A Close Look at the Imitation Performance of Children with Autism and Typically Developing Children Using a Robotic System

Alireza Taheri[1] · Ali Meghdari[1] · Mohammad H. Mahoor[2]

Abstract
Deficit in imitation skills is one of the core symptoms of children with Autism Spectrum Disorder (ASD). In this study, we have tried to look closer at the body gesture imitation performance of 20 participants with autism, i.e. ASD group, and 20 typically developing subjects, i.e. TD group, in a set of robot-child and human-child gross imitation tasks. The results of manual scoring by two specialists indicated that while the TD group showed a significantly better imitation performance than the ASD group during the tasks, both ASD and TD groups performed better in the human-child mode than the robot-child mode in our experimental setup. Next, to introduce an automated imitation assessment system, we present different mathematical models of the children's imitation performance using some State-Image based algorithms including Acceptable Bound, Mahalanobis Distance, and Signals' Cross-Correlations as well as Hidden Markov Models based on the time-dependent kinematics data of the participants' joints. Among the different studied models, we observed that the "State-Image Acceptable Bound method with position, velocity, and acceleration features" is the best one. This method has a mean Pearson correlation of ~45%, which is fairly comparable to the related works (out of autism field) in assessing the quality of dynamic actions. Finally, for a treatment application of using artificial intelligence algorithms in automated evaluation of children's behaviors as an unbiased and quantifiable measurement in HRI, we propose a reciprocal gross imitation human–robot interaction platform with the potential to aid in the cognitive rehabilitation of children with autism.

1 Introduction

Impaired imitation and motor skills are among the skills that are considered as core symptoms in children with autism spectrum disorder (ASD) [1, 2]. These impairments have been studied comprehensively by researchers in the autism area [1–5]. As mentioned in the literature, the comprehensive Reciprocal Imitation Training (RIT) programs and its principles and basic factors, such as being face-to-face with patients, imitating movements of the body, sounds, and postures, how to work with dolls and toys, and imagination and pretended play and symbolic games, affect not only on the imitation skills of children with ASD but also can positively affect individual with autism's improvement in social and cognitive skills [3]. Moreover, it could be stated that the following functions of imitation can improve social skills of children with ASD by (I) Using the movements of body to develop social interaction and reciprocity with others in society [4], and (II) Empowering the patient to be aware of his or her interlocutor's activities and also intentions which can potentially lead to social learning [5].

In addition, in the last 15 years it has been widely shown that conducting robot-assisted interventions as assistant tools in autism clinical sessions can significantly improve the impact of the rehabilitation process such as imitation skills of individuals with autism [1, 2, 6–10]. From the early days

✉ Alireza Taheri
artaheri@sharif.edu

Ali Meghdari
meghdari@sharif.edu

Mohammad H. Mahoor
mohammad.mahoor@du.edu

1 Social and Cognitive Robotics Laboratory, Center of Excellence in Design, Robotics and Automation (CEDRA), Sharif University of Technology, Tehran, Iran

2 Ritchie School of Engineering and Computer Science, Computer Vision and Social Robotics Lab, University of Denver, Denver, CO, USA

of applying this technology in the autism area, the hypothesis that "children with autism performed significantly better in robot-child mode than the human-child mode (e.g. in imitation tasks) in the intervention sessions" arose among researchers (mostly engineers) and has been widely argued and summarized in related review papers in this area [1, 6, 7, 10–12]. Studying this hypothesis has been a motive for further research on social robots as assistant tools in this field during the last years [13]. In this study, we report our observations regarding the confirmation or rejection of this hypothesis in the pre-set imitation tasks in our clinical experimental setup. While most published studies have reported a qualitative or descriptive assessment of autistic subjects imitation performance during robot-assisted games or have analyzed imitation games as some sort of classification or clustering problem [1, 12, 14, 15], this paper investigates the kinematics of children with ASD in imitation tasks to assess their quality of actions and undertakes a mathematical comparison with their TD peers in order to fill a serious gap in the literature. In fact, the use of unbiased and quantifiable measurement in HRI is important to advance the field of Socially Assistive Robotics (SAR) and the support for children with ASD. Ali et al. [16] proposed a mathematical model based on the adaptive multi-robot treatment of children with autism called MRIS, focusing on joint attention (JA) and imitation. The authors used a Kinect sensor to evaluate the success rate of the imitation actions. They mainly focused on their proposed cognitive architecture's strategies for scoring JA and tasks' adaptation with the participants' imitation performance scored as 0 or 1 in each experiment, but the assessment of the imitation tasks' quality was not analyzed systematically based on all the participants' joints' kinematics. Aly [17] used image processing techniques amplitude, such as normalized histogram intersections throughout the video frames or producing the characteristic vector including the center of gravity, mean absolute deviation of a pixel, etc. to measure gestures of small and large in dynamic gesture recognition for application in autism therapy. However, this study also reported its results as a classification problem. In another study, Fujimoto et al. [18] published techniques for imitation assessment and evaluating the motion of a human mediator in real time by their humanoid robot for use in applications in ASD treatment. The authors used some body angles as the features as well as the Q-Learning approach as a method of choosing the key frames for the movements' signal de-noising. They proposed a method of a cluster-based Mixture Gaussian framework and an Expectation-Maximization algorithm using parameters which are converted by Principal Component Analysis (PCA) to do the clustering of each user movement in real time. Since their work had the potential for dynamic assessment of the quality of actions, they used a very small number of children with autism in their study to validate their study's algorithm as well as there is a lack of providing

quantitative data regarding their subjects with ASD. None of the mentioned studies have compared the kinematics of typically developing children with those of children with autism, which is one of the noticeable points of our current study. While the focus of most studies has been on recognizing the actions instead of studying the quality of the actions, there are some noticeable papers in the field of automated assessment of the quality of dynamic actions (out of the autism area). In [19], Pirsiavash et al. used algorithms including Space Time Interest Points (STIP), Hierarchical Network Features, Discrete Fourier Transform (DFT), and Fast Fourier Transform (FFT) to evaluate the quality of some actions in skating and diving sports. They reported that after testing all of their different algorithms on the data set, the best calculate mean rank correlation between the human judges and their automated scores were 0.41 in diving evaluation (applying the DCT algorithm) and 0.45 in figure skating evaluation (applying the Hierarchical algorithm). The authors of [19] mentioned that they aimed to propose a general method for assessing the quality of actions and tested their proposed algorithms on two different sports' datasets: diving and figure skating.

Two most common methods in assessing the quality of imitation actions for children with ASD are: first, different questionnaires filled in by the parents, teachers, etc., and second, child psychologists conducting some imitation tasks for children with skill impairment and usually discretely scoring the subjects' performance based on their observations and experiences [1, 2]. This is done to predict the severity of the children's deficits as well as to design the next rehabilitation tasks for the participants. These kinds of approaches are known as "clinical methods" [20–22]. Although clinical methods are commonly used worldwide, their main limitations, which could seriously affect the methods' reliability, include the qualitative or descriptive nature of the assessment, dependence of the results on the evaluator's personal viewpoint, the child's mental state, and environmental impacts that should be considered and handled. Moreover, the limited interval and discrete nature of the scoring reduce the accuracy and measurement sensitivity of assessing the quality of actions; hence, possible small changes in patient performance might not be detected by the evaluators. Therefore, the serious issue in using such clinical methods is possible biases when caregivers fill out surveys. To overcome this problem, engineers are trying to use the kinematics and dynamics characteristics of the movements alongside artificial intelligence and control theory to propose quantitative and more general criteria, i.e. independent of the evaluators, as modern Health Outcome Measures for better accuracy and sensitivity in diagnostic and treatment evaluations. More generally, in the literature of biomechanics there are some clinical methods used as criteria for assessing actions which consider the motions' range, speed, accuracy, smoothness, and used energy (e.g., for physical rehabilitation of individuals with

strokes) [20–25]. An overview of such works inspired us to propose the automated assessment algorithms in this study. The readers can also find some restrictions and requirements of the available methods in [22–25].

A popular recent application of artificial intelligence in studies containing SAR for education and therapy of children is focused on the "automated evaluation of behaviors" and "adaptive teaching". Engineers would like to empower the robots and machines to automatically assess the behaviors of their audiences and then report back to teachers, human-mediators, psychologists, and parents. Moreover, they try to empower the robots to adapt the treatment and teaching protocols based on both the current status and history of the users [13, 16, 26]. Achieving these goals will bring us one step closer to the practical use of social robots in education and treatment centers and to assist in the study of their impact on society. In this regard, a robotic platform for reciprocal gross imitation training is proposed at the end of the paper.

1.1 Research Purposes and Hypothesis

This paper focuses on finding a preliminary exploratory answer for the following main research question: *How different is the performance of children with autism in gross imitation movements in comparison to typically developing children?* To this end, we recruited 20 children with autism referred to as the ASD group, and 20 typically developing children, i.e., the TD group, to participate in different robot-child and teacher-child body gesture imitation exercises. This paper is an initial attempt to find the differences within and between the TD and ASD groups in dynamic imitation tasks. We divided the mentioned main research question into the following sub-questions:

1. What are the preliminary estimates of the acceptance rates of a NAO humanoid robot as a companion in gross imitation games for TD and ASD groups?
2. Is the imitation performance of the ASD and TD group in the robot-child mode better than the teacher-child mode?
3. What is the difference between the movements' kinematics of TD and ASD subjects in the imitation tasks considering agility and performance's accuracy?
4. How accurate can a robotic system score the imitation performance of a participant in dynamic tasks?

We are interested to study whether the two following common hypotheses in the literature are confirmed: hypotheses-I) "children with autism have a deficit when doing imitation tasks in comparison with their typically developing peers" and hypotheses-II) "the quality of imitation performances of children with autism in robot-assisted imitation games is higher than their performances in human-based non-robotic similar tasks". To this end, the results of manual

assessments of robotic and non-robotic, i.e. robot-assisted and teacher-assisted, gross imitations are compared. Then, a set of algorithms, including State-Image based and Hidden Markov Model-based assessment algorithms with the capability of automated scoring during the robot-assisted imitation actions, is presented and each of the subset's performance on automated assessment is investigated. In addition, we explored the accuracy of the robot automated assessments using manual scores provided by two human mediators as reference scores. We assessed the Pearson correlation rate of machine scores and human scores via the presented set of algorithms to see how consistent the machine outputs are with the equivalent manual scores generated by our two human coders in the Likert scale of 0 to 4 as the ground truth scores which is described in Sect. 2.5. In order to provide a fast guideline for readers to easily follow, an overview flowchart of this study is presented in Fig. 1. The reader may ask, "why have the authors tested or proposed different categories of intelligent algorithms for scoring the performance of the imitations?" Further clarification is needed to answer to this question. Just as there is only one way to get 100 out of 100 in an exam while there are numerous ways to get 75 and even more possible ways to get 0 out of 100; there is only one correct time-series solution for the "assessment of action's quality" which is much more complicated than the actions' classification or recognition for each dynamic action that receives full marks; but, there are an infinite number of methods to obtain an incomplete task solution. Therefore, having access to a number of finite incomplete scores is not necessarily an appropriate tool for applying common machine learning methods. With this introduction, one can see that empowering a machine to act like a trained judge in scoring an action similar to clinical evaluators is a very interesting and useful problem in artificial intelligence. The readers can also find other uses and challenges of assessing the quality of dynamic actions in [19, 27, 28].

Lastly, we have proposed an architecture for a human–robot interaction platform with the ability of automated imitation assessment for reciprocal body gesture imitation training programs. The current research is built on our previous studies on developing reciprocal human–robot-interaction platforms for children with autism [13].

2 Research Methodology

2.1 Participants

.

