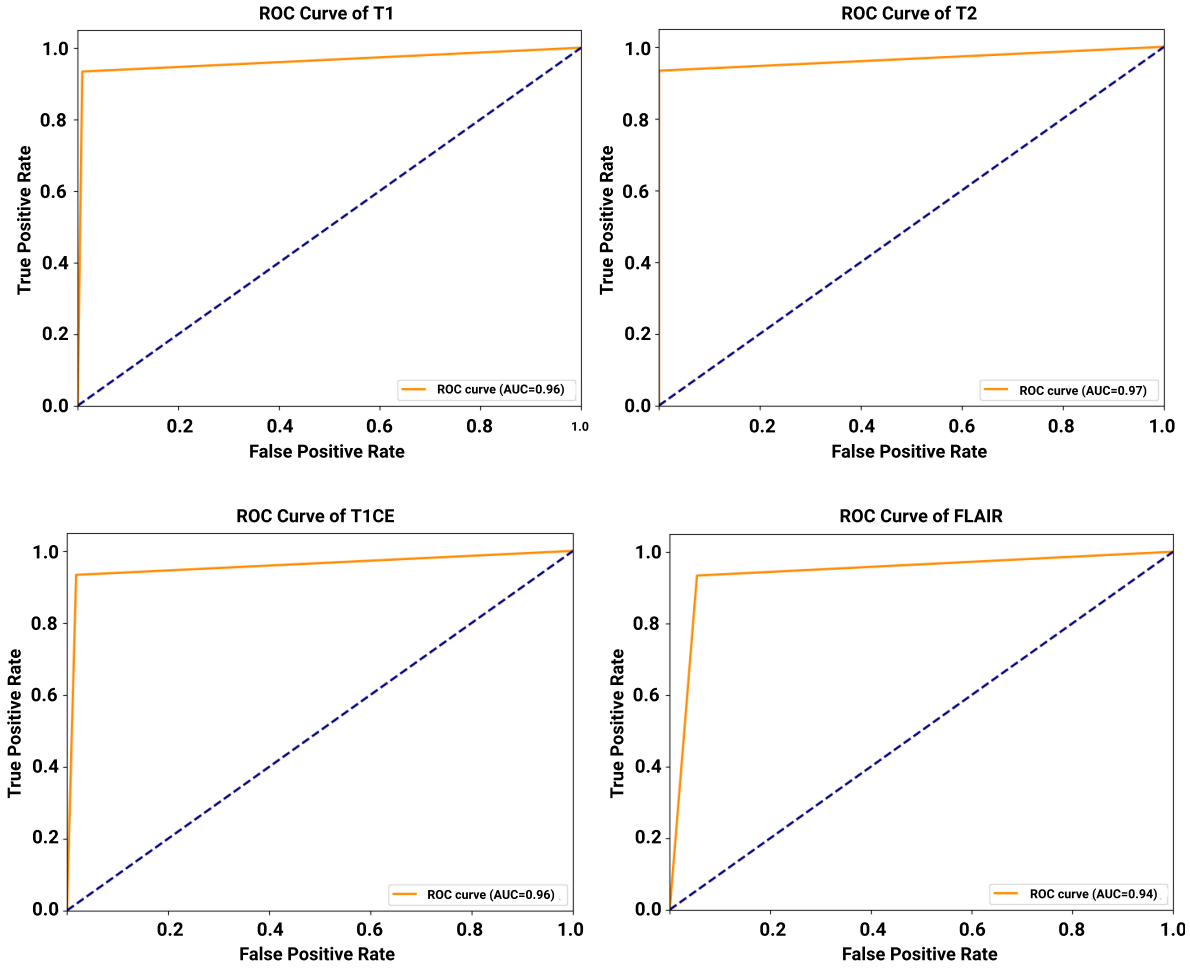

Figure 8: Performance Analysis of RGNN3D Model for 4 MRI Stages in ROC Curves

As shown in Fig. 9, the prioritization of specific radiomic features in LIME-based interpretations is driven by contribution to model performance where all radiomic features are clinically relevant. Still, the top three most essential features make a significant contribution to the grading of glioma in these specific data points. For GBM, the features "Mesh Volume", "Maximum 2D diameter", and "Maximum 3D diameter" are crucial factors among the 112 features, underscoring their significant roles in distinguishing glioma. These features provide critical information about the tumor's size, shape, and spatial dimensions. For LGG, the features "Minor Axis Length", "Maximum 2D diameter", and "Gray Level Non-Uniformity" are the top features, emphasizing their interpretability and actionable insights in clinical decision-making. The shape and texture features captured by these metrics are vital in differentiating between glioma grades and understanding the heterogeneity within the tumor. While the remaining features also contribute valuable information, they are comparatively less significant for several reasons, such as redundancy, specificity, statistical significance, clinical validation, and model performance.

