Video Based Computational Coding of Movement Anomalies in ASD Children

Priya Singh[1], Abhishek Pathak[3], Umer Jon Ganai[2], Braj Bhushan[2], Venkatesh K. Subramanian[2]

[1]TCS Research, India
[2]Indian Institute of Technology Kanpur, India
[3]Airbase Labs India Pvt Ltd, India

Abstract

Autism spectrum disorder (ASD) is a neurodevelopmental disorder. Early detection and diagnosis are instrumental in early intervention, yet diagnosis often remains delayed due to the limited availability of clinical practitioners and specialists. We propose a Computer Vision and Machine Learning based novel framework for quantitative screening of autism spectrum disorder (ASD). This is aimed to minimize the need for trained professionals at the initial stage but not substitute for it. We designed simple activities in consultation with ASD clinical psychologists and therapists for children in the 3-7 years age group that could be performed in their natural environment (home). The temporal features extracted from these activities encode the behavioral differences between Autism Spectrum Disorder (ASD) and Typically Developing (TD) control groups. Due to the unavailability of a public dataset of children performing the designed task, we created our own video dataset of 210 videos taken in unconstrained natural settings. The dataset was collected from a single RGB camera. The proposed vision and learning-based algorithms extract features from the collected data for a comprehensive set of indicators including the visual attention span, name-calling response, neck pose of the subjects, gross motor movement and establish a parametrized automated protocol for early detection without the need to take the subjects out of their natural daily environment. This forestalls the possibility of misperformance by the subject out of nervousness due to unfamiliar surroundings. Results show that our ASD screening methodology can achieve superior performance compared to the single phenotype approaches, and thus has a prognostic value that could be helpful for both clinical and research applications.

1. Introduction

Autism spectrum disorder (ASD) is one of the most disabling neurodevelopmental disorders that occurs in childhood [1, 39, 41] with an estimated worldwide prevalence of 1 in 160 among children (WHO, 2019). The recently estimated prevalence of ASD in India ranges from 0.15% to 1.01% in various studies, depending on the screening method used and the areas surveyed [32, 34]. There was a study conducted on ninety-five ASD children in four major metropolitan cities of India. Through the study, it was found that the time between their first visit to a doctor and their first diagnosis of autism averaged two and a half years. The parents on average visited four doctors before their child was diagnosed with autism, while some visited as many as ten to twelve doctors[8, 10].

The DSM-V (American Psychiatric Association, 2013) and ICD-11 (WHO, 2019) are the two international gold-standard classification manuals that provide criteria for diagnosing ASD and the main symptoms used for the diagnosis of ASD are impairments in social and interaction abilities and the presence of restrictive interests and repetitive behaviors.

Over the past few years, researchers have attempted to identify more quantifiable and objective biological markers of ASD. Objectively quantifying the biomarkers could provide a more systemic diagnosis and earlier detection of ASD, allowing for more customized and earlier interventions [18]. Research has demonstrated the effectiveness of the early intervention that occurs between the age of birth and four years [10]. The clinical tools used for early diagnosis require interpretation by specialists. The number of experts available is a challenge in low- and middle-income countries (LMIC). Most families lack easy access to ASD specialists. The diagnosis takes place mostly in cities where there is a limited number of knowledgeable professionals. Small towns and rural areas are mostly outside the ambit of diagnosis. Therefore there is a need for automatic and quantitative analysis tools that can be used by clinicians, general practitioners or parents of the children to identify children at risk of ASD. The computational approaches have been extensively used to aid in diagnosing ASD. It has countered the limitations in the diagnostic process and emerged as an efficient alternative to rating-based assessment tools. There has been a shift from observation schedules and rating scales to computer vision and machine learning approaches for assessing ASD markers [18, 33]. Therefore, the major contributions of this work are:

- We designed a set of tasks for feature selection with assistance from ASD clinical psychologists and therapists.
- Due to the unavailability of a public dataset of children performing the designed task in an unconstrained environment, we created our video dataset collected in unconstrained natural settings.
- Our computational coding of spatio-temporal features differentiates the behavioral phenotype of our two study groups.
- We propose a Computer Vision and Machine Learning based novel framework for early detection of ASD in children in their natural environment with a performance accuracy of 95%

2. Related Work

A study by Dawson et al.[9] showed that toddlers with ASD exhibited a significantly higher rate of head movement than non-ASD toddlers, indicating difficulties in maintaining midline head position while engaged in a task. Bradshaw et al[3] showed that while pull-to-sit trajectories did not differ between groups, infants with ASD were more likely to exhibit head lag by 4 months. Furthermore, pull-to-sit trajectories were found to be predictive of social and speech skills 2 years later. The infants later diagnosed with ASD also demonstrate a lower percentage of midline head position PMHP compared to typically developing infants [15]. These findings suggest that early atypical postural development, persistent head lag and midline head position may be early indicators of ASD. Research has shown that toddlers with ASD often have a reduced response to their name (RTN), which may be a critical indicator for ASD. Various screening and diagnostic instruments, such as the M-CHAT and ADOS, include RTN as an important item [22, 24, 26]. To be attentive to the social world, one must engage in cognitive behaviors that involve capturing social cues and processing social information, which requires directing attention to pertinent stimuli. It has been suggested that the social deficits seen in individuals with autism spectrum disorder (ASD), such as reduced attention to relevant social information, may be partly due to abnormal attention processes, as indicated by various studies [7, 17, 31].

