Improving Autism Detection with Multimodal Behavioral Analysis

William Saakyan[1], Matthias Norden[1], Lola Eversmann[2], Simon Kirsch[3], Muyu Lin[2], Simon Guendelman[2], Isabel Dziobek[2], and Hanna Drimalla[1]

[1] Center for Cognitive Interaction Technology (CITEC), Bielefeld University
Corresponding: drimalla@techfak.uni-bielefeld.de
[2] Institute of Psychology, Humboldt University of Berlin
[3] Department of Psychiatry and Psychotherapy, Medical Center-University of Freiburg

Abstract. Due to the complex and resource-intensive nature of diagnosing Autism Spectrum Condition (ASC), several computer-aided diagnostic support methods have been proposed to detect autism by analyzing behavioral cues in patient video data. While these models show promising results on some datasets, they struggle with poor gaze feature performance and lack of real-world generalizability. To tackle these challenges, we analyze a standardized video dataset comprising 168 participants with ASC (46% female) and 157 non-autistic participants (46% female), making it, to our knowledge, the largest and most balanced dataset available. We conduct a multimodal analysis with a primary focus on gaze behaviour, complemented by facial expressions, voice prosody, head motion, heart rate variability (HRV). Addressing previous limitations in gaze modeling, we introduce novel statistical descriptors that quantify variability in eye gaze angles, improving gaze-based classification accuracy from 64% to 69% and aligning computational findings with clinical research on gaze aversion in ASC. Using late fusion, we achieve a classification accuracy of 74%, demonstrating the effectiveness of integrating behavioral markers across multiple modalities. Our findings highlight the potential for scalable, video-based screening tools to support autism assessment. To facilitate reproducibility, we share our code on GitHub: https://github.com/mbp-lab/miccai25_sit_autism_classification.

Keywords: Autism Detection · Machine Learning · Multimodal Analysis · AI in Healthcare · ASC

1 Introduction

Impairments in social communication are a key characteristic of Autism Spectrum Condition (ASC), affecting an individual's ability to interpret and respond to non-verbal cues such as facial expressions, eye contact, and vocal tone [1]. The clinical diagnosis of ASC currently relies on subjective assessments, including standardized instruments (e.g., Autism Diagnostic Observation Schedule (ADOS), the Autism Diagnostic Interview-Revised (ADI-R) [9]), which are time

consuming and depend on the availability of experts. This often leads to long waiting times and potential misdiagnoses, particularly in adults and females who may present atypical symptoms [23,19]. This highlights the urgent need for objective, scalable, and accessible tools to support autism screening and behavioral assessment.

Recent advancements in machine learning (ML) have enabled the automated analysis of non-verbal behaviors to assist with ASC detection via video-based approaches [5]. Studies have analyzed facial expressions [8], gaze behavior [4,20], and prosodic features [7] to identify ASC-related patterns. These approaches can complement existing diagnostic frameworks and facilitate early detection, remote screening, and large-scale behavioral studies. However, despite this progress, significant challenges remain.

Most datasets used for computer-aided ASC assessment include children, aiming for early diagnosis [21,6]. However, ASC is often undiagnosed until adulthood, particularly in females and individuals with milder traits [16]. Datasets with adult populations remain scarce, and existing ones, such as [15], have limited sample sizes (58 individuals) and feature sets. Moreover, most datasets are collected in highly controlled laboratory environments, limiting model robustness in naturalistic settings. Additionally, many data types, such as neuroimaging, molecular, and genetic data, require expensive equipment, creating development and deployment challenges of AI-based tools for ASC assessment.

Social interaction is multimodal, involving gaze, facial expressions, vocal prosody, and body movements [25]. While multimodal ML models have improved classification accuracy [28,10], many over-rely on facial features, underutilizing other modalities, such as gaze behavior.

Atypical gaze is a well-established ASC marker [24,18]. Reduced eye contact and increased gaze aversion have been observed across lab-based and naturalistic settings [26]. While some studies have successfully used eye-tracking devices to analyze gaze behavior in ASC [17], these approaches often rely on specialized hardware or controlled experimental tasks rather than naturalistic social interactions. Webcam-based interaction analyses, such as [12,27], have reported poor performance for gaze-based ASC classification, likely due to simplistic descriptors that fail to capture gaze behavior in relation to social stimuli.

