Cognitive Computation (2022) 14:1773–1800
https://doi.org/10.1007/s12559-021-09895-w

Automated Detection Approaches to Autism Spectrum Disorder Based on Human Activity Analysis: A Review

Sejuti Rahman[1] · Syeda Faiza Ahmed[1] · Omar Shahid[1] · Musabbir Ahmed Arrafi[1] · M. A. R. Ahad[2,3]

Received: 26 August 2020 / Accepted: 29 March 2021 / Published online: 27 July 2021

Abstract

Autism Spectrum Disorder (ASD) is a neuro-developmental disorder that limits social and cognitive abilities. ASD has no cure so early diagnosis is important for reducing its impact. The current behavioral observation-based subjective-diagnosis systems (e.g., DSM-5 or ICD-10) frequently misdiagnose subjects. Therefore, researchers are attempting to develop automated diagnosis systems with minimal human intervention, quicker screening time, and better outreach. This paper is a PRISMA-based systematic review examining the potential of automated autism detection system with Human Activity Analysis (HAA) to look for distinctive ASD characteristics such as repetitive behavior, abnormal gait and visual saliency. The literature from 2011 onward is qualitatively and quantitatively analyzed to investigate whether HAA can identify the features of ASD, the level of its classification accuracy, the degree of human intervention, and screening time. Based on these findings, we discuss the approaches, challenges, resources, and future directions in this area. According to our quantitative assessment of the dataset Zunino et al. (IEEE: 3421–3426, 2018 [1]), Inception v3 and LSTM Zunino et al. (IEEE: 3421–3426, 2018 [1]) give the highest accuracy (89%) for repetitive behavior. For abnormal gait-based approach, the multilayer perceptron gives 98% accuracy based on 18 features from dataset Abdulrahman et al. (COMPUSOFT: An International Journal of Advanced Computer Technology 9(8):3791–3797, 2020 [2]). For gaze pattern, a saliency-metric feature-based learning Rahman et al. (Int Conf Pattern Recognit, 2020 [3]) gives 99% accuracy on dataset Duan et al. (Proceedings of the 10th ACM Multimedia Systems Conference: 255–260, 2019 [4]), while an algorithm involving statistical features and Decision Trees yields an accuracy of 76% on dataset Yaneva et al. (Proceedings of the Internet of Accessible Things. W4A '18, Association for Computing Machinery, New York, NY, USA, 1–10, 2018 [5]). In terms of the state of the art, fully automated HAA systems for ASD diagnosis show promise but are still in developmental stages. However, this is an active research field, and HAA has good prospects for helping to diagnose ASD objectively in less time with better accuracy.

Keywords Autism spectrum disorder · Activity analysis · Automated detection · Repetitive behavior · Abnormal gait · Visual saliency

Introduction

Autism Spectrum Disorder (ASD) is a complicated developmental disability. The signs and symptoms typically appear in early childhood and affect the child's ability to communicate (verbally and non-verbally) and to interact with others [6]. ASD is a diverse developmental disorder. It is described in terms of a set of behaviors and affects individuals in unusual ways and to varying levels. Studies have found that worldwide, 1 in 68 children are affected by ASD [7]. This estimate is an average and developing countries have a higher number of autistic children compared to

✉ Sejuti Rahman
sejuti.rahman@du.ac.bd

Syeda Faiza Ahmed
syeda.faiza.ahmed@gmail.com

Omar Shahid
omarshahid232@gmail.com

Musabbir Ahmed Arrafi
musabbir.arrafi@gmail.com

M. A. R. Ahad
atiqahad@du.ac.bd

1 Department of Robotics and Mechatronics Engineering, University of Dhaka, Dhaka, Bangladesh

2 Osaka University, Osaka, Japan

3 Department of Electrical & Electronic Engineering, University of Dhaka, Dhaka, Bangladesh

developed countries [8]. Based on epidemiological studies carried out over the last 50 years, the incidence of ASD 400 appears to be growing globally [9]. Until now, there has been no research that can establish a single cause for ASD [10]. Some research suggests that various types of environmental and genetic factors are behind the condition [11]. However, heightened awareness of ASD, early intervention, and access to appropriate services can stimulate positive outcomes [12] and can lead to noticeable improvements [12].

According to [13], ASD can be broadly categorized into 500 5 different types: Classical Autism or Kanner's Syndrome, Asperger's Syndrome, Pervasive Developmental Disorder (PDD-NOS), Rett's Syndrome, and Childhood Disintegrative Disorder (CDD). The level of severity is often determined based on the impairments noted. Classical autism causes mild to severe levels of functional disability, with Asperger's syndrome having less severe effects. PDD-NOS can affect the well-being of individuals to a mild to moderate degree, whereas Rett's Syndrome and CDD cause severe dysfunctional impairment and patients require a substantial support for their daily routines. All these neurodevelopmental disorders exhibit different physical and behavioral impairments which can disrupt an individual's cognitive, social, and physiological development in different ways. Rett's syndrome predominantly affects girls and one of its major indications includes repetitive hand and arm movements. In contrast, individuals with Kanner's syndrome demonstrate atypical eye patterns and PDD-NOS patients have difficulties walking and related physical activities due to gait abnormalities. In CDD patients, no significant functional change is noticed in the first two years, but they then start to show developmental abnormalities and the situation steadily degrades.

The health-care of children with ASD is complex and 600 involves several approaches, including promoting fitness, care, and rehabilitation along with special education, employment, and social care [14]. Early detection is the first and foremost requirement to ensure that ASD patients receive appropriate attention and care [15]. State-of-the-art ASD detection techniques call for expert physician knowl- 700 edge. A single screening takes a significant amount of time as the whole process is manual and depends largely on the co-operation of the individual. A complicating factor is that the parents of suspected ASD child often do not want to accept the fact that their off springs has symptoms of a developmental abnormality and so fail to take prompt action. 800 Medical professionals currently identify ASD by following the DSM-5 guidelines [16]. This is a costly and lengthy process that involves several tasks requiring the active participation of health professionals. It includes observing a patient's behavior, interviewing the parents, screening different visual and hearing parameters, and genetic and neurological testing. However, DSM-5 is only capable of detecting autism in children older than four years [17]. This can delay 800 detection of ASD and can adversely affect children who have impaired basic abilities and social skills, and who exhibit abnormal behavioral patterns or activities which are habitual and repetitive. Appropriate monitoring of development of children with ASD is an essential precondition of early detection, ideally within the first 18-24 months of life [18]. For this purpose, a physician often use the Modified Checklist for Autism in Toddlers (M-CHAT) [19] - a screening test that determines if a child is at risk of having ASD. However, further testings is required to confirm the suspicion. According to recent findings [20], magnetic resonance imaging (MRI) can accurately diagnose 80% of babies with ASD at the age of 2. Although this method allows early diagnosis of autism, capturing brain images using MRI of younger children or people with a high level of anxiety, hyperactivity, or sensitivity to noise can be very discomforting and stressful [21].

To fill the gaps in existing ASD detection techniques, researchers have started looking for alternative methods whose outcomes strongly correlate with the medical diagnosis. With the recent advances in the field [22–24], Human Activity Analysis (HAA) has been proposed as a potential tool for the development of an automated ASD detection system. HAA is a field in computer science that involves recording sensor data from somebody movement, gesture, or motion and analyzing it to detect abnormal behavior [25]. An autistic child may show some abnormal facial expression 900 [26, 27], unusual behavior [28], repetitive action [29], atypical walking pattern [30] or reveal an irregular salient region in an image [31, 32]. All of these symptoms can be taken as potential classifiers for ASD. To develop an automated ASD detection system, these distinctive features of autistic individuals are used to build a model. The relevant data is fed into learning algorithms, which can be broadly categorized into two types: Feature Engineering (FE) and End to End (E2E) learning. The FE-based models require a feature set and a classifier, unlike E2E learning which uses raw data as the input and predicts the class of the input without requiring any external feature extraction.

Another line of research employs the latest learning technologies to analyze fMRI data to detect ASD. The core idea 11o is use resting state fMRIs, both from normal individuals and from autistic subjects and identify significant variations in brain activities to detect ASD [33]. To this end, the work in [34] proposed a scalable machine-learning framework that is claimed to be able to identify significant biomarkers of ASD based on fMRI data. The authors used a single layer perceptron and auto-encoders in the learning procedure to extract features and optimize parameters so as to predict whether and individual can be classified as having ASD. Other work in this direction [35] has built a Convolutional Neural Network (CNN)-based model for distinguishing ASD

120 children from typically developing (TD) children based on functional connectivity patterns in brain images.

Another avenue has used Virtual Reality (VR) technology
130 as an intervention platform for ASD patients. It has already shown good potential in improving the social functioning of ASD people by using virtually simulated interactive environments [36, 37]. In [38], the authors proposed a virtual environment to simulate a life-like driving experience for ASD individuals. The study finds the optimal difficulty level for the participant using prediction modelling based on SVM and ANN algorithms. Some VR technologies have also been used to develop therapies for ASD patients [39–41].

This paper gives a comprehensive overview of how automated detection systems have been used to diagnose ASD through analysis of a patient's motion.

Related Surveys

Previously, Hyde et al. [42] provided a review of supervised learning related to ASD. Their review was based on two perspectives. The first focused on 35 papers on ASD diagnosis using standard assessment criteria (e.g., ADI-R, ADOS, etc), genetic factors, neural imaging, kinematic data, and eye movement data. Among these 35 cases, only 3 used HAA to detect ASD. As a comparison, the authors looked at 10 other text related to ASD intervention but none of these had any significant connection with HAA-based detection.

Thabtah et al. [43], however, provided a review focused exclusively on studies which had used machine learning techniques to identify ASD the aim being to reduce screening time. reduce human intervention and improve overall performance. However, this review did not look any specific mode of HAA-based automated ASD detection.

Song et al. [44] gave a comprehensive survey focusing on the use of Artificial Intelligence (AI) in existing ASD assessments as well as with novel observational data. Existing methods place an emphasis on strengthening correlations between behavioral observations and ASD as well as highlighting the gaps in current clinical diagnosis methods and ways of improving them. In terms of novel observational data, the authors reviewed five studies which explored ASD classification based on behavioral features, Eye-tracking, Hand and Upper limb movement. They found that these data are very much associated with HAA, but their work on ASD detection did not discuss the HAA method based on gait, a lack which our review intends to fill.

Boucenna et al. [45] included a review of information communication technology (ICT)-based ASD intervention. Focusing on human-computer-interaction (HCI), this work surveyed a wide range of recent works devoted to effective therapeutic solutions, touch screen-based interactive devices, educational platforms, tele-rehabilitation, and avatar games for autistic children. It also highlighted areas related to the early development of imitation and joint attention in children with ASD. Therefore, this work was more focused on ICT interventions for ASD individuals rather than on ASD detection techniques in general

Sevin et al. [46] focused on solutions for repetitive behavior-based impairments in ASD individuals. The problem with autistic children is the repetitive pattern of their behaviors, and finding an effective solution requires a detailed activity analysis. However, this work did not include a survey of ASD classification techniques.

Reinders et al. [47] discussed the interesting bidirectional nature of physical activity (PA) and social functioning (SF). They concluded that PA might be a feasible solution for intervening in the different social, cognitive and physiological impairments affecting ASD individuals.

Finally, Scharoun et al. [48] looked at the effects on ASD individuals of PA interventions provided by experts in the field. As well as the current situation, the authors also briefly described future possibilities and the challenges of PA interventions.

Table 1 provides a summary of all the papers we surveyed.

Contributions

To the best of our knowledge, this paper is the very first review with an extensive and systematic study on automated ASD detection approaches using human activity analysis.

This work is not circumscribed to any particular learning, rather a whole range of algorithms presented in all the existing works that we found with a systematic literature search for every HAA-based ASD detection approach was explored in the paper. A thorough study of the state of the art methodologies, research challenges in this domain with their probable solutions and available resources pertaining to HAA-based ASD detection are provided here.

main objective of our paper

140

In addition to a qualitative assay, we provided a quantitative analysis by evaluating several methods with publicly available datasets. We have regenerated the results of those methods with the same train-test split as well as a different validation to objectively assess the performance. Furthermore, we compared each method against several alternative algorithms using the same scale. Throughout this quest, we address the following key research questions: 1) What are the main cues that have been utilized by the HAA-based approaches to detect ASD? 2) Which algorithms have been used in state-of-the-art methodologies in different approaches? 3) To what extent human intervention is required for the automated diagnosis? 4) How closely does the diagnosis of an automated method align with the physician's assessment? 5) What are the key challenges in HAA-based automated ASD detection? These questions have been

Table 1 A comparison between our study and the existing surveys on ASD detection

No.	Reference	Field of Study	Key topics/ Research questions	No. of viewed papers[a]	No. of discussed dataset[b]	Quantitative assessmentmethods[c]
01	Hyde et al. [42]	Automated ASD detection	Supervised learning in ASD classification	3	3	No
02	Thabtah [43]	Automated ASD detection	Machine learning in ASD diagnosis for development of novel methods and improving the current ones with less screening time and human intervention	-	-	No
03	Song et al. [44]	Automated ASD detection	AI Technologies in ASD Classification with existing tools and developing novel methods with observational data	5	5	No
04	Boucenna et al. [45]	ICT-based intervention of ASD	Use of ICT technology as an assistive tool for ASD individuals	-	-	No
05	Sevin et al. [46]	Medical intervention	Assistive strategies for transitions between daily tasks and activities of ASD individuals	-	-	No
06	Reinders et al. [47]	Social & physical influence of ASD	Relationship between physical activity and social functioning for ASD individual	-	-	No
07	Scharoun et al. [48]	Medical intervention	Impact of Physical Activity in ASD individuals	-	-	No
08	Ours	Automated ASD detection	The approaches, state-of-art methods, extant of human intervention, coherence with the existing tools and key challenges of Automated ASD detection with HAA	45	40	Yes

[a]HAA-based automated diagnosis-related papers

[b]Datasets for automated HAA-based diagnosis

[c]HAA-based automated diagnosis-related methods

conducive for the exploration of different aspects of the topic throughout the study.

