Advancements in vision & speech analyses of ASD symptom domains

Contributors: Zoltán Tősér[1,2], Zachary Arnold[3], Richard Bellon, Máté Csákvári[2], Edith Ocampo[3], Allison Wainer[3], Mark Pollack[3], András Lőrincz[2], Latha Soorya[3]
1. Argus Cognitive, Inc. 2. Eotvos Lorand University, Budapest, Hungary, 3. Rush University Medical Center, Chicago, IL

Introduction

Objective: The mission of Argus Cognitive is to develop a non-intrusive, fully automated system that increases accessibility, affordability and objectivity of behavioral health care for common mental health and neurological conditions. We present advancements in our application of AI methods to detect and monitor social-communication behaviors including speech and gestural communication. We propose a minimally intrusive ML model to simultaneously monitor multiple key social behaviors in individuals with ASD and related neurodevelopmental disorders (NDDs).

Procedures

Argus-MDS system architecture:

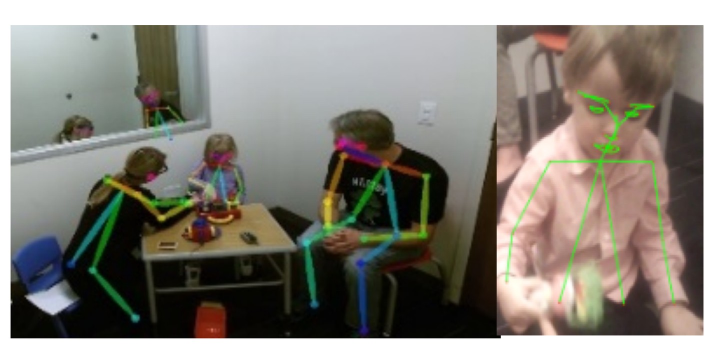

Figure 1. Video feeds from one of the cameras (left) and Tobii2 smart glasses (right). Colored lines show body and face estimations.