Several studies suggest heart rate variability (HRV) as a biomarker for ASC due to its link to autonomic nervous system function. Research indicates that individuals with ASC often exhibit autonomic dysregulation, characterized by both hyperarousal and hypoarousal states at rest, which may impact their ability to engage with social environments and regulate sensory input[2]. [30,31] found significantly lower resting-state HRV in both adults and children with ASC. [14] applied machine learning to HRV data, achieving an AUC of 0.89 for ASC classification. While HRV shows promise as a non-invasive biomarker, its role in ASC remains complex, with findings influenced by measurement context, heterogeneity in study methodologies, and individual variability[2].

To address the challenges of subjective ASC diagnosis, limited dataset availability for adults, and the underutilization of key social interaction markers, we

improve computer-aided autism detection by evaluating multimodal behavioral markers using the largest adult social interaction dataset to date. To establish this dataset, we used the Simulated Interaction Task (SIT) [11,12], a standardized paradigm designed to elicit social behaviors by presenting participants with a video-recorded conversational partner. Unlike prior studies using lab-based, child-focused datasets, we included adult participants in both clinical and home settings, providing a more valid dataset. To our knowledge, this is the largest and most balanced dataset available.

Our approach primarily introduces enhanced gaze behavior descriptors alongside facial expression, head movement, voice prosody, and heart rate variability (HRV) features to improve diagnostic performance. We systematically evaluate unimodal and multimodal feature combinations, emphasizing gaze features due to their significant improvement and clinical relevance in ASC detection. By combining behavioral analysis with computational modeling, our work helps bridge the gap between computer-aided ASC detection and clinical practice, contributing to the development of scalable, non-invasive screening tools for diverse populations.

2 Methods

2.1 Dataset

We expanded our video dataset using the Simulated Interaction Task (SIT) paradigm [12], combining newly collected participants with those from a previous study. The final dataset includes 168 individuals with ASC (46% female) and 157 controls (46% female), recorded across clinical ($n = 254$) and home ($n = 71$) environments, representing one of the largest and most gender-balanced collections for adult social interaction analysis to date. Lab study participants were recruited for a larger research project (number DRKS00017817). Inclusion criteria: age 18–65, IQ $\geq$80, fluency in German, and stable or no pharmacotherapy. Exclusion criteria included psychiatric comorbidities of schizophrenia, psychosis, severe depression, acute manic episodes within bipolar disorder, and acute suicidality. ASC diagnoses were confirmed by licensed clinicians using ICD-10 criteria, while the non-autistic participants reported no psychiatric diagnoses. Home-study participants were recruited via clinics, therapy groups, and online postings. ASC diagnoses were confirmed via medical records. Comorbid conditions (e.g., ADHD, Social Anxiety Disorder, Major Depression) were documented but not excluded, allowing us to capture ASC-related interaction traits in a clinically representative sample. The studies were approved by the respective ethical committee of Humboldt University of Berlin and the Medical Center-University of Freiburg (approval 20-1144_3 and 2021-20, https://doi.org/10.1186/s13063-021-05205-9).

Procedure Participants completed the SIT application on a computer in lab or home settings. The fully automated procedure began with head positioning for facial landmark calibration. The conversation scenario included three phases:

Table 1. Gender (male, female, diverse) and age distributions by setting and group.

Study Setting	ASC			Non-ASC		ASC Age		Non-ASC Age		Total
	M	F	D	M	F	Median	Range	Median	Range	
Home	13	13	-	24	21	36	18-57	27	18-60	71
Lab	74	65	3	60	52	34	18-63	35	18-64	254
Total	87	78	3	84	73	-	-	-	-	325

Meal preparation ("Neutral"), Favorite foods ("Joy"), and Disliked foods ("Disgust"). Each phase consisted of two interactions: the actress speaking while the participant listened, followed by the participant speaking while the actress displayed empathic listening.

