

**Artificial Intelligence Methods for Improved Diagnosis
of Autism Spectrum Disorder**

AN HONORS THESIS
SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE
AT STANFORD UNIVERSITY

Maya Varma
Advised by Dennis P. Wall and James Landay
May 2020

Abstract

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impaired social skills and challenges with communication. Traditional diagnostic methods for ASD include lengthy parent questionnaires and extensive clinical observation, which are often subjective, expensive, and time-consuming. The rising prevalence of ASD and increasing demand for diagnostic tests have placed an enormous strain on the healthcare system, resulting in a large backlog of patients waiting to secure appointments with clinicians. Since delayed diagnosis is associated with poorer therapeutic outcomes, the lack of timely access to diagnostic tests for ASD has become a critical healthcare issue in recent years.

This thesis explores the utility of artificial intelligence techniques in improving the efficiency and accessibility of diagnostic tests. In this work, I present computational methods for obtaining diagnostic insights from two distinct modalities: (1) *the genome* and (2) *the phenome*.

The first part of this thesis focuses on the human genome, which consists of 3.2 billion DNA base pairs that define human traits, behaviors, and risk for diseases. The genome has been shown to play a major role in the manifestation of ASD, but the contribution of noncoding regions remains unknown. I introduce a novel machine learning methodology and dimensionality reduction algorithm based on maximum flow in order to elucidate the contribution of the noncoding genome to ASD. My work identifies a set of 55 high-confidence genetic variants in simple repeat sequences likely to be correlated with ASD, which can be prioritized for future targeted study.

The second part of this thesis focuses on the phenome, which is comprised of all observable behaviors that arise from the interplay between an individual's genes and surrounding environment. Children with ASD exhibit behavioral and social impairments, giving rise to the possibility of utilizing computational techniques to evaluate a child's social phenotype from home videos. I utilize computer vision and deep learning techniques to analyze over 8 hours of video footage depicting children engaged in gameplay with parents in a natural home environment, and I identify differences in gaze fixation and visual scanning behaviors between individuals with ASD and neurotypical individuals. My results show the utility

of game-based approaches in collecting valuable medical data and serving as a potential diagnostic aid.

Overall, the computational methods presented in this thesis provide insights into the complex workings of ASD and pave the way towards the creation of genetic screening tools and social phenotyping methods that can enable automated and accessible diagnoses.