Twenty children with autism, including 14 boys and 6 girls (mean age: 4.95 years, standard deviation: 2.01 years), and twenty typically developing children, including 10 boys and 10 girls (mean age: 5.30 years, standard deviation: 1.95 years)
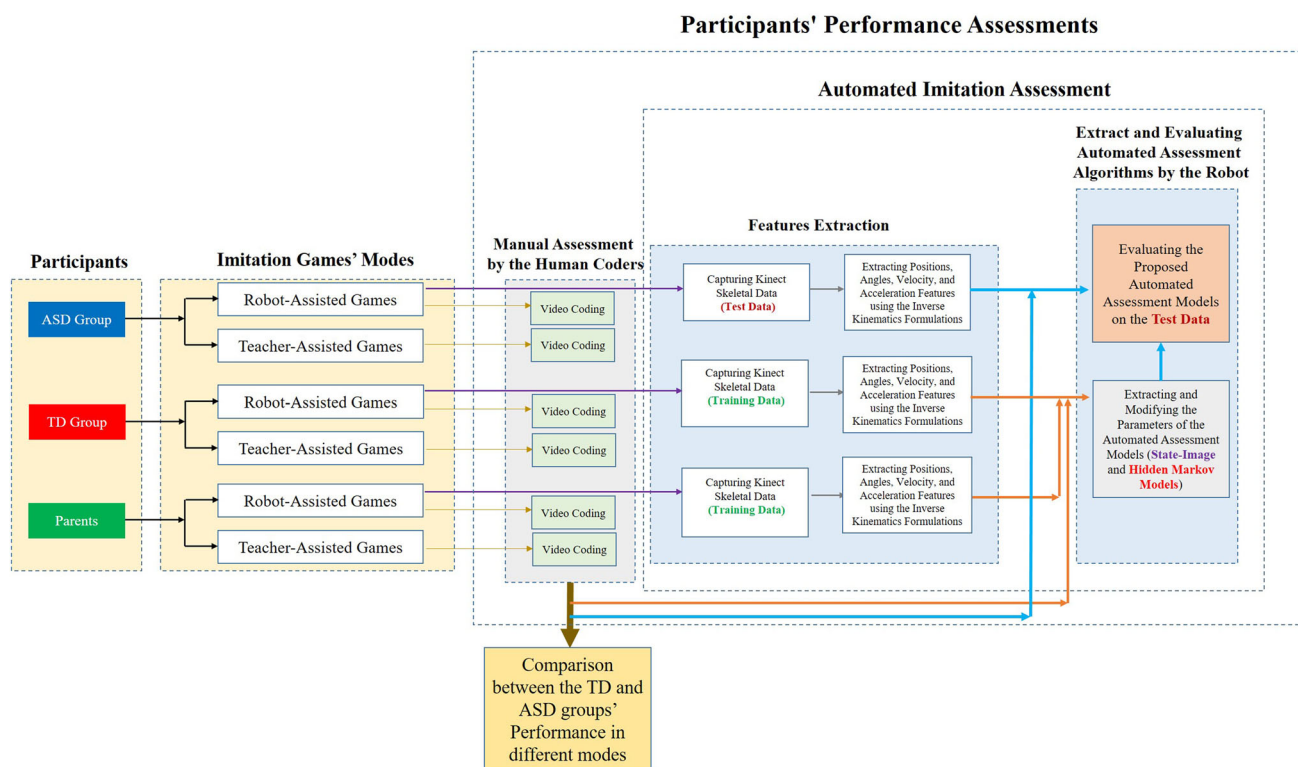
Fig. 1 An overview flow-chart of this study containing the participants, imitation games' modes, manual assessment, automated assessments procedures, groups' comparisons, and algorithms' evaluation

voluntarily enrolled in this study. The children with autism were diagnosed by the doctors and psychologists of the Center for the Treatment of Autistic Disorders (CTAD) in Iran and have received some Applied Behavioral Analysis (ABA) clinical interventions before attending this study.

Regarding the gender ratio in autism prevalence, i.e. about four males for every one female [1, 6], gaining access to the same number of girls with autism was quite difficult for us, and we were unable to recruit ten girls for this study. It should be added that in this study, we did not consider any possible gender differences between the imitation performances of the recruited participants.

2.2 Humanoid Robot

The robot used in this study is a NAO H-21 commercial humanoid robot [29]. To help to make a better connection and further communication with our participants the robot was given the Iranian male name "Nima".

2.3 Experimental Setup

The study was conducted in $10 \times 5 \times 3$ m^3 game room at the Social and Cognitive Robotics Lab., Sharif University of Technology, Iran. Two cameras and one Microsoft Kinect2

sensor were used to capture the subjects' skeletal kinematics data. A human mediator, a robot operator, each participant, and his or her parent(s) were also present in the game room. The children's tasks were preset and the games session was structured. The robot was commanded by an operator in a wizard of Oz style control. During the tasks, a person was always standing near the robot for taking care of Nima to avoid any possible falls or for the cases that the participants tend to touch the robot; however, this person was not involved in the imitation tasks at all. It should be noted that for machine assessment of the participants' movements, the x, y, and z positions of each subject's skeletal joints versus time were captured in the robot-assisted modes of this study using the Kinect sensor. The robot operator captured the Kinect's data of the subjects from one second before each robot's action until the end of participants' performances. The details of the used algorithms and obtained results for automated assessments are presented in Sects. 2.6 and 3.2, respectively.

2.4 Designed Imitation Games

A set of body gesture imitation actions were designed to be conducted in three different modes: (1) Imitation of Nima by each participant in a dyadic robot-child interaction (mode *A*); (2) Imitation of the human-mediator by all subjects in a
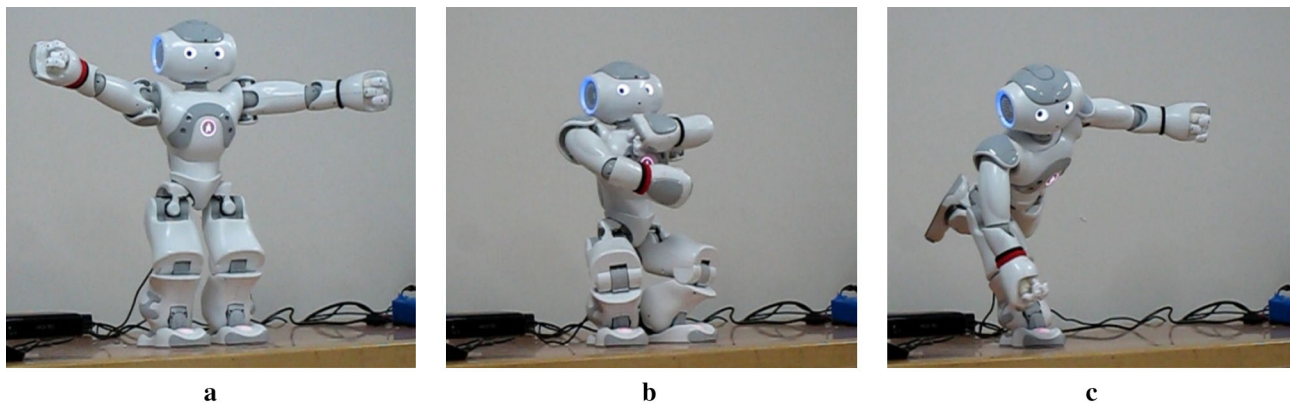
Fig. 2 Snapshots of one frame in the designed movements for the robot-child mode: **a** task #1: Nima moves its arms open in the straight positions in the middle of the action, **b** task #4: The robot moves the arms cross each other on the chest and bends both knees, and **c** task #7: The robot balances on one leg while bending the torso, the knees, and moving its both arms

dyadic teacher-child interaction (mode *B*); and (3) Simultaneous imitation of the robot by the child and the parent in a triadic robot-child-parent interaction (mode *C*). It should be noted that after conducting the study, we realized that similar to mode *B*, we could possibly add and conduct a mode *D* to our study containing the imitation of teacher by both child and parent, simultaneously which could enable us to have a more comprehensive analysis between and within the child-robot and teacher-child modes.

We designed 10 gross imitation tasks which included a combination of arms, feet, neck, and torso movements for each mode and the participants were expected to imitate the robot or teacher simultaneously. Having a hierarchical trend, our dynamic tasks started with symmetric, visible, and easy exercises and got progressively more complicated and asymmetric as they continued (such as one-leg balance) in each mode (Fig. 2). All of the actions were intentionally started from the neutral standing position of the robot or teacher. The minimum and maximum time length of the movements were 4 and 15 seconds, respectively. In both groups half of the participants were randomly chosen to be involved first in the robot-child interaction (mode *A*), and then in the teacher-child interaction (mode *B*); and vice versa for the second half of the subjects (i.e. counterbalance condition). Finally, mode *C* was conducted for each participant and one of his or her parents. Short descriptions of what robot does in each task are presented in Table 1.

2.5 Manual Rating of the Movements

Two video coders manually and independently rated the performance of the participants in all tasks and modes using the Likert scales 0 (no action done by the child or completely wrong performance) to 4 (excellent imitation). The average of their scores were also used as the reference grades for the robot's automated imitation assessment. A comparison

between the performances of the ASD and TD groups was done and is reported in the results section.

2.6 Automated Imitation Assessment

As one of the paper's contributions, we empower the robotic system to automatically score the quality of the participants' gross imitations. As mentioned above, a set of reasonable and standard algorithms were studied as possibilities for the robot-assisted automated assessment to see which one produced the most compatible outputs to the equivalent human scores based on the Person correlation coefficient. The appropriate features and automated assessment algorithms were chosen so that they could handle the following important challenges: (1) the dynamics nature of body gesture imitation actions, (2) the possible misalignment in the participant's starting time and time duration of performing an action, (3) different subjects' height and dimensions, and (4) different initial participant position with respect to the Kinect sensor. Considering that the robot performs the tasks quite similar for all the participants, to have more reliable and comparable mathematical models of the ASD and TD groups, we only considered the robot-assisted modes' kinematics data for our automatic assessments. On the other hand, due to any potential variations in movements' exact timing and spatial properties in the teacher performances, the Kinect's data of mode B has not been considered for extracting the mathematical models in the automated assessment section in this study.

To conduct the automated assessment of the subjects' imitation performance, we have chosen State-Image based algorithms [30, 31] including the Acceptable Bound, Mahalanobis distance, and signals' cross-correlations as well as the Hidden Markov Model (HMM) algorithm combined with a Gaussian Mixture Model (GMM), which is used as a tool for

Table 1 Imitation tasks for both robot-child and teacher-child modes

Task no.	Descriptions	Status of symmetry	Main involved components
1	Move hands forward, up, open, and down in the straight positions	Symmetric	Left arm and right arm
2	Move hands up, folding in chest, open, folding on upper-arm, and down	Symmetric	Left arm and right arm
3	Stretch one hand and bend the torso while the other one is on waist, and next, its mirror symmetry	Asymmetric	Left arm, right arm, and torso
4	Move the arms cross each other on the chest and bend both	Symmetric	Left arm, right arm, left knee, and right knee
5	Move one hand up and the other one down while making right angles between lower-arm and upper-arm for both arms, as well as the head movement	Asymmetric	Left arm, right arm, and head
6	Bend one knee, bend the torso, and move both hands cross each other in a pendulum pattern	Asymmetric	Left arm, right arm, torso, left foot, and right foot
7	Balance on one leg while bending the torso, the knees, and moving both arms in an asymmetric pattern to keep the balancing situation	Asymmetric	Left arm, right arm, torso, left foot, and right foot
8	Bend both knees, move the arms and head during a symmetric dance pattern	Symmetric	Left arm, right arm, torso, left foot, right foot, and head
9	Move the arms, head, and feet during an asymmetric dance pattern	Asymmetric	Left arm, right arm, torso, left foot, right foot, and head
10	Balance on one leg from the beginning of the movement, moving the head, and moving both arms in an asymmetric circular and curved pattern	Asymmetric	Left arm, right arm, torso, left foot, right foot, and head

recognizing sequential patterns [32–34] which are described in this section.

