Classifying Individuals with ASD Through Facial Emotion Recognition and Eye-Tracking

Ming Jiang[1], Sunday M. Francis[2], Diksha Srishyla[2], Christine Conelea[2], Qi Zhao[1,*] and Suma Jacob[2,*]

Abstract— **Individuals with Autism Spectrum Disorder (ASD) have been shown to have atypical scanning patterns during face and emotion perception. While previous studies characterized ASD using eye-tracking data, this study examined whether the use of eye movements combined with task performance in facial emotion recognition could be helpful to identify individuals with ASD. We tested 23 subjects with ASD and 35 controls using a Dynamic Affect Recognition Evaluation (DARE) task that requires an individual to recognize one of six emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) while observing a slowly transitioning face video. We observed differences in response time and eye movements, but not in the recognition accuracy. Based on these observations, we proposed a machine learning method to distinguish between individuals with ASD and typically developing (TD) controls. The proposed method classifies eye fixations based on a comprehensive set of features that integrate task performance, gaze information, and face features extracted using a deep neural network. It achieved an 86% classification accuracy that is comparable with the standardized diagnostic scales, with advantages of efficiency and objectiveness. Feature visualization and interpretations were further carried out to reveal distinguishing features between the two subject groups and to understand the social and attentional deficits in ASD.**

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a broad spectrum neurodevelopmental disorder characterized by difficulties in social interaction and communication, as well as restricted, repetitive, and stereotyped behaviors and interests [1]. As a core deficit, social impairment is one of the most studied aspects of ASD [2]. It is theorized that atypical viewing of socially relevant stimuli may contribute to these deficits. The addition of eye-tracking technologies and techniques within the field has dramatically enhanced our ability to investigate atypical visual attention in ASD objectively. Starting in infancy, the high saliency and importance of social information processing and emotion recognition can be observed [3]. Indeed, atypical visual processing was noted in toddlers (10–49 months) with ASD [4]. Similar findings that individuals with ASD dwelled longer on non-social versus social stimuli was also noted in adolescents with the disorder [5]. As noted in a systematic review by Black et al. [6], results have varied across studies. Previous research has reported decreased gazing upon the eye region either with or without an increase in attention towards other facial regions, such as the mouth [7]–[11]. In 2007, Spezio and colleagues combined the "Bubbles" task and eye-tracking [8], [9]. They were able to determine in an age- and IQ-matched sample of adults that individuals with ASD increased their gaze towards the mouth region and utilized this information to determine the emotion presented. Other studies reported a general decrease in gazing upon socially relevant areas with an increase in gazing upon areas outside of the face [12]. These differences in visual behavior do not necessarily reflect the ability of individuals to identify emotions or match facial feature [7]. Although some studies have found indices that may discriminate between individuals with and without ASD (e.g., accuracy, atypical visual behaviors, limbic activity [11]), the results of eye-tracking studies can be impacted by sample size, age, developmental level and IQ, variation in stimuli, and task effort and objective (reviewed in [6], [13]). Therefore, the ability to objectively and quickly differentiate between those with and without ASD independent of these factors may take combining eye-tracking with other techniques from other fields.

At present, ASD diagnosis is primarily based on behavioral criteria. This approach introduces subjectivity, is time-consuming, and sometimes inaccessible. Recent studies have shown the potential to characterize autism with gaze patterns, which makes eye-tracking ASD diagnosis highly desirable and feasible. Many studies have successfully applied eye-tracking and machine learning algorithms for classifying individuals with ASD. Liu et al. [14] investigated the eye movements of children with ASD in a face recognition task, and proposed a support vector machine (SVM) to classify children with ASD from matched controls. Jiang and Zhao [15] made use of deep learning to extract features from fixated image regions automatically, and achieved a 92% accuracy in classifying adults with ASD. Canavan et al. [16] combined gaze features with demographic features and tested three classifiers on individuals with low, medium, and high ASD risks. Nevertheless, these studies only considered the eye-tracking data, mainly what was looked at, without jointly taking into account the task performance. A possible reason is that the behavioral tasks are general but not specific enough to test the performance difference between individuals with ASD and controls.