2.2 Preprocessing

To ensure consistency with previous research [27], we applied the same video preprocessing pipeline: For videos with frame rates between 15 and 25 FPS, we applied linear interpolation to achieve a uniform 30 FPS frame rate. Videos with frame rates below 15 FPS were excluded from analysis due to motion artifacts.

2.3 Features

We extracted non-verbal features from following modalities: facial expressions, gaze, head movement, paralinguistics, and HRV, using open-source libraries.

Feature extraction was performed across six interaction phases, corresponding to three emotion-specific segments (neutral, joy, disgust) and two speaking roles (participant speaking and participant listening). Audio features were extracted only from participant-speaking segments, while HRV signals were extracted primarily during listening segments to minimize motion artifacts.

Visual Visual features were extracted using OpenFace 2.2 [3], which detects Action Units based on the Facial Action Coding System, head position (rotation angles: pitch, yaw, roll), and gaze direction (angle x and y). Frames with a detection confidence below 75% were excluded. Participants with more than 10% invalid frames were removed (five in total).

Facial expression: OpenFace computes 18 AUs, each with a presence score (binary) and an intensity score (0–5). We calculated the mean, median, and standard deviation of the AU intensities in each of the six interaction phases, as well as the mean AU presence. Additionally, we computed the AU onset frequency, which is the number of activation onsets (transitions from 0 to 1) per phase.

Head movement metrics, including velocity, acceleration, and stability durations, were calculated as in previous work [27], along with the mean and standard deviation of yaw and roll rotation angles. Additionally, we included the pitch angle and, to mitigate potential gender and height biases, normalized the values

based on each participant's median pitch angle. To provide a robust measure of spread, we computed the interquartile range (IQR) alongside other summary statistics. Furthermore, we implemented a nod detection algorithm that scans the pitch trajectories for characteristic downward-upward sequences occurring within 1.5-second windows.

Gaze behavior: Following [27], we computed gaze movement velocity, acceleration, saccade amplitude, and fixation duration, as well as the mean and standard deviation of the x angle. Additionally, we included statistics for the y angle, which was normalized relative to each participant's median y angle to mitigate height-related biases. To move beyond angle-based gaze descriptors and capture socially meaningful gaze behavior, we implemented a geometric transformation that maps gaze angles from the webcam coordinate system to a screen-centered coordinate space. For lab participants, this transformation involved measuring camera-to-screen distance, eye-to-screen distance, and screen resolution to determine gaze position relative to the screen. For home participants, gaze position was approximated. From these transformed coordinates, we computed: Mean screen fixation time (proportion of frames with gaze directed toward the screen area containing the actress's face), Number of off-screen fixations (frequency of gaze shifts beyond screen boundaries, indicating gaze aversion), Euclidean distance from the screen center (mean, standard deviation, skewness, kurtosis, minimum, and maximum distance from the actress's face).

Audio Paralinguistic features from the voice were extracted using OpenSmile 3.0 [13] with the eGeMAPSv02 feature set [33]for each participant-speaking phase. Extracted features include pitch (mean, variance), intensity, shimmer, jitter, harmonic-to-noise ratio (HNR), and formant frequencies.

Heart rate HR features were extracted using the rPPG Toolbox [22], applying face cropping and the Plane-Orthogonal-to-Skin (POS) method to estimate remote photoplethysmographic (rPPG) signals. HeartPy was used to extract HR features from the rPPG signals. HR signals were extracted during listening phases to minimize motion artifacts. The rPPG Toolbox failed to extract certain feature values for 18 participants. We imputed the missing values using the median from the training set. Features included mean HR, HR variability (HRV), root mean square of successive differences (RMSSD), and low-frequency to high-frequency power ratio (LF/HF).