	ASD	TD
No of video data	116	94
No of participants	17	18
Gender distribution	5 females 12 males	10 females 8 males
Age distribution	Mean age: 4 Std Dev: 1.5	Mean age: 4.7 Std Dev: 0.64
Data collection timeline	59 days	

Table 1. Participants demography

3. Dataset and Data Collection Protocol

The database of Autism Spectrum Disorder (ASD) and Typically Developing (TD) children performing the set of activities in an unconstrained environment required for our study is not available in the public domain. Our best option was to build our own dataset. Therefore, we built a new dataset of Indian children, taken in unconstrained natural settings, where the subjects (ASD and TD children) perform regular day-to-day activities in school. The dataset consists of a total of 210 videos, of which 116 videos are of ASD children, and 94 videos are of TD children. The data of 35 children (18 neurotypical and 17 with ASD) was collected over a period of two months. This was done to bring variations in the same subject data. Table 1 presents the participants' demography. Confirmed diagnosis of ASD was the inclusion criteria for the clinical group while absence of neurodevelopmental disorder was the inclusion criteria for the TD group. The ASD group was drawn from a special school. These children were already diagnosed by in-house licensed clinical psychologists according to the diagnostic criteria for ASD in the DSM-5 (APA, 2013). Participants in the ASD group who had received a diagnosis other than ASD were excluded from the group. The TD and ASD groups were matched on age. Informed consent was obtained from the parents of the children before the initiation of the study. The study protocol was approved by the Institutional Ethics Committee of the Indian Institute of Technology Kanpur. The access to the blurred dataset to maintain the anonymization of the participants can be shared with the academic community for research purposes on request.

4. Features

The study aims to investigate the movement biomarkers of autism, i.e., eye movements, head movements and neck movement to discriminate between ASD and neurotypical (TD) children and characterize autism-specific motor differences. To investigate these biomarkers, we have designed simple activities for children in the 3-7 years of age group such that even the low-functioning ASD participants could comply and will be able to perform it in their natural environment. The activities have been designed based on the feedback received from clinical psychologists, and therapists who have expertise in the area of neurodevelopmental disorders found in children. To extract the features corresponding to each of the above markers, we have used OpenPose [6] a deep-learning network for human pose estimation.

4.1. Eye movement

Studies show that ASD children show reduced attention span i.e. low eye fixations and high saccade movements [2, 11, 20, 40]. A common methodology for the analysis of attention span behavior is eye-tracking [36]. The above studies have used either expensive eye-tracking devices or complex advanced methodologies, such as dark pupil-corneal reflection using IR cameras to study the gaze pattern. Such acquisition methods are expensive, inaccessible and are not suitable for natural settings data acquisition. However, atypical eye movement patterns demonstrated by individuals with ASD can serve as biomarkers for ASD diagnosis. To explore this, we study eye movement velocity while watching an engaging stimulus using an RGB camera. We have carefully designed stimuli and proposed a computationally simple computer vision algorithm to extract the features.

4.1.1 Data acquisition

Research shows that ASD children display an attentional preference for non-social stimuli compared to social stimuli [21, 37]. They are attracted to colors and music. Therefore for studying eye movement, we picked the video from the popular Sesame Street series for children. The language of the video is Hindi, the most commonly spoken language in Kanpur, where the whole research is conducted. The participants were requested to watch a video over the course of five minutes on a computer screen. Participants were seated approximately 15 inches from a 13-inch computer screen. The stimuli were presented at eye level. The external camera of resolution 4K recorded the participants while watching the video in an unconstrained environment as shown in Figure 1.

4.1.2 Data Pre-processing

The length of the time-series data for the eye movement task is five minutes. For evaluation, we have trimmed the initial and end 30 seconds from the video data. The intervals containing blinks were discarded. The missing data was extrapolated using the Kalman filter.

4.1.3 Feature Extraction

Since the video is taken in the participant's natural environment, multiple heads are also present in the data. To eliminate the head movement of the non-participants we apply a tracking algorithm. We are using an optical flow-based tracking algorithm by Bruce D. Lucas and Takeo Kanade [25] for our feature tracking. The algorithm works under the following assumptions:

- Two consecutive frames are separated by a small time increment dt
- Neighbouring pixels have similar motion.

Taking u the movement w.r.t. x and v the movement w.r.t. y and $I(x, y)$ the pixel intensity at pixel (x, y). The optical flow equation is:

$$\frac{\partial I}{\partial x}(x, y, t) \cdot u + \frac{\partial I}{\partial y}(x, y, t) \cdot v = -\frac{\partial I}{\partial t}(x, y, t) \tag{1}$$

which can be further written as,

$$I_x v_x + I_y v_y = -I_t \tag{2}$$

where $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$ and $\frac{\partial I}{\partial t}$ are the derivatives of $I(x, y, t)$ with respect to x, y and t. Using the bounding box, we select the part or patch of the image that includes our subject as shown in Figure 8. We assumed that the displacement of the image pixels between two consecutive frames is small and approximately constant within a neighborhood of the pixels under consideration. For a patch of n neighborhood points, the equation can be expressed as,

$$\begin{aligned} I_x(p_1)v_x + I_y(p_1)v_y &= -I_t(p_1) \\ I_x(p_2)v_x + I_y(p_2)v_y &= -I_t(p_2) \\ &\vdots \\ I_x(p_n)v_x + I_y(p_n)v_y &= -I_t(p_) \end{aligned} \tag{3}$$

where $p_1, p_2, ..., p_n$ are the pixels of the patch, and $I_x(p_i), I_y(p_i), I_t(p_i)$ are the partial derivatives of the image I with respect to position x, y, and time t, evaluated at the pixel p_i. We now face the problem of solving two unknowns v_x and v_y from n equations, which is an over-determined case and can be solved using the least squares fitting solution, expressed as,