In this work, to study the atypical visual attention of people with ASD, we investigate the predictive power of eye-tracking data under a facial emotion recognition task. It has been shown that individuals with ASD act differently when perceiving or responding to others' emotions. Therefore, we collected eye-tracking data from individuals with ASD and

[1]M. Jiang and Q. Zhao are with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN.

[2]S. M. Francis, D. Srishyla, C. Conelea, and S. Jacob are with the Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN.

*Q. Zhao and S. Jacob contributed equally to this paper.

TD controls under the Dynamic Affect Recognition Evaluation (DARE) task [7], [17]–[19]. We used task performance, eye movements, and face features in conjunction with state-of-the-art machine learning techniques, to tackle the ASD classification problem.

This work carries two major contributions:

- We tested 23 subjects with ASD and 35 TD controls in a facial emotion recognition task and recorded their eye movements. Statistical analyses were carried out to identify various between-group differences in both task performance and eye-movement patterns.
- We proposed a machine learning method to classify subjects based on how they performed and where they looked in the emotion recognition task. Instead of relying on annotated areas of interest, we extracted automatically learned high-dimensional face features from a deep neural network and combined it with gaze and task information for classification. We also visualized and interpreted the feature importance to understand the social and attentional deficits in ASD.

II. METHOD

We conducted an eyetracking experiment to investigate the atypical gaze patterns of individuals with ASD. Statistical analyses were carried out before the main study.

A. Subjects

Fifty-eight subjects with and without ASD completed the following study. They were recruited from the University of Minnesota (UMN) clinics, UMN websites, local and regional registries, local advertising as well as, the 2015 Minnesota State Fair. All subjects were recruited with the approval of the UMN Institutional Review Board. Twenty-three individuals with ASD (20 males; age range: 8–17 years; mean±SD: 12.74±2.45 years; IQ score range: 58–137) were diagnosed with ASD based on testing with standardized instruments, review of diagnostic history and evaluations, as well as DSM criteria confirmed by clinical research staff. Thirty-five TD controls (25 males; age range: 8±34 years; mean±SD: 14.11±5.09 years) participated who did not have a psychiatric history, developmental delay, or previous special education. The age difference between the groups was not statistically significant (t-test p=0.241), neither was the difference in the frequency of the sexes (p=0.171).

B. Procedure

The visual stimuli utilized in this study was a modified Dynamic Affect Recognition Evaluation (DARE) task [17]. Each trial is a series of still facial images, from the Cohn-Kanade Action Unit-Coded Facial Expression database [20], [21], displayed sequentially to create a video without audio cues. The videos are of a face starting with a neutral expression and slowly transitioning into one of six emotions: anger, disgust, fear, happiness, sadness, or surprise (see Figure 1). Video lengths varied from 19 to 33 seconds (mean±SD: 23.33±3.75 s), and all videos had the same
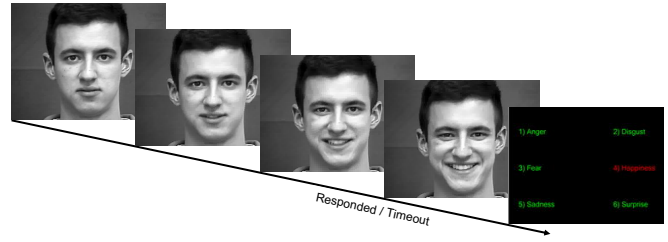

Fig. 1. The DARE task procedure. The example stimulus presents a face transitioning from neutral to an emotion (happiness).

640×480 resolution. The stimuli were presented on two displays (19 inches with a 1680×1050 resolution or 27 inches with a 1920×1080 resolution) depending on the location of collection, and uniformly scaled to fit the height of the screen. Subjects were seated approximately 65 cm from the screen (camera range: 50–70 cm). They were then instructed to watch the videos and press the spacebar to halt the video upon recognition of the emotion presented. Next, six emotion labels were displayed, and the subjects were asked to identify the emotion that had been recently presented. The entire protocol consisted of two phases, a practice (two trials and two choices) and a test phase (12 trials and six choices), which lasted approximately 10 minutes. The videos used in the practice phase were not repeated in the test phase.