2.4   Classification and analysis

We used XGBoost, a gradient boosting decision tree model, which shows its effectiveness in handling structured tabular data[29]. Model parameters were set to default XGBoost settings, to ensure comparability with [27]: gbtree as booster, learning rate of 0.3, maximum depth of tree was 6, XGBoost version 2.0.3. The classification task aimed to differentiate between individuals with ASC

and non-autistic individuals using the extracted multimodal features. We calculated performance metrics, including accuracy, precision and recall and compared unimodal models, where each modality was evaluated separately, with multimodal fusion approaches. Early fusion involved concatenating all extracted features into a single feature vector. In the late fusion approach, we combined the probability scores from unimodal models and classified them using logistic regression, incorporating polynomial features (degree 2) to capture non-linear relationships and interactions. All analyses reported in this paper were evaluated using a participant-based Leave-One-Out Cross-Validation [32] approach.

To gain deeper insights beyond overall model performance, we conducted several follow-up analyses. First, we applied SHapley Additive exPlanations (SHAP) to identify the key features contributing to ASC classification. Second, we performed a statistical analysis of misclassifications to determine whether errors were influenced by dataset source (home vs. lab), participant gender, or Autism Spectrum Quotient (AQ) score. Lastly, we assessed the contribution of each modality by systematically excluding them in late fusion setting.

3   Results and Discussion

3.1   Model performance

We conducted a comprehensive uni- and multimodal video analysis using machine learning on a large dataset of 325 participants, balanced by gender and collected across diverse settings (home and clinical environments). Table 2 presents classification performances, while Figure 1 illustrates the corresponding ROC curves. Our refinement of gaze features led to the most notable improvement, increasing accuracy by 5 percentage points compared to previous works using a similar approach. Additionally, enhanced feature engineering for the head modality resulted in a 2-point accuracy gain. Multimodal late fusion achieved the highest classification accuracy of 74%, outperforming the feature set of prior work by 6 percentage points.

Table 2. ASC classification performance for different modalities using the previous[27] and current feature sets. Bold values indicate the highest accuracy per modality.

Modality	Previous Feature Set			Current Feature Set		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Multimodal (late)	0.68	0.69	0.68	**0.74**	0.76	0.74
Multimodal (early)	0.67	0.67	0.71	**0.68**	0.68	0.71
Facial expression	**0.70**	0.70	0.73	**0.70**	0.71	0.73
Audio	**0.66**	0.67	0.66	**0.66**	0.67	0.66
Gaze	0.64	0.66	0.64	**0.69**	0.70	0.68
Head	0.63	0.64	0.66	**0.65**	0.66	0.68
HR	-	-	-	**0.57**	0.58	0.63

3.2   Gaze Behavior Analysis

Prior computational studies have reported inconsistent results regarding gaze behavior in ASC [12,27], despite psychological research suggesting its diagnostic relevance [18,20]. These discrepancies are likely due to simplistic angle-based descriptors that fail to fully capture gaze aversion patterns. In our analysis, gaze behavior exhibited the largest performance gain among unimodal models, demonstrating the effectiveness of our new screen-centered descriptors in capturing gaze aversion. Figure 2 presents the SHAP analysis, identifying gaze distance variability from the screen center as the most influential feature. On the right side, we visualize eye gaze angles projected onto the screen surface, further illustrating these differences. Our results confirm that ASC participants exhibit significantly greater gaze variability than non-autistic individuals ($+63.5\%$), particularly in the "Disgust" phase ($p = 0.003, r = -0.215$).
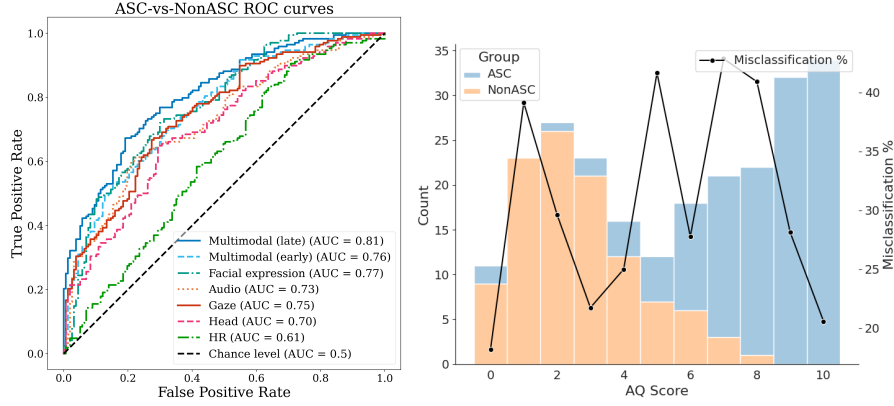

Fig. 1. Left: ROC curves. **Right**: Distribution of AQ scores and misclassification rates.