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(p_i)^2 & \sum_i I_x(p_i)I_y(p_i) \\ \sum_i I_y(p_i)I_x(p_i) & \sum_i I_y(p_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_x(p_i)I_t(p_i) \\ -\sum_i I_y(p_i)I_t(p_i) \end{bmatrix} \tag{4}$$

The output from the face tracking algorithm is then used as input to the feature extraction algorithm. The OpenPose is the state-of-the-art approach for real-time human pose estimation [5]. The OpenPose network is trained on combinations of datasets comprising of CMU Multi-PIE Face [16], Face Recognition Grand Challenge (FRGC) [30], and i-bug [35] datasets. As shown in Figure 2, the model estimates

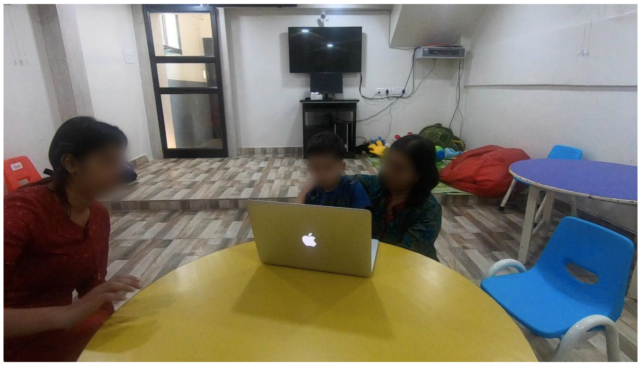
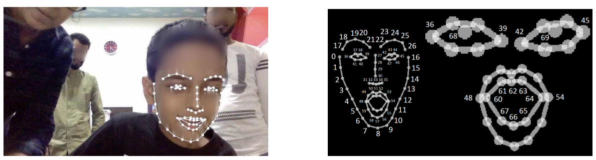


Figure 1. Data collection setup



(a) (b)

Figure 2. Extracted facial keypoints

70 facial landmarks. For the eye movement analysis, we are using keypoint no 68, which is the left eye iris and 69, which is the right eye iris. On the mean Average Precision (AP) metric, the Openpose algorithm has 98.3% of accuracy on FRGC datasets and 96.3% on the Multi-PIE dataset. Since we are analyzing relative eye position across the video data, the stated error in the position estimation can be tolerated.

4.1.4 Evaluation Metrics

To analyse the atypical eye movement patterns of our study groups, we plot eye coordinates along the x-axis in terms of image pixels against the time axis for a period of four minutes or 240 seconds, as shown in Figure 3 and 4. We compare the eye velocity of the two study groups while watching the stimulus. The standard deviation(SD) in eye coordinates is a measure of the dispersion of eye coordinate pixels with respect to their mean pixel location over time. A high standard deviation suggests higher eye movement during the task.

$$SD_x = \sqrt{\frac{(\sum(x_i - \mu)^2}{n}}, SD_y = \sqrt{\frac{(\sum(y_i - \mu)^2}{n}} \quad (5)$$

where, x_i and y_i are eye pixel coordinates of each frame, μ is the mean of all the data points and n is the total number of data points or frames in this case.

4.1.5 Results and Analysis

The left eye movement plot of each ASD and TD child across frames is shown in Figure 3 and Figure 4 respectively. The plot in the figures suggests that the ASD study