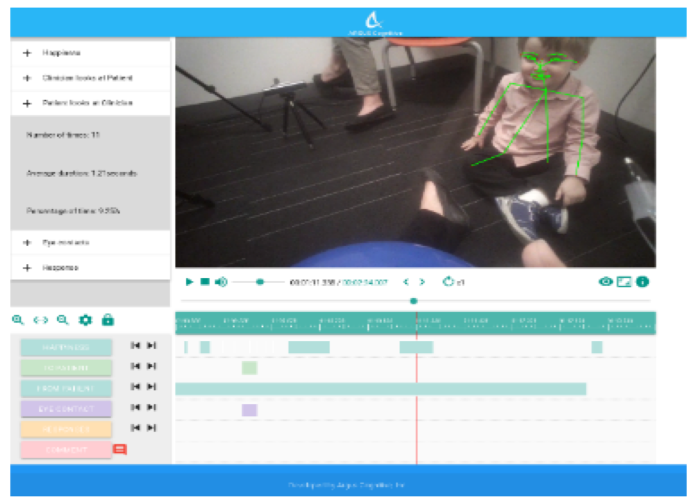

Figure 2. Data analysis via ARGUS-MDS interactive report

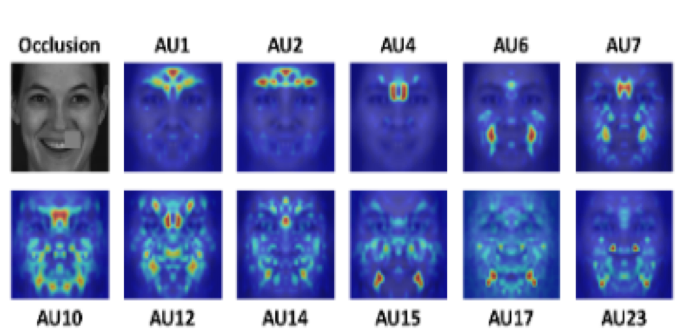

Figure 3. Facial action unit sensitivity maps

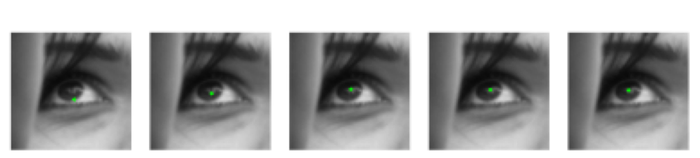

Figure 4. Eye gaze: Iterative gaze refinement process

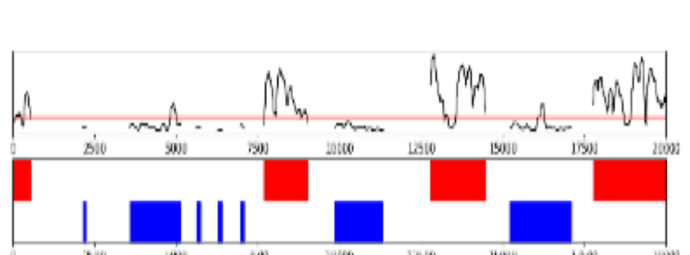

Figure 5. Speech analysis: Pitch variance and pause length

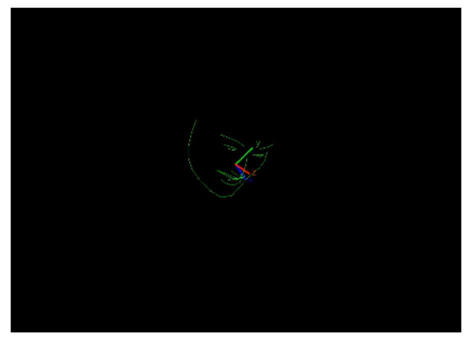

Figure 6. 3D head pose detection for head gestures (e.g. nodding)

Table 1: Reliability of biometric data with manual coding

ASD behaviors	Quality metric
Facial emotions	0.65 mean skew-normalized F1 scores for Action Units 1, 2, 4, 6, 7, 12, 15, and 23 (frame by frame) 0.85 mean skew-normalized F1 score for Action Units 6 and 12 alone (frame by frame)
Gaze estimation	3.5 degree angular RMSE for the patient
Visual patient detection	0.95 precision in a frame-by-frame manner
Acoustic patient detection	0.88 F1 score (0.86 precision, 0.89 recall)

Table 2: Participant Demographics

	Comparative Analysis	
	TD n=8	NDD n=16
Age in Years (2-10 years)	6.6 (3.0)	6.7 (2.9)
Sex (N, % male)	3, 63%	8, 50%
ADOS-2 CSS (mean, SD)	2.0 (2.2)	5.8 (2.9)

2:1 Matching: NDD (Neurodevelopmental disorders including ASD, PMS) & TD (Typically Developing)

