Explainable AI for Autism Diagnosis
Identifying Critical Brain regions using fMRI data

Suryansh Vidya, Kush Gupta, Amir Aly, Andy Wills, Emmanuel Ifeachor, Rohit Shankar

NEURAL INFORMATION PROCESSING SYSTEMS

UNIVERSITY OF PLYMOUTH

THE CHALLENGE OF EARLY AUTISM DIAGNOSIS

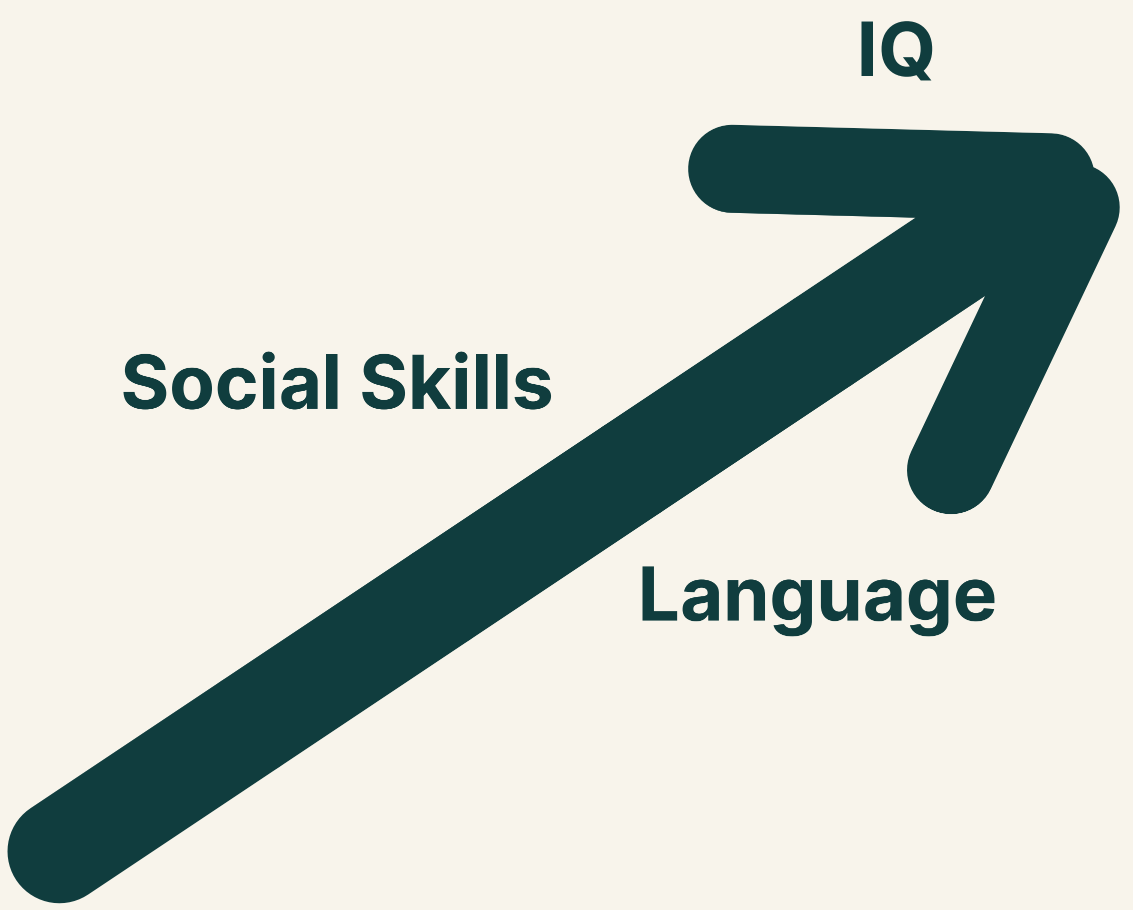
- Lengthy wait times, expensive and subjective assessments
- Potential bias in current diagnostic methods, affecting diverse groups
- Critical early intervention windows often missed due to delayed, uncertain diagnoses

UNIVERSITY OF PLYMOUTH

WHY EARLY DIAGNOSIS MATTERS

- Early (age 2-3) identification of Autism Spectrum Disorder (ASD) can significantly improve long-term quality of life
- Timely interventions enhance language, social skills, and cognitive development
- Reducing diagnostic delays helps harness critical windows for optimal support

(Estes et al., 2015; Okoye et al., 2023)



UNIVERSITY OF PLYMOUTH

THE BRAIN AND AUTISM

- Resting-state fMRI measures how different parts of the brain are active over time at rest.
- Functional connectivity looks at which areas of the brain's activity rise and fall together, indicating they are "connected."
- Differences in functional connectivity can serve as "biomarkers" for ASD.