3.3   Misclassification Analysis

We analyzed **misclassifications** for potential biases. Chi-square tests revealed no significant differences in classification errors based on gender or recording environment ($p \geq 0.05$). There was no evidence of an influence of gender or recording setting (lab vs. home) on model performance, suggesting a degree of generalizability and potential real-world applicability.

Given the overlap between ASC and non-autistic individuals in middle AQ score ranges, we expected more frequent misclassification for participants with intermediate scores. However, a Mann–Whitney U test comparing the AQ score distributions of correctly and incorrectly classified participants revealed no significant difference ($p \geq 0.05$). Figure 1 (right) shows the distribution of late-fusion model misclassifications across AQ scores.
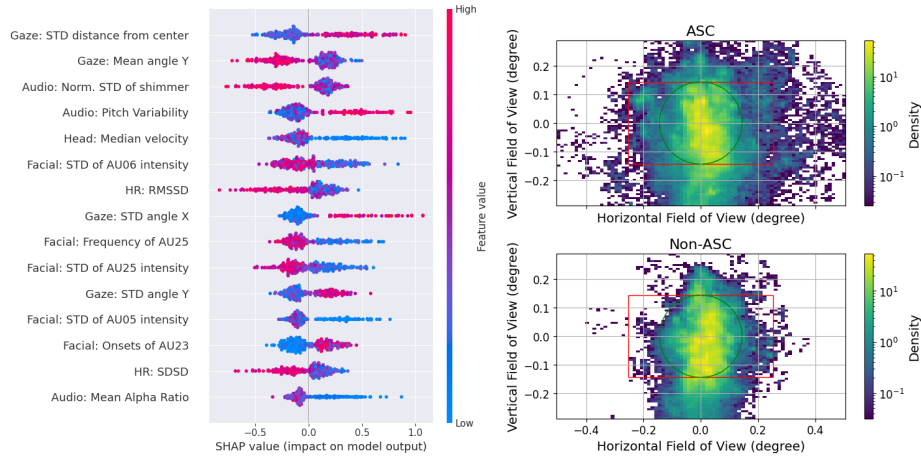
Fig. 2. Left: SHAP values for the multimodal early fusion model. Each dot represents a feature's impact on an individual prediction. **Right**: Eye gaze angle projections on the screen surface for ASC and non-autistic participants during the passive parts. The red rectangle represents the screen area, where the actress was displayed. Vertical and horizontal axes represent fields of view.

3.4   Contribution of each modality

We evaluated each modality's contribution by measuring performance drops after removal. Excluding eye gaze or facial expressions led to the largest accuracy reduction (-4 percentage points), highlighting their importance in the model's predictions. This aligns with prior research emphasizing atypical gaze behavior [18,4,20] and reduced facial expressivity [8] as core characteristics of ASC. Despite its low standalone accuracy, removing HR variability still decreased performance by -2 percentage points, supporting evidence that autonomic regulation differs in ASC individuals [7]. A similar drop for audio features reinforces the relevance of prosodic differences for ASC detection [7]. In contrast, removing head pose data had no effect, suggesting minimal contribution. Future work should further explore modality interactions and more sophisticated fusion strategies to enhance the interpretability and robustness of multimodal assessments.

3.5   Conclusion

This study contributes to computer-aided ASC assessment by introducing refined gaze descriptors as our primary technical contribution, alongside complementary facial expression, head movement, voice prosody, and HRV features within a multimodal fusion framework. We also provide an in-depth analysis of multimodal behavioral markers in standardized interactions within a large-scale, adult-focused dataset. Our results highlight the value of precise gaze features in enhancing ASC classification accuracy.