At this point, we concentrate on the previous studies utilizing radiomic features for glioma tumor grading. From the current literature, we observe that most of the related studies focused on statistical ML models such as LR [10], SVM [11, 12], and RF [2, 13, 14]. Consequently, we aim to evaluate the performance of statistical ML models on our dataset. Consequently, we implement five ML models (K-NN, LR, DT, RF, and SVM) with the acquired radiomic features using our dataset. Here, we use T1 stage MRI images as we got good results during our prior analysis. Table 2 presents a comparative summary of the above-mentioned

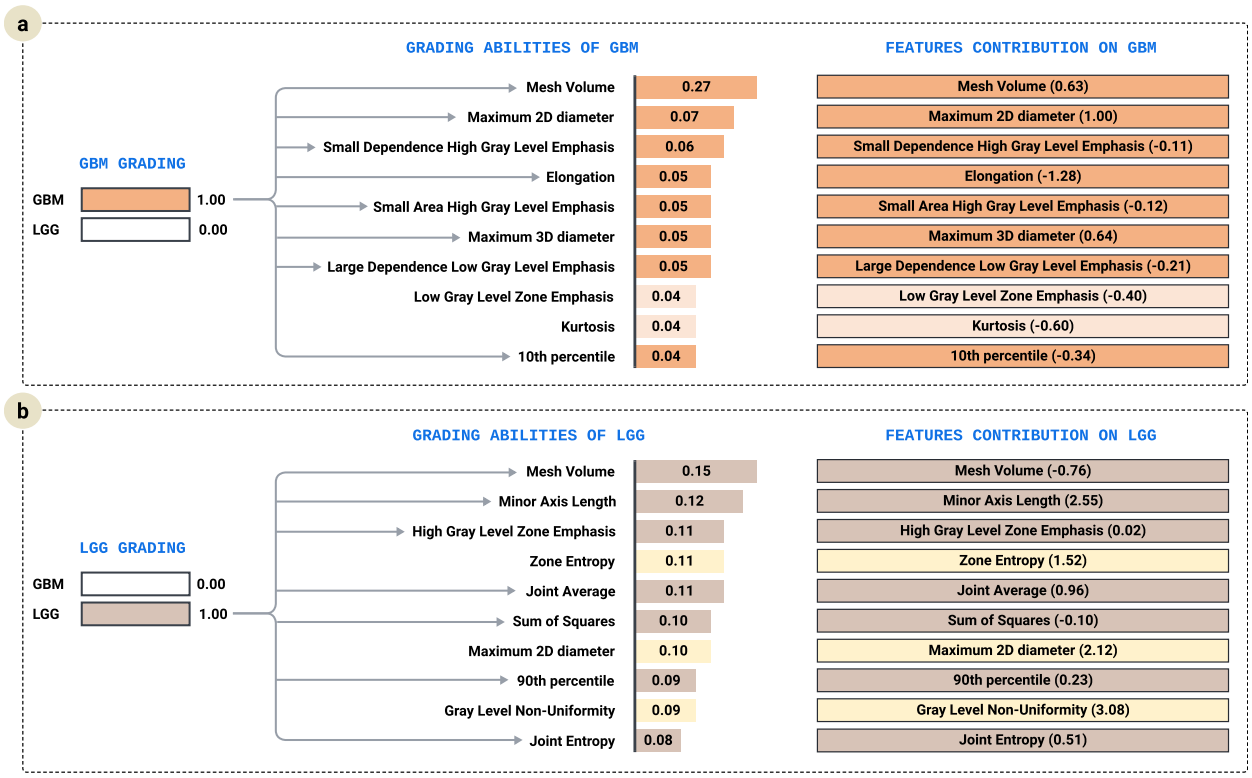

Figure 9: Analysis of Interpretability and Contribution of Radiomic Biomarkers for Glioma Grading using LIME: a) Explainability on grading GBM, b) Explainability on grading LGG

Table 2: Comparative Performance Analysis of Conventional ML Models with RGNN3D for Glioma Grading.