Review Techniques

The existing ASD detection techniques are substantially focused on medical diagnostic approaches that require the involvement of an expert. This paper corroborated the effectiveness of automated HAA approaches for ASD detection with the recent researches conducted in this field. A systematic search was conducted to compile the publications which included different methodologies of ASD detection with data-driven techniques. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [49] approach. The search progressed with 4 phase flow of PRISMA outlined review search pattern. The pipeline labeled with steps "Identification", "Screening", "Eligibility", "Included" narrowed down the primarily acquired large number of papers. The publications under the timeline of 1st January 2011 to present were searched in the Mendeley database. And only Journal, Conference papers and Book chapters were taken into consideration for review in this study

The initial identification of articles was done using combinations of the keywords "autism", "ASD", "Repetitive Behaviours", "Gaze Pattern", "Visual Saliency", "Gait Pattern", "Gait Analysis" . 1719 papers were accumulated from the first phase and were screened for their relevance to the subject of this review based on the insights obtained from the title and abstract. It was a part of the second step "Screening" of the PRISMA flow which also includes the duplication removal from the papers of the first stage and the number tallied to 215 after this stage. In the third phase, the screened paper went through a series of inclusion-exclusion criteria to be chosen for the next step. For instance, many articles included "gaze" and "ASD" but their work was not related to ASD detection. The full text of the papers was studied to decide if it satisfies the conditions of the screening and 104 papers were selected after this phase. For final inclusion, the qualitative and

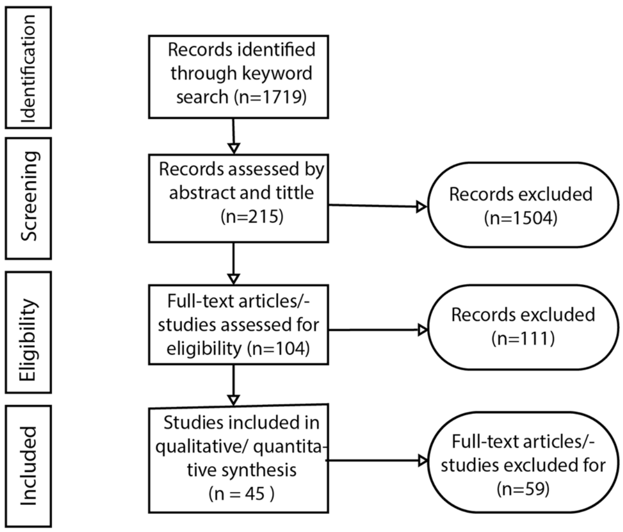


Fig. 1 Flow diagram of the method selection process based on the PRISMA approach

quantitative analysis of the papers were taken into consideration. With all conditions satisfied, only 45 papers were finally selected for this review literature. The complete flow diagram of method selection process with the PRISMA approach is presented in Fig. 1. So with this process finally publications were extracted from numerous initially identified papers. Only journals, conference papers and book chapters were picked to be reviewed in this paper. Thus a methodical approach was used to scheme out the most relevant papers for this review.

The rest of the paper is structured as follows: Dataset and Performance Metrics exhibits some of the publicly accessible data-sets with high relevance to our topic and performance metrics that are frequently used in describing the result of a learning technique. Ideal Experimental Setup sheds light on viable experimental environments to carry out the methods of data collection. In HAA-based ASD Detection Approaches, we described the state-of-the-art approaches by the researchers to develop an automated system of ASD screening. A quantitative analysis of methods with publicly available data-set has been shown in Quantitative Analysis of HAA-based ASD Detection Approaches. Challenges in this Domain is formalized with the challenges to proceed with research with the underlying hindrances for the practical experimentation of the research work. Discussion is a concise overview of our findings on this review paper and future scopes with real-life implementation possibilities. Finally, we have drawn the conclusion in Conclusion.

Dataset and Performance Metrics

Automated HAA-based approaches require dataset comprising different ASD-TD activities to train the FE or E2E learning-based models. In this field, most of the relevant datasets are not publicly available or open for research community. In this section, we will discuss the datasets that are publicly available for the detection of ASD using activity analysis.

The PRISMA [49] approach was followed for conducting the systematic search for datasets. The finally included datasets were found by going through the four steps record inclusion-exclusion screening method described in the PRISMA methodology. For the Identification phase, research papers were compiled by searching the keywords: “autism”, “ASD”, “Repetitive Behaviors”, “Gaze Pattern”, “Visual Saliency”, “Gait Pattern”, “Gait Analysis” and “Dataset”. Articles published within the last ten years were filtered and the Mendeley database was chosen for carrying out the search. The search initially produced only 30 in the Identification step. Then for screening the relevance to subject criteria and duplicate removal narrowed down the number to 16. And after a full-text assessment for eligibility, the relevant dataset count reached 7. For final inclusion, 6 datasets with public accessibility satisfied all the criteria and were selected to be included in the paper. The complete flow diagram of dataset selection process with the PRISMA approach is presented in Fig. 2. A dataset was selected on the basis of the following criteria: (a) The article is either a journal, Conference paper or book chapter, (b) The publication satisfied the conditions of PRISMA statements, (c) The data collection methods were compliant with the privacy-preserving rules, (d) The collected data included a group of individuals with ASD, (5 of the dataset were made up of real autistic people while one paper had involved volunteers who mimicked the movement of ASD patients) (e) It was also assured the annotations were made by expert physicians.

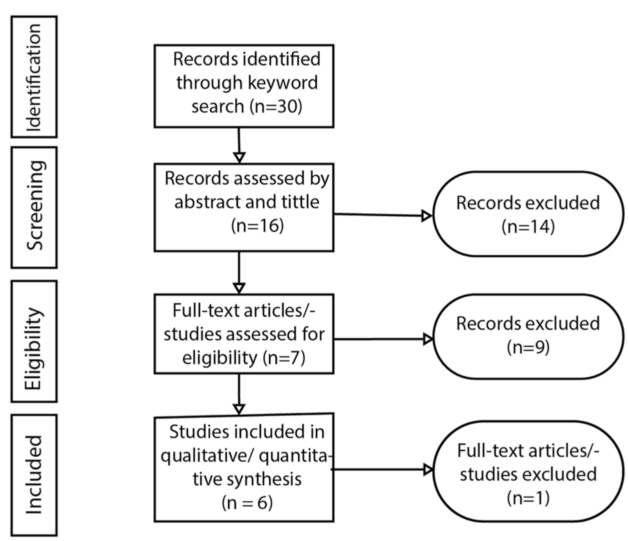


Fig. 2 Flow diagram of the dataset selection process based on the PRISMA approach

Skeleton Dataset

Skeletal data encode the human body posture in terms of relative 3D coordinates of different body joints and the joints' orientation angle. Rihawi et al. [50] developed the first publicly available 3D dataset named '3D-AD' based on ASD subjects using the Kinect-v2 camera. This dataset includes depth maps, which have been captured at 33 frames per second. The sequence of skeleton joint features was collected for ten different actions, e.g., hands-on the face, hands back, hand moving front of the face, headbanging (or rocking back and forth), tapping ears, flicking, hands stimming, toe walking, playing with a toy, and walking in circles. This research reported Dynamic Time Wrapping (DTW) distance as a distinguishable feature to detect ASD and TD.

In another study, Al-Jubouri et al. [2] have published a publicly available 3D skeleton-based gait dataset of 50 children with Autism and 50 typically developing (TD) child. In the video data captured in Kinect v2, the participant walked along a line at their normal speed. The skeletal data included the joint position of 25 three-dimensional joint positions, sixteen angles and on gait cycle from 2 or more feature extracted from the data. Different data augmentation techniques: Jittering, Scaling, translation, Flipping and slicing were used to increase the diversity of data.

Video Dataset

To the best of our knowledge, Zunino et al. [1] developed the only available video dataset to analyze the action style of an autistic child. The dataset includes activities such as the task of placing, picking, passing, and grasping a bottle of a particular size by an autistic child. A Vicon VUE video camera with resolution: 1280 x 720 pixels and a frame rate of 100 frames/sec was used to devise the whole experimentation. The study includes 20 TD and 20 ASD children as subjects, whose state of autism was confirmed by the DSM-5 method. All participants were in between 7 to 12 years of age.

Gaze Dataset

Eye-tracking refers to the process of measuring visual attention. These measurements are captured using an eye-tracking device that records the positions and movements that our eye makes while viewing a scene. Most of the methods in the current literature utilized eye-tracking data in two forms: fixation data and scan path data. To this end, Duan et al. [4] developed an eye movement dataset named 'Saliency4ASD' from 14 ASD and 14 typically developed children. All the participants were in between 5 to 12 years of age. During the eye-tracking process, the participants viewed 300 images, including natural scenes, animals, the human body, and objects. This dataset provided fixation points, fixation maps, heat maps, and scan path data of the participants.

Yaneva et al. [5] also developed a gaze dataset, which included autistic adults instead of children. The participants were in between 30 to 40 years of age. The experiment included 30 participants, where 15 were diagnosed with high functioning autism or Asperger's syndrome, while the rest were non-autistic individuals. In the data collection process, the participants were asked to perform visual tasks such as web browsing and searching using a mouse and keyboard according to the given instructions. The study provided fixation time on the area of interest in images, duration of searching task in a web page, and the number of fixation points. This work utilized this dataset to classify ASD individuals and healthy people applying simple logistic regression.

In recent work, Shihab et al. [51] provided a gaze dataset that comprises face-scanning data of adults and children diagnosed with ASD. The participants were in between 4 to 60 years of age. Participants were asked to view human face images and movie clips involving social interactions on the laptop screen while the eye-tracking data were recorded using two analog cameras placed in front of the laptop. This research studied the difference in the face-scanning pattern of ASD and TD individuals and classified ASD and TD using the PCA.

This paper reviews the existing methods of automatic ASD detection with HAA, and in most cases, the methods produce the classification result by feeding the data to different learning algorithms. Therefore, most of the methods reported their performance in terms of one or multiple performance metrics. Most commonly, the efficiency of a method is assessed with accuracy (i.e., ratio of the correctly labeled samples to a total number of samples), recall or sensitivity (i.e., the proportion of actual positive cases that got predicted as positive, in other words, true positive), specificity (i.e., the proportion of actual negatives, which got predicted as the negative or true negative), precision (i.e., the number of positive class predictions that actually belong to the positive class) and F1 score (i.e., the harmonic mean of the precision and recall.). Besides, visual or graphical tools like receiver operating characteristic (ROC) curve and area under the curve (AUC) are used in measuring the performance of models.

Table 2 presents a summarized description of the available datasets with information about the samples, capturing devices, scenarios, and limitations.

Table 2 Detailed description of publicly available datasets for ASD detection through activity analysis

Id	Reference	Data Type	No. of Subjects	Approach	Capturing Device	No. of Instances	Age (years)	Scenarios	Limitations
1.	Riwahi et al. [50] weblink	Skeleton	ASD-15 TD-14	Repetitive behavior	Kinect v2	1200	7 to 16	Data were recorded in a closed environment. Instructors demonstrated the actions (e.g., hand moving in front of the face, toe walking, walking in circles, etc.) first, and then the participants followed.	No medical diagnosis was performed to confirm the state of the participants (i.e., ASD or TD).
2.	Zunino et al. [1] weblink	Video	ASD-20 TD-20	Repetitive behavior	Vicon VUE camera	1100	7 to 12	Participants performed activities such as placing, picking, grasping and passing a bottle to the volunteers to the other end of the table.	The number of video instances is not good enough to train a DNN.
3.	Al-Jubouri et al. [2] weblink	Skeleton	ASD-50 TD-50	Gait Pattern	Kinect v2 motion capture	800	4 to 12	Participants were asked to walk at normal speed for 2 gait cycles in the range of 1.5 to 4 meters in front of the camera.	- Incomplete lateral body movement capturing due to self occlusion -The range of skeletons tracking by Kinect v2 prevented the capturing of two gait cycles -The reduction in frame rate while recording the skeleton data
4.	Duan et al. [4] weblink	Gaze	ASD-14 TD-14	Gaze pattern	Tobi Pro X3-120	600	5 to 12	Participants viewed the images shown on the screen.	This dataset did not provide individual subject's eye-tracking data.
5.	Yaneva et al. [5] weblink	Gaze	ASD-20 TD-16	Gaze pattern	GP3	1200	30 to 40	Participants performed visual tasks, i.e., browsing and searching in web pages, using mouse and keyboard according to the instructions.	Recorded the eye-tracking data of six images only.
6.	Shihab et al. [51][a]	Gaze	ASD-44 TD-58	Gaze pattern	Analogue camera	800	4 to 60	Participants were asked to view images of human face or a movie involving social interactions on a laptop screen.	All participants did not view same images.

[a]The dataset is available upon request

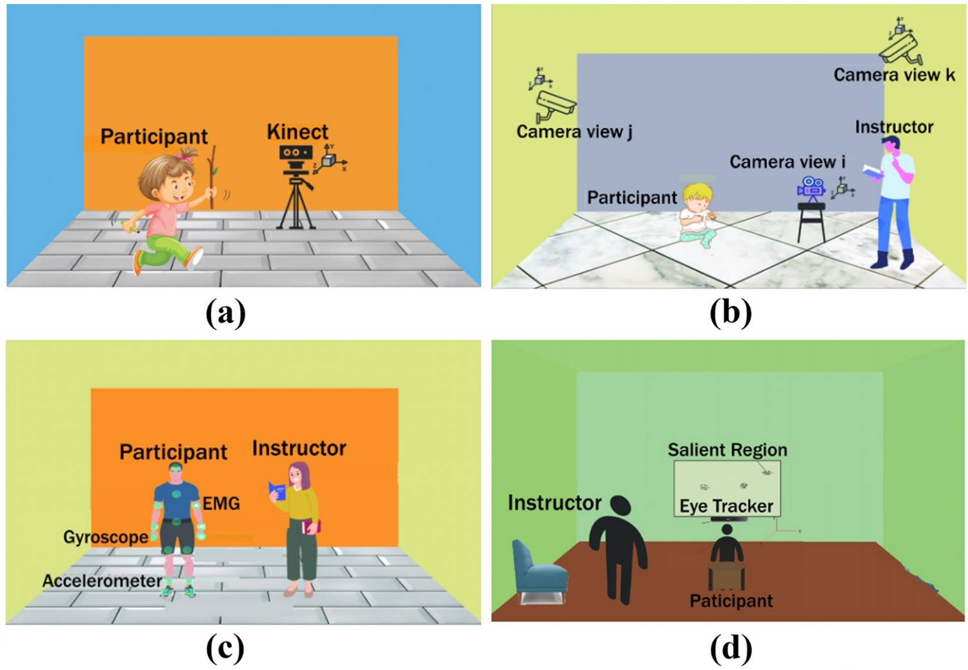


Fig. 3 Sample experimental setups for capturing data using (**a**) Kinect, (**b**) video camera, (**c**) wearable sensors, and (**d**) eye-tracker for the detection of ASD

Ideal Experimental Setup

The experimental setup for data collection in activity analysis is preferred to be performed in constrained environments. It minimizes the additional features introduced in computational modeling due to different backgrounds or surroundings [52]. This makes the model more focused on the specific activity's dynamic features, which is to be analyzed. Consequently, this compels it to be well trained in detecting the desired activity primarily. Further computational development to this basic model with complex data can improve its capabilities to perform well in different environments. Figure 3 provides experimental setups for capturing skeleton, video, wearable sensors, and eye-tracking data for the detection of ASD.