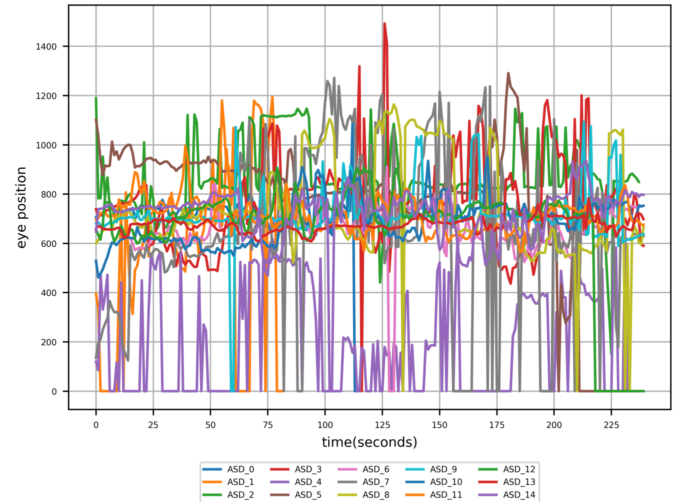


Figure 3. Eye position of ASD children across 240 seconds

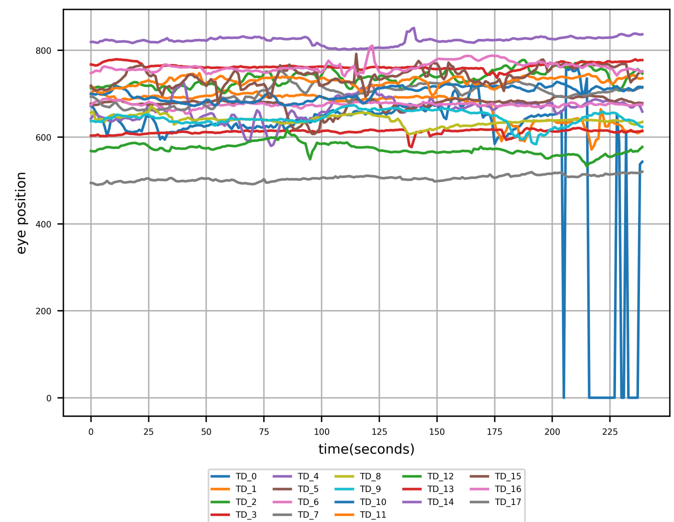


Figure 4. Eye position of TD children across 240 seconds

group has higher eye velocity than TD group under identical tasks and conditions. The mean and standard deviation plot of each child is shown in Figure 5. The SD of the ASD group varies between 37 pixels to 367 pixels over a period of four minutes. In contrast, the TD group's SD varies majorly between 5 pixels to 32 pixels except one child demonstrates a deviation of 181 pixels. Figure 6 compares the mean and standard deviation of eye position to the non-social stimulus of the two study groups. For the TD group, the mean pixel location=678.13, SD=87.88, and for the ASD group, the mean pixel location=675.7, SD=240.01.

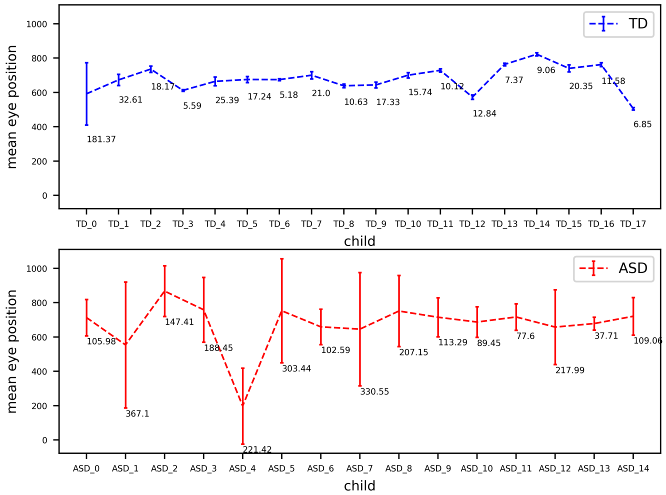


Figure 5. Mean and Standard Deviation of eye movement of each ASD and TD child.

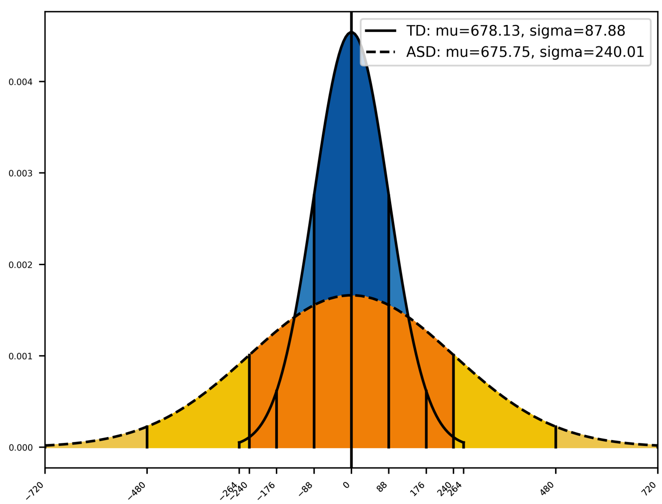


Figure 6. Standard Deviation comparison of eye movement data of ASD and TD children

4.2. Name-calling response

ASD children have difficulty responding or orienting to a name-call when engaged in activity [27]. Studies have evaluated the response to name call patterns from 6 to 24 months and found that toddlers who fail to respond to their name call may be at risk for ASD [23, 28]. The study suggests that consistent failure to name-call response in early life may be an important indicator for developmental disorders. Screening and diagnostic instruments such as M-CHAT and ADOS include response to name calls as an important checkpoint. These methods rely on either the parent's report, with the subjective error or the evaluator's screening which may be a time-consuming process. It is important to quantify

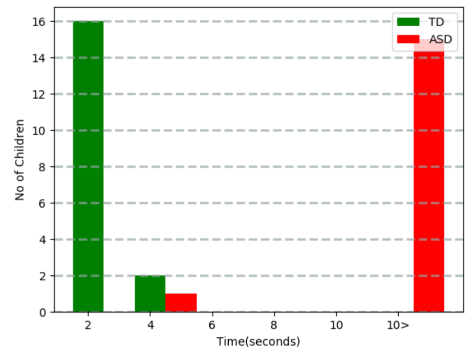


Figure 7. Name-calling response time of children

the name-calling response via an automated and accurate method.

4.2.1 Data Acquisition

While engaged in the task of watching the video, each participant is called by their name at a pre-defined fixed point in time which is two minutes after the beginning of the task. Their response is registered by capturing their behavior data in the form of a video by the camera.

4.2.2 Feature Extraction

In the video data of four minutes, the head pose of each participant is extracted using the OpenPose. The facial landmarks are extracted and based on the facial landmarks, a triangulated plane connecting the left, right eye center and nose tip is formulated. The yaw and pitch angle of this triangulated plane are calculated to determine the head pose. Since the direction of name-calling takes place behind each subject, an abrupt change in head pose in response to their name is expected.