Model	Precision	Recall	Specificity	F1-Score	AUC	Accuracy
K-NN	96.30%	86.67%	99.10%	91.23%	92.88%	95.72%
LR	96.00%	80.00%	99.10%	87.27%	89.55%	92.69%
DT	90.32%	93.33%	97.30%	91.80%	95.32%	95.55%
RF	100.00%	93.33%	100.00%	96.55%	96.67%	96.79%
SVM	89.66%	86.67%	97.30%	88.14%	96.67%	92.87%
RGNN3D	**100.00%**	**98.23%**	**100.00%**	**98.10%**	**97.00%**	**98.58%**

ML models and RGNN3D. After a thorough investigation, we noticed that RGNN3D outperformed all other ML models across nearly all metrics, demonstrating superior performance in glioma grading. These results indicate the efficacy of our proposed model in accurately grading glioma tumors compared to other ML models.

As evidenced in Table 3, the proposed system achieves a testing accuracy of 98.58% on the BraTS 2019-20 dataset, ensuring head-to-head competency. For image-based data-learning tasks, deep convolutional neural networks remain the gold standard. Still, our proposed neural network, trained on only radiomic features, showcases comparable performance, narrowly missing out on the highest accuracy by less than 0.5%. In fact, RGNN3D outperforms the only other graph-based model in the table by around 5% in accuracy. The comprehensive comparative analysis establishes RGNN3D as a reliable, accurate, and interpretable alternative to the existing state-of-the-art.

3.1. Ablation Study

To validate the design of our RGNN3D model, we performed an ablation study that focuses on optimizers, activation functions, and GCN block configurations (model depth). As shown in Table 4, the Adam optimizer

Table 3: Comparative Analysis of State-of-the-Art Glioma Grading Methods with the Proposed RGNN3D Architecture.

Architecture	Dataset	Result
MM-XGB, 2023 [35]	BraTS 2020	93.00%
SASG-GCN, 2023 [36]	TCGA-LGG	93.62%
SGD with ADASYN, 2023 [37]	BraTS 2020	96.00%
CNN Based, 2023 [38]	BraTS 2017-19	97.85%
MMD-VAE, 2022 [39]	BraTS 2019	98.46%
TEWMA-CNN, 2022 [40]	BraTS 2015,21	98.76%
TD-CNN-LSTM, 2022 [41]	BraTS 2019-21	98.90%
Proposed RGNN3D, 2025	**BraTS 2019-20**	**98.58%**

not only achieved the best accuracy, but it also ensured the fastest operation (33s x 50). It proves the effectiveness of performance and efficiency. While Admax and Nadam achieved competitive accuracies of 95.03% and 97.87% respectively, they took longer computational times (52–83s for every epoch). In contrast, SGD optimizer is both slower and less accurate (39.83%) compared to others. For activation functions, ReLU

Table 4: Ablation Study: Impact of Optimizers, Activation Functions, and GCN Block Configurations on RGNN3D Performance.

No	Variant	Training Time × Epoch	Test Accuracy	Findings
		Optimizers		
1	Adam	33 s × 50	**98.58%**	Best accuracy
2	Adamax	83 s × 50	95.03%	Good accuracy
3	Nadam	52 s × 50	97.87%	Good accuracy
4	RMSprop	59 s × 50	97.16%	Good accuracy
5	SGD	62 s × 50	39.83%	Poor accuracy
		Activation Functions		
6	Sigmoid	64 s × 50	95.03%	Good accuracy
7	Elu	67 s × 50	97.87%	Good accuracy
8	ReLU	33 s × 50	**98.58%**	Best accuracy
9	Tanh	60 s × 50	97.16%	Good accuracy
10	Leaky ReLU	62 s × 200	97.16%	Good accuracy
		Model Blocks		
11	1 GCN block	42 s × 50	93.17%	Poor accuracy
12	2 GCN blocks	50 s × 50	97.87%	Good accuracy
13	3 GCN blocks	62 s × 50	97.16%	Good accuracy
14	RGNN3D (3 GCN + LSTM)	33 s × 50	**98.58%**	Best accuracy
		LSTM Cell Sizes		
15	LSTM (8 units)	59 s × 50	97.87%	Good accuracy
16	LSTM (16 units)	30 s × 50	97.16%	Good accuracy
17	LSTM (32 units)	33 s × 50	**98.58%**	Best accuracy
18	LSTM (64 units)	63 s × 50	98.58%	Best accuracy but time consuming

activation achieves 98.58% test accuracy with the lowest training time (33s in every single epoch). Although Elu, Tanh, and Leaky ReLU achieved good accuracy, Sigmoid has shown poor performance.

