Remote Detection and Severity Classification of Parkinson's Disease Using Markerless Videos in Uncontrolled Settings

Dhoot, Arav (School: Neerja Modi School)

Parkinson's Disease (PD) is the second most common neurodegenerative disease. Yet, it lacks a cure. Timely detection and treatment can significantly slow the progression of symptoms, yet access to neurologists is a challenge, particularly in rural and remote regions. Therefore, this study aims to develop a system for the detection and classification of PD using in-the-wild videos for early-stage diagnostic evaluation. For this research, a novel dataset consisting of videos from open-source domains with varying PD intensities and gait abnormalities was constructed. Using a human pose estimation (HPE) model, HPE images were generated from the videos. Scaled and centered HPE images were used to extract joint coordinates to train a 10-tree Random Forest Classifier (RFC), and generate spatiotemporal images—images produced by the superposition of the average of 17 HPE images. 7 modified classification models (ViT-Tiny, ViT-Small, ViT-Base, ResNet-18, ResNet-34, 5-layer ConvNet, and 6-layer ConvNet; each incorporating the previously trained RFC as an additional layer) were trained on 19218 spatiotemporal images. Using shallower models required less computational resources for training and inference compared to state-of-the-art models. Training on spatiotemporal images communicated greater computational information to the models, while incorporating the RFC provided greater global context during the classification. Upon testing on 2402 images, the models surpassed state-of-the-art performance, with ResNet-18 and ResNet-34 both achieving 99.88% accuracy. The models were tested in a clinical trial (n=27). All models displayed an accuracy close (within 5%) to the performance observed on the testing dataset, with ViT-Base and ResNet-34 achieving an accuracy of 96.30%.