4.2.3 Results and Analysis

The change in the head pose in the temporal video data and the response time corresponding to each child is plotted, as shown in Figure 7. All TD children except for one responded to their name call in less than 4 seconds. In contrast, none of the ASD children reacted to their name calls except for one. To validate the results we also used the manually annotated data for change in head pose. It was found to be aligned with the output of the feature extraction algorithm.

4.3. Motor Movement Abnormalities

Along with behavioral phenotypes, ASD children have been observed to show motor abnormalities. The early studies

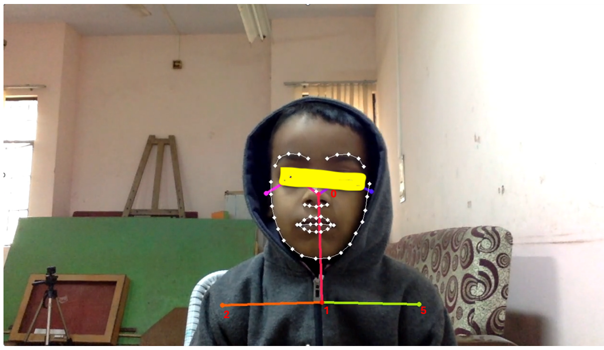


Figure 8. Extracted keypoints

support asymmetrical upper limb movement [12, 38] and difficulty maintaining the midline neck position [14, 15] as early motor signs of ASD. Motor abnormalities may occur very early in development and be apparent over time [38]. Prior work demonstrates such abnormalities in ASD children in the form of postural asymmetries during lying, head lag during the pull-to-sit transition [13, 14], and motor developmental delay in early infancy [15]. In this part of our work, we examine motor abnormalities in the form of irregularities in the neck alignment or movement while performing a task. And establish whether these differences in neck movement could be used as features to classify children with ASD from TD children computationally.

4.3.1 Data Acquisition

For the neck movement analysis, we are using the data acquired for our eye movement study in section 4.1.1

4.3.2 Feature extraction

We have used the Lucas Kanade optical flow-based tracking algorithm for feature tracking and OpenPose for feature extraction, which are 25 body keypoints [16]. Since the accuracy of the OpenPose model on the mean average precision metric on COCO keypoint dataset is 70.9%, we are using only those keypoints for which the confidence score of the algorithm is more than 0.8.

As features for our analysis, we are using four key points, which are -the left and right shoulder joints, the midpoint of the shoulder joints and nose key points, as shown in Figure 8. For neck movement analysis we consider the angle formed by keypoints (0, 1) which are the nose tip and midpoint of the shoulder joints and the keypoints (2, 5) which are left and right shoulder joints. Using the following equation, we determine the cosine angles formed by the vectors connecting the feature keypoints:

$$\begin{aligned} ab &= b - a \\ ac &= c - a \\ \angle\theta &= \cos^{-1}\frac{ab.ac}{|ab|.|ac|} \end{aligned} \tag{6}$$

where θ is the neck angle, ab is the vector formed by the coordinates of keypoints (0, 1) and ac is the vector formed by the coordinates of keypoints (1, 5)

4.3.3 Results and Analysis

The mean and standard deviation plot of the neck angle for each child is shown in Figure 9. The SD of the neck angle of the TD group varies between 1.74 to 16.85 degrees, with the exception of one child demonstrating a deviation of 28.97 degrees. Whereas, the ASD group's SD varies majorly between 18.33 to 49.3 degrees. The SD of neck angle data and the variance plot as shown in Figure 10 computationally supports our observation and previous studies that ASD children have difficulty in maintaining the mid-line neck position. The neck angle SD value is used as one of the features of our machine learning algorithm for the final classification.

5. Methodology

We propose a computer vision and machine learning-based framework as shown in Figure 11 for the classification of the our two study groups. As a first step, the input video is fed to the network as a sequence of frames at 15 fps. So, for a one-minute video, we are feeding 900 frames to the network. Using OpenPose the facial and body landmarks are identified. In the data denoising step, we drop the frames where all the keypoints are not present and the confidence score is below 0.8. The missing data is extrapolated using the Kalman filter. In the feature extraction step, the features are derived as explained in section 4. We consider the feature matrix (M) as being formed of three parts representing: eye movement feature (E), neck angle feature (N) and name-calling as headpose feature (H), which can be represented as M ={E, N, H}. We define E as having dimension (2 x n), N as having dimension (1 x n) and H as having dimension (3 x n), where n is the number of frames in a sequence of frames given as input to the model. These feature matrices are concatenated with equal weights. The classifier models are then trained using these M (6 x n) matrices.

5.1. Data Preparation

Based on the variation in the dataset, we have normalized the data using the Min-Max method. The mathematical for-