Figure 4(a) shows a simple setup for skeleton data recording using a single view Kinect device where a subject is viewed to be executing some actions in motion. The problem of self-occlusion happens when image data of a desired object is obstructed by itself [53]. Therefore, to reduce this drawback, the camera needs to be positioned to capture the individuals' front view images. The quality of the data can be more subserved by introducing more cameras in the motion extraction system [54]. Also, the background needs to be uncluttered and static in this kind of experimental setup.

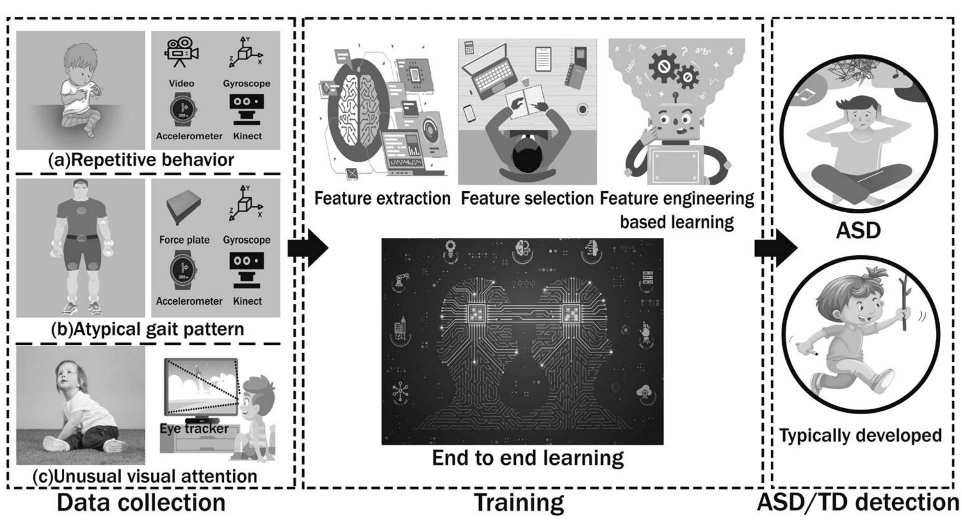


Fig. 4 The activity analysis-based automated detection of ASD comprises three main steps: data collection, training, and classification. The data collection block shows three core approaches; each of them utilizes different sensors to capture different exclusive characteristics of autistic individuals: repetitive behavior, atypical gait pattern, and unusual visual saliency. Next comes the training phase, which utilizes feature engineering-based or end to end learning to learn discriminative features of the data to classify ASD and TD in the final step of the automated detection

The setup for video data collection should consider several issues. It has a stipulation for constant illumination as variable illumination in a scene may create complexity to the environment [55]. Moreover, a dataset created from a single observational viewpoint tends to address the problem of featuring less specifications of the captured data. In contrast, a multiview based video dataset facilitates the operations for activity analysis by attributing better clarity and comprehensibility for the system [56]. In Fig. 4(b), an individual exhibits a specific action for a video clip, and every single action is necessitated to be clicked off in correspondence with a fixed viewpoint and against a quiescent background. The entire experimental setup should be held up as such the autistic children do not feel any discomfort in the whole process. Apart from these setup arrangement complications, a skilled instructor is ought to monitor the ASD individuals to perform the desired task.

The sensor-based data collection methods are applicable in different contexts due to its flexibility of having both the options of the wired or wireless flow of data from the environment to the data processing units [57]. Seemingly for integrating this technology into our portion of the research, it was requisitioned to form a wireless setup of the sensor for the involvement of autistic children. The data streamed from the sensors is often fraught with noise and inferences, so the structure of data classification should be made by considering these in appropriates [58]. In Fig. 4(c), the developed wearable sensor-based network is embedded with accelerometers, gyroscopes, and EMG (electromyography) sensors. They are placed in different body parts like chest, hand, wrist, torso, kneecap, legs, and feet of the participant who is instructed to perform certain activities. The sensors track the positional data of those parts during different activities.

Eye-tracking technologies offer great conveniences in data collection for research studies where experimentation includes participants with impaired organ functionality. So, this technology provides great ease to perform different empirical observations of ASD detection using activity analysis [59]. As demonstrated in Fig. 4(d), a participant is only required to be seated and follow the instructions to perform a visual task. His visual attention has been captured through eye-tracking devices attached to the screen. Though some head-mounted eye-tracking devices with higher accuracy of data extraction are available, to capture ASD individuals' data, eye trackers without any body contact are mostly used [60].

HAA-based ASD Detection Approaches

Studies on autistic children have found some unusual behavior and response by an ASD child [61]. As automated detection aims to automate the whole diagnosis process, so the system needs to focus on those identifying characteristics, e.g. repetitive behavior, atypical walking style and particular visual saliency. In this section, we have explored the recent literature on activity analysis-based ASD detection by arranging them in three main groups. Figure 4 shows the overall process of the activity analysis-based automated detection, followed by the existing approaches in this domain.

Study of Repetitive Behavior

Repetitive behavior refers to showing abnormality in behavior, characterized by repetition, inappropriateness, less adaptability, and rigidity [62]. Repetitive behavior is a common observation of the parents of autistic and normally developing children; however, the pattern of repetition of the autistic individuals significantly differs from the TD children [63]. However, there have been noticeable differences between repetitive behaviors of ASD and TD children in qualitative and quantitative characteristics [64, 65]. It is important to note that repetitive behavior, being a core symptom of ASD and a prominent cue to identify autism, requires an expert eye for the diagnosis. Therefore, automated detection of ASD based on the repetitive behavior cue has been a potential research avenue. Figure 5 provided a visual representation of how an ASD individual can express repetitive behavior.

Goodwin et al. [66] represented a pattern recognition algorithm with three-axis accelerometer data to automatically detect hand flapping and body rocking movement, which got approximately 90% accuracy in both classroom and laboratory environment. Therefore, this study included both controlled and real-world experimental results. Goncalves et al. [67] proposed automatic detection of hand flapping movement using two systems- a Kinect sensor and a watch with an accelerometer device. This study found better results using the watch with an accelerometer attached to the wrist by applying statistical analysis on the collected data. The study of Zhao et al. [68] demonstrated the significance of kinematic features to distinguish ASD from TD using FE-based learning. The authors used hand flapping activity recorded

Fig. 5 A visual representation of some of the repetitive behaviors showed by autistic individuals

with a LeapMotion device for the classification. Five classifiers: Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Decision Tree (DT), Random Forest (RF), and K Nearest Neighbor (KNN) have been applied on four kinematic features (i.e., maximal amplitude range, minimal velocity range, median velocity entropy, and minimal velocity entropy), where KNN gives the best accuracy of 88.37%.

Meanwhile, Jazouli et al. [53] focused on real-time automatic identification of five repetitive behaviors, i.e., hand flapping, hand on the face, hands behind back, fingers flapping, and body rocking, in individuals with ASD. In their work, they had used a 3D skeleton dataset to characterize repetitive behavior and reported an above 93.3% recognition rate using Artificial Neural Network (ANN). Later on, Jazouli et al. [69] proposed a point-cloud recognizer method based on the nearest-neighbor classifier with the Euclidean distance function that classifies five repetitive behaviors of their previous study.

Coronato et al. [70] demonstrated an E2E learning-based real-time method that detects ASD by tracking repetitive behaviors that are symptoms of a status of anxiety or of isolation from the surrounding environment. The authors used 3D-axis accelerometer data and analyzed quite different cues, such as hand hitting against the ear, arm flapping, hand rotation in both up and down directions. It reported an average of 92% online and 99% offline classification accuracy.

In another work, Sadouk et al. [71] proposed convolutional neural networks (CNN)-based algorithm for detecting repetitive behaviors within and across subjects. Jaiswal et al. [72] suggested an E2E dynamic deep learning method using 3D analysis of behavior for ASD-TD classification, and this study attained a classification rate of 96%. Moreover, their study used RGB-D data for classification purposes. Riwahi et al. [50] provided the first-ever publicly available 3D skeleton dataset of repetitive behavior that ASD individuals usually show. This study used data recorded by volunteers who were not diagnosed with ASD that indicated a drawback. However, their study provided a wide range of repetitive actions, including hand moving front of the face, toe walking, walking in circles, and playing with a toy, which is not common in most studies.

Zunino et al. [1] analyzed hand gestures during a bottle grasping task with four different underlying intentions, i.e., picking, placing, pouring, and passing a bottle. The study found that the execution of this simple task is different between ASD and TD individuals and, therefore, can be effectively utilized to discriminate between these two groups. Their study reported 82% classification accuracy using an LSTM network followed by a VGG-16 architecture. Later on, Tian et al. [73] used temporal pyramid network, and Sun et al. [74] used SA-B3D with the LSTM network on the same dataset [1]. These latter studies reported 87.17% and 95.2% classification accuracy, respectively. In another study, Kumdee et al. [75] proposed an E2E learning framework to classify ASD/TD using video-based hand flapping movement. This system ensures its robustness by recording data from multiple views and reported an average accuracy of 91.15%. However, the method is yet to be tested on people with ASD. In addition to that, Albinali et al. [76] and GroBekhatofar et al. [77] provided a descriptive study on hand flapping and body rocking using various FE-based learning techniques and found better results using decision tree algorithm.

Abnormal Gait Recognition

Abnormal gait is defined as an atypical style of walking that may cause stagnation in occupational and other ample ranges of daily activities. Numerous studies have reported that analyzing abnormal gait features is a powerful tool for early diagnosis with proper treatment planning for individuals with ASD [78]. In general, gait recognition techniques use temporospatial features like stride length (distance between successive ground contacts of the same foot), step length, step width, cadence (steps per minute), velocity, stance time (the duration that passes during the stance phase of one extremity in a gait cycle), and double support of both legs [79]. In Fig. 6, we have shown an illustration of an abnormal gait pattern by an ASD child.

Calhoun et al. [80] provided a comprehensive study to compare kinetic and kinematic gait parameters of ASD and TD children. This study found noticeable differences between cadence, peak hip, and ankle kinematics by applying sound approaches of statistics and principal component analysis. Hasan et al. [81] proposed an automated technique for the classification of gait patterns of ASD individuals based on the kinetic and kinematic features with the aid of advanced machine learning approaches. Here, impactful features were selected statistically using the t test and Mann–Whitney U test. However, this approach suggested Linear Discriminant Analysis (LDA) with kinetic gait features as input, and this provided 91.7% classification accuracy.

Fig. 6 An illustration of abnormal gait patterns of people with ASD

In another study, Hasan et al. [82] presented an approach of automatically identifying the gait patterns of the ASD using three-dimensional ground reaction forces (3D-GRF). This study was a binary classification of ASD and TD children, whereas time series parameterization was applied to find important gait features. The authors used k-nearest neighbor (KNN) as classifier, which leads to having 83.33% performance accuracy. On the contrary, Hasan et al. [83] analyzed vertical ground force (VGF) in terms of walking speed, frequency, and push rate by providing a statistical analysis with a p value to represent the differences between ASD/TD walking style. In a different study, Hasan et al. [84] investigated on 3D Kinematic data of ASD individuals. They used statistical features and reported 91.7% accuracy using ANN classifier. Ebrahimi et al. [85] proposed a markerless approach using a 3D sensor to distinguish between tip-toe walking and regular walking. This study used the time domain and frequency domain feature to get the best set of feature combination for classification. However, it got 86% accuracy with the linear support vector machine (SVM) technique. Ilias et al. [86] used a fusion of tempo-spatial and kinematic features to classify gait patterns, and this approach got 87% accuracy with neural network classifier and SVM with polynomial kernel individually. With SVM polynomial as kernel attains 100% sensitivity and 85% specificity, show the efficacy of their approach in utilizing the SVM to identify autism. In a different study, Ilias et al. [87] used a motion analysis system to capture 3D features of gait data, and their analysis used 21 gait features : 4 features from basic temporal spatial, 5 kinetic parameters, 12 features from kinematic, which finally reported 98.44% classification accuracy using FE-based NN classifier, while [88] proposed a feature engineering-based method using three types of gait characteristics (basic, kinetic and kinematic) of ASD/TD to distinguish them, and their study reported 87% classification accuracy.

Henderson et al. [89] made comprehensive research to determine the effect of intra-subject variance in gait to distinguish ASD and TD. Their analysis was based on the time duration between two-foot step of the participants to complete a gait cycle, and they applied several FE-based learning algorithm. Finally, they reported KNN algorithm has a better performance of 76.67% accuracy in spite of having an intra-subject variance, and SVM has the least performance as it can not deal with intra-subject variance properly. Shigeta et al. [90] designed an experiment to analyze walking patterns of participants in pair using IMU-motion capture system. Their study used the angular velocity of four body parts (waist, left/right foot, and head) and performed multiple regression analysis. Their findings reported ASD individuals showed distinctive walking patterns when they pass another person during walking rather than walking alone. Dufek et al. [91] provided a comprehensive study on gait analysis of children aged between 5 to 12 years old. Their method investigated significant differences in terms of the p value of the hip, knee, and ankle joint positions as well as VGF. Their study reported that children with ASD exhibited abnormal differences in impact force attenuation and walking step control strategies when compared to a TD individual. In another study, Al-Jubouri et al. [2] proposed the classification of ASD and normal children with 3D Skeleton-based gait data captured with Kinect V2. They demonstrated 95% classification accuracy using Principal Component Analysis (PCA) and Multilayer Perceptron (MLP) in the skeleton data collected from 50 ASD and 50 TD children. Different data augmentation techniques were also applied to the dataset that increased the diversity of their data for training the model.

Analyzing Gaze Pattern

A gaze pattern refers to the viewing style of an individual towards a scenario or an image. Irregular gaze pattern decodes the abnormality of the activation of the social brain network, the inability of social communication, and atypical brain response [92]. Individuals with ASD show unusual gaze patterns, which can be determined through some eye-tracking matrices like heatmap, scan path, fixation point, fixation map, and attention time. Heatmap is a visualization tool that demonstrates a general distribution of gaze points in an image. It is illustrated as a color-gradient overlay on the stimulus, e.g., an image. Figure 7 is a representation of the heatmap of an image, where we can notice that TD and ASD individuals are different as their visual attention is not the same. In general, the gaze points or fixations (i.e., points in the visual field that are fixated by the two eyes) are distributed more densely towards the area that gets more visual attention than other areas of the image.