While our study analyzed aggregated behavioral summaries across interaction phases, future work could use recurrent neural networks (LSTMs) or transformer-based architectures to capture fine-grained time-dependent social interaction dynamics in gaze shifts, facial expressivity, and vocal prosody. Improving webcam-based HR extraction, including rPPG advancements, is another promising avenue. Although our model distinguishes behavioral differences between individuals with and without ASC, clinical use involves differentiating ASC from other conditions with overlapping social impairments, such as social anxiety or personality disorders. Given the prevalence of comorbidities in mental health patients, as seen in our dataset, future research should explicitly extend into transdiagnostic settings. This will involve examining how multimodal interaction markers vary across overlapping conditions, thereby enhancing diagnostic specificity and applicability. Nevertheless, our findings demonstrate the potential of multimodal behavioral analysis for ASC detection.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders: DSM-5. American Psychiatric Association, 5th edition edn.
2. Arora, I., Bellato, A., Ropar, D., Hollis, C., Groom, M.J.: Is autonomic function during resting-state atypical in Autism: A systematic review of evidence **125**, 417–441
3. Baltrusaitis, T., Zadeh, A., Lim, Y.C., Morency, L.P.: OpenFace 2.0: Facial Behavior Analysis Toolkit. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). pp. 59–66
4. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I.: The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism **42**(2), 241–251
5. Belen, Ryan Anthony J., p.u., Bednarz, T., Sowmya, A., Del Favero, D.: Computer vision in autism spectrum disorder research: A systematic review of published studies from 2009 to 2019 **10**(1), 1–20
6. Billing, E., Belpaeme, T., Cai, H., Cao, H.L., Ciocan, A., Costescu, C., David, D., Homewood, R., Garcia, D.H., Esteban, P.G., Liu, H., Nair, V., Matu, S., Mazel, A., Selescu, M., Senft, E., Thill, S., Vanderborght, B., Vernon, D., Ziemke, T.: The DREAM Dataset: Supporting a data-driven study of autism spectrum disorder and robot enhanced therapy **15**(8), e0236939
7. Bone, D., Black, M.P., Ramakrishna, A., Grossman, R., Narayanan, S.S.: Acoustic-prosodic correlates of 'awkward' prosody in story retellings from adolescents with autism. In: Interspeech 2015. pp. 1616–1620. ISCA

8. Briot, K., Pizano, A., Bouvard, M., Amestoy, A.: New Technologies as Promising Tools for Assessing Facial Emotion Expressions Impairments in ASD: A Systematic Review **12**, 634756

9. Bölte, S., Poustka, F.: [psychodiagnostic instruments for the assessment of autism spectrum disorders] **33**(1), 5–14

10. Derbali, M., Jarrah, M., Randhawa, P.: Autism Spectrum Disorder Detection: Video Games based Facial Expression Diagnosis using Deep Learning **14**(1)

11. Drimalla, H., Landwehr, N., Baskow, I., Behnia, B., Roepke, S., Dziobek, I., Scheffer, T.: Detecting Autism by Analyzing a Simulated Social Interaction. In: Berlingerio, M., Bonchi, F., Gärtner, T., Hurley, N., Ifrim, G. (eds.) Machine Learning and Knowledge Discovery in Databases. pp. 193–208. Lecture Notes in Computer Science, Springer International Publishing

12. Drimalla, H., Scheffer, T., Landwehr, N., Baskow, I., Roepke, S., Behnia, B., Dziobek, I.: Towards the automatic detection of social biomarkers in autism spectrum disorder: Introducing the simulated interaction task (SIT) **3**(1), 1–10

13. Eyben, F., Wöllmer, M., Schuller, B.: Opensmile: The munich versatile and fast open-source audio feature extractor. In: Proceedings of the 18th ACM International Conference on Multimedia. pp. 1459–1462. MM '10, Association for Computing Machinery

14. Frasch, M.G., Shen, C., Wu, H.T., Mueller, A., Neuhaus, E., Bernier, R.A., Kamara, D., Beauchaine, T.P.: Brief Report: Can a Composite Heart Rate Variability Biomarker Shed New Insights About Autism Spectrum Disorder in School-Aged Children? **51**(1), 346–356

