Gaze-Wasserstein: A Quantitative Screening Approach to Autism Spectrum Disorders

Kun Woo Cho¹, Feng Lin¹, Chen Song¹, Xiaowei Xu¹, Michelle Hartley-McAndrew², Kathy Ralabate Doody³, and Wenyao Xu¹

Department of Computer Science and Engineering, University at Buffalo (SUNY)¹
Children’s Guide Foundation Autism Spectrum Disorder Center²
Department of Exceptional Education, Buffalo State University (SUNY)³
Autism Spectrum Disorder

A neurodevelopmental condition that is defined by concerns in three major domains:

- **Social Interaction**
  - Abnormality in nonverbal behaviors related to eye-to-eye gaze, facial expression, and body gesture.

- **Communication**
  - Lack of spontaneous social imitative play

- **Behavior**
  - Persistent and intensive preoccupation with parts over whole
Estimated Prevalence of Child Autism

- According to CDC, about 1 in 68 children in the US has been diagnosed with ASD in 2016.
- Significant gap between the number of children diagnosed with autism and the number receiving services.
- Limitations of current diagnosis approach for ASD.
Current Diagnostic Practice

- **Low accessibility**
  - Autism Diagnostic Observation Schedule (ADOS)
    - Specialized clinical setting.
      - High costs: $3,095 per kit.
      - Long duration: 30-45 minutes per modules.
    - Trained professional.

- **Subjectiveness**
  - Dependence on the expertise and experience of physicians

---

Late Screening

Delayed Treatment
Why early intervention of ASD is needed?

- **Early delivery of appropriate medical care**
  - Early treatment can encourage child’s brain to reroute around faulty neural pathways.

- **Pre-emptive education planning**
  - Improvement in autism symptoms, particularly with building social and communication skills.

- **Provision for family support**
  - Martial: High Divorce Rate of 23.5%.
  - Financial: 14% more family expense.
Prior works show gaze is promising marker of ASD.

Behavior:
- Abnormality in visual function of individuals with ASD on the visual perception stimulus (Dakin et al).

Social Interaction:
- Abnormality in dynamic gaze pattern when social scene is shown (Song et al).

References:
Rationale of our system: *Gaze-Wasserstein*

- We propose a gaze pattern-based ASD screening method.

**Rationale:**

- Can provide **objective** gaze-pattern-based measurements for home-based ASD screening.
- Can be easily **deployed on any mobile technologies** with a front camera.
- Cost-effective, fast, and highly accessible.
Experimental Details

- **Eye tracker:**
  - Tobii EyeX
  - Sampling frequency: 80Hz
  - Angel resolution: 0.1 degree

- **Participants:**
  - 32 children ranging in age from 2 to 10
  - 16 children with ASD and 16 typically developing children

- **Experimental Setting:**
  - 8 visual stimuli: 4 social scene and 4 non-social scene
  - 5 sec duration for each stimulus
  - 2 sec black screen on the monitor between two consecutive stimuli
    - To refresh the participant’s gaze response
    - To make each stimulus independent
Experiment Method

- **Leave-one-out-cross-validation** used to estimate how accurately a predictive model will perform in practice.

- **Classification with k nearest neighbors**
  - Modified 1\textsuperscript{st} Wasserstein is our distance function.
  - K equals 3.
Social(S)-Tech(T) Challenges

- How to design visual stimuli? (S)
- Which is the best similarity matching metric?
- How long is the exposure duration for visual stimuli?
Two categories of visual stimuli are designed based on previous literature review [1].

- **Social Scene (SS)**: Images that are illustrating social interactions of more than one human figure.
- **Non-social Scene (NSS)**: Images that are depicting a single non-human object.
- Background is visually clean to avoid any unintentional distraction.

Impact of Stimuli Type on System Performance

- **F-score Accuracy**
  - Harmonic mean of precision and recall used to avoid unbalanced accuracy measurement.