Scanpaths are also used to visualize an individual's gaze pattern through a series of dots and fine lines. Figure 8 represents scan paths for both TD and ASD children in a social scene. In the figure, fixation points and saccades (rapid movement of the eye between fixation points) are represented by dots and lines, respectively. Besides, the size of the dots represents the duration of corresponding fixation points.

Wang et al. [31] experimented on the gaze pattern of the TD and ASD individuals of mean age 32.3 years and 30.8 years respectively. The subjects' gaze patterns were converted to probabilistic heatmaps according to their visual attention in a video. Eventually, a heatmap score, calculated based on the statistical features, used to differentiate a healthy child from an individual with ASD. In another work, Jiang and Zhao [93] proposed a method based on Deep Neural Network (DNN) to differentiate the gaze pattern of regular people and ASD individuals. This work selected images based on the Fisher score [94], which ensures that those images would allow the learning

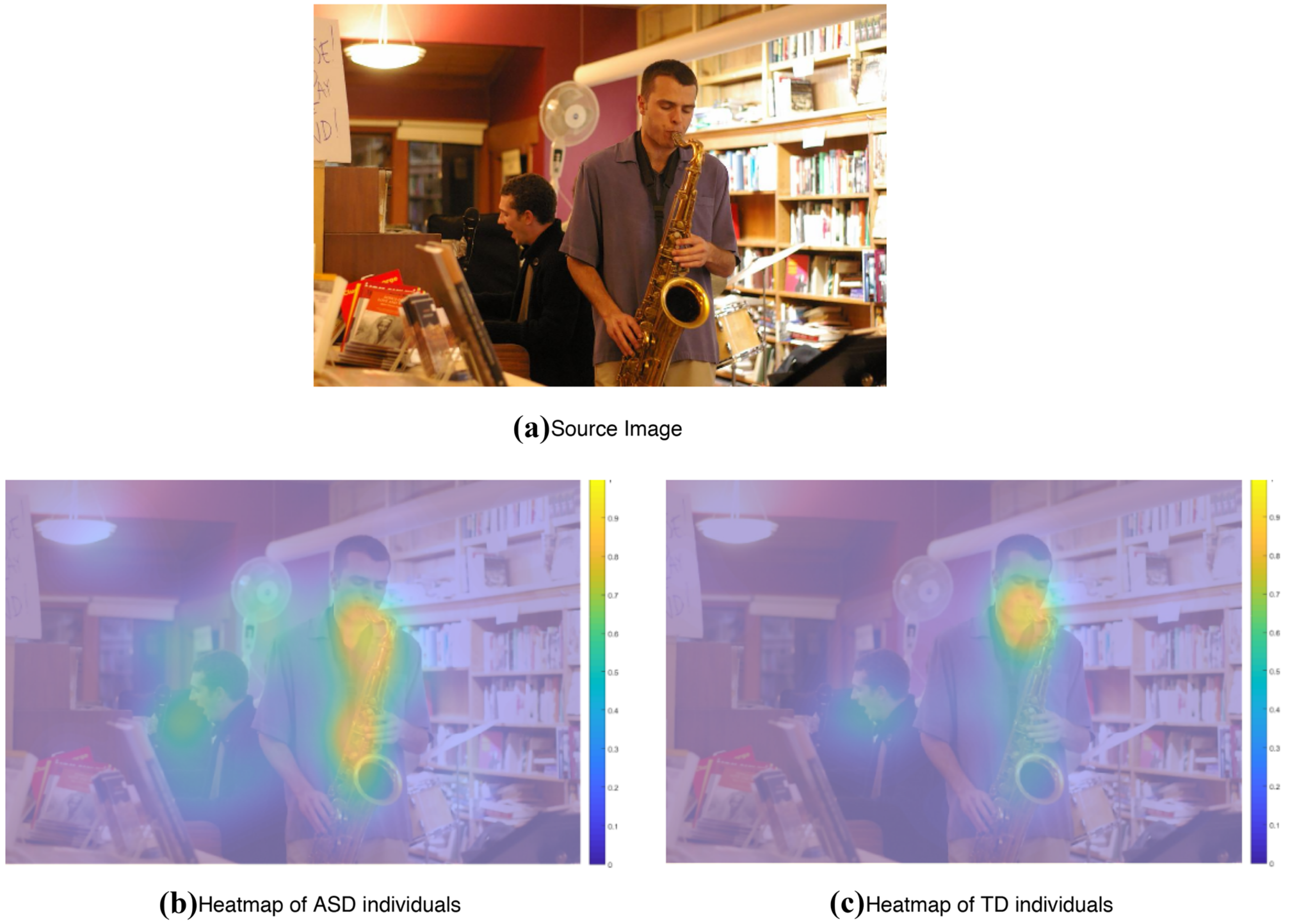


Fig. 7 Heatmaps of ASD and TD shown superimposed on the input image [4], where ■ is the most viewed region, and ■ the least viewed region. (**a**) The source image, which was used as visual stimuli to analyze the gaze pattern; (**b**) An ASD individual shows more attention to the objects rather than the human face; (**c**) A TD individual gives more attention to the human face

of discriminative features from image contents. Furthermore, this developed model evaluated multiple matrices, e.g., attention time, pupil movement, expression change, which are used in diagnostic tests, and it got an overall accuracy of 92%.

In a related study, Cho et al. [95] used fixation points for ASD detection and reported 93.96% classification accuracy using the KNN algorithm. Startsev et al. [96] reported a comprehensive study on scan path data and fixation maps using the Random Forest algorithm. This study used statistical features for classification and reported 76.5% accuracy. In another study, Chen and Zhao [97] proposed a deep learning approach using an LSTM network followed by a ResNet-52 architecture and achieved 93% classification accuracy. Moreover, this study provided a novel method of ASD/TD classification using a photograph-taking task where the participants were asked to take photos of their region of interest in a natural scene. Furthermore, Tao et al. [98] proposed an LSTM architecture followed by a CNN network to classify ASD/TD using their scanpath in an image set. Their proposed method achieved 74.22% classification accuracy, moreover, in this paper authors have shown some visualization of data to represent clear difference between scanpath of ASD and TD children. In a recent study, Rahman et al. [3] proposed a novel feature extraction technique using saliency map to classify eye-tracking data through fixation points. The authors compared several conventional saliency map with fixation maps of the participants to calculate the feature set, and reported 99% classification accuracy using XGBoost classifier.

Wan et al. [99] analyzed the fixation time of six different areas of interest and able to discriminate ASD from TD with a classification accuracy of 85.1%. Yaneva et al. [5] provided a study on adult ASD individuals by analyzing their capability in tasks such as web browsing and searching. This study investigated the area of interest, fixation time, and fixation points for classification. They reported 75% classification accuracy using simple logistic regression. On the contrary, Babu et al. [100]

Fig. 8 Scanpaths of ASD and TD shown superimposed on an input image [4]. (**a**) The scanpath of an ASD individual concentrates more on the background scene including objects than the human face, whereas, (**b**) a TD individual's scanpath concentrates mostly on the human face

(a) Scanpath of an ASD **(b)** Scanpath of a TD

analyzed the responses of ASD and TD children in a virtual reality-based social scene and designed an experiment to collect the gaze-related physiological indices (PIs) and behavioral looking pattern indices (BIs) of a virtual social scene. Their method used statistical analysis to extract feature, used Classification Tree (CT), Regression Tree (RT), Bayesian Network Tree (BNT) and Support Vector Machine (SVM) for designing their classifier where SVM classifier reported 97% accuracy in average. Nebout et al. [101] designed a coarse-to-fine convolutional neural network (CNN) to predict saliency maps for ASD children that provides better results than 6 of existing saliency models. This study reported that no centre bias is applicable for the visual attention of individuals with ASD, which contradicts the findings of other studies in [31, 102]. Dris et al. [103] proposed a method of classifying ASD using fixation duration on different regions of interest in the image and got 88.6% accuracy; 92.31% specificity; 86.63% sensitivity using an SVM classifier.

Syeda et al. [104] provided a comprehensive study of face scanning and emotion recognition of ASD and TD children. The study found that individuals with ASD show less attention to prime features of faces like eyes, nose, and mouth during face scanning and experience more difficulty in perceiving basic human emotions. Sadira et al. [105] analyzed the face-scanning pattern of ASD and TD people. Their fixation time analysis found that ASD individuals spend more time looking at the mouth rather than the eyes or nose in a human face. Arru et al. [106] analyzed the scan path of both ASD and TD children and extracted fixations as well as a center bias to classify them. This study developed a decision tree-based classifier and reported 60% accuracy on the test set. Liu et al. [107] proposed a machine learning-based approach to classify ASD/TD children by analyzing their face-scanning pattern in a face recognition task. It reported 88.85% classification accuracy. Recently, Shihab et al. [51] provided a comprehensive study on children and adults with ASD. It analyzed the scan path data with principal component analysis (PCA) and developed an unsupervised classifying method, which gives a sensitivity of 78.6% and specificity of 82.47%. Apart from this, Alie et al. [108] used variable markov model (VMM) to detect abnormal eye gaze pattern of 6-month-old children who had a risk of developing ASD in the future. The experimentation included 32 infants, where 18 had siblings with autism and 14 had normal older siblings. Their method actually classifies whether the child have risk or not using their gaze pattern, and reported 93.75% accuracy.

The automated HAA-based ASD detection needs human intervention to some extent, however it prospects to develop a fully automated diagnosis system for ASD detection. Experimentation of each of the approaches has some common as well as special requirements that set out to limit the fully automated operationalization of the systems. The general issues may include establishing a perfect setup with the accurate organization of equipment, ensuring the standard of the required annotating steps, calibrating the devices several times for unerring data without overlooking the fact than an autistic individual is participating in the process, conducting data analysis and reckoning for decisive statements and many more. In repetitive behavior and gait pattern-based methods, the necessitated instruments like ground attached sensor, motion capture camera, wearable sensors, etc. need dexterous supervision which may demand considerable human intervention in the processes. The gaze data-based studies involve eye tracker and other vision feature tracking measures in their data collection, which is also associated with operational complexities. The selection of the visual stimuli also requires substantive human interference. Following the qualitative assay presented in this section, we see that the whole pipeline of HAA-based automated ASD detection demands human involvement in two different phases: 1) Data collection and annotation, and 2) Data analysis. The majority of the existing methods require human intervention in the first phase. On the other hand, the statistical analysis and FE learning-based systems need human intervention in the second phase as well.

In Fig. 9, we provide a tree diagram of HAA-based automated ASD detection approaches. Starting from the root,

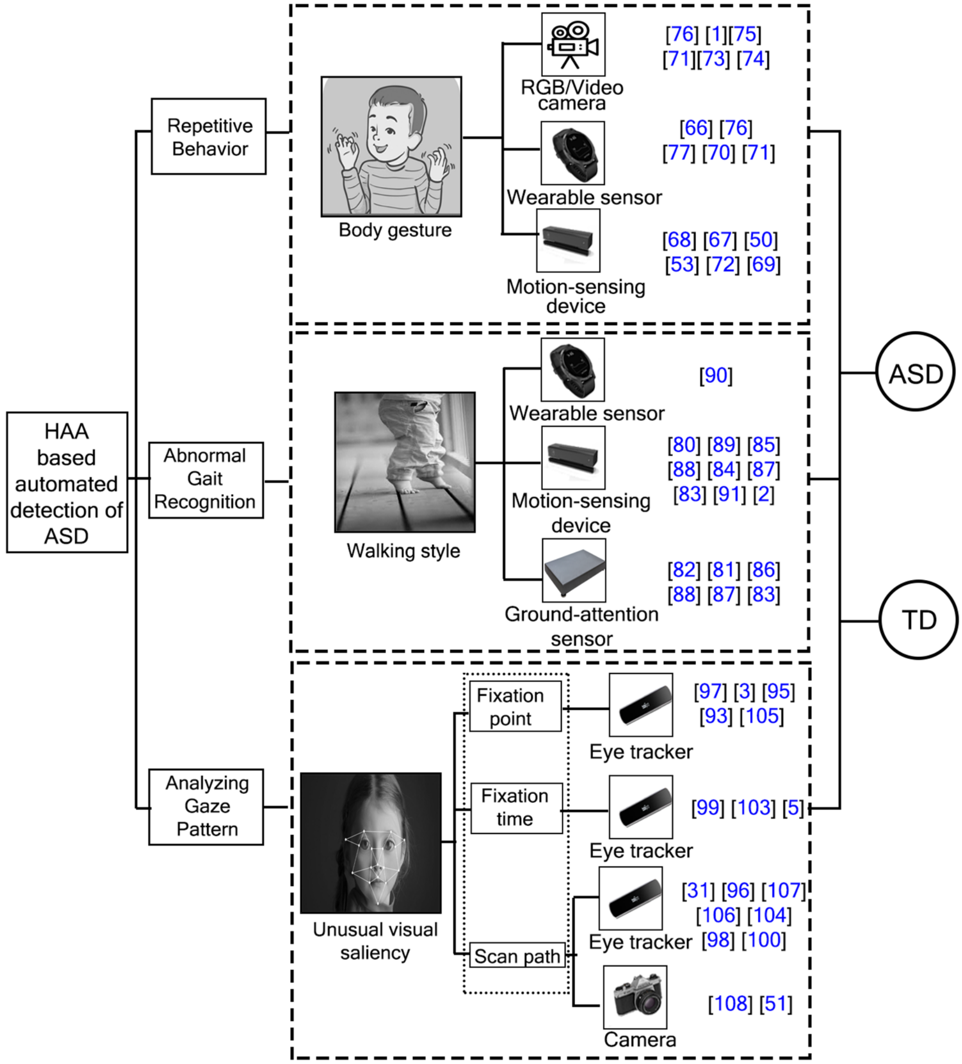


Fig. 9 A tree diagram of HAA-based automated ASD detection approaches. Starting from the root, three branches are out to represent the main approaches: repetitive behavior, abnormal gait, and gaze pattern. Then, each category is subdivided based on the devices they utilized for sensing the activities

three branches are out to represent the main approaches: repetitive behavior, abnormal gait and gaze pattern. Then, each category is subdivided based on the devices they utilized for sensing the activities. A detailed description of activity analysis-based studies of ASD detection is given in Table 3.