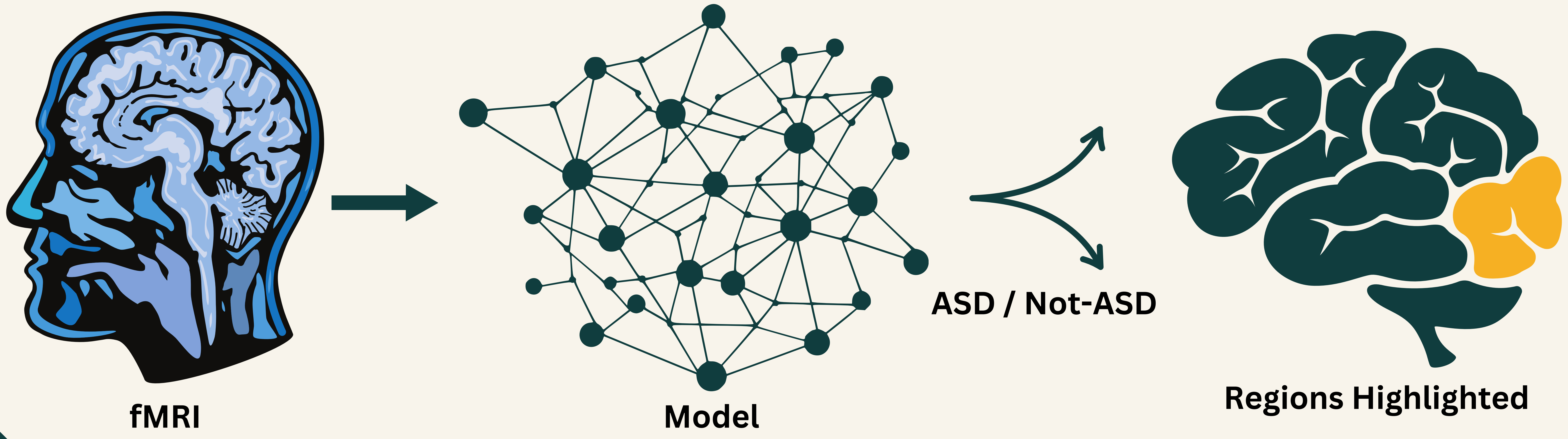
UNIVERSITY OF PLYMOUTH

THE POTENTIAL OF AI

- Deep learning detects subtle fMRI patterns linked to autism.
- But most AI models are "black boxes," offering no explanation.
- We build upon previous research to improve MRI-based models and add transparency, so we know why a decision is made.

UNIVERSITY OF PLYMOUTH

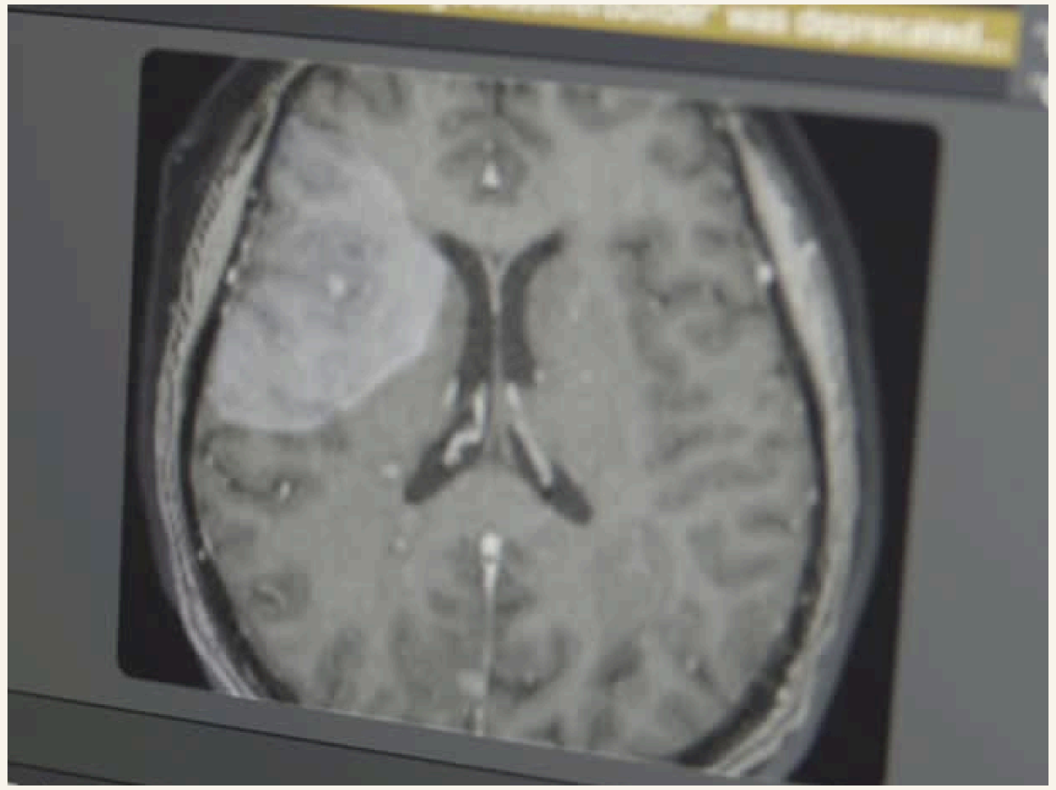
HIGH LEVEL APPROACH



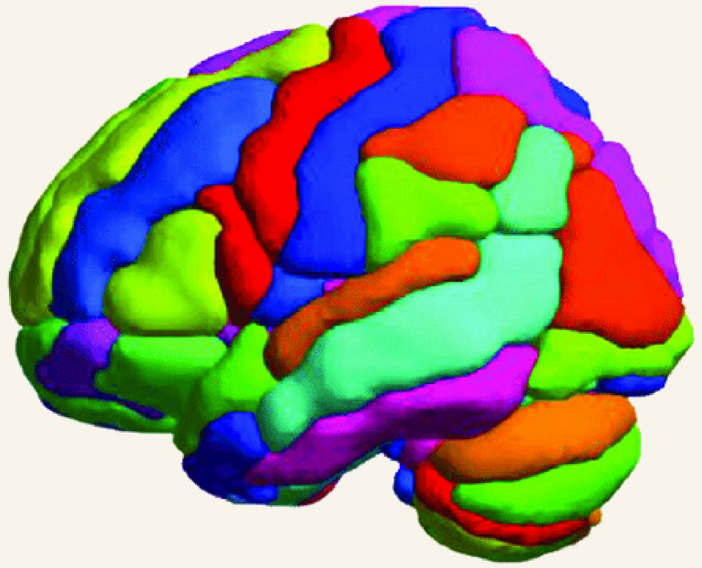
UNIVERSITY OF PLYMOUTH

DATA AND PREPROCESSING

- Utilized the largest publicly available dataset: ABIDE (1112 samples)
- Ensured data quality by removing scans with excessive head movement, down to 884 samples.
- Applied standard preprocessing steps (e.g., motion correction, normalization)
- Parcellated the brain into 116 regions using the AAL atlas for consistent feature extraction



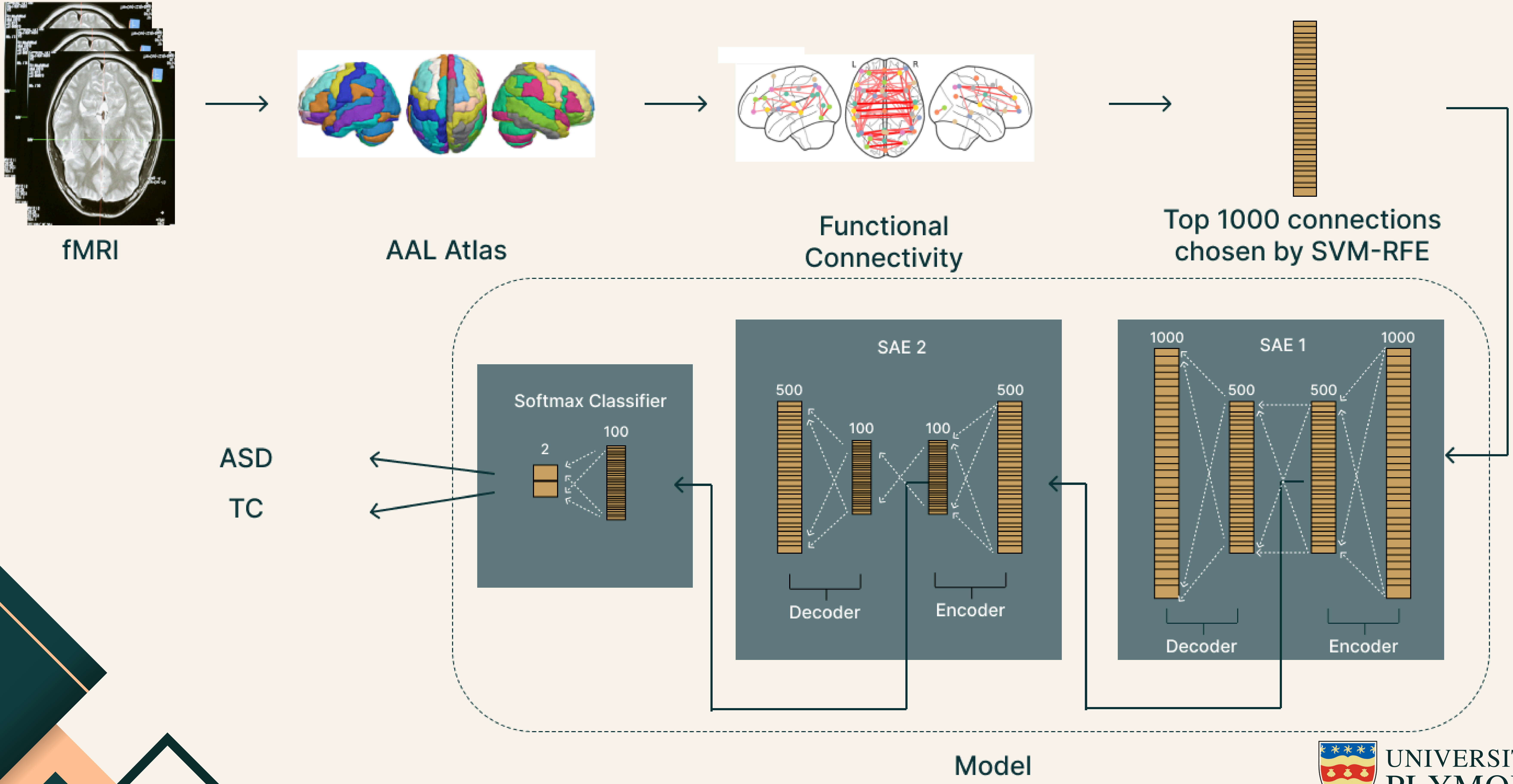
ABIDE I (Di Martino et al., 2013)



AAL atlas (Tzourio-Mazoyer et al., 2002)

UNIVERSITY OF PLYMOUTH

THE MODEL ARCHITECTURE



UNIVERSITY OF PLYMOUTH

RESULTS

Model	Accuracy
Pavithra et al. (2023)	85.0%
Bhandage et al. (2023)	92.4%
Wadhera et al. (2023)	88.1%
Wang et al. (2019)	93.5%
Our Model	**98.2%**

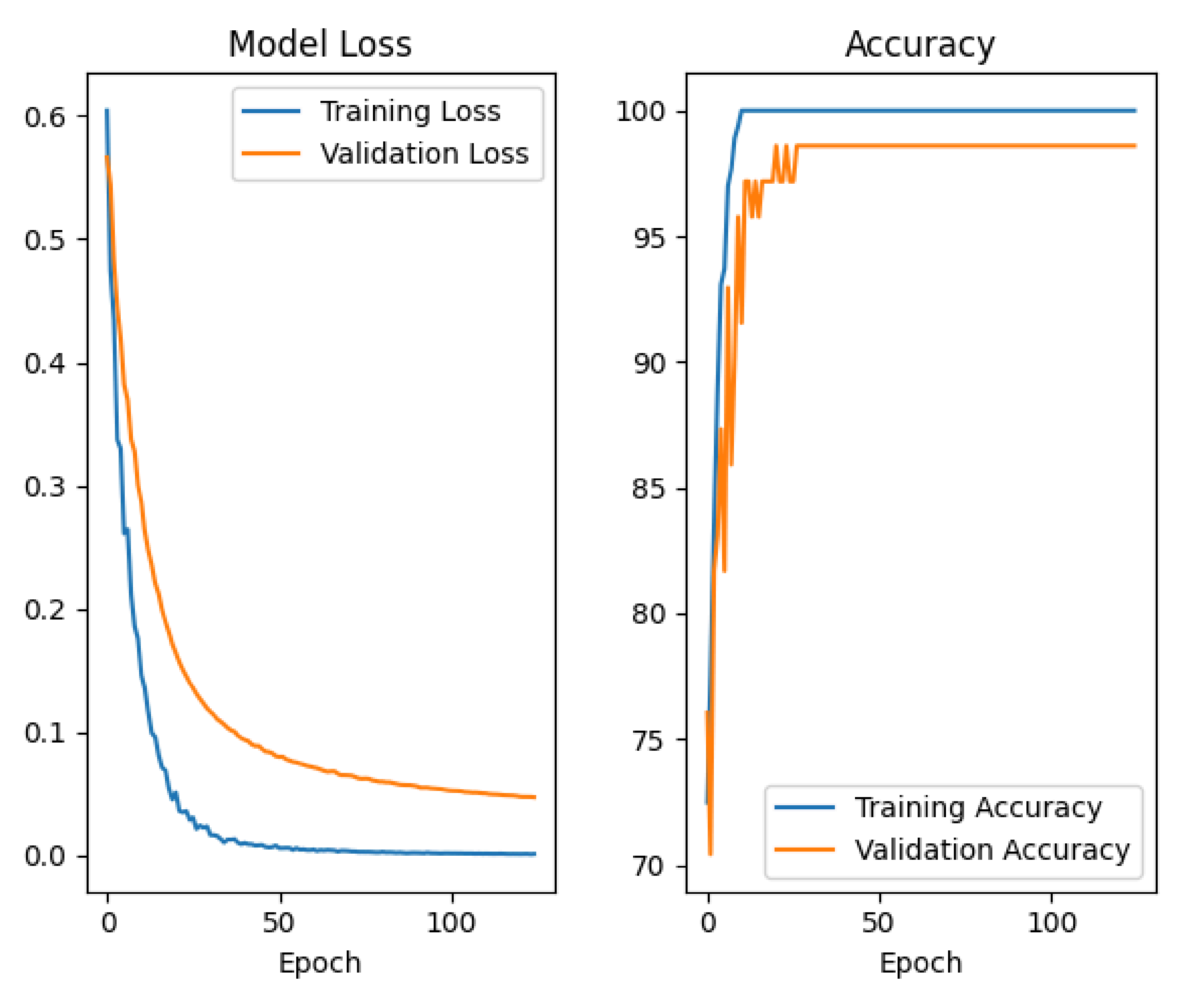
Sensitivity: 0.99
Specificity: 0.98
Precision: 0.98
F1 Score: 0.98

Overfitting?

UNIVERSITY OF PLYMOUTH

LOSS MONITORING

- Training and validation loss both decrease steadily over time
- Validation loss remains stable, not diverging from training loss
- Indicates good generalization



UNIVERSITY OF PLYMOUTH

WHAT METHOD TO USE?

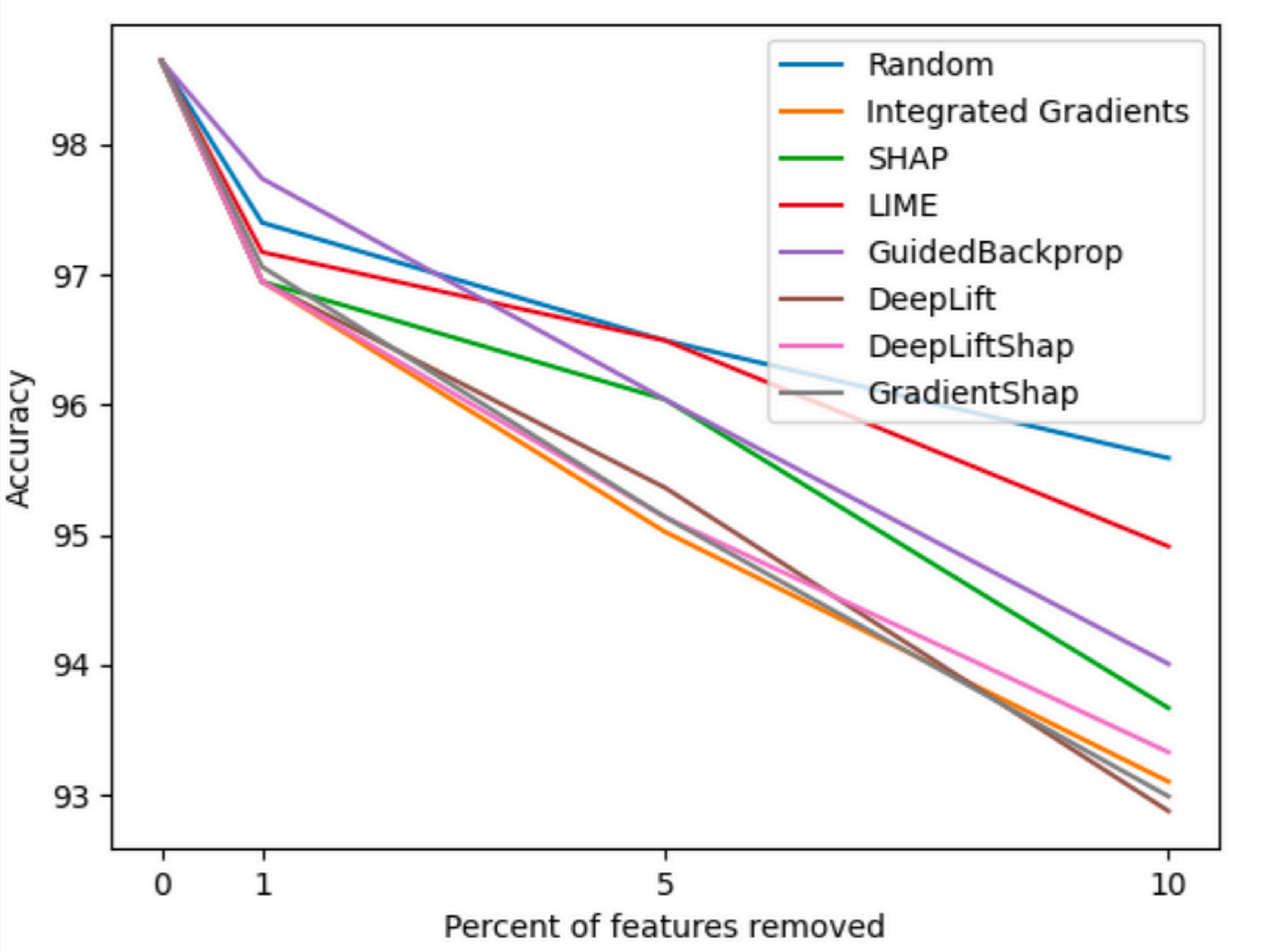
List of recent interpretablility methods we implemented from literature:

- Integrated Gradients
- LIME
- SHAP
- DeepLift
- DeepLiftShap
- GradientShap
- GuidedBackprop

UNIVERSITY OF PLYMOUTH

BENCHMARKING METHODS

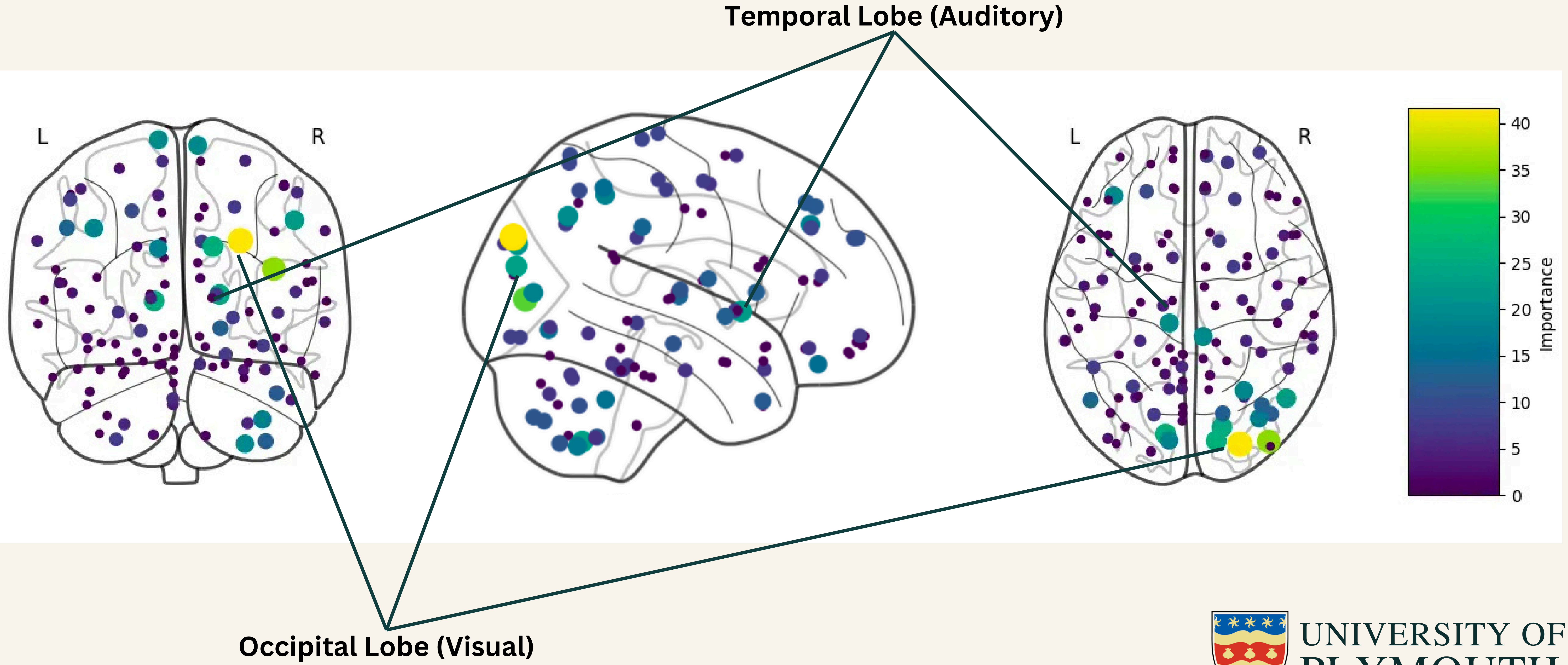
- The RemOve And Retrain (ROAR) benchmark (Hooker et al., 2019).
- Gradient-based methods performed best
- Perturbation methods performed the same as a random baseline
- Integrated Gradients was the best performing method



ROAR Analysis

UNIVERSITY OF PLYMOUTH

IDENTIFIED BRAIN REGIONS



UNIVERSITY OF PLYMOUTH

LITERATURE ANALYSIS

Consistent findings across multiple research approaches:

- Genetic studies highlighting importance of visual cortex in autism (Gandal et al., 2022)
- Other directions like global motion perception deficit research and atypical gamma oscillation studies also show links to visual and auditory cortex (Orekhova et al., 2023; Robertson et al., 2014)

Demonstrates potential biomarkers common across ASD severity levels

UNIVERSITY OF PLYMOUTH

LIMITATIONS AND FUTURE DIRECTIONS.

- Tested only on ABIDE so we do not know performance on OOD data. There is a need for larger datasets.
- Need a dataset with other forms of neurodivergents such as ADHD, otherwise we might just be building a neurodiverse classifier instead of an Autism one.
- Use of multi modal data. Autism symptoms show in other modalities of data as well such as eye tracking, EEG etc.
- Also need to explore alternative imaging methods like optical imaging, which offer greater accuracy and speed than fMRI

UNIVERSITY OF PLYMOUTH

QUESTIONS