Name: Qing Lu, Ph.D. Email: lucienq@ufl.edu

Deep Transfer Learning for Substance Use Disorders Genetic Research

Our capacity to detect novel substance use disorders (SUD) genes in small-scale studies or minority populations (e.g., African-ancestry populations) is often limited by the small sample size. The vast amount of genetic data collected from biobank projects and the Caucasian population provides us with an additional resource for genetic research in small-scale studies or minority populations. In this project, we develop a transfer learning–based deep neural network (TL-DNN) framework that leverages the rich genetic data available from large biobank datasets to enhance SUD genetic research in small-scale studies or minority populations. To identify genes significantly associated with SUD, we further integrate a permutation-based association test within the TL-DNN framework. Through both cross-project and cross-ethnicity studies, we demonstrate the advantages of the proposed TL-DNN approach. In the cross-ethnicity study, knowledge learned from the White British population in the UK Biobank (UKB) was transferred to the Black British population. By leveraging information from the White British dataset, we observed improved statistical power in the Black British population. In the cross-project study, we transferred knowledge learned from the large-scale UKB dataset to the relatively smaller SAGE dataset, leading to the detection of more significant associations for nearly all candidate genes except CHRNB3.

Authors:
Submitter: Qing Lu, Professor, University of Florida, Gainesville
Kejin Dong, BS, University of Florida
Pei Geng, PhD, University of New Hampshire
Jinling Liu, PhD, University of Florida
Shan Zhang, Ph.D., University of Florida
Yuan Zhou, MS, University of Florida

DRAFT