Quantitative Analysis of HAA-based ASD Detection Approaches

In HAA-based ASD Detection Approaches, we presented a qualitative analysis of 45 methods from the literature where the datasets, number of subjects, and evaluation protocols varied from method to method. This non-uniformity does not allow us to assess the performance of different methods on the same scale. Therefore, being motivated by the urge to present an objective assessment, here we conduct a quantitative analysis where methods under each of the three approaches are evaluated using the same standard. We have used 4 publicly available datasets, one for the repetitive behavior [1], one for the gait pattern [2], and 2 datasets for the saliency analysis [4, 5]. To compare the performance of each method under study, we consider frequently used architectures, and feature extraction methods that are compliant with the type of the dataset, to act as alternative methods for the comparative analysis. Also, for consistency in the evaluation protocol, the performance metrics are computed with two cross-validations: one that is used by the authors to report the results in the original paper, and the other is selected by us for a confirmatory check. The results of the quantitative analysis are presented in Table 4.

For the repetitive behavior-based approach, we have evaluated the performance of the method proposed by Zunino et al. [1]. In that paper, the authors proposed an E2E learning using the Inception v3 framework [109] followed by the LSTM network [110]. Since the method used a video dataset, we have compared its performance against three

Table 3 A detailed description of activity analysis-based studies for ASD detection, grouped by core approaches (i.e., study of repetitive behavior, recognition of abnormal gait, and analysis of gaze behavior). Description of each study includes a method (Feature Engineering-based learning (FE), end to end learning (E2E), or statistical analysis (SA)); the name of the specific algorithm and cue/feature used for learning; followed by details of dataset (capturing sensor, age-range and no. of subjects), results in terms of ASD-TD classification acc(accuracy), sen(sensitivity), spe(specificity), FS(F1 score), and Time (realtime/ offline). Here, "-" denotes "not mentioned in the paper"

Id	Reference	Approach	Method	Algorithm	Cue/Feature Used	Sensor	No. of Subjects	Age (years)	Result	Time
1.	Goodwin et al. [66]	Repetitive behavior	FE	Decision Tree	Hand flapping, body rocking	Accelero meter	6	12-20	Mean acc- 91.1% [a] Mean acc- 90.4%[b]	-
2.	Albinali et al [76]	"	"	Decision Tree	Hand flapping, body rocking	Video camera, Accelerometers	6	12-20	acc- 88.6%	Real-time Off-line[c]
3.	Gro-ßekathöfer et al [77]	"	"	DT, SVM, RF	Hand flapping, body rocking	Three- axis sensor	6	12-20	acc- 83%, 86%, 86% [d]	-
4.	Zhao et al[68]	Repetitive behavior	FE	SVM, LDA, DT, RF, KNN	Hand gesture	Leap-Motion device	43	6-13	acc- 88.37% spe- 91.3% sen- 85% AUC- 88.1%	-
5.	Jaiswal et al. [72]	"	"	SVM+CNN	Head pose, hand moving front of the face	Kinect	55	18+	acc- 96.4%[e] acc- 93.9%[f]	-
6.	Jazouli et al. [69]	"	"	Nearest Neighbor	Hand flapping, hand on the face, hands behind back, fingers flapping, stimming and body rocking	Kinect	5	5-10	acc- 94%	real-time
7.	Coronato et al [70]	Repetitive behavior	E2E	SA+NN/NBN/BN	Hand hitting against the Ear, arm flapping, hand rotation up and hand rotation down	Accelerometers	-	-	acc- 92.16% - 99.13%[g]	Real-time Off-line[h]
8.	Jazouli et al. [53]	"	"	MLP	Hand flapping, hand on the face, hands behind back, fingers flapping, and body rocking	Kinect	10	-	acc- 93.3%	-
9.	Kumdee et al [75]	"	"	MLP	Hand flapping	Video camera	8	-	acc- 91.15% spe- 83.35% sen- 97.80%	Real-time
10.	Zunino et al. [1]	"	"	VGG-16 LSTM	Hand gesture	Video camera	40	7-12	acc- 82%	-
11.	Tian et al. [73]	Repetitive behavior	E2E	TPN	Hand gesture	RGB camera	40	7-12	acc- 87.17% spe- 89% sen- 88.2%	-

Table 3 (continued)

Id	Reference	Approach	Method	Algorithm	Cue/Feature Used	Sensor	No. of Subjects	Age (years)	Result	Time
12.	Sun et al. [74]	Repetitive behavior	E2E	SA-B3D LSTM	Hand gesture	RGB camera	40	7-12	acc- 95.2% spe- 93% sen- 96.1%	-
13.	Sadouk et al [71]	”	”	CNN	Body rocking, mouthing, complex hand and finger movements	Video camera, accelerometer	6	-	acc-(SMM)-96.55% acc-(HAR)-98.29%	-
14.	Gonçalves et al. [67]	”	SA	SA	Hand flapping	Kinect	4	10-12	acc- 76%	real-time
15.	Rihawi et al. [50]	Repetitive behavior	SA	SA	Head banging flipping Hand stimming Hand moving front of the face, toe walking, walking in circles, play with a toy.	Kinect	29	-	DTW distance is an unique feature to detect ASD/TD.	-
16.	Calhoun et al. [80]	Gait pattern	FE	PCA	Peak hip and ankle kinematics	Motion capture camera	34	5-9	acc- 81% spe- 83% sen-78%	-
17.	Al-Jubouri et al. [2]	”	”	PCA MLP	3D skeleton-based gait data	Kinect	100	4-12	acc- 95%	-
18.	Ilias et al. [88]	”	”	PCA LDA SVM	21 Gait features	Motion Camera, Force Plate	44	4-12	acc- 87.5%	-
19.	Ilias et al. [87]	”	”	LDA PCA NN	21 Gait features	Motion capture camera	44	4-12	acc- 98.4%	-
20.	Hasan et al. [81]	Gait pattern	FE	SA LDA	Walking speed, transverse joint angles	Force sensor	60	4-12	acc- 91.7% spe- 90% sen- 93.3%	-
21.	Henderson et al. [89]	”	”	RF	Gait cycle duration	Motion capture camera	21	-	acc 76.67% spe- 72.2% sen -86.67% FS- 78.8% (feature set B)	-
22.	Ebrahimi et al. [85]	”	”	SVM	Tip-toe walking	Kinect	75	12-20	acc- 86%	-
23.	Ilias et al. [86]	”	”	SVM	Stride length, cadence, stance time	Force plate	16	10-15	acc- 95%	-
24.	Hasan et al. [82]	”	”	KNN	3D Ground reaction forces	Force sensor	60	4-12	acc- 83.33%	-

Table 3 (continued)

Id	Reference	Approach	Method	Algorithm	Cue/Feature Used	Sensor	No. of Subjects	Age (years)	Result	Time
25.	Hasan et al. [84]	Gait pattern	FE	SA+MLP	3D Kinematic data	Motion-camera	60	4-12	acc- 91.7% spe- 90% sen- 93.3%	-
26.	Hasan et al[83]	Gait pattern	SA	SA	Walking speed, push rate	Motion-camera, force-plate	60	4-12	ASD children have lower push rate	-
27.	Dufek et al [91]	”	”	SA	Ground reaction forces, Joint Angles	Motion capture system	20	5-12	Gait cycle is unique for ASD children	-
28.	Shigeta et al. [90]	”	”	Multiple Regression	Angular velocity, Acceleration	Accelerometer	28	Mean-21.68	Angular velocity during passing another person is distinct for ASD children	-
29.	Wang et al. [31]	Gaze pattern	FE	Random Forest	Heatmap	Eye tracker	39	Mean age ASD - 30.8 ± 11.1 years TD 32.3 ± 10.4 years	acc- 81%	-
30.	Startsev et al. [96]	”	”	Random Forest	Scanpath	Eye tracker	28	Mean-8	acc- 76.5%	-
31.	Arru et al. [106]	”	”	Decision Tree	Face scanning	Eye tracker	32	7-16	acc- 60% spe- 50.5% sen- 56.9% AUC- 59.5% FS- 61.6%	-
32.	Babu et al[100]	”	”	SVM & Classification tree	Face scanning	Eye tracker	18	10-20	acc- 97%	-
33.	Wan et al. [99]	Gaze pattern	FE	SVM	Fixation time	Eye tracker	74	4-6	acc- 85.1% spe- 83.8% sen- 86.5%	-
34.	Dris et al. [103]	”	”	SVM	Fixation time	Eye tracker	28	8-15	acc- 88.6% spe- 92AUC- 96%	-
35.	Liu et al. [107]	”	”	SVM	Face scanning	Eye tracker	87	4-11	acc- 88.5% spe- 86.2% sen- 93.1% AUC- 89.6%	-
36	Alie et al. [108]	”	”	VMM	Face scanning	Camera	32	6 months	acc- 93.75% spe- 92.3% sen- 100%	-
37.	Yaneva et al. [5]	”	”	Logistic Regression	Fixation Time	Eye tracker	36	30-40	acc- 75%	-
38.	Cho et al. [95]	Gaze pattern	FE	KNN	Fixation point	Eye tracker	32	2-10	acc- 93.9% FS- 89.5%	-

Table 3 (continued)

Id	Reference	Approach	Method	Algorithm	Cue/Feature Used	Sensor	No. of Subjects	Age (years)	Result	Time
39.	Rahman et al. [3]	Gaze pattern	FE	XGBoost	Fixation point	Eye tracker	28	5-12	acc- 99% spe- 99% sen- 100% AUC- 99%	-
40.	Shihab et al. [51]	”	”	PCA	Fixation point	Camera	102	4-60	acc- 78.6% spe- 82.4% sen- 80%	-
41.	Jiang et al. [93]	”	E2E	VGG-16+SVM	Fixation point	Eye tracker	39	7-14	acc- 92% spe- 92% sen- 93% AUC- 92%	-
42.	Chen et al. [97]	Gaze pattern	E2E	Resnet & LSTM	Fixation point	Eye tracker	39	5-12	acc- 93% spe- 93% sen- 93% AUC- 93%	-
43.	Tao et al [98]	”	E2E	CNN & LSTM	Scanpath	Eye tracker	28	5-12	acc- 74.22%	-
44.	Syeda et al. [104]	”	SA	SA	Face scanning	Eye tracker	42	5-17	acc- 86%	-
45.	Sadira et al. [105]	”	SA	SA	Fixation point	Eye tracker	40	6-30	Between -ness centrality is distinct for ASD and TD.	-

[a]In laboratory settings

[b]In classroom settings

[c]The authors tested their method in both realtime and offline

[d]Accurecy for DT, SVM and RF respectively

[e]For classification into Control and Condition group

[f]for classification into Comorbid(ASD+ADHD) and ASD only group

[g]92% in real-time, whereas 99% in offline

[h]The authors tested their method in both realtime and offline

Table 4 A quantitative analysis of different methods of HAA-based automated detection of ASD.The bold entries signify the best-performing method for each approach under study.

Approach	Dataset	Method Type	Algorithm	Learning Type	Cross-validation	Performance Metrics			
						Acc.	Sen.	Spec.	FS
Repetitive Behavior	Zunino et al. [1]	Best-performing method	Inception v3 + LSTM [1]	E2E	LOO	0.82	0.83	0.81	0.81
					10-fold	0.89	0.91	0.91	0.88
		Alternative methods	Resnet-50 +LSTM	„	LOO	0.75	0.74	0.75	0.74
					10-fold	0.80	0.79	0.81	0.79
		”	VGG-16 +LSTM	„	LOO	0.65	0.67	0.64	0.61
					10-fold	0.72	0.75	0.70	0.69
		”	3D-CNN	„	LOO	0.58	0.58	0.58	0.53
					10-fold	0.64	0.67	0.62	0.59
Abnormal Gait pattern	Al-Jubouri et al. [2]	Best-performing method	F_{18}+MLP	FE	LOO	0.97	0.97	0.98	0.97
					10-fold	0.98	0.97	0.98	0.98
		Alternative methods	F_{18}+RF	„	LOO	0.96	0.96	0.96	0.96
					10-fold	0.96	0.95	0.96	0.95
		”	F_{11}+RF	„	LOO	0.95	0.93	0.96	0.94
					10-fold	0.94	0.93	0.95	0.94
		”	F_{11}+MLP[2]	„	LOO	0.94	0.93	0.95	0.94
					10-fold	0.94	0.93	0.95	0.94
		”	F_{18}+DT	„	LOO	0.93	0.92	0.93	0.93
					10-fold	0.90	0.89	0.91	0.90
		”	F_{18}+SVM	„	LOO	0.92	0.99	0.84	0.92
					10-fold	0.91	0.99	0.84	0.92
		”	F_{11}+SVM	„	LOO	0.91	0.97	0.84	0.91
					10-fold	0.90	0.97	0.84	0.91
		”	F_{11}+DT	„	LOO	0.90	0.90	0.90	0.90
					10-fold	0.89	0.88	0.90	0.89
		”	F11+LR	„	LOO	0.84	0.84	0.84	0.84
					10-fold	0.84	0.83	0.84	0.84
Gaze Pattern	Duan et al. [4]	Best-performing method	Saliency-metric feature + XGBoost [3]	„	LOO	0.99	0.99	0.99	0.99
					10-fold	0.99	0.99	1.00	0.99
		Alternative methods	Resnet-50 + LSTM [97]	E2E	LOO	0.93	0.93	0.93	0.93
					10-fold	0.94	0.93	0.94	0.94
		”	Resnet-50 + RF	FE	LOO	0.90	0.89	0.91	0.90
					10-fold	0.90	0.90	0.89	0.90
		”	GIST +RF	„	LOO	0.69	0.70	0.68	0.68
					10-fold	0.73	0.76	0.71	0.72
		”	VGG-16 + RF	„	LOO	0.63	0.65	0.62	0.61
					10-fold	0.66	0.68	0.64	0.64
		”	HOG + RF	„	LOO	0.57	0.57	0.57	0.56
					10-fold	0.63	0.64	0.63	0.63
Gaze Pattern	Yaneva et al. [5]	Best-performing method	SF + DT	„	10-fold	0.76	0.84	0.70	0.75
					70-30	0.72	0.77	0.67	0.71
		Alternative methods	SF + LR [5]	„	10-fold	0.75	0.79	0.72	0.74

Table 4 (continued)

Approach	Dataset	Method Type	Algorithm	Learning Type	Cross-validation	Performance Metrics			
					70-30	0.70	0.76	0.66	0.66
		Alternative methods	SF + SVM	FE	10-fold	0.73	0.81	0.68	0.71
					70-30	0.70	0.78	0.65	0.68
		”	SF + RF	„	10-fold	0.68	0.73	0.64	0.67
					70-30	0.66	0.71	0.62	0.65
		”	SF + NN	„	10-fold	0.61	0.65	0.58	0.60
					70-30	0.61	0.65	0.58	0.60