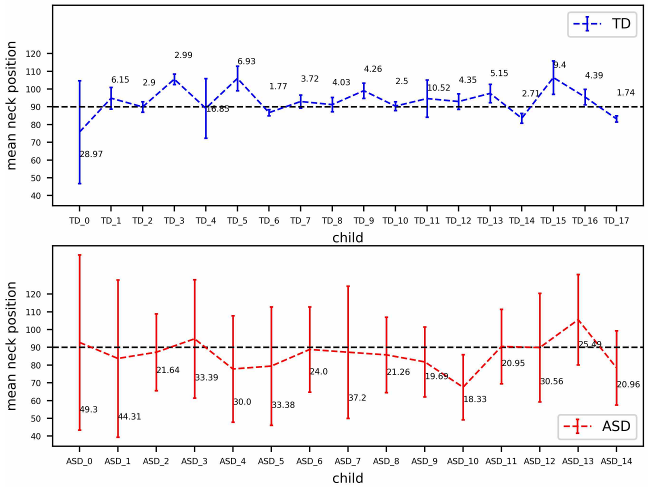


Figure 9. Mean and SD plot of neck angle of ASD and TD children

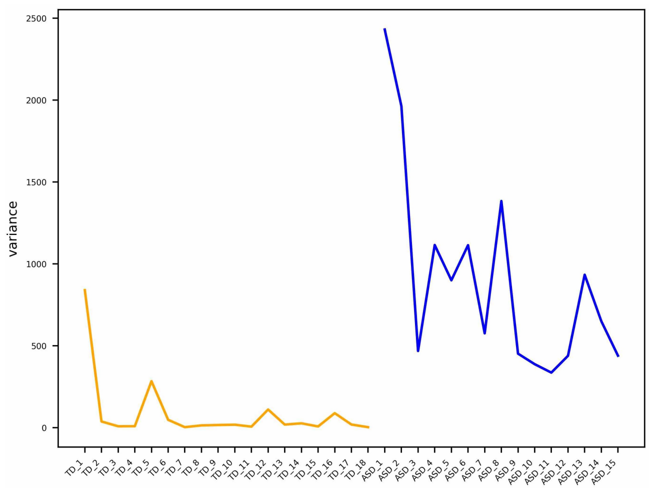


Figure 10. Variance plot of neck angle of TD(yellow curve) and ASD children(blue curve)

mulation is,

$$x_{scaled} = \frac{x - min(x)}{max(x) - min(x)} \tag{7}$$

An imbalanced dataset results in classifiers having more sensitivity for the dominant class. Our data has nearly equal class distribution. we are using only those data points for which all feature values are available and the confidence score of the key points is more than 0.8. The missing data points are extrapolated using the Kalman filter. The video data for which the missing data points are more than 30% was dropped. The final video data for evaluation was reduced to 200 videos from 210. We split 80% of our data into a training set and 20% into a test set i.e. 160 videos for training and 40 videos for test. For cross-validation, the training set data is further split into 80:20 as training and validation sets respectively.

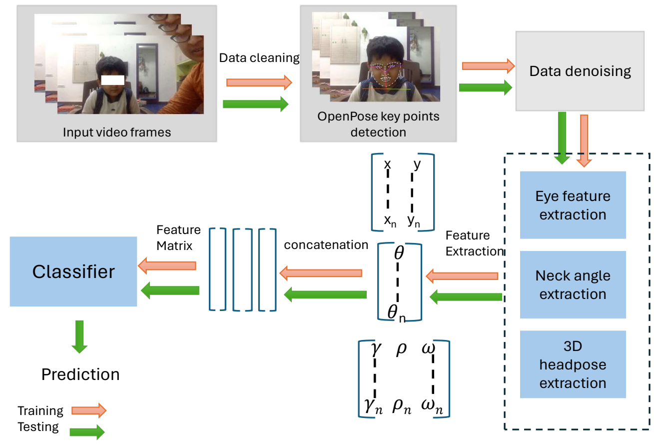


Figure 11. Model architecture of our proposed framework

5.2. Classification Model

We trained and tested our dataset on three popular machine learning algorithms -Decision Tree, K-Nearest Neighbor and SVM classifier. The comparison and evaluation of the models are done using the metrics -Accuracy, Precision, Recall and F1 Score.

- **Decision Tree**: We used Decision Tree, a supervised machine learning algorithm for our study group classification [29]. The algorithm is inexpensive and very fast at classification for this type of data set, and the accuracy is comparable to other complex classification algorithms.
- **K-Nearest Neighbor**: KNN classification [19] tests the closeness of unclassified data k in the dataset using Euclidean distance given as,

$$distance(p, q) = \sqrt{(\sum_{i=1}^{k}(p_i - q_i)^2} \tag{8}$$

- **SVM classifier**: SVM [4] uses hyperplane for data classification in n-dimensional space. It is a supervised algorithm that is best suitable for smaller but complex datasets.

6. Result and Discussion

The model is tested on 40 videos. The classification performance is assessed using the metrics: -Accuracy, Precision, Recall and F1 Score. The confusion matrix of our classifier on the test data is shown in Figure 12. Table 2 presents the performance of the three classifiers. From Table 2 it can be concluded that SVM classifier effectively classified ASD children with an accuracy of 95 percent.

$$Accuracy = \frac{TP + TN}{(TP + FP + FN + TN)} \tag{9}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{10}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{11}$$

$$F1Score = 2 * \frac{Precision * Recall}{(Precision + Recall)} \tag{12}$$

where, TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative

Model	Accuracy	Precision	Recall	F1 Score
Decision Tree	0.9	0.86	0.95	0.94
KNN	0.925	0.90	0.95	0.926
SVM	0.95	0.95	0.95	0.95

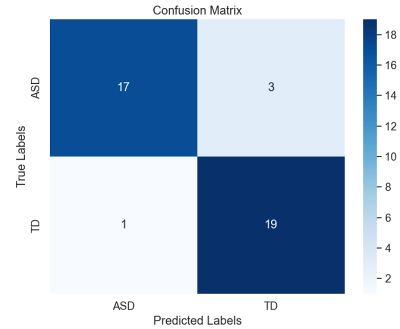
Table 2. Evaluation of machine learning models on ASD-TD children's dataset.