2.6.1 Skeletal Data and the Kinematics Formulations

The main tool for proposing the automated assessment system is extracting body angles as the features from the skeletal data of the participants using kinematics formulations. To this end, we used a Microsoft Kinect Sensor, which made it possible to record the position of the subject's joints in each frame, and then the features were calculated using the appropriate forward kinematics formulations. The x, y, and z positions of each participant's skeletal joints versus time were captured at the rate of 30 frames per second during the imitation games unless the child moved out of the sensor's view field which was common because of his or her low instruction perception, wandering around the room, maladaptive behaviors, etc. The Kinect's data of mode A and mode C were captured from one second before each robot's action until the end of subject's performance. In order to reduce the effect of the Kinect's noise, a 4th order Butterworth filter with a cutoff frequency of 6 Hz was applied to the data before calculating the features. The extracted features for machine assessment are introduced in the next section; however, to avoid interrupting the paper's flow, the details on extracting the kinematics formulations are presented in the Appendix.

One of the main points of the automated assessment section is that the used method could be considered as a measure to enable comparison between studies and reduced potential evaluator biases.

2.6.2 Features

Considering the mentioned concerns, in this study (except for the Mahalonobis distance method which is based on the 20 skeletal joint positions of the subjects), 19 direction invariant angles have been selected and calculated from the skeletal data as the independent position-based features to describe the movements. These features are presented in Table 2. It was assumed that the value and range of these angles is approximately the same for all subjects, which is not a flawless but acceptable assumption. Using a numerical derivatives technique, we also calculated and added the velocity and the acceleration of the angles called the velocity-based and acceleration-based features to the features' space alongside their positions to improve the actions' automated analysis. Hence, there was a total of 57 features considering the position, velocity, and acceleration, simultaneously. In our study, the robot was aware of the participants' action numbers.

We assumed that the upper body of the users has three degrees of freedom (DOF), including F15, F16, and F17, which represent the rotation of the body about the three coor-

Table 2 List of the extracted position-, velocity-, and acceleration-based features for movement analysis in this study: the position features (i.e. the angles) were calculated from the Kinects skeletal data using the forward kinematics formulations (presented in the Appendix)

Description	Position-based features (X)	Velocity-based features (V)	Acceleration-based features (A)
Right shoulder pitch	F1	F20	F39
Right shoulder roll	F2	F21	F40
Right elbow pitch	F3	F22	F41
Right elbow roll	F4	F23	F42
Left shoulder pitch	F5	F24	F43
Left shoulder roll	F6	F25	F44
Left elbow pitch	F7	F26	F45
Left elbow roll	F8	F27	F46
Right hip pitch	F9	F28	F47
Right hip roll	F10	F29	F48
Right knee pitch	F11	F30	F49
Left hip pitch	F12	F31	F50
Left hip roll	F13	F32	F51
Left knee pitch	F14	F33	F52
Euler ZXY:θ_z	F15	F34	F53
Euler ZXY:θ_x	F16	F35	F54
Euler ZXY:θ_y	F17	F36	F55
Head yaw	F18	F37	F56
Head pitch	F19	F38	F57

The velocity and acceleration-based features were calculated from the position-based features using numerical derivatives techniques

dinate axes. Only the rotation angle about the z-axis (F15) was considered for the lower body of the subjects. Regardless of the initial orientation of the subjects toward the Kinect sensor, we set the Euler angle of the upper body about the z-axis (F15) to zero by multiplying appropriate rotation matrices for all the users in every frames of the actions to omit the effect of participants' standing orientations.

The dimensions of the input data for our automated assessment algorithms is $T \times F$ where T and F are the number of frames and features, respectively. Due to a possible difference in the time duration of each subject's action, the signals are first normalized in dimensions; therefore, a pre-set N data sampling is done on the signals during the whole time duration of the movement by fitting appropriate spline curves [32]. In this paper, $N=50$ points have been extracted from each signal to build the modified $50 \times F$ input signal for each captured data. It should be noted that by aligning the raw participants' skeletal signals, the information regarding the velocity and time-rate of the actions vanishes from the modified input signals, and the "agility" of the real actions could not be studied as only the positions-based features are considered as the features' space. To this end, by calculating the first and second derivative on the position data with respect to time, i.e. numerical derivative, the velocity-based and acceleration-based features are also added to the features' space. Due to the dynamic nature of the actions, it was estimated that adding such features could improve

the performance of the assessment algorithms. Therefore, each presented model was tested three times considering (1) the position features only, (2) the position and velocity features, and (3) the position, velocity, and acceleration features altogether as the feature space. Moreover, we present a subsection in the Results section in which the tasks' agility of the ASD and TD groups is compared.

2.6.3 State-Image Method

One of the automated algorithms used in this manuscript is called the State-Image (SI) method [30, 31]. This technique codes every state of the data into pixels of a 2-D image and converts a set of time-series into a single image. Each element in the data, i.e. the value of each feature in each frame, is coded as the brightness of the image pixel ranging from 0 to 255. Using this method, we can convert a dynamic time-series of movements into a set of static equivalent images, and then study the produced images based on common available techniques. By using the State-Image method in this study, each action signal with the dimension of $T \times F$ was converted to an image with $T \times F$ pixels. Just observing the produced images can also give appropriate information regarding the subjects' performance and their probable errors in the actions.

This section clarifies how the pixels' values are calculated and assigned from the input data. To do so, we applied three

different criteria to produce the images using the State-Image method as follows. It should be noted that in each criteria, the mean and standard deviation of the TD subjects' performances with a full score (i.e., 4 out of 4) have been used as the *reference signals* for each action.

I, SI-MH) *Mahalanobis Distance* of each participant's signal from the 20 body joints' positions (Eq. 1) of the reference signal extracted from the Kinect sensor is used as the criteria for producing the static black and white image. The formulation for assigning the value to each pixel is presented in Eq. 2. It should be noted that the distribution of the reference signals in each frame and feature has been assumed to be normal; therefore, the n and m parameters in Eq. 2 are set to 2 and 3, respectively.

$$Mahalonobis\, Distance = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (1)$$

$$pixel\, value = 255 \times \begin{cases} 1 & if\ \text{Mahalnobis Distance} \leq n \\ \frac{1}{(n-m)} \times \text{Mahalnobis Distance} - \frac{m}{n-m} & if\ n \leq MahalnobisDistance \leq m,\ n = 2,\ m = 3 \\ 0 & if\ \text{Mahalnobis Distance} \geq m \end{cases} \quad (2)$$

In Eq. 1, x is the vector of the studied observation, μ is the mean vector of a set of observations, and Σ is the covariance matrix of the data.

II, SI-AB) *Acceptable Bound* In this sub-method, the acceptable bound for each feature in each frame is the interval [(*mean of the reference signals*) $\pm n \times$ (*standard deviation of the reference signals*)]. If the studied participant's performance lies in the mentioned bound, the equivalent pixel value for his or her action in each frame and feature is set to 255 (i.e., the performance of the child in that feature and frame is assumed to be perfect). If the studied signal lies out of the interval [(*mean of the reference signals*) $\pm m \times$ (*standard deviation of the reference signals*)] ($m > n$), his or her related pixel value is set to 0 which means that the performance of the studied subject in that feature and frame is completely wrong with respect to the original task. We have considered a linear formulation for the performance changes between the above mentioned intervals. Again, assuming that the reference signals are distributed normally, in this article n and m are assumed to be 2 and 3, respectively (Fig. 3).

III, SI-CC) *Signals' cross-correlations* One of the automated assessment algorithms used in this study is calculating the correlation between the reference signal and the test signal. In this method, a moving window of length w (e.g., 1/10 of the total frames of the signal), simultaneously goes through the two mentioned signals and the time dependent values of their correlation for each feature and frame are calculated using the Pearson correlation formulation of the pairs of $w \times 1$

vectors. The correlation values are between -1 and 1 which are linearly mapped to the interval [0-255] in order to make the static images.

After producing the corresponding image by each of the mentioned image-state method, we have a matrix, i.e. image pixels, with a value between 0 and 255 (or better said, 0–1). One way to extract a score as an automated score from each image is to calculate the Root Sum Square (RSS) of the pixel values and dividing the result into the number of the pixels to produce a number between 0 and 1. However, considering that each joint's movement can have a different importance in every action, we assigned a weighted or correction factor between 0–1 for each feature's elements of the actions by calculating the Pearson correlation between that feature's pixel values and the human scores for the TD and Parents' groups,

i.e., using weighted mean strategy. Finally, the automated score related to that algorithm is extracted by calculating the root sum square of the corrected pixel values. We also calculated the mean values of all pixels as a simpler method of scoring; however, using RSS led to a slight better result in automated scoring.

Therefore, in the automated scoring process of this paper, the data are separated in three groups: first, the performances of the TD participants with a score of 4 out of 4 are considered as the reference signals; second, the other TD group's data as well as the Parent's group data is used for calculating the weighted factor regarding each feature's importance during the actions; and third, the ASD group's data is used as the test set. The results of using the above mentioned methods for evaluating the robotic system's ability to assess the quality of children's actions are presented in the Results section.