Sensory, touching self & others

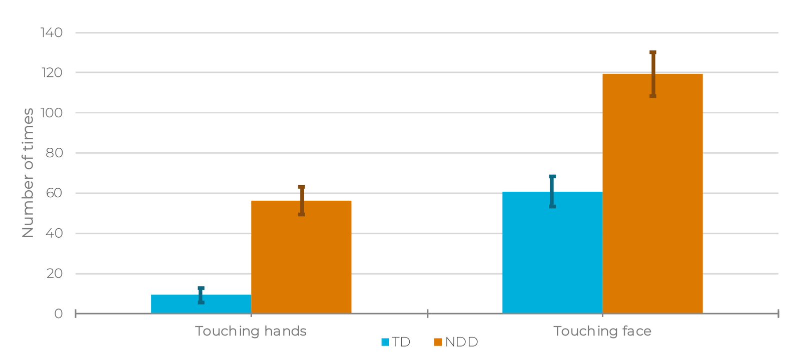

Figure 7. Sensory motor behaviors

Gestures

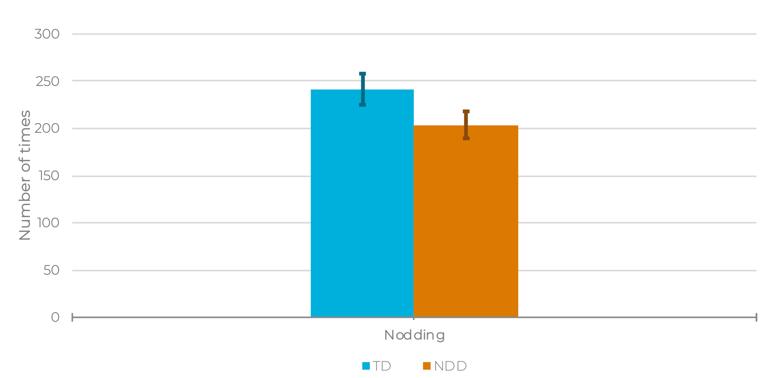

Figure 8. Nodding

Eye gaze

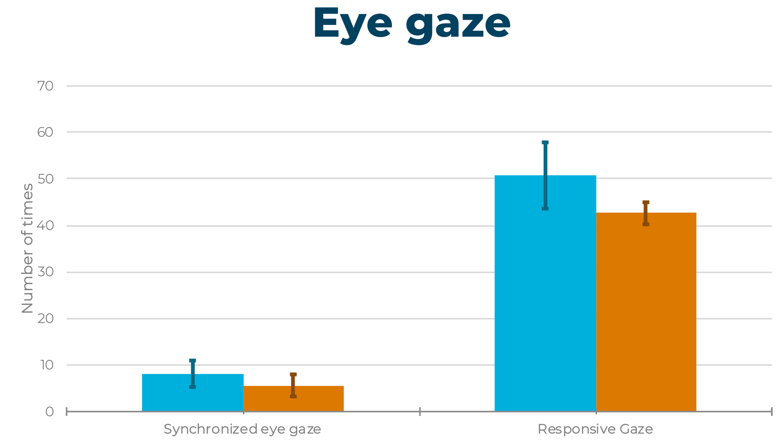

Figure 9. Interpersonal Gaze Metrics

Vocalizations

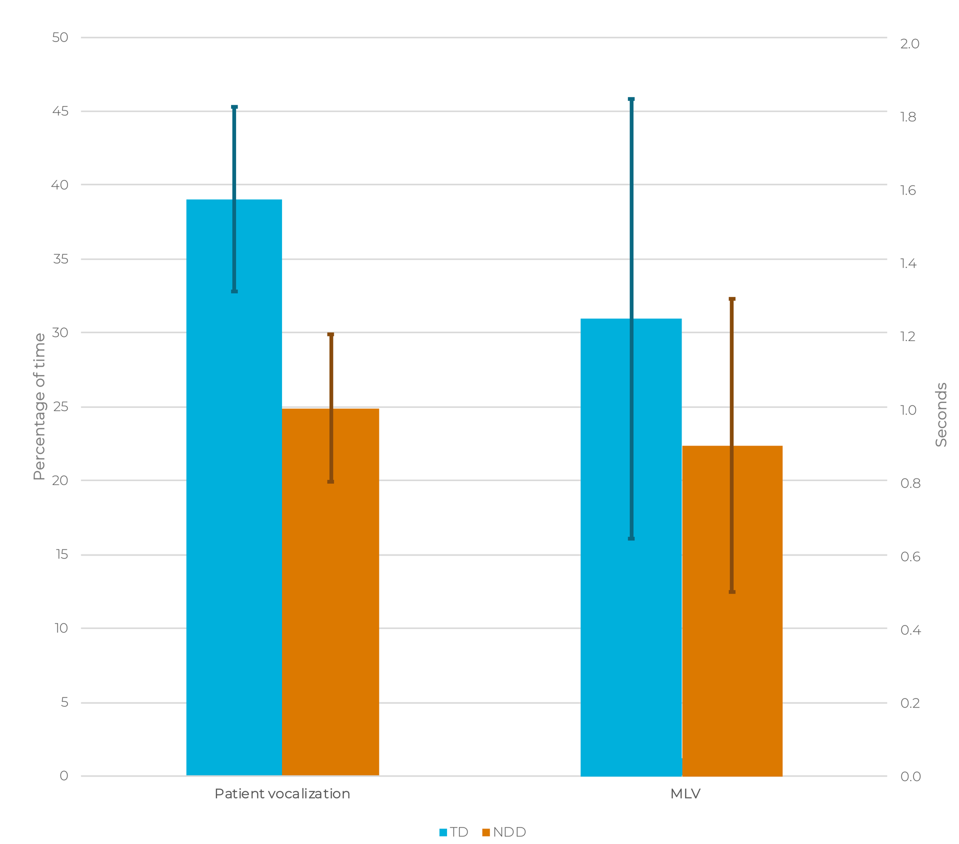

Figure 10. Patient vocalization and speech mean length vocalizations

Vocal exchanges

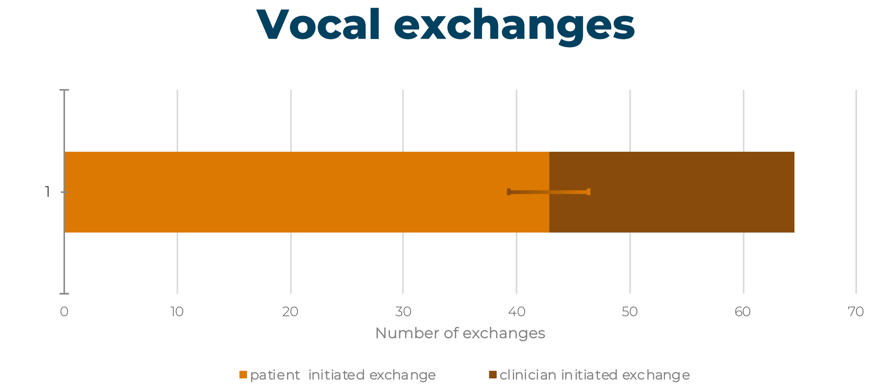

Figure 11. Vocal exchanges in NDD patients

Analyses

- Biometric data analyses:
  - Our system works as a pipeline. We start with estimating body poses and facial landmarks (Fig 1.)
  - We automatically detect the patient based on age (Fig 2.)
  - We measure the following behaviors for the detected patient:
    - Facial expressions (positive and negative valence) (Fig 3.)
    - Head and hand gestures, and sensory motor behaviors (Fig 6.)
    - Gaze behavior (eye contact, synchronized gaze, response time) (Fig 4.)
    - Speech (pitch variance, length of pauses, length of vocalizations, vocal exchanges) (Fig 5.)
- We validate the quality of the system's components and reported behavior against manual coding and external data sets (Table 1)
- Our system is non-obtrusive, we do not attach any devices to the patient and we observe natural interaction with a clinician (ADOS-2)