7. Case Study

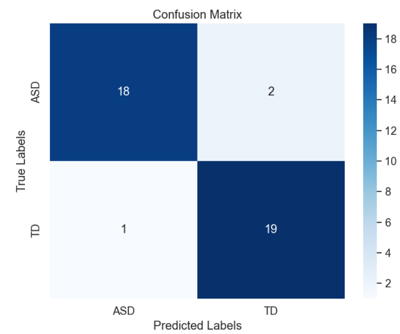
The accuracy of the classifiers will be higher without any outliers. Our data set has two outliers -one from each study group. The ASD outlier, misclassified as TD, had an early intervention at the age of two. Due to early detection and intervention, the child's clinical therapist considers her condition as a borderline case. The TD outlier, misclassified as ASD, seems to have an attention deficit hyperactivity disorder based on the observational data. When inquired with the school teachers, they affirmed that the child's academic performance is poor, suffers from speech irregularities, and lacks concentration and interest in studies and other social classroom activities. From the case study data, we can conclude that the reason for their misclassification is their ambiguous behavior or phenotype.

8. Conclusion and Future Scope

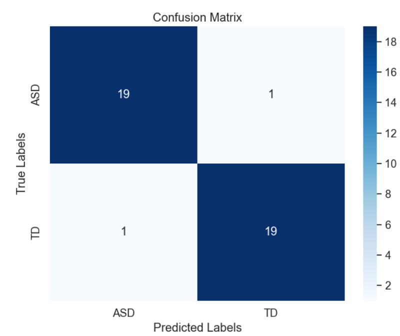
In this work, we explored the behavioral phenotypes of ASD children. We proposed a novel framework for early detection using computer vision and machine learning algorithms. We demonstrated that the features of ASD children can be detected by setting up the experiments described. Our model's accuracy in the classification of ASD children from TD justifies our feature selection. As the first milestone in this long-term goal, we want to explore the scope of adding more features like stimming and speech irregularities. We want to compare the accuracy of neural network models. We are working towards designing an automated platform, such as a mobile application that will assist in the detection of ASD in children. This will be based on videos taken, on a camera phone, of children performing simple activities as designed and described in this work. We believe that the proposed framework will be helpful in early detection It will not only accelerate the required early intervention but also create some much-needed awareness about the disorder itself.



(a)



(b)



(c)

Figure 12. Confusion matrix of machine learning models(a) Decision Tree (b) KNN Classifier (c)SVM tested on the dataset.

Acknowledgement

The author carried out this work as a part of the MS Research program in the Electrical Engineering department of Indian Institute of Technology Kanpur, under the guidance of Prof K S Venkatesh and Prof Braj Bhushan.

References

[1] American Psychological Association. Diagnostic and statistical manual of mental disorders (dsm-5®). *American Psychiatric Publishing*, 2013. 1

[2] M. D. M. J. Bovery, G. Dawson, J. Hashemi, and G. Sapiro. A scalable off-the-shelf framework for measuring patterns of attention in young children and its application in autism spectrum disorder. *IEEE Transactions on Affective Computing*, doi: 10.1109/TAFFC.2018.2890610., 2018. 3

[3] Jessica Bradshaw, Dexin Shi, Alexis Federico, Cheryl Klaiman, and Celine Saulnier. The pull-to-sit task: Examining infant postural development in autism spectrum disorder. *The Journal of Pediatrics*, 2022. 2

[4] Cortes C and Vapnik V 1995 Support vector networks Mach. Learn. 20 273–97. 7

[5] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019. 3

[6] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7291–7299, 2017. 3

[7] Meia Chita-Tegmark. Social attention in asd: A review and meta-analysis of eye-tracking studies. *Research in developmental disabilities*, 48:79–93, 2016. 2

[8] Tamara C Daley. From symptom recognition to diagnosis: children with autism in urban india. *Social Science and Medicine*, 58(7):1323–1335, 2004. 1

[9] Geraldine Dawson, Kathleen Campbell, Jordan Hashemi, Steven J Lippmann, Valerie Smith, Kimberly Carpenter, Helen Egger, Steven Espinosa, Saritha Vermeer, Jeffrey Baker, et al. Atypical postural control can be detected via computer vision analysis in toddlers with autism spectrum disorder. *Scientific reports*, 8(1):17008, 2018. 2

[10] Dr. Tamara C. Daley and Ms. Merry Barua. Autism. http://www.rehabcouncil.nic.in/writereaddata/autism.pdf, 2020. [Online]. 1

[11] M. Elsabbagh, J. Fernandes, S. J. Webb, et al. Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry*, 74(3):189–194, 2013. 3

[12] Gianluca Esposito, Paola Venuti, Sandra Maestro, and Filippo Muratori. An exploration of symmetry in early autism spectrum disorders: Analysis of lying. *Brain and Development*, 31(2):131–138, 2009. 6

[13] Gianluca Esposito, Paola Venuti, Sandra Maestro, and Filippo Muratori. An exploration of symmetry in early autism spectrum disorders: Analysis of lying. *Brain and Development*, 31(2):131–138, 2009. 6

[14] Joanne E. Flanagan, Rebecca Landa, Anjana Bhat, and Margaret Bauman. Head Lag in Infants at Risk for Autism: A Preliminary Study. *American Journal of Occupational Therapy*, 66(5):577–585, 09 2012. 6

[15] H. Gima, H. Kihara, H. Watanabe, et al. Early motor signs of autism spectrum disorder in spontaneous position and movement of the head. *Experimental Brain Research*, 236(4):1139–1148, 2018. 2, 6

[16] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-pie. *Image and Vision Computing*, 28(5):807–813, 2010. 3, 6

[17] Quentin Guillon, Nouchine Hadjikhani, Sophie Baduel, and Bernadette Rogé. Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience & Biobehavioral Reviews*, 42:279–297, 2014. 2

[18] Jordan Hashemi, Thiago Vallin Spina, Mariano Tepper, Amy Esler, Vassilios Morellas, Nikolaos Papanikolopoulos, and Guillermo Sapiro. A computer vision approach for the assessment of autism-related behavioral markers. In *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, pages 1–7. IEEE, 2012. 1, 2

[19] NS. A 1992 An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 43(3) 175–85. 7

[20] D.Y. Isaev, S. Major, M. Murias, et al. Relative average look duration and its association with neurophysiological activity in young children with autism spectrum disorder. *Scientific Reports*, 10:1912, 2020. 3

[21] J. C. Kirchner, A. Hatri, H. R. Heekeren, and I. Dziobek. Autistic symptomatology, face processing abilities, and eye fixation patterns. *Journal of Autism and Developmental Disorders*, 41(2):158–167, 2011. 3

[22] Kenneth Larsen, Astrid Aasland, Synnve Schjølberg, Ulla I Hansen, and Trond H Diseth. Piloting the use of a short observation list for asd-symptoms in day-care: challenges and further possibilities. *Journal of Autism and Developmental Disorders*, 50:3413–3423, 2020. 2

[23] Kenneth Larsen, Astrid Aasland, Synnve Schjølberg, Ulla I Hansen, and Trond H Diseth. Piloting the use of a short observation list for asd-symptoms in day-care: challenges and further possibilities. *Journal of Autism and Developmental Disorders*, 50:3413–3423, 2020. 5

[24] Catherine Lord, Susan Risi, Linda Lambrecht, Edwin H Cook, Bennett L Leventhal, Pamela C DiLavore, Andrew Pickles, and Michael Rutter. The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, 30:205–223, 2000. 2

[25] Bruce D Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI'81: 7th international joint conference on Artificial intelligence*, volume 2, pages 674–679, 1981. 3

[26] Meghan Miller, Ana-Maria Iosif, Monique Hill, Gregory S Young, AJ Schwichtenberg, and Sally Ozonoff. Response to name in infants developing autism spectrum disorder: A prospective study. *The Journal of pediatrics*, 183:141–146, 2017. 2

[27] A. Nadig, S. Ozonoff, G. Young, A. Rozga, M. Sigman, and S. Rodgers. A prospective study of response to name in infants at risk for autism. *Archives of Pediatrics and Adolescent Medicine*, 161(4):378–383, 2007. 5

[28] Aparna S Nadig, Sally Ozonoff, Gregory S Young, Agata Rozga, Marian Sigman, and Sally J Rogers. A prospective study of response to name in infants at risk for autism. *Archives of pediatrics & adolescent medicine*, 161(4):378–383, 2007. 5

[29] Quinlan J R 1986 Induction of decision trees Mach. Learn.

1 81–106. 7

[30] P. Jonathon Phillips. Face Recognition Grand Challenge (FRGC). https://www.nist.gov/programs-projects/face-recognition-grand-challenge-frgc, 2010. [Online]. 3

[31] Eleni A Papagiannopoulou, Kate M Chitty, Daniel F Hermens, Ian B Hickie, and Jim Lagopoulos. A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social neuroscience*, 9(6):610–632, 2014. 2

[32] S K Raina, Vishav Chander, Ashok K Bhardwaj, Dinesh Kumar, Seema Sharma, et al. Prevalence of autism spectrum disorder among rural, urban, and tribal children(1-10 years of age). *Journal of Neurosciences in Rural Practice*, 8(3):368–374, 2017. 1

[33] Shyam Sundar Rajagopalan and Roland Goecke. Detecting self-stimulatory behaviours for autism diagnosis. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1470–1474. IEEE, 2014. 2

[34] A Rudra, M K Belmonte, P K Soni, S Banerjee, S Mukerji, and B Chakrabarti. Prevalence of autism spectrum disorder and autistic symptoms in a school-based cohort of children in kolkata, india. *Autism Research*, 10(10):1597–1605, 2017. 1

[35] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in the-wild challenge: Database and results. *Image and Vision Computing*, 47:3–18, 2016. 3

[36] N. J. Sasson and J. T. Elison. Eye tracking young children with autism. *Journal of Visualized Experiments*, 61:3675, 2012. 3

[37] L. Shi, Y. Zhou, J. Ou, et al. Different visual preference patterns in response to simple and complex dynamic social stimuli in preschool-aged children with autism spectrum disorders. *PLoS One*, 10(3):e0122280, 2015. 3

[38] P Teitelbaum, O Teitelbaum, J Nye, et al. Movement analysis in infancy may be useful for early diagnosis of autism. *Proceedings of the National Academy of Science of the United States of America*, 95(23):13982–13987, 1998. 6

[39] F R Volkmar, C Lord, A Bailey, R T Schultz, and A Klin. Autism and pervasive developmental disorders. *J Child Psychol Psychiatry*, 45(1):135–170, 2004. 1

[40] E. Werner, G. Dawson, and N. Osterling, J.and Dinno. Recognition of autism spectrum disorder before one year of age: a retrospective study based on home videotapes. *Journal of Autism and Development Disorders*, 30(2):157–162, 2000. 3

[41] Lorna Wing and Judith Gould. Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of autism and developmental disorders*, 9(1):11–29, 1979. 1