2.6.4 Hidden Markov Models (HMM)

While using HMM, considering the independence assumptions of the body parts, the automated assessment could be studied in different situations such as (1) fully independent assumptions for all angles, velocities, or acceleration: extracting 19, 38, or 57 one-variable parallel independent Hidden Markov Models for each action, (2) considering all features at once to extract one HMM model for each task, and (3) different combination of body parts, e.g. assuming the movement of left side and right side of the body independently, assuming each foot and arm's movements independent, etc. Due to the small amount of the avail-
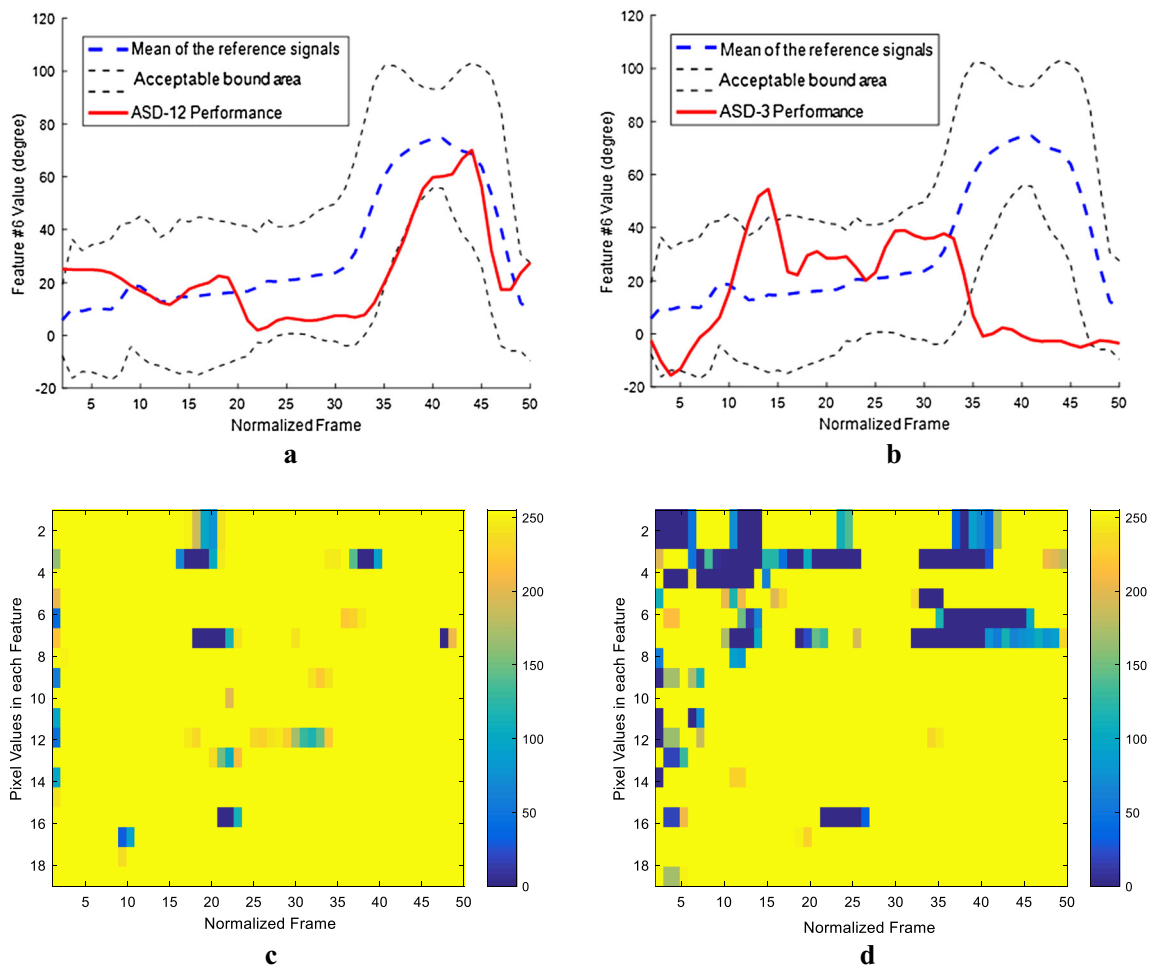
Fig. 3 a, **b** The mean (blue dash line), acceptable bound (the area between the black lines) and the performance of two participants with ASD (red solid line) in feature #6 (Left Shoulder Roll) versus the normalized or aligned time in movement #1. The participants' IDs are ASD-12 and ASD-3 in a and b, respectively. The human coders scored these ASD participants 4 and 1, respectively, in movement #1. As it can be seen, the red lines show that the left shoulder angle for the ASD-12 lies in the acceptable bound during the action while highly follows the pattern of the reference signals' mean extracted from the TD subjects; but for the ASD-3, the performance lies out of the acceptable bound in a noticeable time duration especially in the last frames of the actions which indicates that this participant did not perform the studied action, well; **c**, **d** The equivalent images produced by the Acceptable Bound sub-method for the participants with ASD in 2-a and 2-b, respectively. In these images, only the 19 position features, i.e. Features 1–19, are considered. The value of each feature in each frame is coded as the brightness of the image pixel ranging from 0 to 255. Higher pixel intensities or values in images show the higher performance in each feature and frame. Regarding these images and the used color bars, deviating from the yellow color shows that the performance of the subject is taking distance from the perfect criteria of this sub-method which means that the features are getting out of the acceptable bound. Related to this sub-method's pixel evaluation, readers can visually observe the difference between the overall performance of the selected participants and the reference signal in each frame and feature of movement #1. A look at 2-c and 2-d shows us that in the viewpoint of the Acceptable Bound assessment method, ASD-12 gets higher score than ASD-3 in this task which is in line with the assessments of the human coders

able data in our experiments and the gross imitation tasks' design which includes feet, arms, head, and torso movements, instead of assessing an action all at once or choosing the fully independent assumption, we have divided each task into six sub-actions by considering independence assumptions for the movements of the participants' left arm, right arm, left foot, right foot, head, and torso; therefore, six independent Hidden Markov Models have been trained regarding each imitation task based on the TD participants with a full score

[32, 33]. Then, similar to the State-Image part, the rest of the TD data as well as the Parent group's data is used to calculate the weighted factor for each sub task. The features for each Hidden Markov Models are defined in Table 3. The Left-to-Right (LTR) method for initializing the models' parameters was used in training the sets of Hidden Markov Models [32, 33]. Selecting the HMM state numbers was done using the Bayesian Information Criterion (BIC) [32–34].

Table 3 The features for the Hidden Markov Models in this study

No.	Body parts	Involved features in each extracted HMM model		
		Position features (X)	Velocity features (V)	Acceleration features (A)
1	Right Arm	F1, F2, F3, F4	F20, F21, F22, F23	F39, F40, F41, F42
2	Left Arm	F5, F6, F7, F8	F24, F25, F26, F27	F43, F44, F45, F46
3	Right Foot	F9, F10, F11	F28, F29, F30	F47, F48, F49
4	Left Foot	F1, F13, F14	F31, F32, F33	F50, F51, F52
5	Torso	F15, F16, F17	F34, F35, F36	F53, F54, F55
6	Head	F18, F19	F37, F38	F56, F57

Table 4 The summary of automated assessment algorithms used in this study with short descriptions

Method	Sub-method	Used features
State-Image (SI)	Acceptable bound	X
		X and V
		X, V, and A
	Mahalanobis distance	X
	Cross-correlation of signals	X
		X and V
		X, V, and A
HMM	LTR initiation and BIC	X
		X and V
		X, V, and A

The log likelihood or probability of each sub-action's test signal using the trained models, i.e. HMM's output, could be used as the criteria for the robot's automated assessment of gross imitation movements. We used the weighted mean of six HMM's outputs considering the correction factors as the automated score of this algorithm. To this end, the TD and parents groups was selected and used as the training set for each action and the equivalent average of manual ratings by the video coders was used as the labels. Therefore, appropriate Hidden Markov Models were trained and the features' importance factors were calculated based on the training set. Applying the trained model on the data of the ASD groups as the test set, the six static outputs as well as their weighted means were calculated and considered as the score by machine. The Pearson correlation of the test set's outputs with the human scores of ASD group are presented in the Results section of the article.

The summary of automated assessment algorithms used in this study are presented in Table 4.

3 Results and Discussion

In data collection, a 15-min introduction session was held for each participant. In this session, the robot introduced itself

Table 5 ANOVA analysis of the TD and ASD groups in the gross imitation tasks in child-robot mode considering the children's diagnosis and action number

Factor	DOF	F value	P value
Covariate: age	1	207.59	0.000
Independent factor 1: diagnosis	1	322.09	0.000
Independent factor 2: action number	9	10.39	0.000
Interaction of diagnosis and action number	9	2.05	0.032

and tried to communicate with the child via singing songs and dancing. After the introduction session, the imitation games were run for the cooperative children. Figure 4a–f show some snapshots of the imitation games in different modes. Among the TD and ASD groups, 19 and 14 participants, respectively, showed a tendency to imitate the robot's tasks. Hence, to answer the first sub research question, in these experimental conditions, 95% of the TD participants and 70% of the ASD group members showed an interest in performing the tasks of the Nima robot in their first child-robot interaction, which gave us a preliminary estimate of the robot's acceptance for TD children and children with ASD. Similar to the observations in [13], it is interesting to note that all the uncooperative children were less than 5 years old.

3.1 Manual Assessment of the Imitation Performance

The Pearson's correlation coefficient (r) of our two coders' scores were 0.847 (p value = 0.000). This indicates a strong positive correlation between our two judges' scores. We have considered the mean of the coders' scores as the participants' performance in each action and mode.

Applying a two-way ANOVA test on the manual scores of the children considering "age" as the covariate and the two independent factors: Diagnosis (including 2 levels: TD or ASD) and Action Numbers (including 10 levels: 1–10) led to the results found in Table 5. In Fig. 5, the mean of the manual scores for both groups of the participants in each imitation task are presented.
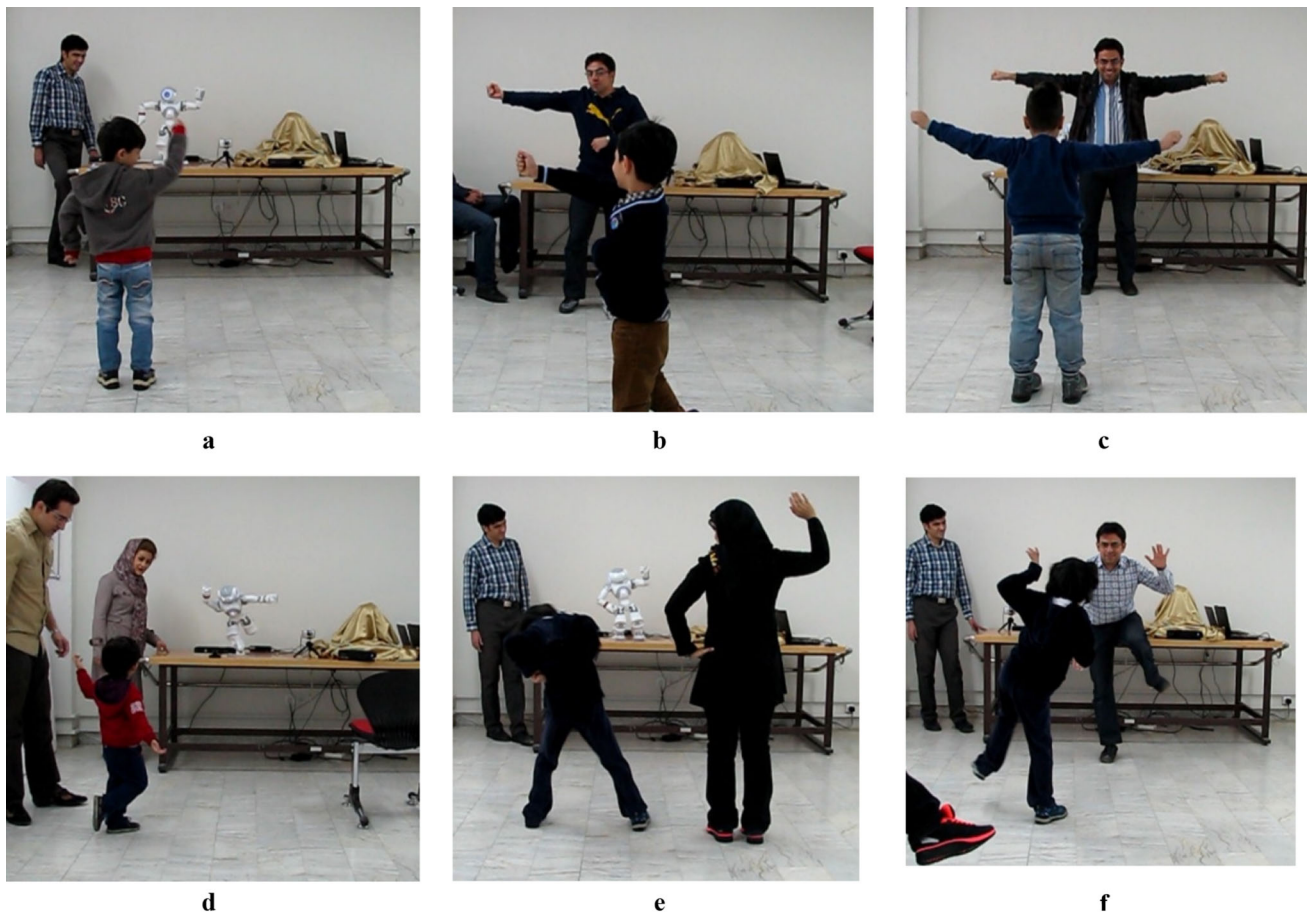
Fig. 4 Snapshots of the children's performance in different modes of the imitation games: **a** robot-child mode in task #5, **b** teacher-child mode in task #9, **c** teacher-child mode in task #1, **d** robot-child mode in task #7; in this case, the subject with ASD was 3 years old and had difficulty in instruction-perception. His parents highly encouraged him during the games in order to involve their child in the games; however, we asked them not to do the imitation tasks alongside the robot and teacher in the robot-child and teacher-child modes to avoid any possible cofounding factor in our study regarding the person that is imitated, **e** robot-child-parent mode in task #3, and **f** teacher-child mode in task #7. It should be noted that during the tasks, a human was always standing near the robot for taking care of Nima if any possible falls happen or for the cases that the participants tend to touch the robot; however, this person was not involved in the imitation tasks at all.
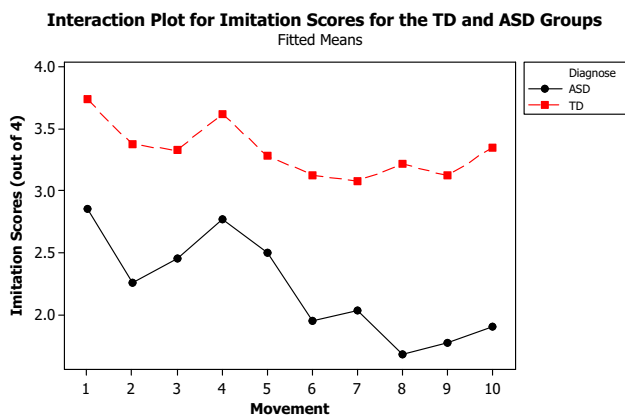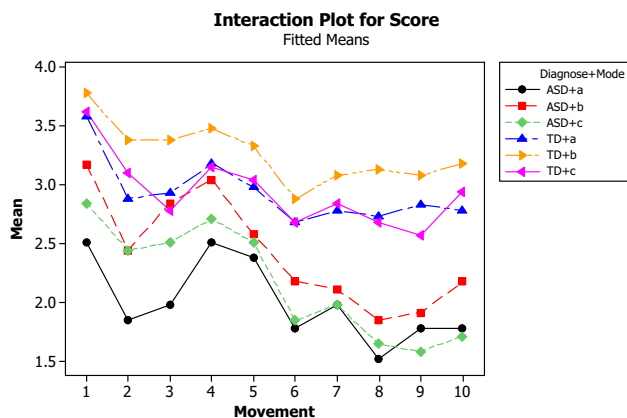

Fig. 5 Interaction plot for the imitation scores of the ASD and TD groups considering the children's diagnosis and action number

To have a closer look at the performance of the subjects in each mode, we performed another two-way ANOVA analysis on the manual imitation scores considering "age" as the covariate and two independent factors: "Diagnosis + Game Mode" including 6 different levels and Action Number including 10 levels. The results are presented in Table 6; Fig. 6.

According to Table 6, the effect of the independent factors "Diagnosis + Game mode" and "Action Number" on the gross imitation scores of the participants are significantly different, separately (p values < 0.05). Moreover, in our research, no interaction was observed between the factors "Diagnosis + Game mode" and "Action Number" (p value $= 0.945 > 0.05$). As expected, we observed that the performance of the TD group in imitation movements was significantly better than the ASD, which confirms the deficit of children with ASD in imitation tasks in comparison with

Table 6 ANOVA analysis of the TD and ASD groups in gross imitation tasks in the robot-child mode considering the children's diagnosis and game modes as the first independent factor and action number as the second factor

Factor	DOF	F value	P value
Covariate: age	1	211.07	0.000
Independent factor 1: diagnosis	1	70.36	0.000
Independent factor 2: action number	9	10.39	0.000
Interaction of diagnosis and action number	9	0.68	0.945

Fig. 6 Interaction plot for the imitation scores of ASD and TD groups considering the children's diagnosis and game modes as the first independent factor and action number as the second factor

their TD peers (Table 5; Figs. 5, 6). Interestingly, having analyzed TD and ASD groups independently, the performance of both TD and ASD groups in teacher-child mode (mode *B*) was significantly better than the robot-assisted modes, i.e. modes *A* and *C*, in our experimental setup with a 0.000 and 0.0001 *p* values, respectively. Therefore, the hypothesis that "children performing gross imitations tasks better in the robot-child mode than in the human-child mode especially for children with autism" is not confirmed in this research. We had observed a similar result in our previous study in facial imitation tasks for children with ASD [13]. Costa et al. [14] also reported that, on average, their participants more often imitated the human-mediator than their robot; however, their observations indicated that there was no statistically significant difference between the two performances ($p = 0.180$). Unlike the observation in this study, the authors in [35, 36] claimed that their participants with autism had a lower performance in the facial expressions imitation tasks in the human mediator-child mode than the robot–child mode. As can be seen, different studies reported various findings regarding the mentioned hypothesis. We believe that the main reasons for the significant difference in the subjects' performances in the robot-child and human-child modes of the gross imitations in this study could be as follows: first, the previous Applied

Behavioral Analysis (ABA) treatment program received by the ASD groups in this study in autism clinics could positively affect their performance in the imitation tasks in the human-child mode. It should be noted that the participants have been somehow taught how to communicate and work with human mediators, especially in the similar environments in the clinics, in performing some of the imitation tasks such as whole body movements, facial expressions, and children's activities with objects. Second, the robot's different degrees of freedom and range of motions in comparison to the human; third, the robot's small size and probable low visibility of the movements' details; and fourth, the robot's lower verbal communication with the participants in comparison to the human teacher during the tasks, which might affect the subjects' obedience especially for the TD participants.

One of the qualitative observations in the current study was the performance of a child with low-functioning autism who, because of fatigue and low instruction perception, refused to complete the last three tasks of mode *C*. Nima's ability to give some short verbal communication with the participants was one of the advantages of this study in comparison to [13], and led to more effective communication between the children and the robot. Some of the participants asked the teacher if they could touch the robot at the end of the game sessions. However, regarding the robot's charm, we observed that some of the participants with ASD either attacked the robot or moved out of the Kinect's field of view, which led to data recording corruption. It should be noted that such behaviors were not observed from the TD participants.

3.2 Automated Assessment of the Imitation Tasks

We have used the introduced State-Image and HMM based algorithms as the automated algorithms candidates in this paper. As previously mentioned, in each algorithm, the ASD group's data was considered as the test set while the other data (i.e., the TD and the Parents' groups) was chosen as the training or ground truth set for calculating the importance factors of each feature in each task. We used each algorithm three times (except for the Mahalonobis Distance): (1) only considering the position angles as the features, i.e. 19 features F1–F19, (2) considering the position and velocity features, i.e. 38 features including F1–F38, and (3) considering all the position, velocity and acceleration features, i.e. 57 features including F1–F57 as the feature space. Table 7 presents the Pearson correlation coefficient between the machine and human scores of the ASD group's evaluation as the test set for each proposed automated algorithms. As it can be seen in Table 7, the best performance in this study is observed by the State-Image Acceptable Bound method considering all 57 features. Therefore, we prefer to introduce this method for the automatic assessing system for our robotic platform in this manuscript. The results indicate that the acceleration

Table 7 The Pearson correlation coefficient of the proposed automated assessment algorithms of the ASD group's imitation performance in different modes in each action

Method	Sub-method	Features	Action #										Mean in method
			1	2	3	4	5	6	7	8	9	10	
State-Image	Acceptable Bound	X	**0.72**	0.15	0.57	0.61	0.02	**0.49**	0.40	0.14	0.35	0.05	0.35
		X and V	0.62	0.30	**0.59**	**0.63**	0.11	0.27	0.45	0.51	0.33	0.15	**0.40**
		X, V, and A	*0.59*	*0.35*	*0.59*	*0.63*	*0.11*	*0.25*	*0.51*	*0.56*	*0.36*	*0.25*	*0.42*
	Mahalanobis distance	X	0.20	−0.31	0.06	0.27	−0.19	0.28	0.41	0.13	**0.50**	−0.07	0.13
	Cross-Correlation of signals	X	0.61	0.22	0.01	0.44	0.08	−0.26	0.32	0.44	−0.01	0.24	0.21
		X and V	0.62	0.37	−0.01	0.46	**0.15**	0.23	0.08	0.37	−0.05	0.16	0.24
		X, V, and A	0.62	**0.44**	0.00	0.46	0.14	−0.31	0.00	0.31	−0.05	0.22	0.18
HMM	LTR Initiation	X	0.21	0.14	0.24	0.44	−0.19	0.39	0.02	0.26	−0.04	−0.13	0.13
		X and V	0.12	0.26	0.46	0.32	−0.04	0.32	0.09	0.45	−0.04	0.21	0.22
		X, V, and A	0.17	0.29	0.47	0.37	−0.08	0.34	0.39	0.36	0.40	**0.30**	0.30

The highest correlation coefficient in each action is shown in bold in the table. The method with the highest performance is italicized

of the joints which is not easily calculated, understood, or sensed by the human evaluators could be an important factor in assessing the quality of actions, and this is where the machine could be great help to the human therapists. Table 8 shows the details of the best algorithms in each movement as well as the Acceptable Bound method with the position, velocity, and acceleration features. In addition, the mean and standard deviation of errors for each algorithm in all of the actions are also presented in Table 8. Here, we define an error which is equal to the absolute of difference between the human's score and the automated score by the machine (i.e. Error = |Human's Score – Automated Score|). It should be noted that to avoid needlessly prolonging the article, we declined to provide the detailed results of all of the algorithms.

3.3 Graphical Comparison of the Imitation Performance Between ASD and TD Groups Based on the *x, y*, and *z* Positions

In order to have a better sense of the subjects' performances in the games as well as doing the comparisons, the readers may be interested to observe graphical pictures of the participants' body skeletons. Therefore, instead of the introduced angles, we also take a look at the *x*, *y*, and *z* positions of the body joints in this subsection which was used in Mahalonobis distance sub-method. To this end, in the figures based on the spatial positions of the participants' joints, the position of the point between the two ankles of the user is set as the origin of the coordinate system, i.e. (*0, 0, 0*). Therefore, the effect of the initial distance between the users and the Kinect sensor vanishes. Moreover, for height normalization, the height of each user is set to 1 by dividing the *x*, *y*, and *z* coordination values to the user's real height. Here, it is assumed that the "ratio of body lengths to the subject's height" is the same for all participants, which is also not a flawless but acceptable assumption. Figure 7 shows the normalized spatial data of the bodies for the ASD, TD, and Parent groups for action #1 simultaneously in four different frames and time instances (after time normalization). In these figures, the blue, red, and green ellipses represent the spatial wrist distributions of the ASD, TD, and Parent groups, respectively. Regarding Fig. 7, the impairment of the children with autism in doing the imitation tasks is obvious compared to their TD peers and the Parents (for example, see Fig. 7c), which is in line with the results of the manual scoring by the evaluators. Similar results have also been observed in the other actions. In Figs. 8 and 9, we have presented the data without aligning the signals, called the situation of real unchanged or unwarped time, to present a better simultaneous sense of the performances' accuracy and agility which is described in the next subsection. As an example, Fig. 8 also shows the y-position of the left wrist versus time for each participant in action #1. The

Table 8 The correlation coefficient and error's mean and standard deviation of the proposed algorithms for the automated assessment of the ASD Group's imitation performance for the situations with the highest correlation coefficient

Action no.	Sub-method	Features	Correlation coefficient	p value for correlation factor	Mean of error (out of 4)	Std of error (out of 4)
1	Acceptable bound	X	0.72	0.00	1.21	0.85
1	Acceptable bound	X, V, and A	0.59	0.00	1.22	0.90
2	Cross-correlation of signals	X, V, and A	0.44	0.04	1.18	0.71
2	Acceptable bound	X, V, and A	0.35	0.11	1.19	0.80
3	Acceptable bound	X, V, and A	0.59	0.00	1.02	0.64
4	Acceptable bound	X, V, and A	0.63	0.00	0.94	0.71
5	Cross-correlation of signals	X and V	0.15	0.53	1.55	0.95
5	Acceptable bound	X, V, and A	0.11	0.64	1.03	0.72
6	Acceptable bound	X	0.49	0.05	1.11	0.65
6	Acceptable bound	X, V, and A	0.25	0.34	1.05	0.64
7	Acceptable bound	X, V, and A	0.51	0.02	0.77	0.57
8	Acceptable bound	X, V, and A	0.56	0.01	1.49	0.69
9	Mahalanobis distance	X	0.50	0.07	1.27	0.87
9	Acceptable bound	X, V, and A	0.36	0.21	0.75	0.54
10	HMM	X, V, and A	0.30	0.18	1.09	1.11
10	Acceptable bound	X, V, and A	0.24	0.29	1.41	1.00

solid lines show the mean values of the participants' performance while the blue regions represent of the mean value plus or minus one standard deviation (Std) of each group's performance. As we mentioned in the State-Image Acceptable Bound sub section, we chose the TD subjects with full mark's $Mean \pm 2 \times Std$ as the acceptable boundary for each movement. It should be noted that assuming normal distributions for a group of users' performances in such actions would lead us to expect that about 68% of that group's distribution would fall in the shaded area, which is $Mean \pm 1 \times Std$. Equivalent to Fig. 8d, the participant's left shoulder pitch angle versus time for action #1 is also presented in Fig. 9 in two situations: unchanged time and normalized time, which definitely gives us a less visual sense of the quality of the action or the subjects' faults; however, valuable for the robotic system. To avoid prolonging the text, the graphs for the rest features and actions are not presented in the manuscript.

