An Analytical Study of Optimized EfficientNet CNNs, Vision Transformers, and Hybrid Transformers for State-of-the-Art MRI-Based Brain Tumor Diagnosis With Integrated Explainable AI

Singh, Ishmeet (School: Stockdale High School)

Brain tumors affect thousands annually, with approximately 95,000 primary brain tumor cases diagnosed each year in the U.S. Diagnostic imaging errors affect an estimated 40 million patients annually worldwide, with 3–5% error rates contributing to 795,000 deaths or permanent disabilities. Brain tumor classification remains complex due to high intra-class variability and interclass similarity. To address this, this study developed and compared three optimized deep learning architectures—EfficientNet CNNs (b5 model), Vision Transformers (ViTs) (Swin Transformer model), and Hybrid Transformers—for classifying 15 brain tumor types from a dataset of 4,479 MRI images. EfficientNet models from my prior baseline study were optimized and re-evaluated. ViTs were introduced for their ability to model long-range spatial dependencies. Additionally, hybrid transformers—which combine convolutional feature extraction with transformer-based global attention—were implemented. All architectures were enhanced using an optimization framework incorporating Stochastic Weight Averaging, MixUp/CutMix augmentation, Soft Target Cross-Entropy Loss, learning rate scheduler (cosine annealing), gradient clipping, and AdamW optimization. Among the transformer-based models, one achieved up to 99.00% validation accuracy, indicating a potential reduction in diagnostic error rates from 5% to 0.05%. To ensure clinical applicability, Explainable AI (XAI) methods were applied to visualize tumor localization and model decision pathways. Models were further optimized for inference time and computational efficiency to maximize deployment readiness. This study presents an interpretable, state-of-the-art model for MRI-based brain tumor diagnosis, with strong potential for clinical integration.

Awards Won: