

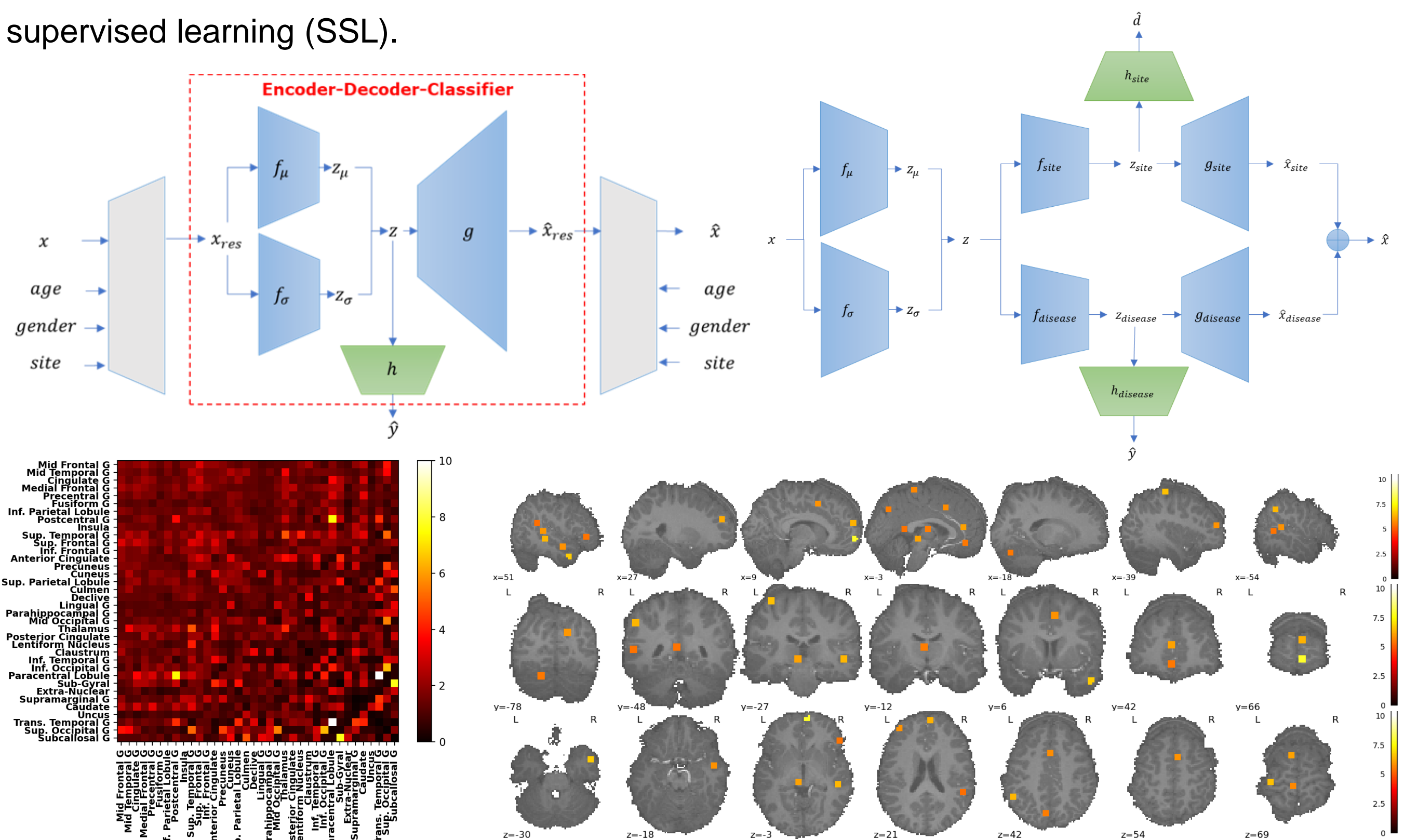
Semi-supervised Learning of Functional Connectome for Disease Classification

Student: Yew Wei Chee

Supervisor: Prof Jagath Chandana Rajapakse

Project Objectives:

Overfitting is a common problem when computational models are applied on neuroimaging datasets, which are high-dimensional and small in terms of sample sizes. One way to overcome this is to pool datasets of similar diseases to augment the small dataset. However, such efforts may introduce undesirable variations due to site effects and inconsistent labeling. To mitigate these issues, two encoder-decoder-classifier architectures were proposed to carry out semi-supervised learning (SSL).



The use of SSL led to a consistent increment in the model accuracy for the task of classifying between healthy subjects and patients with diseases including autism spectrum disorders (ASD) and attention-deficit hyperactivity disorder (ADHD). In addition, performing data harmonization simultaneously with SSL led to even greater improvements (+5.5% for ASD and +3.3% for ADHD in average). Biomarkers generated from the proposed method could potentially represent site-invariant biomarkers as they were shown to place more emphasis on a subset of previously discovered site-specific biomarkers. The findings in this report emphasize the importance of taking both site effects and labeling inconsistencies into account when gathering datasets from multiple sites to overcome neuroimaging data paucity.