C. Eye-Tracking

Eye-tracking data were collected with two Tobii Pro eye-tracker devices utilizing Tobii Studio (version 3.3.2; Tobii, Stockholm, Sweden; http://www.tobii.com). Data from all subjects with ASD and six TD controls were collected on the Tobii Pro TX300 at a 300 Hz sampling rate. The remaining TD samples were collected on the Tobii X2-60 with a sampling rate of 60 Hz. The precisions of the two devices were similar, and their differences in the rate of data collection did not factor into the analysis of the data. Eye trackers were calibrated using a standard 9-point grid, and calibration error for all subjects was less than 0.5 degree on the horizontal or vertical axis. All TD controls had at least one fixation detected in each trial, whereas data from five subjects with ASD were excluded because of failure to capture their eye movements in at least six trials. Figure 2 compares the
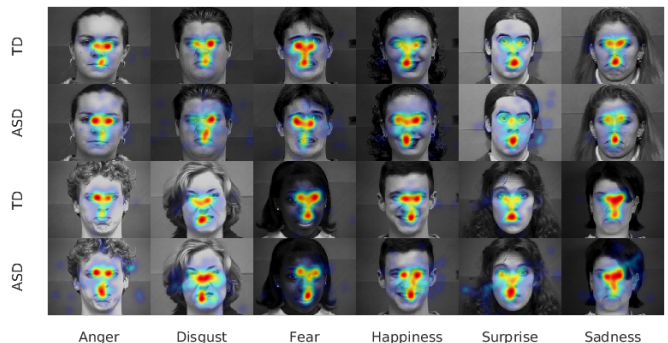

Fig. 2. Fixation density maps of the ASD and TD groups.

Fig. 3. Regions where ASD and TD subjects had significantly different fixation densities (t-test p<0.05). Red indicates a higher fixation density of the ASD group, while blue indicates a higher fixation density of the TD group.

Fig. 4. Comparisons between ASD and TD groups on task performance and gaze patterns.

fixation density maps of the two groups overlaid on example video frames. The fixation density map was initialized by setting the values of fixated pixels to 1 and the others to 0, and blurred using a Gaussian kernel (σ=15 pixels). Finally, as a probability density function, it was normalized to the sum of one. It can be observed that for both groups the fixations are clustered at the eyes and mouth regions, but subjects with ASD appear to have more low-density fixations in other facial regions and the background.

To confirm this observation, we computed fixation density maps for each subject and tested the difference of fixation densities at each pixel. Figure 3 presents the regions where ASD and TD subjects had significantly different fixation densities (t-test p<0.05). The comparison suggests that the TD controls were more attracted to the eyes and mouth, but subjects with ASD have more scattered fixations on the forehead, hair, ears, chin, and other features not strongly associated with emotion recognition.

D. Data Analysis

In 603 of the 696 total trials, subjects responded by pressing the spacebar (responded trials) and subsequently identified the emotion, while in the other 93 trials they identified the emotion after the video completed playing (timeout trials). The accuracy of the responded trials was 77.94%, and the accuracy of the timeout trials was 74.19%. The overall accuracies of the ASD (mean±SD: 77.89±13.90%) and TD (mean±SD: 77.14±14.05%) groups were not significantly different (t-test p=0.841).

