Predicting Future Health Conditions: Studying Brain Scans in Patients Affected by Oxygen Shortage at Birth

Hypoxic-ischaemic encephalopathy (HIE) is characterized by an infant's brain being deprived of oxygen at or near birth, potentially leading to long-term adverse neurological consequences. Predicting these outcomes poses a challenge for clinicians. To address this, a Data Science Clinic team collaborated with Center for the Science of Early Trajectories (SET) to analyze magnetic resonance imaging (MRI) data to identify predictive indicators that could inform clinical decision-making for HIE patients.

The team trained a model to label different regions of the brain using neonatal MRI images with tissue-type annotations from a public dataset. Such a model could mask irrelevant regions to improve classification model performance. While model predictions were highly accurate on the training data (Figure 1), they performed poorly on the SET-provided data (Figure 2).

A classification model was trained to predict the risk of developing adverse outcomes from MRI data based on labels specified by the SET. While the model demonstrated poor performance, some discriminative capacity for different outcome labels was observed. Visualizations (Figure 3) were created to highlight critical regions of input images for the classification task, suggesting the brain's center as a crucial area.

Apparent Diffusion Coefficient (ADC), an MRI-derived metric indicating the average magnitude of water diffusion within the brain, was also analyzed. With the hypothesis of a link between a patient examination’s ADC image and their neurological outcomes, histograms of each ADC image's pixel intensity frequency were created, and correlations between histogram statistics and outcomes were assessed. Analysis revealed weak correlations ($r{\sim}0.17$) that did not reach significance ($0.1 < p < 0.4$) (Figure 4).

The team’s efforts showed potential in predicting adverse neurological outcomes in HIE patients using MRI data. Nonetheless, more data and further optimization of the models are required to enhance model performance and improve generalizability.

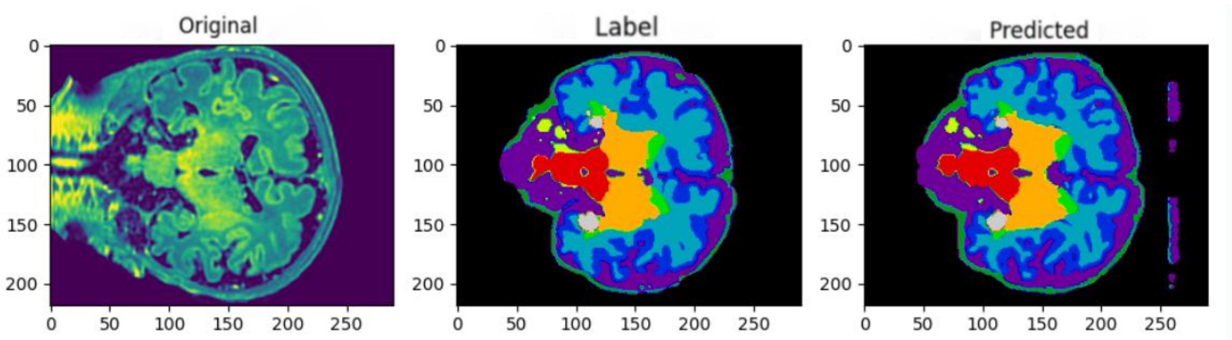

Figure 1. A section from the middle of a subject's original MRI in the training dataset (left), its corresponding annotations, i.e., the ground truth label (center), compared with the model prediction (right).