3.4 Agility

According to our observations in this study, children with autism showed a deficit in the time duration of completing each action in comparison to their TD peers. To have a criterion for studying the agility, we have considered the signals of the typically developing participants who received a full score, i.e. 4 out of 4, in each action as the ground truth or reference data for assessing the agility in that action. There-

fore, we can define an agility factor, which is the ratio of the time duration of each participant to the mean time of the reference data for each action. Figure 10 shows the box plot of the agility factor for the subjects with autism and typically developing children in different imitation tasks. Figure 10a and b could show us the preliminary differences between the ASD and TD groups' time duration of the actions. In Fig. 10a, values less than 1 mean that the child with ASD "finished or left" the action sooner than the reference time extracted from the TD group with a full score for that action. It is clear that in studying agility, there is no need to align the dimension of the signals before the analysis; hence, the real captured signals have been used for plotting Fig. 10. As it can be seen in Fig. 10a, the median of the agility factor regarding the participants with ASD's performance is less than 1 for all of the actions in this study. A detail observation would lead us to find that in five tasks, including actions #2, #5, #6, #7, and #9, the agility factor's median is less than 0.9. This means that the subjects with ASD in this study finished (rarer) or incompletely left (usually) the imitation tasks sooner than their TD peers. This observation brought up some interesting unanswered questions in the area of autistic children' imitations performance that should be studied deeper: "Do children with autism receive the inputs in the same way as the TD subjects but are unable process or perform it as well? Or is their poorer performance in gross imitation tasks because of their brains' inability to receive the input signals appro-
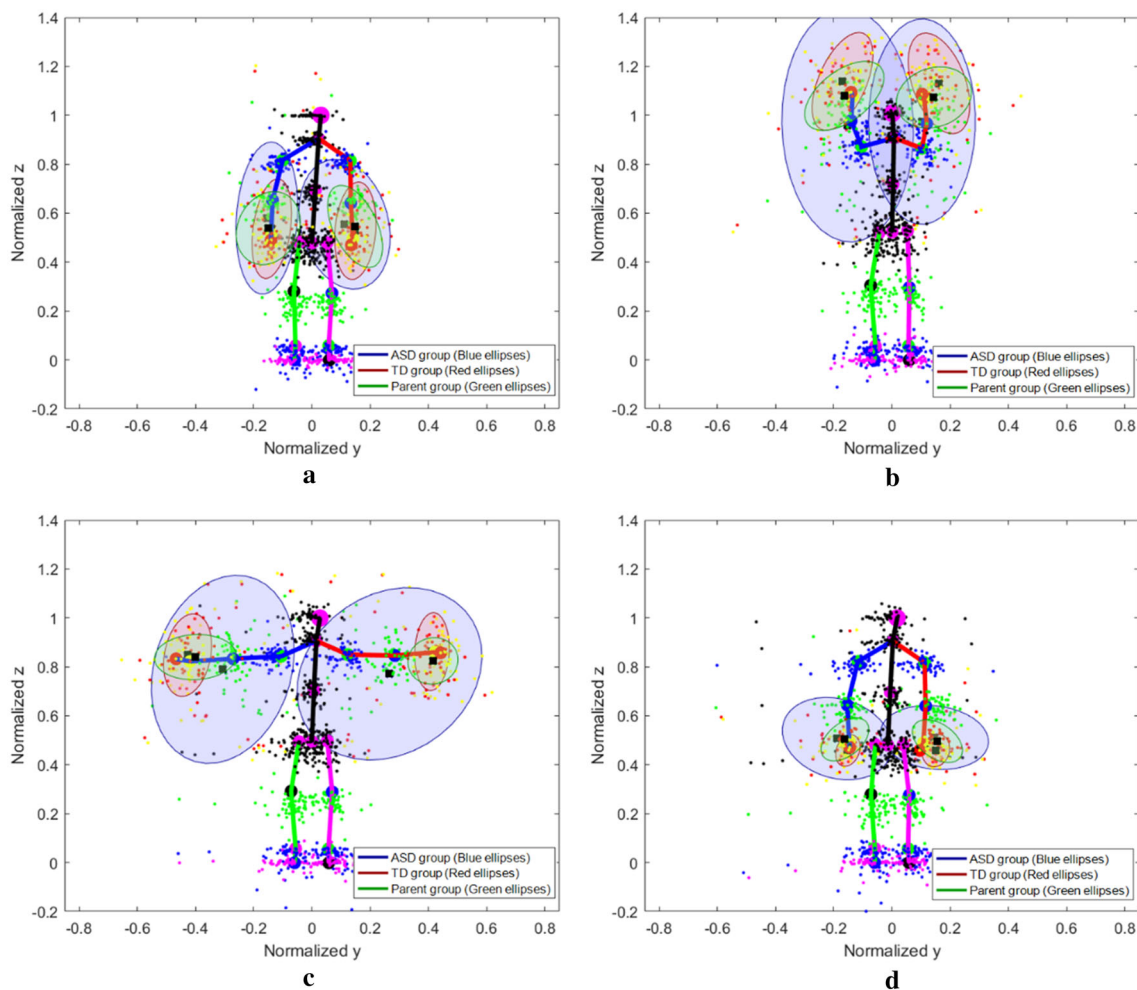
Fig. 7 The graph of wrists' spatial distributions for ASD (blue ellipses), TD (red ellipses), and Parent groups (green ellipses) in action #1 after time normalization at four instants: **a** frame 5, **b** frame 25, **c** frame 40, and **d** frame 50

priately from the beginning?". Considering that the agility factor has been calculated based on the mean time of the TD subjects with the full grade in different imitation actions, the distributions of the equivalent boxes for all the TD participants in comparison to the ASD group have compact ranges and the median are much more closer to 1.

Children with ASD may have other disabilities or deficits alongside their autism which may have possible effects on their agility and imitation performances. For example, a review on the subjects' clinical records showed that two of the ASD participants in this research had Mental Retardation and ADHD, respectively, alongside their autism. However, as a limitation, we did not consider this issue in our study.

3.5 Results Summary

As a summary of Sects. 3.1–3.4, the important points of the different groups' performances are as follows: (1) Confirmation of the first hypothesis by seeing a significant difference in

the ASD group's performance in comparison with the TD and Parent groups both in manual and automatic scoring, which shows the acceptable positive correlation between the manual and automated scores as well as the appropriate features' selection as the main contribution in automated assessment section. The mentioned differences are somewhat visible in Figs. 3, 7, 8, and 9. (2) No significant difference between the TD and Parent groups' imitation performances, and (3) No significant difference between the parents of the ASD group and parents of the TD group in the games. It should be noted that the second and third findings were not discussed in details in the text in order to avoid distracting the readers from the main contributions of the paper.

Although the current results could be improved significantly by improving the number of data set's signals as well as the algorithms such as using Deep Neural Networks (DNN) or Recurrent Neural Networks (RNN), the required time and resources for performing such an automated assessment are quite affordable. Using machines in assessment of the qual-
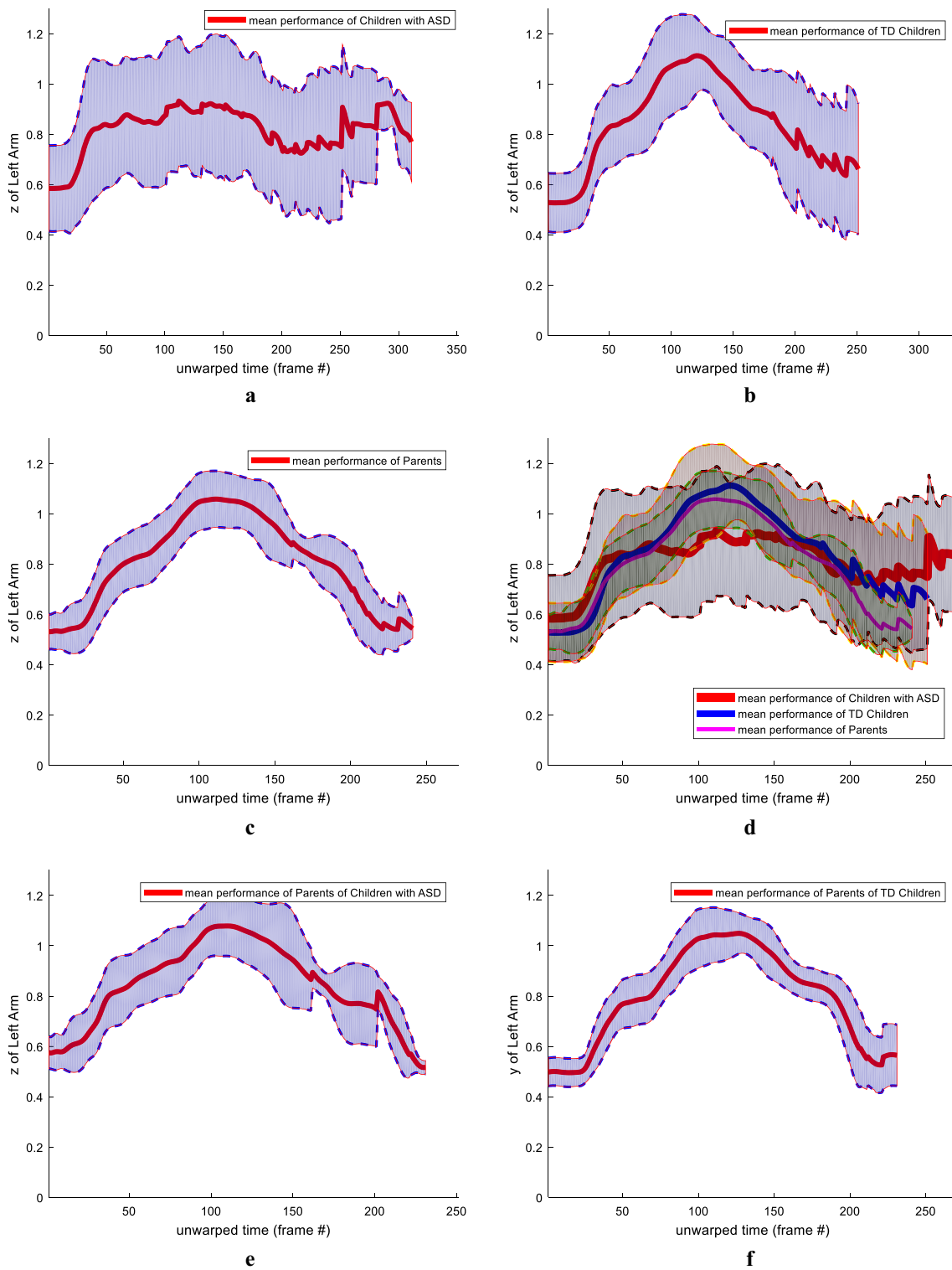
Fig. 8 The mean (solid line) and the standard deviation around the mean values (dashed line) of the participants' wrist y-position versus time in action #1: **a** ASD group, **b** TD group, **c** Parent group, **d** comparison of the three groups, **e** parents of children with autism, and **f** parents of typically developing children (data captured at 30 frames per second)
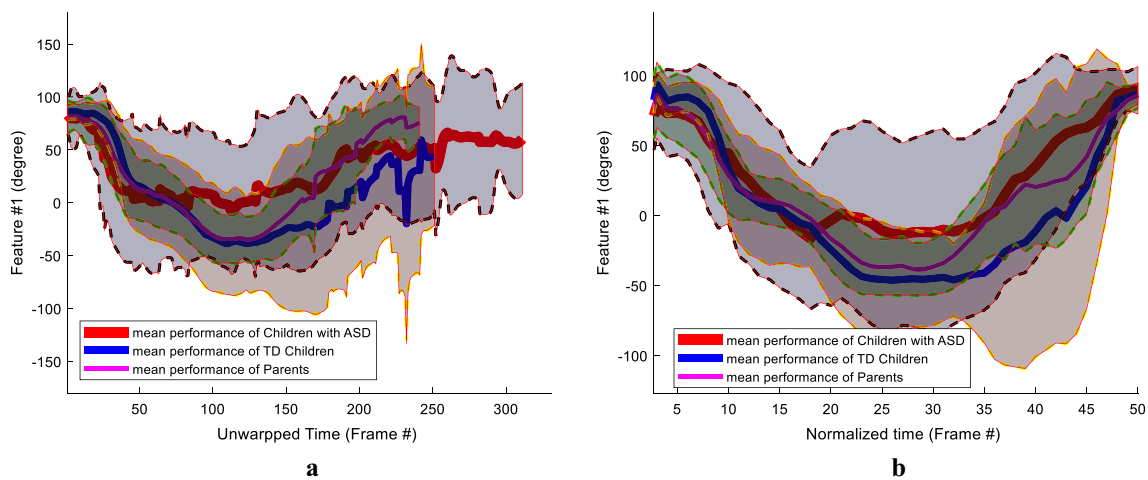
Fig. 9 The mean (solid line) and the standard deviation around the mean values (dashed line) of the participants' left shoulder pitch angle (ASD, TD, and Parents groups) versus time in action #1 (data captured at 30 frames per second), **a** the results before normalizing the time and **b** the results after normalizing the time

ity of actions can definitely be useful as they can analyze some data characteristics such as velocity and acceleration, parameters that are not easily measurable by human therapists. Being able to easily conduct the situations at home and send the data to the psychologists, the proposed method can also decentralize the children's performance in stressful environments. In this study, we proposed an automated assessment protocol which considers participants' range, speed, and accuracy of the motions. The proposed system is able to continuously assess the quality of actions instead of being limited to reporting the results descriptively and/or qualitatively. Moreover, the proposed measure enables reduction in potential evaluator biases and comparison between similar studies.

A clear message in using social robots in autism area is that most studies have tried to conduct some form of imitation therapy [37]. Statements such as spontaneous imitation of children with autism during HRI [1] have probably motivated the researchers to study and report whether or not their participants with ASD perform better in robot-assisted mode than human-mediator modes during the imitation tasks. As a summary, we have found a variety of works regarding this issue such as no significant difference between the two modes [14], better performance of children with ASD in robot-assisted mode in imitating facial expressions [38], and better performance in human mediator-assisted mode [13]. There are also some studies that report the advantages of both modes such as [35]. In [35], the authors engaged four children with ASD and divided them into two groups in which one of them interact with a robot and the other with a teacher. They observed higher performance in imitation of body movements of familiar actions for the teacher-assisted group and better performance of imitation in smile facial expressions

for the robot-assisted group. All in all, it should be noted that in ours as well as the mentioned studies, the sample size is a serious limitation of these studies and the non-homogenous group of participants regarding autism severity, age, and gender. This limitation makes it difficult to generalize the findings and observations as well as make any strong claims for confirming hypothesis-II.

4 Application of the Study: Proposing an Architecture of Human–Robot Reciprocal Interaction Platform

Using the results from this study, we present a human–robot reciprocal interaction platform for gross imitation with an evaluation system as a treatment application of a robotic system with the capability of automated imitation assessment in the autism area (Fig. 11). It should be noted that reciprocal imitation programs, a well-known treatment protocol worldwide, have been used to promote social responsiveness, shared attention, etc. in individuals with ASD [13, 39]. Therefore, knowing that improvement in gross imitation and motor skills of children with autism can positively affect their social and cognitive skills in real life and have possible cognitive rehabilitation purposes [1, 2, 6, 13, 39], reciprocal interaction with a social robot may be an appropriate treatment application of these researches. To design a reciprocal gross imitation training human–robot Interaction (HRI) for children with ASD, we have considered two different modes: the Non-Structured and Structured interaction modes. While some robotic-based cognitive architectures for autism treatment have been proposed in the literature [16, 26], we did
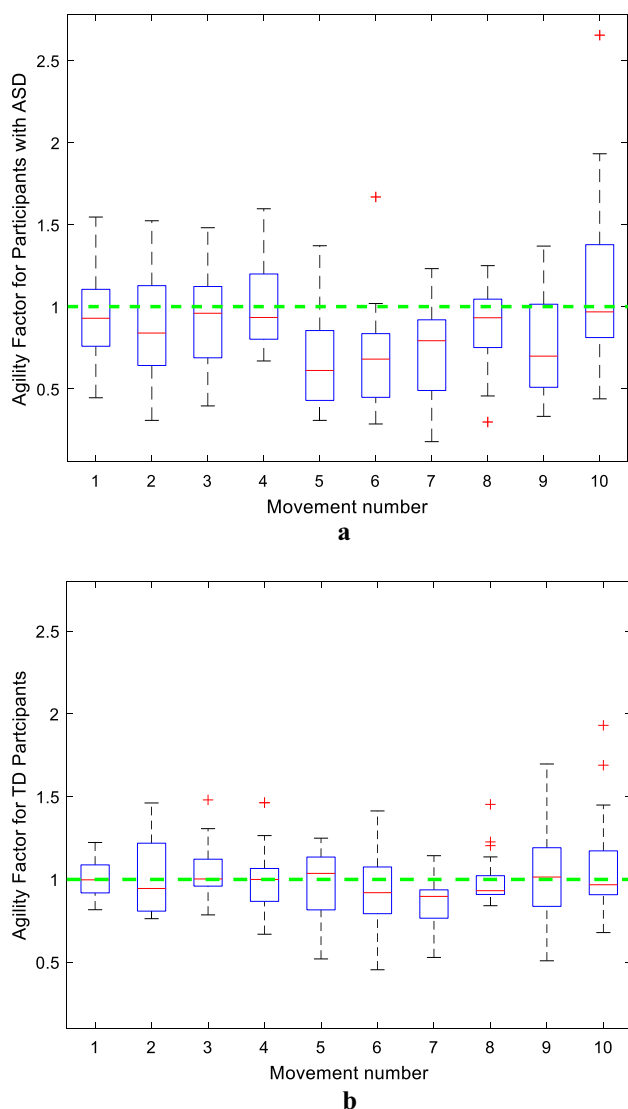
Fig. 10 The box plot for the agility factor of **a** children with autism's performance, and **b** all TD subjects' performance; in comparison to the TD peers with the full grade in different imitation actions

not see a systematic way to assess the quality of imitation actions in dynamic movements.

Imitation of individual with autism's behaviors is an effective way to trigger their attention toward human-mediators [1, 2, 40, 41]. As mentioned in autism literature, conducting imitation and turn-taking games and tasks could be an appropriate way to promote social interactions, creativity, sense of self, leadership, etc. in individuals with ASD [1, 6, 38, 41]. Therefore, in the Non-Structured interaction mode our aim is to empower the proposed platform to have the robot perform real-time gross imitation of the subjects, which is a required capability for a reciprocal HRI platform to be efficient for conducting reciprocal imitation programs. To this end, the kinematics data of the participants' skeletons is captured and used to reproduce their movements in the robot's

upper body using appropriate forward and inverse kinematics formulations. In this mode, the robot imitates and reacts to arbitrary actions of the subjects including their arms and head movements. Such Non-Structured interaction modes have the potential to estimate the acceptability rate of social robots for children as well. The details of how to empower the robot to imitate the upper body movements of the users is presented in the Appendix. It should be noted that in the proposed architecture, Kinect can be replaced with any RGB-D sensor or camera system that provides human skeleton tracking.

On the other hand, in the Structured interaction mode, the automated assessment system based on the State-Image or HMM algorithms introduced and discussed in the previous subsections of the article has been embedded in the proposed HRI architecture, this enables us to assess the subjects' gross imitation performance in imitating the robot's exercises during the child-robot interaction. This mode is the other requirement for reciprocal platforms to be used in imitation training programs' content.

What we have presented in this paper, including the participants' recruitment, conducting the imitation tasks, introducing the automated evaluation algorithm, doing the manual and automated assessments, and finally, the comparison of the results, was an initial attempt to observe and study the performance of the Structured interaction mode of the proposed architecture, which was shown to provide quite acceptable outputs.

Alternatively, manual assessment of gross imitations by specialists is the most reliable and accurate way to score the performance of children with autism; however, it is time-consuming and a very hard task. Therefore, due to the observed performance of our proposed algorithm so far, we can conclude that HRI platforms with an automated assessment capability and real-time evaluation system could handle the process of imitation assessment with admissible accuracy and have the potential to be an efficient assistant to therapists in treatment centers.

5 Limitations and Future Work

Enriching the datasets by increasing the number of participants and considering more effective parameters, such as the homogeneity in age, gender, and autism severity, would definitely lead to more accurate mathematical models of the children's behavior, and therefore a more efficient HRI platform. Currently, the mentioned non-homogeneities in this pilot study make it somewhat impossible to make generalized claims regarding the mathematical modeling accuracy. Having a more considerable number of signals in our dataset could definitely help the researchers to go through the powerful algorithms for dynamic assessment such as Deep Neural Networks (DNN) or Recurrent Neural Networks (RNN). Sig-
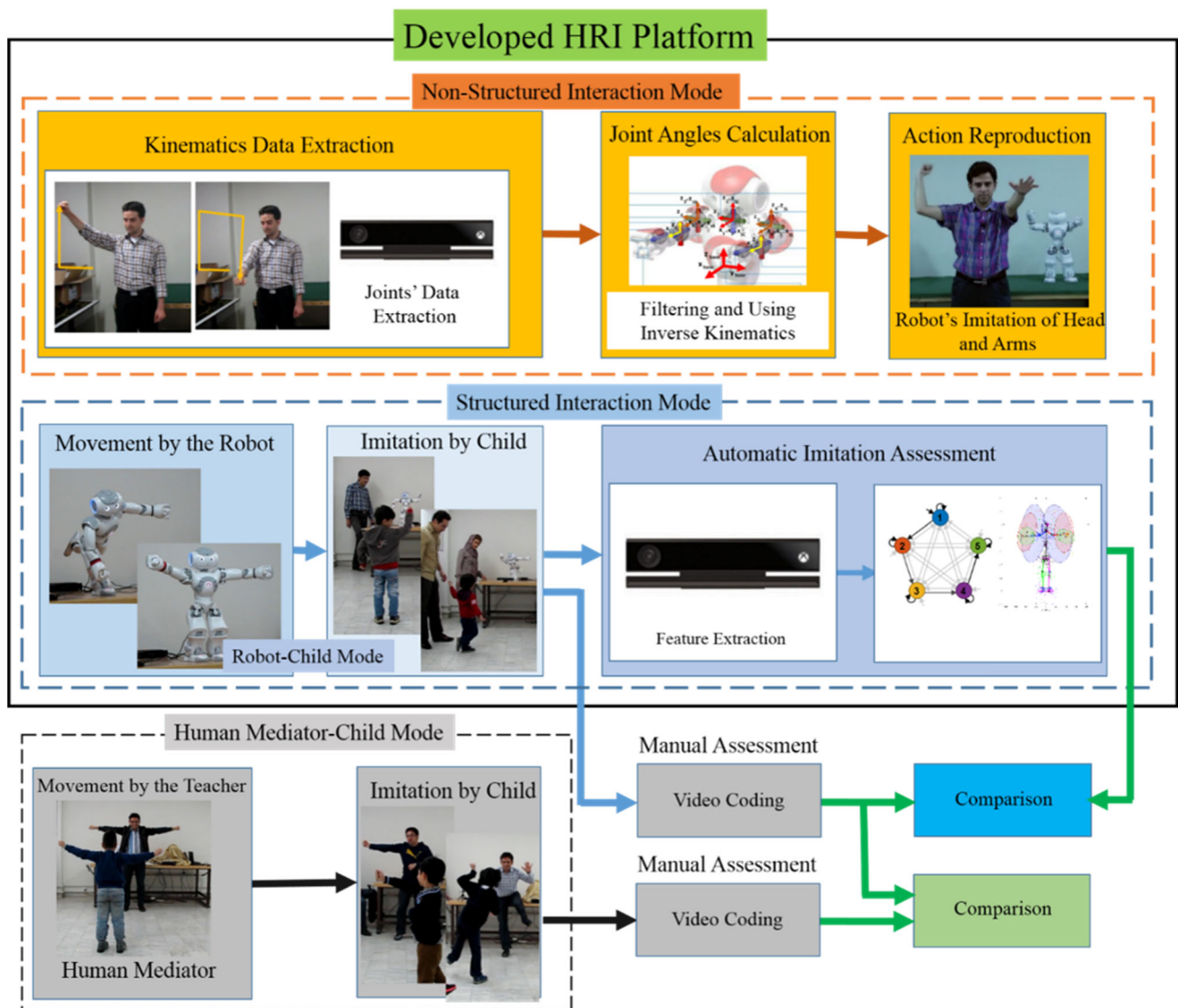
Fig. 11 The proposed architecture of a human–robot reciprocal interaction platform for gross imitations with the evaluation system