- Clinical data is a descriptive comparison between children with NDDs and typically developing children (TD) (Table 2). Nonparametric Mann-Whitney U and Spearman's rho are reported.

Results & Discussion

- Reliability of ML analyses (Table 1)
- In speech analyses, summary variables differentiate between NDD & TD samples including percentage (p=.015) and mean length (p=.027) of vocalizations. (Fig.10.)
- Sensory motor behaviors (i.e. touching) are reliably assessed and differ between TD & NDD (p<.001) (Fig. 7.)
- In this small sample, moderate correlations found between synchronized eye gaze and ADOS2 Calibrated Severity Score (CSS, r==.47, p=.06) (Fig. 9.)
- Recently added gesture, sensory motor, and speech components show promise in discriminating between TD and NDD samples during ADOS2 sessions. Synchronized eye gaze continues to show promise as correlate of functional impairment. (Fig. 8., 11.)

Future Directions

- Training system in broader neurodevelopmental and psychiatric samples (e.g. at-risk schizophrenia, ADHD)
- To establish psychometrics and provide data on sensitivity to change in target behaviors
- Support development of ARGUS-MDS into a comprehensive, scalable, digital treatment progress indicator (TPI, Tuso, 2014) for NDDs

RUSH UNIVERSITY MEDICAL CENTER

ARGUS Cognitive