

Hybrid Approach for Improving Feature Extraction by Integrating Ensemble Learning with Transformers for Brain Tumor MRI Images Classification (#101471)

Reviewer Comments

- The Authors proposed a a hyperd model which integrates of transfer learning and the transformer encoder mechanism.
- six pre-existing deep learning model, both individually and in combination, it was determined that and ensemble of three pretrained models achieved the highest accuracy. This ensemble, comprising DenseNet201, GoogleNet (InceptionV3), and InceptionResNetV2, is selected as the feature extraction framework for the transformer encoder network.
- The transformer encoder module integrates a Shifted Window-based Self-Attention mechanism, sequential Self-Attention, with a Multilayer Perceptron layer (MLP).
- These experiments were conducted on three publicly available research datasets for evaluation purposes. The Cheng dataset, BT-large-2c, and BT-large-4c dataset, each designed for various classification tasks with differences in sample number, planes, and contrast.
- The model gives consistent results on all three datasets and reaches an accuracy of 99.34%, 99.16%, and 98.62%, respectively, which are improved compared to other techniques.

The proposed work is commendable and innovative, but it lacks validation. Numerous explainable AI models, such as LIME and SHAP, are available. I recommend that the authors validate their model for image classification.