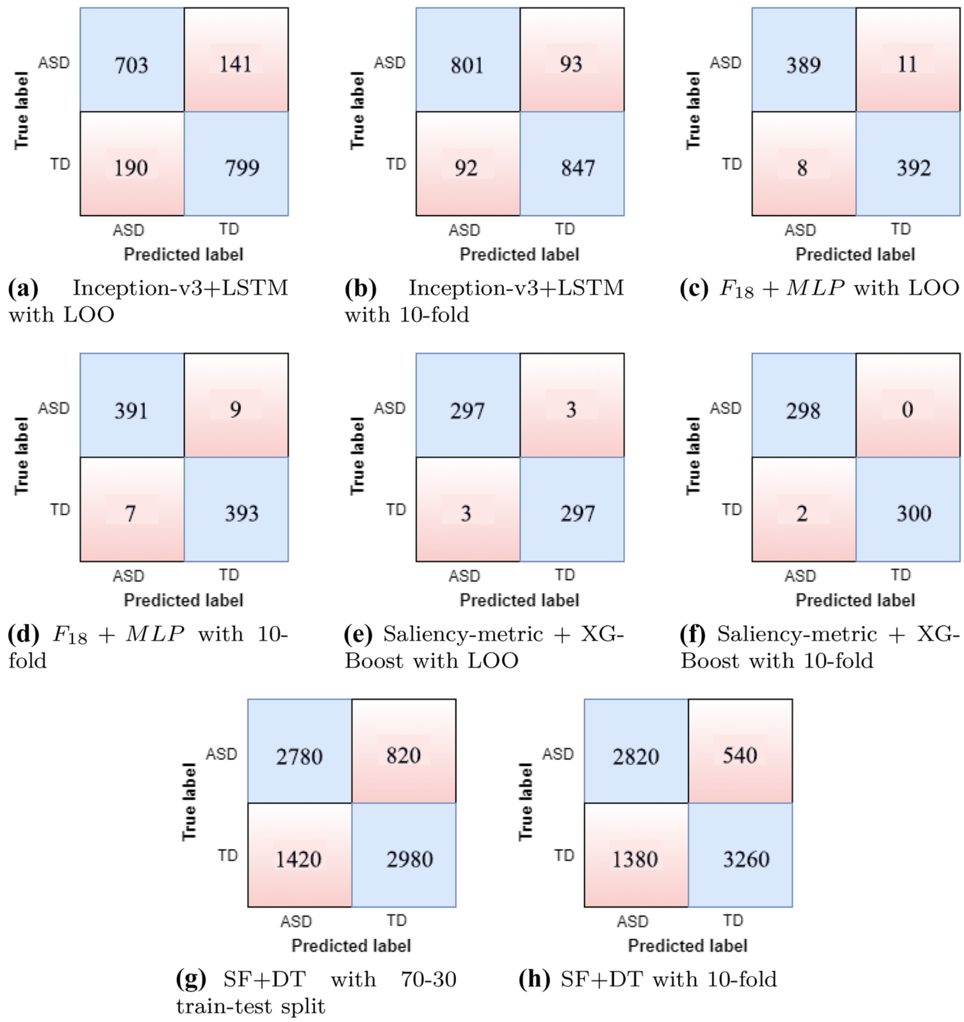
LOO Leave One Out, *MLP* Multi Layer Perception, *RF* Random Forest, *LR* Logistic Regression, *DT* Decision Tree, *SVM* Support Vector Machines, *SF* Statistical Feature, *NN* Neural Network, *LSTM* Long short-term memory, *VGG* Visual Geometry Group, *CNN* Convolutional Neural Network, *ResNet* Residual Neural Network, *HOG* Histogram of Oriented Gradients, F_{11} & F_{18} Top 11 and 18 features extracted from dataset [2]

other alternative architectures (i.e., 3D CNN, VGG-16 [111] + LSTM [110] , and Resnet-50 [112] + LSTM [110] that have been frequently used on video datasets [113–115]. To objectively analyze the efficacy of the proposed method, two evaluation protocols have been used and corresponding performance metrics are reported in Table 4. One is the leave-one-out (LOO) cross-validation in the form of leave-one- subject-out (where one subject is randomly selected for testing, while other subjects are used for training, and the procedure is repeated until every subject appears as a test case once) as used by the authors in [1] and the other is the 10-fold cross-validation. Our study found that the method of Inception v3 + LSTM [1] outperforms other alternative methods in both LOO (82% accuracy) and 10-fold cross-validation (89% accuracy). For the LOO split, the alternative algorithms can be ranked in terms of accuracy as follows: Resnet-50 + LSTM (acc: 0.75), VGG-16 + LSTM (acc: 0.65), and 3D-CNN (acc: 0.58). It is noteworthy that the order remains the same for the 10 fold split but with better accuracy.

Next, we assessed the performance of the method proposed by Al-Jubouri et al. [2] that classifies ASD vs TD based on abnormal gait pattern. An FE-based learning framework is used where the Principal Component Analysis (PCA) [116] is applied to reduce the dimension of the features extracted from the 3D-skeleton-based gait dataset [2]. As reported by the authors, 95% classification accuracy was achieved using the Multi-layer Perceptron (MLP) [117] and the top 11 features from PCA. Unfortunately, the authors did not mention the PCA-details and also the train-test split used for evaluating their model. For the analysis, here we have used PCA to select the top 11 and 18 features with 95% of the variance retained. Several other classifiers, e.g., Random Forest (RF) [118], Decision Tree (DT) [119], and Support Vector Machines (SVM) [120], have also been applied beside MLP to compare the performance using two cross-validations: 10-fold and leave-one-instance-out (i.e., selecting one instance in random order for testing while other instances are used to train the model, and this is repeated until every instance is used as a test case). According to our experimentation, the MLP with 18 features (i.e., 18 input nodes, 10 hidden nodes and 2 output nodes) gives the best accuracy of 98% on the tenfold cross-validation (97% on LOO) and outperforms the method in [2] and the other alternatives.

For the unusual gaze pattern-based detection, we assessed the method of Chen and Zhao [97], where an E2E learning method is proposed using the Resnet-50 [112] architecture followed by an LSTM network. The authors used leave-one-out (LOO) cross-validation in the form of leave-one-instance-out on the dataset in [4]. Here, we also applied the same LOO cross-validation and to further investigate the effectiveness of the method, we performed the 10-fold cross-validation. As the method reports its performance using an eye-tracking dataset of fixation points, we, therefore, selected alternative methods such as HOG [121], GIST [122], VGG-16 [111], and Resnet-50 [112] which have been widely used for feature extraction from fixation maps [101, 123–125]. The random forest [118] classifier is used for the classification step. In addition, we assessed a very recent method proposed by Rahman et al. [3] that utilized the saliency-metrics based-feature for ASD detection using the same dataset and cross-validations. According to our quantitative analysis, the method of Rahman et al. [3] reported a classification accuracy of 99% outperforming that of Chen and Zhao [97] and other alternative methods under study.

In another evaluation of the gaze pattern-based methods, we analyzed the method of Yaneva et al. [5]. The work presented a method of FE-based learning combining statistical features and a logistic regression classifier. To provide an objective analysis, we performed two splits, one is 10-fold cross-validation following the original work in [5] and the other is a 70-30 train-test split. Since the dataset includes statistical features of the samples, we did the analysis on the classification part. As alternative learning methods, we



(a) Inception-v3+LSTM with LOO (b) Inception-v3+LSTM with 10-fold (c) $F_{18} + MLP$ with LOO (d) $F_{18} + MLP$ with 10-fold (e) Saliency-metric + XG-Boost with LOO (f) Saliency-metric + XG-Boost with 10-fold (g) SF+DT with 70-30 train-test split (h) SF+DT with 10-fold

Fig. 10 Confusion matrix of the best-performing method under each approach. (**a**)-(**b**): confusion matrix of (Inception-v3+LSTM) for the repetitive behavior-based approach on dataset [1], (**c**)-(**d**): (F_{18} +MLP) under the abnormal gait pattern-based approach on dataset [2], (**e**)-(**f**): (Saliency-metric+XGBoost) for the unusual gaze pattern approach on dataset [4], and (**g**)-(**h**): (SF+DT) under the unusual gaze pattern approach as computed using the dataset [5]

applied other frequently used classifiers such as: RF, DT, SVM, and a Neural Network (NN) [126] with one hidden layer. Our findings show that the decision tree provides the best classification accuracy of 76% on both splits used for the evaluation.

Figure 10 presents the confusion matrices for the best-performing methods for each approach we analyzed.

Challenges in this Domain

In many ways classifying mental complexities are very different from diagnosing the more familiar physiological problems [127]. Besides detecting the gesture changes and the movements of an autistic child with various sensors precisely is more complicated as they may behave unpredictably with frequent chances of sudden seizures. In this section, we have focused on some of the challenges which seem to arise to proceed with research in this field.

Insufficiency of Standard Dataset

- A dataset is considered to be a high-level tool for commencing study in any field. A major challenge in conducting research in the fields related to autism is the paucity of standard datasets.
- Having sensitive subjects like autistic individuals and, more importantly making them act accordingly to get the required data for any study is quite a big challenge.
- Even though experiments on people with ASD are found to be done, seldom the datasets are made publicly accessible. Most of the observations involving autistic children are funded by NGOs or research institutes that do not make the availability of the dataset public.

Non-homogeneous Symptoms in Subjects

- Due to its heterogeneous condition, each individual with autism has a unique profile of symptoms. There can be an increasing amount of response diversity for a single action among different individuals with autism. A research observed that three typically developing pre-schoolers had unvarying responsive action to a block building task, whereas diverse responses were recorded from individuals with ASD [128].
- Even the responses from the same child with ASD can be different from time to time. It was also noticed that adding concrete reinforcements had effected the response of an individual with autism to get differed from the previous one [129].

Complications in Laboratory Setup and Experimentation

- It gets more difficult to have a setup with accurate orientations of equipment and connectivity as individuals with ASD have to deal with them [130].
- Moreover, individuals with ASD may not have the capacity to adapt in a laboratory environment as their disorder is associated with disability in socializing.
- It is prevalent for autistic children to have disarray in language and cognitive functionalities, which may escalate the chances of difficulties in responding accordingly with the instructions of the experiment director [131].
- They may be very uncomfortable with a certain scent, sounds, pattern, and tastes. Therefore, in many cases, an autistic child may get panicked with sudden audio instruction, or the laboratory environment may not be very favorable to their known surrounding experiences [132]. Hence, the experimental setup should be designed and implemented carefully considering the group's physiology and psychology.

Data Annotation

- Implementing DSM-5 in ASD detection needs a handful of skilled physicians who may need a prolonged period to annotate the result.
- Sometimes expert screening of DSM-5 may vary with different physicians [133, 134] The difference in determining the severity level may also create difficulties in annotating the data.
- Along with these, misclassifying 'late learners' as 'autistic individual' is a common phenomenon in ASD detection [135]. So, annotating data in researches like this is undoubtedly a challenging task with numerous possibilities of ambiguity.

Resource Constraints

- Arrangement of the space for conducting the experimentation with Autistic children is one of the main limitations for the manifestation of ASD detection using activity analysis.
- The machine-based classifier decides for activity analysis by analyzing data from a diverse set of fields. Access to this kind of data is achieved by dense, incessant feedback from sensors with multiple modes and advanced machine learning algorithms [136]. The computaional devices are required to have sufficient memory space to store and analyze the data continuously [137].
- Furthermore, the necessary hardware and software tools for the observations require significant funding and expenditure [138].

Frequent Device Calibration

- Very often, electric components like sensors may not work as expected from the onset and are required to be calibrated for several times to get finer results [139]. For example, visual saliency-based ASD detection approaches use eye trackers that are often required to be calibrated several times to include information about shapes, light reflection, and refraction properties of the different parts of the eyes, and during calibration, participants are asked to look at different positions on the screen [140].
- An autistic individual is very unlikely to cooperate in the time of calibration.

Privacy Issues

- Analyzing facial expressions needs images of participants, which may violate the privacy of that individual. Therefore, most of the time, parents do not allow their children to participate in the research.
- Besides, skeleton data and eye-tracking data do not violate privacy to that extent as it does not require or expose any body parts of that individual [141].

Discussion

The atypical behavioral patterns of autistic children supports the idea that these unique features might be characterized by data-driven algorithms and produce an automated ASD detection technique. In many cases, it might be enough to use an RGB camera or skeleton tracking device to capture videos or images. Gaze abnormalities might also be a good source for identifying unique characteristics that could feed into a feature model. This approach protects patient privacy and although it has less computational needs, it has high operational complexities. An eye-tracking system cannot look at multiple behaviors simultaneously, whereas skeletal or video-based motion detection can. On the other hand, the experimental setup for skeleton tracking is complex and not very practical. This makes the prediction model incompatible with real-world use. Nevertheless, it must be said that making use of a degree of intuition about the dataset before beginning to train the model can minimize methodological difficulties.

Since the automated detection of ASD requires an annotated dataset, supervised learning algorithms are preferable. Moreover, for datasets containing sequential instances like videos or images, the E2E learning approach is more suitable as it can learn temporal or sequential information better than FE-based learning algorithms.

The range of human intervention in the described methods of different approaches is varied. In most cases, the classification model needs to be trained with labeled data. So, expert physicians' involvement is required for annotating the data with the current medical standards. Then, for data collection in repetitive behavior, gait pattern and gaze analysis-based methods, the necessitated instrument like ground attached sensor, motion capture camera, wearable sensors, eye-tracking devices need dexterous supervision to calibrate and operate. In the eye-tracking-based systems, the visual stimulus should be selected in accordance with the age, gender, and other attributes of the individuals. Moreover, in some cases, the collected data are directly fed into different classifier algorithms and the result is generated whereas many methods need further analysis and assessment from experts for a definitive conclusion. However, with the recent advancement in E2E learning, the data analysis part has achieved a substantial amount of automation. Whereas the existing medical diagnosis requires multiple visits with prolonged sessions for a decisive conclusion from expert physicians, the HAA-based techniques are quite a lot undemanding. As the required data is fed into the E2E learning-based systems, the result acquisition is more straightforward without the need for data analysis by any other sources.

A real-time automated ASD detection methodology is still to be achieved but there has been considerable success in determining the specific features of behavioral inconsistencies of ASD individuals. To be more specific, instances like stereotyped hand-flapping movements [67], body rocking, fingers flapping, hand on the face and hands behind back [70, 75, 76] have been identified in an almost real-time and offline manner. Moreover, [70] is even deployed in a hospital for its satisfactory result from the laboratory. However, the complete functional real-time entity of the methods posses some impediments like less accuracy, conflict in annotation result, failure to identify the swift and small movements, the beginning and end of the actions. In [76], the real-time implementation included online human intervention only on the training set, where considerable disagreement between the annotators had impacted the overall classification accuracy. But in the offline annotation phase, the more well observed insights from the annotators were very likely to have accordance with each other and resulted in better accuracy from the classifier algorithms. The real-time detection methods do also have the requirement of a large amount of data for training the model with a wide range of unusual behavioral samples from extensive varieties of specimens to have reliable results. However, the quest for redressing these challenges remains in a very active field of research and the findings are very promising for the contrivance of the desired system in the near term.