Regarding model architecture, a single GCN block is not good (93.17%) in terms of accuracy but it consumes reasonable training time (42 s per epoch). Afterward, we added another GCN block to our model. By doing it, performance was improved considerably (97.87%), while it took more training time, around 50 s in every epoch. Then, we integrated another GCN block into our method, but the performance fell slightly and the time also increased. To decrease the computation power and increase robustness, we integrated the LSTM layer into our model. The proposed RGNN3D model (3 GCN + LSTM) not only achieved the highest accuracy (98.58%) but also trained efficiently (33s × 50 epochs), which offered the benefit of combining spatial and temporal learning. It is clear that LSTM enhances temporal learning with the highest performance, and it also maintains computational power efficiently.

To experiment on the impact of recurrent layers, we conducted an ablation study on different LSTM cell sizes (8, 16, 32 and 64 units). Smaller LSTM configurations with 8 and 16 units achieved slightly lower

accuracies of 97.87% and 97.16% respectively, while they are computationally efficient. In contrast, the unit size 64 also achieves the best accuracy, but it takes the highest time as computational power. It means a larger LSTM cell does not offer performance benefits. We noted that using LSTM with a 32 cell size gives us the best accuracy (98.58%), aslo the least time consuming (33 s per epoch). That makes the most balanced configuration for our proposed (RGNN3D) model. So, these findings support our architectural choices and demonstrate the strength of RGNN3D for glioma grading.

3.2. Limitations

This study has a few limitations. The model was trained on a limited dataset (BraTS 2019-20), which may affect its generalizability across diverse patient groups. Performance on FLAIR images was lower, suggesting room for improvement in preprocessing. While LIME helps explain predictions, deeper clinical interpretability remains a challenge. Lastly, the model has not yet been tested in real clinical environments, which is essential for confirming its practical utility.

4. Conclusion

This study innovated a novel hybrid architecture that could be used to grade glioma tumors reliably and accurately utilizing 3D MRI data. Hence, one of the key strengths of the RGNN3D model is its ability to effectively integrate and utilize medical radiomic features, which are critical in the accurate characterization of tumor heterogeneity and progression. We believe the RGNN3D architecture would serve as an intuitive decision support system for medical experts by significantly improving diagnostic precision. Our evaluation across four MRI stages (T1, T1CE, T2, and FLAIR) reveals that both T1 and T2 stages can be utilized to get the highest performance in terms of grading accuracy, demonstrating their robustness in capturing critical tumor characteristics. Future research should focus on integrating multi-parametric MRI data to leverage the strengths of each modality. Additionally, improving preprocessing and feature extraction techniques for FLAIR images may assist in mitigating current limitations and enhancing their utility in grading. The model's robustness could be further enhanced with a larger dataset that includes diverse patient glioma grading reports, ensuring broader applicability and accuracy. We also believe that our proposed system could be applicable in clinical settings and digital healthcare, especially in rural or isolated places with limited access to specialist physicians.

References

[1] D. Ricard, A. Idbaih, F. Ducray, M. Lahutte, K. Hoang-Xuan, J.-Y. Delattre, Primary brain tumours in adults, The Lancet 379 (9830) (2012) 1984–1996.

[2] R. Kumar, A. Gupta, H. S. Arora, G. N. Pandian, B. Raman, Cghf: A computational decision support system for glioma classification using hybrid radiomics-and stationary wavelet-based features, IEEE Access 8 (2020) 79440–79458.

[3] E. B. Claus, K. M. Walsh, J. K. Wiencke, A. M. Molinaro, J. L. Wiemels, J. M. Schildkraut, M. L. Bondy, M. Berger, R. Jenkins, M. Wrensch, Survival and low-grade glioma: the emergence of genetic information, Neurosurgical focus 38 (1) (2015) E6.

[4] Q. T. Ostrom, L. Bauchet, F. G. Davis, I. Deltour, J. L. Fisher, C. E. Langer, M. Pekmezci, J. A. Schwartzbaum, M. C. Turner, K. M. Walsh, et al., The epidemiology of glioma in adults: a "state of the science" review, Neuro-oncology 16 (7) (2014) 896–913.

[5] H. Hyare, S. Thust, J. Rees, Advanced mri techniques in the monitoring of treatment of gliomas, Current treatment options in neurology 19 (2017) 1–15.