- **Equal Error Rate**
  - Rate at which acceptance error (FPR) and rejection error (FNR) are equal.
  - Accuracy of the system is high when EER is low.

### F-score for Social and Non-Social Scene (%)

<table>
<thead>
<tr>
<th>Scene Type</th>
<th>ASD</th>
<th>TD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Scene</td>
<td>93.8</td>
<td>94.12</td>
</tr>
<tr>
<td>Non-social Scene</td>
<td>88.84</td>
<td>90.21</td>
</tr>
</tbody>
</table>

### EER for Social and Non-Social Scene

<table>
<thead>
<tr>
<th>Scene Type</th>
<th>ASD-SS</th>
<th>TD-SS</th>
<th>ASD-NSS</th>
<th>TD-NSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Scene</td>
<td>0.034</td>
<td>0.032</td>
<td>0.092</td>
<td>0.075</td>
</tr>
<tr>
<td>Non-social Scene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Gaze Distribution on Social Scenes

TD subjects

ASD subjects
Social(S)-Tech(T) Challenges

- How to choose visual stimuli? (S)
- Which is the best similarity matching metric? (T)
- How long is the exposure duration for visual stimuli?
Gaze Distribution in Each Group

- **Similar Distribution** Among TD Subjects
- **Dissimilar Distribution** Among ASD Subjects

Gaze pattern of ASD subject is either very narrow or widely distributed.
- Persistent preoccupation with parts over whole ➔ Fixed Range of View
- Inability to focus on the social scene ➔ Wide Distribution
**Modified 1\textsuperscript{st} Wasserstein Distance**

- M to N matching.
- Modified to avoid **partial matching** by adding a constraint [1], otherwise some ASD may misclassify as TD.

  ![Diagram](image)

- Insensitive to noise [2] caused by accidental eye movement or possible error from eye tracking device.

---


Social(S)-Tech(T) Challenges

- How to choose visual stimuli? (S)
- Which is the best similarity matching metric? (T)
- How long is the exposure duration for visual stimuli? (S/T)
Social-Tech dilemma

- Longer duration → better classification
- Low user compliance → shorter duration
Optimization of Screen Time Efficiency

- It is important because child participant may have low compliance and limited attentions in the study.

- Accuracy at 26-second screen time for Social Scene reaches over 93%.
  - Compared to duration of ADOS (30-45 min), our method is 69-104 times faster.
Evaluation Results

- Using gaze pattern and social scene stimuli, our method can achieve f-score accuracy of 93.96% and equal error rate of 3.30%.

<table>
<thead>
<tr>
<th>All Stimuli</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>EER (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD</td>
<td>92.19 ± 7.86</td>
<td>95.41 ± 3.07</td>
<td>3.40 ± 2.67</td>
</tr>
<tr>
<td>TD</td>
<td>95.31 ± 3.13</td>
<td>92.93 ± 6.52</td>
<td>3.19 ± 2.06</td>
</tr>
<tr>
<td>F-measurement</td>
<td>93.75 ± 2.55</td>
<td>94.17 ± 2.08</td>
<td>3.30 ± 1.28</td>
</tr>
</tbody>
</table>

- In order to perform Gaze-Wasserstein, only two technologies are needed:
  - Monitor to display visual stimuli.
  - Hardware containing the eye-tracking system.

Both can be provided by smartphones [1].

Acknowledgement

- UB CSE
  - Feng Lin
  - Chen Song
  - Wenyao Xu

- UB Pediatrics
  - Michelle Hartley-McAndrew

- UB ECRC
  - X. Christine Wang

- Children’s Guild Foundation Autism Spectrum Disorder Center
  - Jana Mertz

- Buffalo State
  - Kathy Ralabate Doody

- Google
  - Changzhan Gu

- UVa
  - Yan Zhuang

- Texas Tech U
  - Changzhi Li

- U Notre Dame
  - Xiaowei Xu