Though the accuracies were similar, the ASD and TD groups showed significant differences regarding their response time and eye movements. As shown in Figure 4, we investigated various dependent variables including

1) *Response Time (RT)*: the length of time spent observing each video before hitting the button or timing out;
2) *Relative RT*: the proportion of time spent observing each video;
3) *Fixation Number*: the number of fixations subjects made in each trial;
4) *Fixation Frequency*: the average number of fixations subjects made in each second of a trial;
5) *Fixation Duration*: the average length of time subjects fixated in each trial;
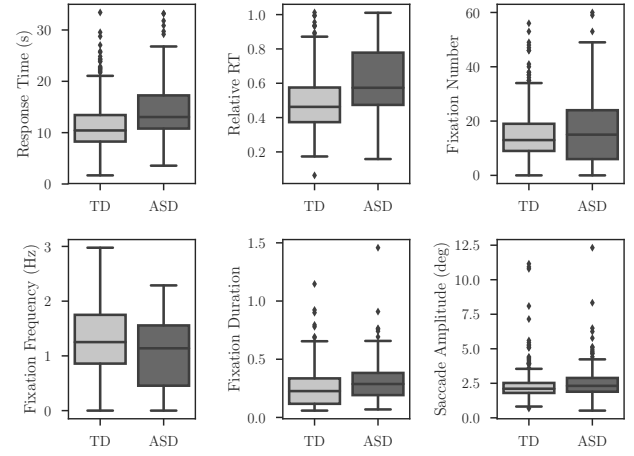6) *Saccade Amplitude*: the average saccade amplitude in each trial.

To determine whether these variables differed across subject groups and trials, we performed two-way mixed-design analyses of variance (ANOVAs) with the subject group (ASD or TD) as between-subjects variable and the experiment trials as a repeated-measures variable.

The ASD group spent more time observing the stimuli (mean±SD: ASD=14.37±5.45s, TD=11.43±4.57s, main effect of subject group: p<0.001). This difference remained significant after normalizing with the video length (mean±SD: ASD=0.61±0.19, TD=0.49±0.16, main effect of subject group: p<0.001). Due to their slower responses, the ASD group had more fixations (mean±SD: ASD=16.15±12.16, TD=14.41±8.91, main effect of subject group: p=0.002). However, their fixation frequencies were lower (mean±SD: ASD=1.06±0.65, TD=1.28±0.60, main effect of subject group: p<0.001) and their fixations lasted longer (mean±SD: ASD=0.31±0.17s, TD=0.25±0.16s, main effect of subject group: p<0.001). The ASD group also had greater saccade amplitudes (mean±SD: ASD=2.54±1.64°, TD=2.30±1.09°, main effect of subject group: p=0.016). The *Response Time*, *Relative RT*, and *Fixation Number* were all significantly different across trials with p<0.001, but no difference was observed in *Fixation Frequency*, *Fixation Duration* or *Saccade Amplitude* (all p>0.05). Interactions were not significant either (all p>0.05).

E. Feature Description

Based on the above observations, we combined task performance, eye movements, and the stimuli for the classification of ASD. Features extracted from these data were categorized as follows:

1) Task Features: The first category of features described the behavioral performance in the facial emotion recognition task. As observed in the statistical analyses, response time and relative response time are significantly different between ASD and TD groups. Therefore, we described a subject's task performance in a trial as a two-dimensional feature vector.

The task features were repeated for all the fixations in the same trial when used for classifying fixations.

2) Gaze Features: The atypical visual attention and oculomotor control in ASD can be described by where and how they looked at the stimuli. Therefore, the second category of features consisted of five primary characteristics of eye movements: the fixation location contains the x (i.e., horizontal) and y (i.e., vertical) coordinates indicating where the subject's attention was focused; the fixation time, fixation frequency, fixation duration, and saccade amplitude that may demonstrate the altered oculomotor function in ASD. Such information obtained from the eye-tracking data forms a six-dimensional feature vector of each fixation.

3) Face Features: We extracted face features from the stimuli using OpenFace [22], a deep neural network model with a human-level performance in the face recognition task. As a feed-forward network, OpenFace was composed of 37 layers of convolutional filters and a final linear projection layer. The convolutional layers were interconnected and grouped into eight Inception blocks. To represent *what* the subjects looked at, given a fixation's spatial coordinate and time, we first extracted the corresponding video frame, and then detected the face region with the Viola-Jones detector [23]. The detected face was scaled to 96×96 pixels and processed with OpenFace. At the fifth Inception block (i.e., inception-4a) the OpenFace network computed 640 activation maps in 6×6 resolution, which resulted in a 640-dimensional feature vector at each fixation location. The OpenFace network was pre-trained on the LFW dataset [24]. Due to the generality of the dataset, we directly took the pre-trained features without fine-tuning the network on our dataset.