15. Georgescu, A.L., Koehler, J.C., Weiske, J., Vogeley, K., Koutsouleris, N., Falter-Wagner, C.: Machine Learning to Study Social Interaction Difficulties in ASD **6**

16. Huang, Y., Arnold, S.R., Foley, K.R., Trollor, J.N.: Diagnosis of autism in adulthood: A scoping review **24**(6), 1311–1327

17. Islam, M.F., Manab, M.A., Mondal, J.J., Zabeen, S., Rahman, F.B., Hasan, M.Z., Sadeque, F., Noor, J.: Involution fused convolution for classifying eye-tracking patterns of children with Autism Spectrum Disorder **139**, 109475

18. Jones, W., Klin, A.: Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism **504**(7480), 427–431

19. Kirkovski, M., Enticott, P.G., Fitzgerald, P.B.: A Review of the Role of Female Gender in Autism Spectrum Disorders **43**(11), 2584–2603

20. Klin, A., Jones, W., Schultz, R., Volkmar, F., Cohen, D.: Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism **59**(9), 809–816

21. Li, J., Chheang, V., Kullu, P., Brignac, E., Guo, Z., Bhat, A., Barner, K.E., Barmaki, R.L.: MMASD: A Multimodal Dataset for Autism Intervention Analysis. In: Internation Conference on Multimodal Interaction. pp. 397–405. ACM

22. Liu, X., Narayanswamy, G., Paruchuri, A., Zhang, X., Tang, J., Zhang, Y., Sengupta, R., Patel, S., Wang, Y., McDuff, D.: rPPG-Toolbox: Deep Remote PPG Toolbox

23. McQuaid, G.A., Lee, N.R., Wallace, G.L.: Camouflaging in autism spectrum disorder: Examining the roles of sex, gender identity, and diagnostic timing **26**(2), 552–559

24. Minissi, M.E., Chicchi Giglioli, I.A., Mantovani, F., Alcañiz Raya, M.: Assessment of the Autism Spectrum Disorder Based on Machine Learning and Social Visual Attention: A Systematic Review **52**(5), 2187–2202

25. Nota, N., Trujillo, J.P., Holler, J.: Facial Signals and Social Actions in Multimodal Face-to-Face Interaction **11**(8), 1017

26. Papagiannopoulou, E.A., Chitty, K.M., Hermens, D.F., Hickie, I.B., Lagopoulos, J.: A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders pp. 1–23

27. Saakyan, W., Norden, M., Herrmann, L., Kirsch, S., Lin, M., Guendelman, S., Dziobek, I., Drimalla, H.: On Scalable and Interpretable Autism Detection from Social Interaction Behavior. In: 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). pp. 1–8. IEEE

28. Sarwani, I.S., Bhaskari, D.L., Bhamidipati, S.: Emotion-based Autism Spectrum Disorder Detection by Leveraging Transfer Learning and Machine Learning Algorithms **15**(5)

29. Shwartz-Ziv, R., Armon, A.: Tabular data: Deep learning is not all you need **81**, 84–90

30. Thapa, R., Alvares, G.A., Zaidi, T.A., Thomas, E.E., Hickie, I.B., Park, S.H., Guastella, A.J.: Reduced heart rate variability in adults with autism spectrum disorder **12**(6), 922–930

31. Thapa, R., Pokorski, I., Ambarchi, Z., Thomas, E., Demayo, M., Boulton, K., Matthews, S., Patel, S., Sedeli, I., Hickie, I.B., Guastella, A.J.: Heart Rate Variability in Children With Autism Spectrum Disorder and Associations With Medication and Symptom Severity **14**(1), 75–85

32. Webb, G., Sammut, C., Perlich, C., Horváth, T., Wrobel, S., Korb, K., Noble, W., Leslie, C., Lagoudakis, M., Quadrianto, N., Buntine, W., Getoor, L., Namata, G., Jin, J., Ting, J.A., Vijayakumar, S., Schaal, S., De Raedt, L.: Leave-One-Out Cross-Validation

33. Yang, M., Konan, J., Bick, D., Kumar, A., Watanabe, S., Raj, B.: Improving Speech Enhancement through Fine-Grained Speech Characteristics