Therapist and Child Engagement Analysis

# 

# Assignment Description:

Understanding people's gaze, emotions, and object interaction tracking in a video is a promising method to get insights into human actions and intentions. We are interested in assessing children with Autism Spectrum Disorder by analyzing the gaze, emotion, and interaction of children with objects and therapists.

# Problem statement:

Develop an optimized inference pipeline that given a long-duration video can show the predictions of the child and therapist's gaze, their emotions, the therapist’s and child’s level of engagement in a particular activity, and detect what they are doing with the objects around them. We prefer your code in the Python language. **You should utilize the following works and models to implement the code. Any other work/repository can be used for emotion recognition and analyzing object interaction. You also feel free to use state-of-the-art models that you think would be better.**

1. Grounded RAM-SAM: <https://github.com/IDEA-Research/Grounded-Segment-Anything>
2. Gaze Transformer: <https://github.com/nizhf/hoi-prediction-gaze-transformer>

The pipeline should be tested on the below test videos shared, consisting of a YouTube video list in the Google Drive link, the code should plot/display the predictions of the child and therapist's gaze, their emotions, the therapist’s and child’s level of engagement in a particular activity, and detect what they are doing with the objects around them.

**Expected Output:**

Output Video with the predictions overlaid on the Test Videos - Predictions of the child and therapist's gaze, their emotions, the therapist’s and a measure of the child’s level of engagement in a particular activity, and detection of what they are doing with the objects around them.

**Test Videos:** <https://drive.google.com/file/d/1uXDCgBx3kpecwc2DiVyO70tXNHeX85JI/view?usp=sharing>

# Deliverables:

1. Source code files: Inference scripts, prediction plots or test video outputs, and requirements files.
2. README.md: Detailed description of the logic behind analyzing the model predictions. The descriptions should be easy to follow and help the evaluator easily reproduce the results.
3. All the above are archived in one single .zip or .tar file. Either send a mail or share the link to download.

# Timeline: The deadline for submission is seven days after receiving the assignment.