F. Random Forest Classification

We classified behavioral and eye-tracking data using random forest (RF), an ensemble learning method that constructs a forest of decision trees for classification. With a bootstrap sampling of the training data, each decision tree classified a random subset of the input features. Their predictions were combined based on a majority voting, so that the ensemble could achieve a high classification accuracy. In this study, we considered each fixation as a data sample, and trained an RF to classify fixations of the two groups. The classification scores of all fixations of the same trial or subject were averaged to achieve trial-level or subject-level classification.

G. Performance Evaluation

A leave-one-subject-out cross-validation was used in the experiments. In each run of the cross-validation, one subject was left out as testing data, while the rest were used for training. This process was repeated 60 times so that each subject was tested once. The testing results of all subjects were combined and evaluated in terms of sensitivity, specificity, and overall accuracy as follows:

$$\text{Sensitivity} = \frac{\text{TN}}{\text{TN}+\text{FP}}, \qquad (1)$$

$$\text{Specificity} = \frac{\text{TP}}{\text{TP}+\text{FN}}, \qquad (2)$$

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}. \qquad (3)$$

These evaluation metrics considered receiver operating characteristic (ROC) parameters such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The area under the ROC curve (AUC) was also used as a quantitative measure of the overall classification performance.

III. RESULTS

To investigate the different effects of the proposed features, we trained and evaluated RF classifiers first using the three categories of features independently, and then using all features combined. The performances were also compared across different classification units – fixations, trials, and subjects. The classifiers were implemented in Python with the scikit-learn library [25]. They were trained using balanced class weights to avoid the influence of unequal numbers of subjects. The classification results are reported in Table I. Note that task features are the response time and relative response time per trial, so only trial-level and subject-level results are reported. First of all, a combination of task, gaze, and face features achieved 72.5% classification accuracy for individual fixations. With soft voting, the accuracy reached 75.6% and 86.2% at the trial and subject levels, respectively. The performance is comparable with standardized diagnostic tools [26] and other ASD classification methods based on eye-tracking [14]–[16]. Further, it is noteworthy that the task features had very low sensitivity, but in combination with gaze and face features, the sensitivity increased to 91.3%, which suggests the important role of eye-tracking data for distinguishing subjects with ASD.

RF classifiers calculate feature importance to determine how to split the data into subsets to most effectively help distinguish the classes. We ranked the features by their average importance values of all cross-validation runs, and normalized them to the maximum one. In Figure 5, we present the importance of the task and gaze features first, followed by the top-10 most important face features. To

TABLE I

A COMPARISON OF THE MODELS' PERFORMANCES OF CLASSIFYING SINGLE FIXATIONS, TRIALS, AND SUBJECTS. RF MODELS ARE TRAINED WITH TASK, GAZE, FACE AND COMBINED FEATURES.

