Leveraging Volume Embeddings for Accurate and Efficient Glioma Classification

Xue Li, PhD Candidate, Student, Electrical and Computer Engineering, University of Wisconsin-Madison
Alan McMillan, PhD

Introduction

Foundation models have shown remarkable success in generating embeddings for downstream tasks, like image retrieval and classification. While 2D embeddings are widely utilized, constructing effective representations for 3D datasets often involves aggregating embeddings into volume features. This study explores multiple approaches to construct volume embeddings from brain MRI scans and evaluate their performance in classification of low-grade glioma (LGG) versus high-grade glioma (HGG).

Hypothesis

Volume embeddings are effective and computationally efficient for identifying LGG and HGG from brain MRI scans.

Methods

As Fig. 1 shows, this study utilized the BraTS 2020 dataset, including 369 subjects with four MRI series (T1, T2, T1CE, and FLAIR). Slice-level embeddings were first extracted from each series using the MedImageInsight foundation model and then aggregated into volume embeddings using various pooling methods, including mean, max, and median. All possible combinations of the series were explored by concatenating their respective volume embeddings. A Multi-Layer Perceptron (MLP) classifier, with a hidden layer size of 200, was trained on 276 subjects, while the remaining 93 subjects were reserved for testing.

Results

Performance was assessed using accuracy, precision, recall, and F1-score, as summarized in Table 1. The top section presents the performance using volume embeddings from different MRI series and pooling methods. The best-performing combination of one to four modalities is reported, with the highest accuracy achieved by max pooling on T1CE+FLAIR (96.8% accuracy and F1-score). The bottom section shows the performance of state-of-the-art (SOTA) models trained directly on images, where ResNet-50 attained the best results (98.9% accuracy and F1-score).

Conclusion

This study demonstrates that volume embeddings can accurately distinguish LGG from HGG, with all metrics exceeding 91%. The best embedding-based model lags behind SOTA only by 2.1%, which is inspiring and highlights the exciting promise for the future development of embedding-based methods as an efficient alternative for classification tasks.

Figure(s)

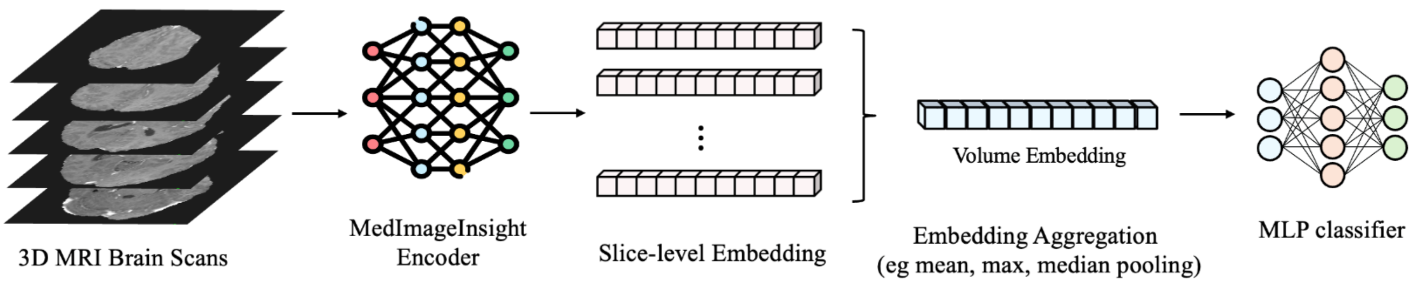

Figure 1. Overview of the proposed pipeline for volume glioma classification. First, 3D MRI scans are processed through the MedImageInsight encoder to produce slice-level embeddings. These are then aggregated using pooling methods (e.g., mean, max, median) to generate a single volume embedding. Finally, an MLP classifier is trained on these volume embeddings to distinguish between low-grade and high-grade glioma.

Volume embedding-based MLP classifier					
Pooling	Series	Accuracy	Precision	Recall	F1 score
Mean	T1CE	91.4%	91.8%	91.4%	91.6%
	T1CE+FLAIR	92.5%	92.6%	92.5%	92.6%
	T2+T1CE+FLAIR	***93.6%***	***93.9%***	***93.6%***	***93.7%***
	T1+T2+T1CE+FLAIR	91.4%	91.8%	91.4%	91.6%
Max	T1CE	93.6%	93.5%	93.6%	93.5%
	T1CE+FLAIR	***96.8%***	***96.9%***	***96.8%***	***96.8%***
	T2+T1CE+FLAIR	94.6%	95.0%	94.6%	94.8%
	T1+T2+T1CE+FLAIR	92.5%	92.5%	92.5%	92.5%
Median	T1CE	91.4%	91.8%	91.4%	91.6%
	T1CE+FLAIR	92.5%	92.6%	92.5%	92.6%
	T2+T1CE+FLAIR	***92.5%***	***93.1%***	***92.5%***	***92.8%***
	T1+T2+T1CE+FLAIR	***92.5%***	***93.1%***	***92.5%***	***92.8%***

SOTA models trained at the image level					
Models	Series	Accuracy	Precision	Recall	F1 score
ResNet18	Fused images (2D+3D)	97.9%	98.3%	97.6%	97.9%
ResNet50	***Fused images (2D+3D)***	***98.9%***	***99.0%***	***98.8%***	***98.9%***

Table 1. Classification performance of the proposed volume embedding-based MLP classifier compared to SOTA image-level models (Janardhan and Kiranmayee, IEEE CONECCT, 2024).

Keywords

Artificial Intelligence/Machine Learning; Imaging Research