In this paper, we have presented a thorough analysis of HAA-based autism detection methods. Some of the methods can be automated without many functional enhancements. However, the next step for automated ASD detection is to drive an easily implementable form of ASD screening that could be used as part of an annual test program in educational institutions.

Conclusion

ASD is a neurological and developmental disorder that severely affects interactions, communication, and learning in a person's life. Numerous reports and studies have shown that early treatment of ASD individuals can suppress its effects and might allow the person to lead a healthy life. Our review has looked at the scope and potential of data-driven activity analysis to automatically detect autism. We have looked at state-of-the-art approaches for detecting ASD that make use of its unique characteristics, e.g. repetitive behavior, atypical gait, and unusual visual saliency as identifying features. Our systematic search used the PRISMA approach to compile all the relevant literature. We reviewed FE and E2E-based learning algorithms for ASD/TD detection and selected 45 papers that examined methods and 6 that focused on datasets. Although the prime goal is to develop fully automated ASD detection a method completely free of human intervention is yet to be achieved. Nevertheless, we believe that the recent advances have shown it is not very far off. Among the 45 papers selected for our review, the SA and some of the FE learning-based methods require significant human involvement in data analysis, although the majority of the methods have been substantially automated. We found that 4 methods reported the ability to perform real-time detection while another 6 used offline detection. To gauge the feasibility of deploying these learning-based methods into mainstream ASD diagnosis, the methods were assessed qualitatively and quantitatively on our chosen dataset. According to our quantitative assessment, the Inception v3 and LSTM-based method [1] gives the best classification accuracy (89%) for repetitive behavior [1]. For abnormal gait recognition, the Multilayer Perceptron gives 98% accuracy using 18 features extracted from the 3D-skeleton-based gait data [2]. For gaze pattern-based approach, a saliency-metric feature-based learning [3] outperforms other methods with 99% accuracy on dataset [4], while a statistical features and Decision Tree-based algorithm yields the best accuracy of 76% on dataset [5]. The dataset used for these studies was labelled by professionals in accordance with the current standard medical diagnosis methods for ASD. The performance metrics achieved with the novel techniques were in line with the labelled data. In other words, in terms of classification accuracy, the automated methods were strongly correlated with current standard methods. Nevertheless, there are constraints which may limit the accuracy and flexibility of automated detection. Any incorrect result will misclassify a normal(non-ASD) child as an autistic child, which commonly happens with late or slow learners who need more time to make decisions and complete a task. In addition, very recent researches are restricted to distinguishing between ASD and TD children and cannot detect the severity of autism. However, we believe that advances in learning algorithms and computational devices will soon pave the way for better and more adaptable data-driven approaches. Parents who are unwilling to accept the fact that their child is displaying some traits of ASD and don't take their child to a physician are unintentionally hindering helpful treatment. This barrier to detecting ASD in early childhood might be lowered by using the automated techniques discussed in this paper. We hope this review might be a useful guide for researchers wanting to explore automated data-driven ASD detection.

Funding Information This research was supported in part by the ICT Division, Ministry of Posts, Telecommunications and Information Technology of the Government of Bangladesh.

Declarations

Ethical Approval This article does not contain any studies with human participants or animals.

Conflict of Interest All of the authors declare that he/she has no conflict of interest.

References

1. Zunino A, Morerio P, Cavallo A, Ansuini C, Podda J, Battaglia F, etal. Video gesture analysis for autism spectrum disorder detection. In: 2018 24th International Conference on Pattern Recognition (ICPR). IEEE; 2018. p. 3421–3426.
2. Abdulrahman A, Hadi I, Rajihy Y. Generating 3D dataset of Gait and Full body movement of children with Autism spectrum disorders collected by Kinect v2 camera. COMPUSOFT: An International Journal of Advanced Computer Technology. 2020;9(8):3791–3797.
3. Rahman S, Rahman S, Shahid O, Abdullah M, Sourov JA. Classifying Eye-Tracking Data Using Saliency Maps. Int Conf Pattern Recognit. 2020.
4. Duan H, Zhai G, Min X, Che Z, Fang Y, Yang X, et al. A dataset of eye movements for the children with autism spectrum disorder. In: Proceedings of the 10th ACM Multimedia Systems Conference; 2019. p. 255–260.
5. Yaneva V, Ha LA, Eraslan S, Yesilada Y, Mitkov R. Detecting Autism Based on Eye-Tracking Data from Web Searching Tasks. In: Proceedings of the Internet of Accessible Things. W4A '18. New York, NY, USA: Association for Computing Machinery; 2018. p. 1–10.
6. World H Organization. Autism spectrum disorders. World Health Organization. 2019 Nov. Available from: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders.
7. Hossain MD, Ahmed HU, Uddin MJ, Chowdhury WA, Iqbal MS, Kabir RI, et al. Autism Spectrum disorders (ASD) in South Asia: a systematic review. BMC Psychiatry. 2017;17(1):1–7.
8. Hamdoun O. Autism Spectrum Disorders, is it Under Reported In Third World Countries. Am J Biomed Sci Res. 2019;4(4):292–3.
9. Meilleur AAS, Jelenic P, Mottron L. Prevalence of clinically and empirically defined talents and strengths in autism. J Autism Dev Disord. 2015;45(5):1354–67.
10. Fakhoury M. Autistic spectrum disorders: A review of clinical features, theories and diagnosis. Int J Dev Neurosci. 2015;43:70–7.
11. Jiang Yh, Yuen RK, Jin X, Wang M, Chen N, Wu X, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am J Human Gen. 2013;93(2):249–63.
12. Zwaigenbaum L, Bauman ML, Choueiri R, Fein D, Kasari C, Pierce K, et al. Early Identification and Interventions for Autism Spectrum Disorder: Executive Summary. Pediatrics. 2015;136(Supplement):S1–9.
13. Volkmar FR, Reichow B, McPartland J. Classification of autism and related conditions: progress, challenges, and opportunities. Dialogues Clin Neurosci. 2012;14(3):229.
14. Hoefman R, Payakachat N, van Exel J, Kuhlthau K, Kovacs E, Pyne J, et al. Caring for a child with autism spectrum disorder and parents quality of life: application of the CarerQol. J Autism Dev Dis. 2014;44(8):1933–45.
15. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. Arch Gen Psychiatry. 2006;63(6):694–701.
16. Association AP. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub. 2013.
17. Maenner MJ, Rice CE, Arneson CL, Cunniff C, Schieve LA, Carpenter LA, et al. Potential impact of DSM-5 criteria on autism spectrum disorder prevalence estimates. JAMA Psychiat. 2014;71(3):292–300.
18. Bryson SE, Rogers SJ, Fombonne E. Autism spectrum disorders: early detection, intervention, education, and psychopharmacological management. Canadian J Psych. 2003;48(8):506–16.
19. Kleinman JM, Robins DL, Ventola PE, Pandey J, Boorstein HC, Esser EL, et al. The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. J Autism Dev Disord. 2008;38(5):827–39.
20. Hazlett HC, Gu H, Munsell BC, Kim SH, Styner M, Wolff JJ, et al. Early brain development in infants at high risk for autism spectrum disorder. Nature. 2017;542(7641):348–51.
21. Tager-Flusberg H. Brain Imaging Studies in Autism Spectrum Disorders. The Asperger / Autism Network (AANE). 2017 Feb.

Available from: https://www.aane.org/brain-imaging-studies-autism-spectrum-disorders/.
22. Zhang S, Wei Z, Nie J, Huang L, Wang S, Li Z. A review on human activity recognition using vision-based method. J Healthcare Eng. 2017.
23. Srivastava AK, Biswas K, Tripathi V. A Robust Framework for Effective Human Activity Analysis. In: International Conference on Innovative Computing and Communications. Springer; 2019:331–337.
24. Wu D, Sharma N, Blumenstein M. International joint conference on neural networks (IJCNN). IEEE. Recent advances in video based human action recognition using deep learning: a review. 2017;2017:2865–72.
25. Aly S, Trubanova A, Abbott L, White S, Youssef A. VT-KFER: A Kinect-based RGBD+ time dataset for spontaneous and non-spontaneous facial expression recognition. In: 2015 International Conference on Biometrics (ICB). IEEE; 2015:90–97.
26. Faso DJ, Sasson NJ, Pinkham AE. Evaluating posed and evoked facial expressions of emotion from adults with autism spectrum disorder. J Autism Dev Disord. 2015;45(1):75–89.
27. Harms MB, Martin A, Wallace GL. Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychol Rev. 2010;20(3):290–322.
28. Rehg JM, Rozga A, Abowd GD, Goodwin MS. Behavioral imaging and autism. IEEE Pervasive Comput. 2014;13(2):84–7.
29. Militerni R, Bravaccio C, Falco C, Fico C, Palermo MT. Repetitive behaviors in autistic disorder. Euro Child Adol Psych. 2002;11(5):210–8.
30. Manicolo O, Brotzmann M, Hagmann-von Arx P, Grob A, Weber P. Gait in children with infantile/atypical autism: Age-dependent decrease in gait variability and associations with motor skills. Eur J Paediatr Neurol. 2019;23(1):117–25.
31. Wang S, Jiang M, Duchesne XM, Laugeson EA, Kennedy DP, Adolphs R, et al. Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. Neuron. 2015;88(3):604–16.
32. Chawarska K, Shic F. Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. J Autism Dev Disord. 2009;39(12):1663.
33. Subbaraju V, Suresh MB, Sundaram S, Narasimhan S. Identifying differences in brain activities and an accurate detection of autism spectrum disorder using resting state functional-magnetic resonance imaging: A spatial filtering approach. Med Image Anal. 2017;35:375–89.
34. Eslami T, Mirjalili V, Fong A, Laird AR, Saeed F. ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data. Frontiers in Neuroinformatics. 2019 Nov;13.
35. Sherkatghanad Z, Akhondzadeh M, Salari S, Zomorodi-Moghadam M, Abdar M, Acharya UR, et al. Automated Detection of Autism Spectrum Disorder Using a Convolutional Neural Network. Frontiers in Neuroscience. 2020 Jan; 13.
36. Kandalaft MR, Didehbani N, Krawczyk DC, Allen TT, Chapman SB. Virtual reality social cognition training for young adults with high-functioning autism. J Autism Dev Disord. 2013;43(1):34–44.
37. Welch KC, Lahiri U, Liu C, Weller R, Sarkar N, Warren Z. An affect-sensitive social interaction paradigm utilizing virtual reality environments for autism intervention. In: International conference on human-computer interaction. Springer; 2009. p. 703–712.
38. Zhang L, Wade JW, Bian D, Swanson A, Warren Z, Sarkar N. Data fusion for difficulty adjustment in an adaptive virtual reality game system for autism intervention. In: International Conference on Human-Computer Interaction. Springer; 2014. p. 648–652.
39. Lahiri U, Warren Z, Sarkar N. Dynamic gaze measurement with adaptive response technology in Virtual Reality based social communication for autism. In: 2011 International Conference on Virtual Rehabilitation. IEEE; 2011. p. 1–8.
40. Lahiri U, Warren Z, Sarkar N. Design of a gaze-sensitive virtual social interactive system for children with autism. IEEE Trans Neural Syst Rehabil Eng. 2011;19(4):443–52.
41. Bekele E, Wade J, Bian D, Fan J, Swanson A, Warren Z. In: 2016 IEEE Virtual Reality (VR). IEEE. Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD). 2016;2016:121–30.
42. Hyde KK, Novack MN, LaHaye N, Parlett-Pelleriti C, Anden R, Dixon DR, et al. Applications of supervised machine learning in autism spectrum disorder research: a review. Rev J Autism Dev Disord. 2019;6(2):128–46.
43. Thabtah F. Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. Inform Health Soc Care. 2019;44(3):278–97.
44. Song DY, Kim SY, Bong G, Kim JM, Yoo HJ. The Use of Artificial Intelligence in Screening and Diagnosis of Autism Spectrum Disorder: A Literature Review. Journal of the Korean Academy of Child and Adolescent Psychiatry. 2019;30(4):145–52.
45. Boucenna S, Narzisi A, Tilmont E, Muratori F, Pioggia G, Cohen D, et al. Interactive technologies for autistic children: A review. Cogn Comput. 2014;6(4):722–40.
46. Sevin JA, Rieske RD, Matson JL. A review of behavioral strategies and support considerations for assisting persons with difficulties transitioning from activity to activity. Rev J Autism Dev Disord. 2015;2(4):329–42.
47. Reinders NJ, Branco A, Wright K, Fletcher PC, Bryden PJ. Scoping review: physical activity and social functioning in young people with autism spectrum disorder. Front Psychol. 2019;10:120.
48. Scharoun SM, Wright KT, Robertson-Wilson JE, Fletcher PC, Bryden PJ. Physical activity in individuals with autism spectrum disorders (ASD): a review. Autism-paradigms, recent research and clinical applications. 2017.
49. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009;62(10):e1–34.
50. Rihawi O, Merad D, Damoiseaux JL. 3D-AD: 3D-autism dataset for repetitive behaviours with kinect sensor. In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE; 2017. p. 1–6.
51. Shihab AI, Dawood FA, Kashmar AH. Data Analysis and Classification of Autism Spectrum Disorder Using Principal Component Analysis. Adv Bioinform. 2020;2020:1–8.
52. Weinland D, Ronfard R, Boyer E. A survey of vision-based methods for action representation, segmentation and recognition. Comput Vis Image Underst. 2011;115(2):224–41.
53. Jazouli M, Elhoufi S, Majda A, Zarghili A, Aalouane R. Stereotypical motor movement recognition using microsoft kinect with artificial neural network. World Acad Sci Eng Technol Int J Comput Electr Autom Control Inf Eng. 2016;10(7):1270–4.
54. Rohrbach A, Rohrbach M, Tandon N, Schiele B. A dataset for movie description. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015:3202–3212.
55. Reddy KK, Shah M. Recognizing 50 human action categories of web videos. Mach Vis Appl. 2013;24(5):971–81.
56. Ryoo MS, Matthies L. First-person activity recognition: What are they doing to me? In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2013:2730–2737.
57. Chen L, Hoey J, Nugent CD, Cook DJ, Yu Z. Sensor-based activity recognition. IEEE Transactions on Systems,