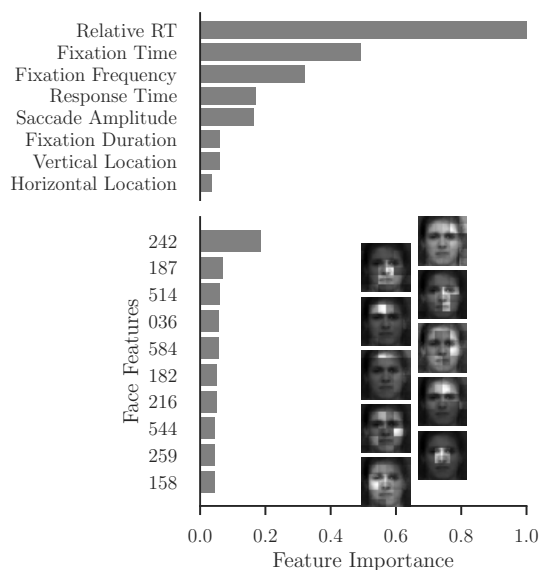| Unit	Features	AUC	Accuracy	Sensitivity	Specificity
Fixation	Gaze	0.699	0.664	0.558	0.740
	Face	0.587	0.619	0.331	**0.827**
	Combined	**0.743**	**0.725**	0.616	0.802
Trial	Task	0.737	0.710	0.424	0.898
	Gaze	0.820	0.714	**0.678**	0.738
	Face	0.741	0.720	0.428	**0.912**
	Combined	**0.824**	**0.756**	0.659	0.819
Subject	Task	0.789	0.810	0.565	**0.971**
	Gaze	0.904	0.845	0.783	0.886
	Face	0.917	0.845	0.870	0.829
	Combined	**0.935**	**0.862**	**0.913**	0.829

Fig. 5. Feature importance and visualization of the most important face features. Activation maps are overlaid on an average face and the brighter regions indicate stronger activation in the corresponding feature channel.

cation, making use of demographic information, data from the autonomic nervous system, functional MRI data, as well as data from other methodologies. While the focus on this work is to classify ASD, similar eye-tracking paradigms and machine learning methods can also be applied to differentiate or classify patients with schizophrenia or ADHD, as they have also demonstrated altered gaze patterns in various visual tasks.

visualize the face features, we averaged the neural network outputs of all video frames for each feature channel. They are overlaid on an average face and presented to the right of the corresponding bars. As shown in the figure, temporal information at a coarse level (e.g., fixation time and frequency, response time and relative response time) were the most important among all features, while the fine-grained fixation statistics played less significant roles. Though independent face features did not strongly contribute to the classification, these face features represent distinctly different facial regions (such as forehead, eyes, cheeks, nose, and ears) suggesting that fixations in these regions were helpful for classification.

IV. DISCUSSION AND FUTURE WORK

In this study, we have investigated the atypical visual attention of individuals with ASD under a facial emotion recognition task. We identified important features in the behavioral and eye-tracking data. Similar to [7], who examined emotion recognition in children (7–17 years) with and without ASD, we observed no difference in accuracy, but a significant increase of response time in ASD. Eye-movement patterns were also significantly different between groups. Based on these observations, a combination of task, gaze, and face features was proposed, leading to an RF classifier that discriminated between ASD and TD subjects. The classification results were encouraging because different features complemented each other in the combined feature domain, making the two groups more separable. These results suggested differences in social information processing that may assist with diagnostic evaluations.

Future research should include extending the study to include more subjects across developmental ages. We also plan to develop a multi-modal approach to ASD classifi-

REFERENCES

[1] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Association Washington, DC, 2013.

[2] L. Kanner *et al.*, "Autistic disturbances of affective contact," *Nervous Child*, vol. 2, no. 3, pp. 217–250, 1943.

[3] C. C. Goren, M. Sarty, and P. Y. Wu, "Visual following and pattern discrimination of face-like stimuli by newborn infants," *Pediatrics*, vol. 56, no. 4, pp. 544–549, 1975.

[4] K. Pierce, S. Marinero, R. Hazin, B. McKenna, C. C. Barnes, and A. Malige, "Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity," *Biological Psychiatry*, vol. 79, no. 8, pp. 657–666, 2016.

[5] H. Crawford, J. Moss, C. Oliver, N. Elliott, G. M. Anderson, and J. P. McCleery, "Visual preference for social stimuli in individuals with autism or neurodevelopmental disorders: An eye-tracking study," *Molecular Autism*, vol. 7, no. 1, p. 24, 2016.

[6] M. H. Black, N. T. Chen, K. K. Iyer, O. V. Lipp, S. Bölte, M. Falkmer, T. Tan, and S. Girdler, "Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography," *Neuroscience & Biobehavioral Reviews*, vol. 80, pp. 488–515, 2017.