[6] Z. Liu, S. Wang, D. Dong, J. Wei, C. Fang, X. Zhou, K. Sun, L. Li, B. Li, M. Wang, et al., The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges, Theranostics 9 (5) (2019) 1303.

[7] M. E. Mayerhoefer, A. Materka, G. Langs, I. Häggström, P. Szczypiński, P. Gibbs, G. Cook, Introduction to radiomics, Journal of Nuclear Medicine 61 (4) (2020) 488–495.

[8] J.-b. Qin, Z. Liu, H. Zhang, C. Shen, X.-c. Wang, Y. Tan, S. Wang, X.-f. Wu, J. Tian, Grading of gliomas by using radiomic features on multiple magnetic resonance imaging (mri) sequences, Medical science monitor: international medical journal of experimental and clinical research 23 (2017) 2168.

[9] G. Cui, J. J. Jeong, Y. Lei, T. Wang, T. Liu, W. J. Curran, H. Mao, X. Yang, Machine-learning-based classification of glioblastoma using mri-based radiomic features, in: Medical imaging 2019: computer-aided diagnosis, Vol. 10950, SPIE, 2019, pp. 1063–1068.

[10] H.-H. Cho, H. Park, Classification of low-grade and high-grade glioma using multi-modal image radiomics features, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2017, pp. 3081–3084.
[11] Q. Chen, L. Wang, L. Wang, Z. Deng, J. Zhang, Y. Zhu, Glioma grade prediction using wavelet scattering-based radiomics, IEEE Access 8 (2020) 106564–106575.
[12] P. Sun, D. Wang, V. C. Mok, L. Shi, Comparison of feature selection methods and machine learning classifiers for radiomics analysis in glioma grading, Ieee Access 7 (2019) 102010–102020.
[13] J. Cheng, J. Liu, H. Yue, H. Bai, Y. Pan, J. Wang, Prediction of glioma grade using intratumoral and peritumoral radiomic features from multiparametric mri images, IEEE/ACM Transactions on Computational Biology and Bioinformatics 19 (2) (2020) 1084–1095.
[14] H.-h. Cho, S.-h. Lee, J. Kim, H. Park, Classification of the glioma grading using radiomics analysis, PeerJ 6 (2018) e5982.
[15] S. Priya, Y. Liu, C. Ward, N. H. Le, N. Soni, R. Pillenahalli Maheshwarappa, V. Monga, H. Zhang, M. Sonka, G. Bathla, Machine learning based differentiation of glioblastoma from brain metastasis using mri derived radiomics, Scientific reports 11 (1) (2021) 10478.
[16] U. Baid, S. U. Rane, S. Talbar, S. Gupta, M. H. Thakur, A. Moiyadi, A. Mahajan, Overall survival prediction in glioblastoma with radiomic features using machine learning, Frontiers in computational neuroscience 14 (2020) 61.
[17] H. Li, Z. Wang, C. Lan, P. Wu, N. Zeng, A novel dynamic multiobjective optimization algorithm with non-inductive transfer learning based on multi-strategy adaptive selection, IEEE transactions on neural networks and learning systems (2023).
[18] P. Wu, H. Li, L. Hu, J. Ge, N. Zeng, A local-global attention fusion framework with tensor decomposition for medical diagnosis, IEEE/CAA Journal of Automatica Sinica 11 (6) (2024) 1536–1538.
[19] T. Xie, X. Chen, J. Fang, H. Kang, W. Xue, H. Tong, P. Cao, S. Wang, Y. Yang, W. Zhang, Textural features of dynamic contrast-enhanced mri derived model-free and model-based parameter maps in glioma grading, Journal of Magnetic Resonance Imaging 47 (4) (2018) 1099–1111.
[20] S. Bisdas, C. Tisca, C. Sudre, E. Sanverdi, D. Roettger, J. M. Cardoso, Non-invasive in vivo prediction of tumour grade and idh mutation status in gliomas using dynamic susceptibility contrast (dsc) perfusion-and diffusion-weighted mri. (2018).
[21] Q. Tian, L.-F. Yan, X. Zhang, X. Zhang, Y.-C. Hu, Y. Han, Z.-C. Liu, H.-Y. Nan, Q. Sun, Y.-Z. Sun, et al., Radiomics strategy for glioma grading using texture features from multiparametric mri, Journal of Magnetic Resonance Imaging 48 (6) (2018) 1518–1528.
[22] C. Su, J. Jiang, S. Zhang, J. Shi, K. Xu, N. Shen, J. Zhang, L. Li, L. Zhao, J. Zhang, et al., Radiomics based on multicontrast mri can precisely differentiate among glioma subtypes and predict tumour-proliferative behaviour, European radiology 29 (2019) 1986–1996.