Man, and Cybernetics, Part C (Applications and Reviews). 2012;42(6):790–808.
58. Liu Y, Nie L, Liu L, Rosenblum DS. From action to activity: sensor-based activity recognition. Neurocomputing. 2016;181:108–15.
59. Duchesnay E, Cachia A, Boddaert N, Chabane N, Mangin JF, Martinot JL, et al. Feature selection and classification of imbalanced datasets. Neuroimage. 2011;57(3):1003–14.
60. Papagiannopoulou EA, Chitty KM, Hermens DF, Hickie IB, Lagopoulos J. A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. Soc Neurosci. 2014;9(6):610–32.
61. Feil-Seifer D, Matarić MJ. In: 2011 6th ACM/IEEE international conference on human-robot interaction (HRI). IEEE. Automated detection and classification of positive vs negative robot interactions with children with autism using distance-based features. 2011;2011:323–30.
62. Bodfish JW. Stereotypy, self-injury, and related abnormal repetitive behaviors. In: Handbook of intellectual and developmental disabilities. Springer; 2007:481–505.
63. Leekam S, Tandos J, McConachie H, Meins E, Parkinson K, Wright C, et al. Repetitive behaviours in typically developing 2-year-olds. J Child Psychol Psychiatry. 2007;48(11):1131–8.
64. Arnott B, McConachie H, Meins E, Fernyhough C, Le Couteur A, Turner M, et al. The frequency of restricted and repetitive behaviors in a community sample of 15-month-old infants. Journal of Developmental & Behavioral Pediatrics. 2010;31(3):223–9.
65. Richler J, Huerta M, Bishop SL, Lord C. Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. Dev Psychopathol. 2010;22(1):55.
66. Goodwin MS, Intille SS, Albinali F, Velicer WF. Automated detection of stereotypical motor movements. J Autism Dev Disord. 2011;41(6):770–82.
67. Gonçalves N, Rodrigues JL, Costa S, Soares F. In: 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. IEEE. Automatic detection of stereotyped hand flapping movements: two different approaches. 2012;2012:392–7.
68. Zhao Z, Zhang X, Li W, Hu X, Qu X, Cao X, et al. Applying Machine Learning to Identify Autism With Restricted Kinematic Features. IEEE Access. 2019;7:157614–22.
69. Jazouli M, Majda A, Merad D, Aalouane R, Zarghili A. Automatic detection of stereotyped movements in autistic children using the Kinect sensor. Int J Biomed Eng Technol. 2019;29(3):201–20.
70. Coronato A, De Pietro G, Paragliola G. A situation-aware system for the detection of motion disorders of patients with autism spectrum disorders. Expert Syst Appl. 2014;41(17):7868–77.
71. Sadouk L, Gadi T, Essoufi EH. A novel deep learning approach for recognizing stereotypical motor movements within and across subjects on the autism spectrum disorder. Computational intelligence and neuroscience. 2018.
72. Jaiswal S, Valstar MF, Gillott A, Daley D. Automatic detection of ADHD and ASD from expressive behaviour in RGBD data. In: 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). IEEE; 2017:762–769.
73. Tian Y, Min X, Zhai G, Gao Z. Video-based early detection via temporal pyramid networks. In: 2019 IEEE International Conference on Multimedia and Expo (ICME). IEEE; 2019:272–277.
74. Sun K, Li L, Li L, He N, Zhu J. Spatial Attentional Bilinear 3D Convolutional Network for Video-Based Autism Spectrum Disorder Detection. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2020:3387–3391.
75. Kumdee O, Ritthipravat P. In: 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). IEEE. Repetitive motion detection for human behavior understanding from video images. 2015;2015:484–9.
76. Albinali F, Goodwin MS, Intille S. Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms. Pervasive Mob Comput. 2012;8(1):103–14.
77. Großekathöfer U, Manyakov NV, Mihajlović V, Pandina G, Skalkin A, Ness S, et al. Automated detection of stereotypical motor movements in autism spectrum disorder using recurrence quantification analysis. Front Neuroinform. 2017;11:9.
78. Kindregan D, Gallagher L, Gormley J. Gait deviations in children with autism spectrum disorders: a review. Autism research and treatment. 2015.
79. Weiss MJ, Moran MF, Parker ME, Foley JT. Gait analysis of teenagers and young adults diagnosed with autism and severe verbal communication disorders. Front Integr Neurosci. 2013;7:33.
80. Calhoun M, Longworth M, Chester VL. Gait patterns in children with autism. Clin Biomech. 2011;26(2):200–6.
81. Hasan CZC, Jailani R, Tahir NM, Yassin IM, Rizman ZI. Automated classification of autism spectrum disorders gait patterns using discriminant analysis based on kinematic and kinetic gait features. Journal of Applied Environmental and Biological Sciences. 2017;7(1):150–6.
82. Hasan CZC, Jailani R, Tahir NM, Sahak R. Autism spectrum disorders gait identification using ground reaction forces. Telkomnika. 2017;15(2):903.
83. Hasan C, Jailani R, Tahir N, Desa H. Vertical ground reaction force gait patterns during walking in children with autism spectrum disorders. Int J Eng. 2018;31(5):705–11.
84. Hasan CZC, Jailani R, Tahir NM. Use of statistical approaches and artificial neural networks to identify gait deviations in children with autism spectrum disorder. Int J Biol Biomed Eng. 2017;11:74–9.
85. Ebrahimi M, Feghi M, Moradi H, Mirian M, Pouretemad H. Distinguishing tip-toe walking from normal walking using skeleton data gathered by 3D sensors. In: 2015 3rd RSI International Conference on Robotics and Mechatronics (ICROM). IEEE; 2015:450–455.
86. Ilias S, Tahir NM, Jailani R, Hasan CZC. In: 2016 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). IEEE. Classification of autism children gait patterns using neural network and support vector machine. 2016;2016:52–6.
87. Ilias S, Tahir NM, Jailani R, Hasan CZC. In: 2017 European Modelling Symposium (EMS). IEEE. Linear Discriminant Analysis in Classifying Walking Gait of Autistic Children. 2017;2017:67–72.
88. Ilias S, Tahir NM, Jailani R. In: 2016 IEEE Industrial Electronics and Applications Conference (IEACon). IEEE. Feature extraction of autism gait data using principal component analysis and linear discriminant analysis. 2016;2016:275–9.
89. Henderson B, Yogarajah P, Gardiner B, McGinnity M, Forster K, Nicholas B. In: 2020 31st Irish Signals and Systems Conference (ISSC). IEEE. Effects of Intra-Subject Variation in Gait Analysis on ASD Classification Performance in Machine Learning Models. 2020;2020:1–6.
90. Shigeta M, Sawatome A, Ichikawa H, Takemura H. Correlation between Autistic Traits and Gait Characteristics while Two Persons Walk Toward Each Other. Advanced Biomedical Engineering. 2018;7:55–62.
91. Dufek JS, Eggleston JD, Harry JR, Hickman RA. A comparative evaluation of gait between children with autism and typically developing matched controls. Med Sci. 2017;5(1):1.
92. Senju A, Johnson MH. The eye contact effect: mechanisms and development. Trends Cogn Sci. 2009;13(3):127–34.
93. Jiang M, Zhao Q. Learning visual attention to identify people with autism spectrum disorder. In: Proceedings of

the IEEE International Conference on Computer Vision; 2017:3267–3276.
94. Perronnin F, Liu Y, Sánchez J, Poirier H. Large-scale image retrieval with compressed fisher vectors. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE; 2010:3384–3391.
95. Cho KW, Lin F, Song C, Xu X, Hartley-McAndrew M, Doody KR. In: 2016 IEEE Wireless Health (WH). IEEE. Gaze-Wasserstein: a quantitative screening approach to autism spectrum disorders. 2016;2016:1–8.
96. Startsev M, Dorr M. Classifying Autism Spectrum Disorder Based on Scanpaths and Saliency. In: 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE; 2019:633–636.
97. Chen S, Zhao Q. Attention-based autism spectrum disorder screening with privileged modality. In: Proceedings of the IEEE International Conference on Computer Vision; 2019:1181–1190.
98. Tao Y, Shyu ML. SP-ASDNet: CNN-LSTM based ASD classification model using observer scanpaths. In: 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE; 2019:641–646.
99. Wan G, Kong X, Sun B, Yu S, Tu Y, Park J, et al. Applying eye tracking to identify autism spectrum disorder in children. J Autism Dev Disord. 2019;49(1):209–15.
100. Babu PRK, Lahiri U. Classification approach for understanding implications of emotions using eye-gaze. J Ambient Intell Humaniz Comput. 2019;11(7):2701–13.
101. Nebout A, Wei W, Liu Z, Huang L, LeMeur O. Predicting Saliency Maps for ASD People. In: 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE; 2019:629–632.
102. Duan H, Zhai G, Min X, Fang Y, Che Z, Yang X, etal. Learning to predict where the children with asd look. In: 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE; 2018:704–708.
103. Dris AB, Alsalman A, Al-Wabil A, Aldosari M. Intelligent Gaze-Based Screening System for Autism. In: 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS). IEEE; 2019:1–5.
104. Syeda UH, Zafar Z, Islam ZZ, Tazwar SM, Rasna MJ, Kise K, etal. Visual face scanning and emotion perception analysis between autistic and typically developing children. In: Proceedings of the 2017 acm international joint conference on pervasive and ubiquitous computing and proceedings of the 2017 acm international symposium on wearable computers; 2017:844–853.
105. Sadria M, Karimi S, Layton AT. Network centrality analysis of eye-gaze data in autism spectrum disorder. Comput Biol Med. 2019;111.
106. Arru G, Mazumdar P, Battisti F. Exploiting Visual Behaviour for Autism Spectrum Disorder Identification. In: 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE; 2019:637–640.
107. Liu W, Li M, Yi L. Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. Autism Res. 2016;9(8):888–98.
108. Alie D, Mahoor MH, Mattson WI, Anderson DR, Messinger DS. In: 2011 IEEE Workshop on Applications of Computer Vision (WACV). IEEE. Analysis of eye gaze pattern of infants at risk of autism spectrum disorder using markov models. 2011;2011:282–7.
109. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016:2818–2826.
110. Gers FA, Schmidhuber JA, Cummins FA. Learning to Forget: Continual Prediction with LSTM. Neural Comput. 2000;12(10):2451–71.
111. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
112. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016:770–778.
113. Ji S, Xu W, Yang M, Yu K. 3D convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell. 2012;35(1):221–31.
114. Zhuang N, Yusufu T, Ye J, Hua KA. Group activity recognition with differential recurrent convolutional neural networks. In: 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). IEEE; 2017:526–531.
115. Lu Z, Zhou W, Zhang S, Wang C. A new video-based crash detection method: balancing speed and accuracy using a feature fusion deep learning framework. J Adv Transpo. 2020.
116. Jolliffe IT. Principal Component Analysis. New York: Springer-Verlag; 2002.
117. Almeida LB. C1. 2 Multilayer perceptrons. Handbook of Neural Computation C. 1997;1.
118. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
119. Swain PH, Hauska H. The decision tree classifier: Design and potential. IEEE Trans Geosci Electron. 1977;15(3):142–7.
120. Chang CC, Lin CJ. Training v-support vector classifiers: theory and algorithms. Neural Comput. 2001;13(9):2119–47.
121. Dalal N, Triggs B. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Histograms of oriented gradients for human detection. 2005;2005:886–93.
122. Oliva A, Torralba A. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. Int J Comput Vision. 2001;42(3):145–75.
123. Pauly L, Sankar D. Detection of drowsiness based on HOG features and SVM classifiers. In: 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). IEEE; 2015:181–186.
124. Ogaki K, Kitani KM, Sugano Y, Sato Y. Coupling eye-motion and ego-motion features for first-person activity recognition. In: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE; 2012:1–7.
125. Wong ET, Yean S, Hu Q, Lee BS, Liu J, Deepu R. Gaze Estimation Using Residual Neural Network. In: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops); 2019:411–414.
126. Müller B, Reinhardt J, Strickland MT. Neural Networks. Berlin Heidelberg: Springer; 1995.
127. Thabtah F, Peebles D. Early Autism Screening: A Comprehensive Review. Int J Environ Res Public Health. 2019;16(18):3502.
128. Napolitano DA, Smith T, Zarcone JR, Goodkin K, McAdam DB. Increasing response diversity in children with autism. J Appl Behav Anal. 2010;43(2):265–71.
129. Whyatt CP, Torres EB. Autism Research: An objective quantitative review of progress and focus between 1994 and 2015. Front Psychol. 2018;9:1526.
130. Sivalingam R, Cherian A, Fasching J, Walczak N, Bird N, Morellas V, et al. A multi-sensor visual tracking system for behavior monitoring of at-risk children. In: 2012 IEEE International Conference on Robotics and Automation. IEEE; 2012:1345–1350.
131. McCann J. Youth and Disability: A Challenge to Mr Reasonable. Int J Disabil Dev Educ. 2017;64(6):668–70.
132. Lord C, Cook EH, Leventhal BL, Amaral DG. Autism Spectrum Disorders. Neuron. 2000;28(2):355–63.
133. Oosterling IJ, Wensing M, Swinkels SH, Van Der Gaag RJ, Visser JC, Woudenberg T, et al. Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach. J Child Psychol Psychiatry. 2010;51(3):250–8.

134. Thabtah F. Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment. In: Proceedings of the 1st International Conference on Medical and health Informatics. 2017:1–6.
135. MacDonald JD. Communicating Partners: 30 Years of Building Responsive Relationships with Late Talking Children including Autism, Asperger's Syndrome (ASD), Down Syndrome, and Typical Devel. Jessica Kingsley Publishers; 2004.
136. Vishwakarma S, Agrawal A. A survey on activity recognition and behavior understanding in video surveillance. Vis Comput. 2013;29(10):983–1009.
137. Marszalek M, Laptev I, Schmid C. Actions in context. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE; 2009:2929–2936.
138. Sunny JT, George SM, Kizhakkethottam JJ, Sunny JT, George SM, Kizhakkethottam JJ. Applications and challenges of human activity recognition using sensors in a smart environment. IJIRST Int J Innov Res Sci Technol. 2015;2:50–7.
139. Wang H, Schmid C. Action recognition with improved trajectories. In: Proceedings of the IEEE international conference on computer vision; 2013:3551–3558.
140. Kasprowski P, Harężlak K, Stasch M. Guidelines for the eye tracker calibration using points of regard. In: Information Technologies in Biomedicine, Volume 4. Springer; 2014:225–236.
141. Guillon Q, Hadjikhani N, Baduel S, Rogé B. Visual social attention in autism spectrum disorder: Insights from eye tracking studies. Neuroscience & Biobehavioral Reviews. 2014;42:279–97.