[7] E. Bal, E. Harden, D. Lamb, A. V. Van Hecke, J. W. Denver, and S. W. Porges, "Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state," *Journal of Autism and Developmental Disorders*, vol. 40, no. 3, pp. 358–370, 2010.

[8] M. L. Spezio, R. Adolphs, R. S. Hurley, and J. Piven, "Abnormal use of facial information in high-functioning autism," *Journal of Autism and Developmental Disorders*, vol. 37, no. 5, pp. 929–939, 2007.

[9] M. L. Spezio, R. Adolphs, R. S. Hurley, and J. Piven, "Analysis of face gaze in autism using "bubbles"," *Neuropsychologia*, vol. 45, no. 1, pp. 144–151, 2007.

[10] A. Klin, W. Jones, R. Schultz, F. Volkmar, and D. Cohen, "Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism," *Archives of General Psychiatry*, vol. 59, no. 9, pp. 809–816, 2002.

[11] K. M. Dalton, B. M. Nacewicz, T. Johnstone, H. S. Schaefer, M. A. Gernsbacher, H. H. Goldsmith, A. L. Alexander, and R. J. Davidson, "Gaze fixation and the neural circuitry of face processing in autism," *Nature Neuroscience*, vol. 8, no. 4, p. 519, 2005.

[12] K. A. Pelphrey, N. J. Sasson, J. S. Reznick, G. Paul, B. D. Goldman, and J. Piven, "Visual scanning of faces in autism," *Journal of Autism and Developmental Disorders*, vol. 32, no. 4, pp. 249–261, 2002.

[13] M. B. Harms, A. Martin, and G. L. Wallace, "Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies," *Neuropsychology Review*, vol. 20, no. 3, pp. 290–322, 2010.

[14] W. Liu, M. Li, and L. Yi, "Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework," *Autism Research*, vol. 9, no. 8, pp. 888–898, 2016.

[15] M. Jiang and Q. Zhao, "Learning visual attention to identify people with autism spectrum disorder," in *IEEE International Conference on Computer Vision*, IEEE, 2017, pp. 3287–3296.

[16] S. Canavan, M. Chen, S. Chen, R. Valdez, M. Yaeger, H. Lin, and L. Yin, "Combining gaze and demographic feature descriptors for autism classification," in *IEEE International Conference on Image Processing*, IEEE, 2017, pp. 3750–3754.

[17] S. W. Porges, J. F. Cohn, E. Bal, and D. Lamb, "The dynamic affect recognition evaluation [computer software]," *Brain-Body Center*, 2007.

[18] K. Heilman, E. Harden, K. Weber, M. Cohen, and S. Porges, "Atypical autonomic regulation, auditory processing, and affect recognition in women with HIV," *Biological Psychology*, vol. 94, no. 1, pp. 143–151, 2013.

[19] G. Domes, A. Steiner, S. W. Porges, and M. Heinrichs, "Oxytocin differentially modulates eye gaze to naturalistic social signals of happiness and anger," *Psychoneuroendocrinology*, vol. 38, no. 7, pp. 1198–1202, 2013.

[20] J. F. Cohn, A. J. Zlochower, J. Lien, and T. Kanade, "Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding," *Psychophysiology*, vol. 36, no. 1, pp. 35–43, 1999.

[21] T. Kanade, Y. Tian, and J. F. Cohn, "Comprehensive database for facial expression analysis," in *IEEE International Conference on Automatic Face and Gesture Recognition*, IEEE, 2000, p. 46.

[22] B. Amos, B. Ludwiczuk, M. Satyanarayanan, *et al.*, "OpenFace: A general-purpose face recognition library with mobile applications," *CMU School of Computer Science*, 2016.

[23] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.

[24] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," in *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, 2008.

[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[26] T. Falkmer, K. Anderson, M. Falkmer, and C. Horlin, "Diagnostic procedures in autism spectrum disorders: A systematic literature review," *European Child & Adolescent Psychiatry*, vol. 22, no. 6, pp. 329–340, 2013.