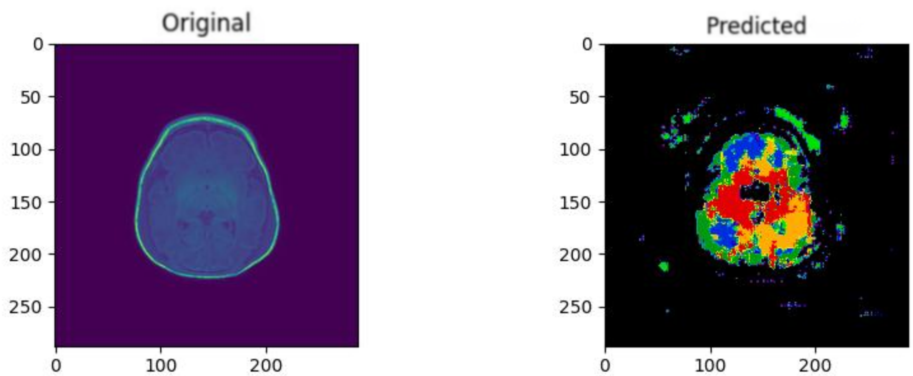

Figure 2. A section from the middle of a patient's original MRI in the SET dataset (left), compared with the model prediction (right).

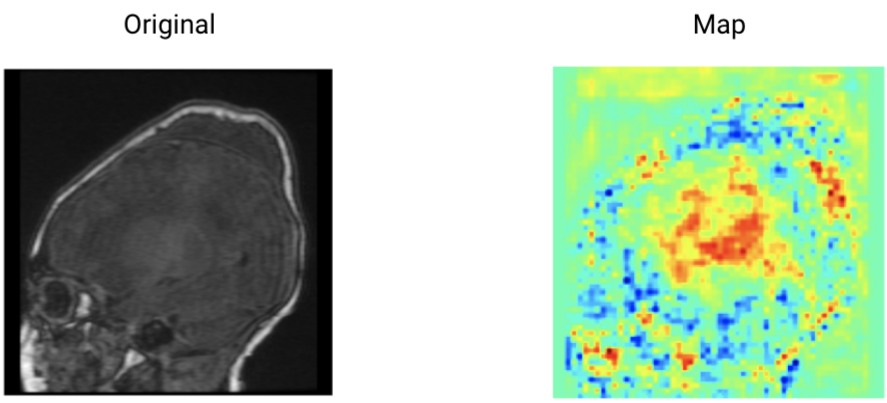

Figure 3. Comparison of a 2-D section of an MRI image from the SET dataset (left) and its corresponding map (right), showing the importance of the center region of the brain in classification on seizure labels.

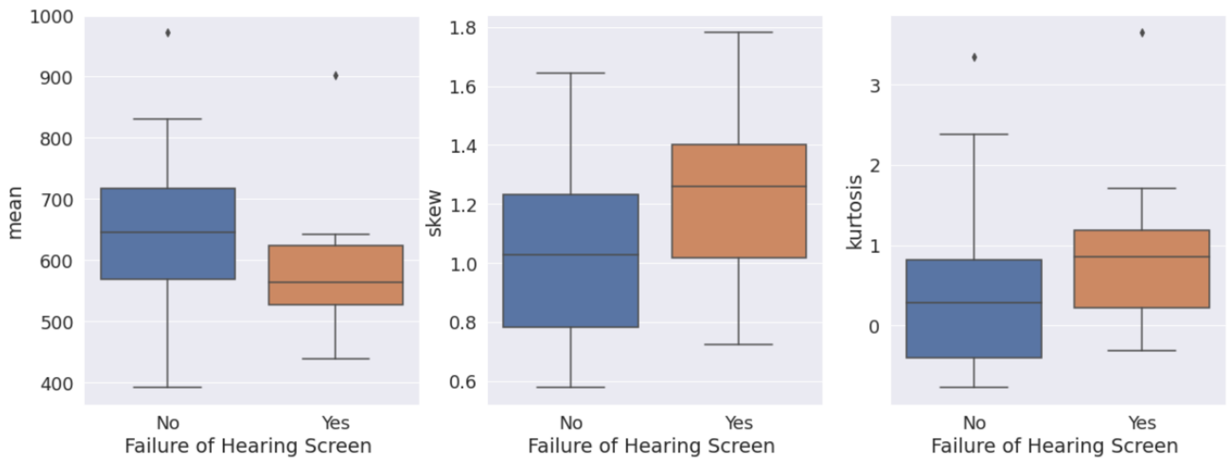

Figure 4. Example box plots showing the relationship between ADC histogram metrics and Failure of Hearing Screen. ADC correlated most strongly with the Failure of Hearing Screen.