Jetwyn Wilson

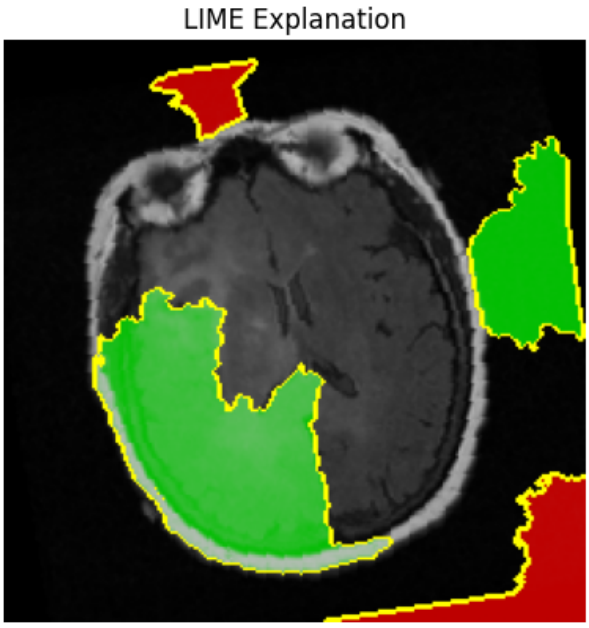
Using ResNet 18 to Apply LIME and SHAP

Dataset Overview

The LGG MRI Segmentation dataset was obtained from Kaggle. The dataset comprises 3900 T1-weighted MRI slices of 110 patients labeled as "Tumor" and "No Tumor". Images were resized to 224x224 pixels and converted to RGB.This dataset is much smaller than typical medical imaging that often contain hundreds of thousands annotated images. This dataset is used to limit the sample size to challenge the model generalization and robustness. Advantage to using a smaller dataset is to try techniques like data augmentation and transfer learning. These are important to improve the model performance. Using data augmentation creates more training samples by challenging the original images by flipping, rotating, or changing brightness. Transfer learning helps to use models like ResNet that were trained on large datasets to have better understanding patterns in the brain scans with limited data.

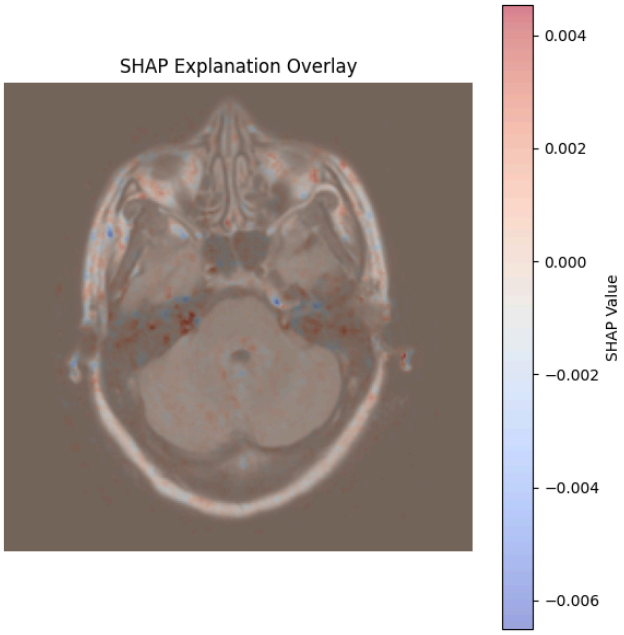
Results and Analysis

To evaluate the fairness of the ResNet-18 model applied to the brain MRI classification dataset. The use of LIME is to help highlight localized regions in green (to show the model prediction) and red (indicating features that were predicted away from the predicate class). This helps interpret the model's reliance on specific texture or edges within a region.



SHAP was able to provide a pixel-wise attribution heatmap to show what pixels contributed to the model's output. The re-blue scale indicates the positive and negative contributions. Shap was able to be more diffuse than LIME, revealing broader areas of influence. The re-shading around the brain suggests that they contribute positively to predicting the tumor.

Jetwyn Wilson



These techniques can help improve the model interpretability by visualizing the model critical areas within medical imaging. These techniques provide fairness by ensuring the model is not just relying on background noise.