[23] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, S. Y. Philip, A comprehensive survey on graph neural networks, IEEE transactions on neural networks and learning systems 32 (1) (2020) 4–24.
[24] A. Graves, A. Graves, Long short-term memory, Supervised sequence labelling with recurrent neural networks (2012) 37–45.
[25] M. T. Ribeiro, S. Singh, C. Guestrin, " why should i trust you?" explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135–1144.
[26] Multimodal Brain Tumor Segmentation Challenge 2019: Data — CBICA — Perelman School of Medicine at the University of Pennsylvania — med.upenn.edu, `www.med.upenn.edu/cbica/brats2019/data.html`, [Accessed 23-06-2024].
[27] Multimodal Brain Tumor Segmentation Challenge 2020: Data — CBICA — Perelman School of Medicine at the University of Pennsylvania — med.upenn.edu, `www.med.upenn.edu/cbica/brats2020/data.html`, [Accessed 23-06-2024].
[28] J. E. Van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, B. Baessler, Radiomics in medical imaging—"how-to" guide and critical reflection, Insights into imaging 11 (1) (2020) 91.
[29] J. J. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H. J. Aerts, Computational radiomics system to decode the radiographic phenotype, Cancer research 77 (21) (2017) e104–e107.
[30] T. T. Truong, D. Dinh-Cong, J. Lee, T. Nguyen-Thoi, An effective deep feedforward neural networks (dfnn) method for damage identification of truss structures using noisy incomplete modal data, Journal of Building Engineering 30 (2020) 101244.
[31] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (8) (1997) 1735–1780.
[32] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, International Conference on Learning Representations (ICLR), 2017.
[33] M. A. Ali, M. S. Hossain, M. K. Hossain, S. S. Sikder, S. A. Khushbu, M. Islam, Amdnet23: Hybrid cnn-lstm deep learning approach with enhanced preprocessing for age-related macular degeneration (amd) detection, Intelligent Systems with Applications 21 (2024) 200334.
[34] A. Palkar, C. C. Dias, K. Chadaga, N. Sampathila, Empowering glioma prognosis with transparent machine learning and interpretative insights using explainable ai, IEEE Access 12 (2024) 31697–31718.
[35] F. Ullah, M. Nadeem, M. Abrar, F. Amin, A. Salam, A. Alabrah, H. AlSalman, Evolutionary model for brain cancer-grading and classification, IEEE Access (2023).
[36] L. Liu, J. Chang, P. Zhang, H. Qiao, S. Xiong, Sasg-gcn: self-attention similarity guided graph convolutional network for multi-type lower-grade glioma classification, IEEE Journal of Biomedical and Health Informatics (2023).
[37] M. Renugadevi, K. Narasimhan, C. Ravikumar, R. Anbazhagan, G. Pau, K. Ramkumar, M. Abbas, N. Raju, K. Satish, S. Prabu, Machine learning empowered brain tumor segmentation and grading model for lifetime prediction, IEEE Access (2023).

[38] H. A. Hafeez, M. A. Elmagzoub, N. A. B. Abdullah, M. S. Al Reshan, G. Gilanie, S. Alyami, M. U. Hassan, A. Shaikh, A cnn-model to classify low-grade and high-grade glioma from mri images, IEEE Access 11 (2023) 46283–46296.
[39] J. Cheng, M. Gao, J. Liu, H. Yue, H. Kuang, J. Liu, J. Wang, Multimodal disentangled variational autoencoder with game theoretic interpretability for glioma grading, IEEE Journal of Biomedical and Health Informatics 26 (2) (2021) 673–684.
[40] S. Divya, L. Padma Suresh, A. John, Enhanced deep-joint segmentation with deep learning networks of glioma tumor for multi-grade classification using mr images, Pattern Analysis and Applications 25 (4) (2022) 891–911.
[41] S. Montaha, S. Azam, A. R. H. Rafid, M. Z. Hasan, A. Karim, A. Islam, Timedistributed-cnn-lstm: A hybrid approach combining cnn and lstm to classify brain tumor on 3d mri scans performing ablation study, IEEE Access 10 (2022) 60039–60059.