nals' alignment can also performed by using the Dynamic Time Warping (DTW) algorithm instead of what we did in this study.

One of our limitations was accessing an adequate number of cooperative female participants with autism which is fairly common due to the small population rate of girls to boys with autism spectrum disorders. Also, more studies could be performed with diverse robots to investigate the possible effects of their appearance, gender, interaction capabilities, etc. on the acceptability rate and children's performance in imitation tasks. Moreover, unfortunately, we did not conduct a mode *D* in this study that contains imitation of teacher by both child and parent simultaneously alongside the mode *C*. A mode *D* experiment could help us to have a more comprehensive understanding and comparison between the performances of the children within and between the robot-child and teacher-

child modes. In addition, in some cases like Fig. 4e in the robot-child-parent mode, the parents were sometimes looking at the participant, probably trying to get their child to imitate or encourage the subject to get involve in the games. As another limitation, it was sometimes hard in this case to identify who (robot or the parent) the participant is really imitating; therefore, this limitation could be considered as a possible interfering factor that should be carefully addressed in next studies.

Based on the observed preliminary acceptance rate of the human–robot interaction alongside the qualitative findings, our next step is to develop and run a set of imitation-based intervention protocols for children with autism. Alternatively, we recommend other researchers to conduct the proposed reciprocal HRI platform with a larger number of participants and compare the results to regular reciprocal

Fig. 12 The Nima robot's kinematic chains and upper body joints
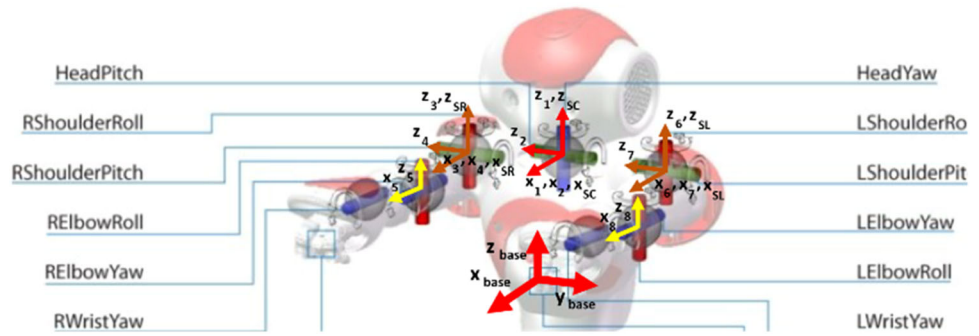
imitation training programs to explore the advantages and disadvantages of the developed architecture. Moreover, to avoid entering extra variables such as possible effect on children's cooperation to the study with the aim of comprising the typical treatment protocols in autism centers, there is no need to have parents in the game room during the intervention sessions.

In the Non-Structured mode of the current version of the proposed architecture, the robot can only imitate the upper body movements including the head and arm movements of the users. Considering the balance conditions for the robots alongside the use of forward and inverse kinematics calculations, our future aim is to empower robots to imitate the feet and waist's bending and rotation of the participants as well.

6 Conclusion

As a preliminary estimation of the robot's acceptability, we observed that 70% of the participants with ASD showed a tendency to imitate Nima's actions. In this study, one of the common hypotheses spread in the literature: "participants with ASD usually show better performance in robotic than non-robotic gross imitation tasks" was investigated for both ASD and TD groups by comparing the scores of the children's performance in mode A and mode B; it is indicated that the mentioned hypothesis is not confirmed in our experimental conditions. In addition, the presented State-Image and HMM-based automated imitation assessment algorithms were tested and we observed that the State-Image Acceptable Bound method's outputs were fairly comparable to the related works out of the autism area in assessing the quality of dynamic actions with the Pearson correlation of $\sim 45\%$ regarding our captured data. Moreover, we tried to follow a scientific approach to propose an HRI platform to be used in reciprocal imitation training programs with the potential for use in social and cognitive rehabilitation for children with ASD. According to the observed acceptable accuracy of the proposed automated assessment algorithm, we concluded that imitation assessment could be handled by human–robot interaction platforms. In a follow up of our past

Table 9 Denavit–Hartenberg parameters for robot's right and left arms and head

	$\alpha_{i-1}(rad)1$	a_{i-1}	θ_i	d_i
1	0	0	θ_1	0
2	$\frac{\pi}{2}$	0	$-\theta_2$	0
3	$\frac{\pi}{2}$	0	θ_4	0
4	α_{RS-RE}	$L_{Right\ Arm}$	θ_5	0
5	$\frac{\pi}{2}$	0	θ_7	0
6	α_{LS-LE}	$L_{Left\ Arm}$	θ_8	0

study [13], the efforts made in this research put us one step further in reaching the ultimate goal of empowering reciprocal human–robot interaction platforms to imitate whole body movements, facial expressions, voice, and even children's activities with objects and toys.

7 Appendix

Figure 12 shows Nima's base frame, zero positions of the upper body joints of the robot, and the attached frames. After recording the user's skeletal data with the Microsoft Kinect Sensor, using the presented Denavit–Hartenberg (DH) parameters of the robot in Table 9 would lead to the calculation of eight joints of Nima's upper body including Head Yaw, Head Pitch, and Shoulder Pitch, Shoulder Roll, and Elbow Pitch of the arms. Therefore, the robot is empowered to perform admissible real time position tracking of the participants' upper body joints. It should be noted that the Microsoft Kinect Sensor sampling rate, i.e. ~ 30 frames per second, is quiet adequate for admissible imitation.

where α_{LS-LE} and α_{RS-RE} are the angles between the z direction of Left and Right Shoulder and Elbow, respec-

Table 10 Inverse kinematics for the head, right hand and left Hand of NAO

Head yaw	$\theta_1 = atan2(H\_SC_y, H\_SC_x)$
Head pitch	$\theta_2 = atan2(\sqrt{H\_SC_x{}^2 + H\_SC_y{}^2}, H\_SC_z)$
Right shoulder pitch	$\theta_3 = atan2(-E\_SR_z, E\_SR_x)$
Right shoulder roll	$\theta_4 = asin\left(\frac{E\_SR_y}{L_{RArm}}\right)$
Right elbow pitch	$\theta_5 = acos\left(\frac{\vec{R}_{RArm}.\vec{R}_{RForehand}}{\left\|\vec{R}_{RArm}\right\|.\left\|\vec{R}_{RForehand}\right\|}\right)$
Left shoulder pitch	$\theta_6 = atan2(-E\_SL_z, E\_SL_x)$
Left shoulder roll	$\theta_7 = asin\left(\frac{E\_SL_y}{L_{LArm}}\right)$
Left elbow pitch	$\theta_8 = -acos\left(\frac{\vec{R}_{LArm}.\vec{R}_{LForehand}}{\left\|\vec{R}_{LArm}\right\|.\left\|\vec{R}_{LForehand}\right\|}\right)$

tively. $L_{Right\ Arm}$ and $L_{Left\ Arm}$ are the length of the arms. Equation 3 shows the homogenous transformation matrix.

$$
{}_i^{i-1}T = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) & 0 & a_{i-1} \\ \sin(\theta_i).\cos(\alpha_{i-1}) & \cos(\theta_i).\cos(\alpha_{i-1}) & -\sin(\alpha_{i-1}) & -\sin(\alpha_{i-1}).d_i \\ \sin(\theta_i).\sin(\alpha_{i-1}) & \cos(\theta_i).\sin(\alpha_{i-1}) & \cos(\alpha_{i-1}) & \cos(\alpha_{i-1}).d \\ 0 & 0 & 0 & 1 \end{bmatrix}
\tag{3}
$$

Combining the DH transformation matrices of the mentioned joints (Eqs. 4–10) makes it possible to calculate Nima's head, left wrist, and right wrist positions.

$$
{}_{Head}^{Shoulder\ Center}T = {}_1^{Shoulder\ Center}T.{}_2^1T
\tag{4}
$$

$$
{}_{Elbow\ Right}^{Shoulder\ Right}T = Rot_y(\theta_3){}_3^{Shoulder\ Right}T
\tag{5}
$$

$$
{}_{Wrist\ Right}^{Shoulder\ Right}T = {}_{Elbow\ Right}^{Shoulder\ Right}T.{}_4^3T
\tag{6}
$$

$$
{}_{Elbow\ Left}^{Shoulder\ Left}T = Rot_y(\theta_6){}_5^{Shoulder\ Left}T
\tag{7}
$$

$$
{}_{Wrist\ Left}^{Shoulder\ Left}T = {}_{Elbow\ Left}^{Shoulder\ Left}T.{}_6^5T
\tag{8}
$$

$$
{}_{Wrist\ Right}^{Shoulder\ Right}P = {}_{Wrist\ Right}^{Shoulder\ Right}T.\left[L_{Right\ Forehead}, 0, 0, 1\right]^T
\tag{9}
$$

$$
{}_{Wrist\ Left}^{Shoulder\ Left}P = {}_{Wrist\ Left}^{Shoulder\ Left}T.\left[L_{Left\ Forehead}, 0, 0, 1\right]^T
\tag{10}
$$

Table 10 also shows the inverse kinematics formulations for the mentioned joints of the robot. We experimentally found that a fraction of the maximum speed of Nima's actuators could be set to 0.8 in order to avoid facing high jerk and high imitation time delays.

Compliance with Ethical Standards

Conflict of interest Authors Alireza Taheri and Ali Meghdari has received research grants from "Sharif University of Technology" (Grant No. G980517) and the "Cognitive Sciences and Technology Council" (CSTC) of Iran (Grant No. 95p22), respectively. The author Mohammad H. Mahoor declares that he has no conflict of interest.

Ethical Approval Ethical approval for the protocol of this study was provided by the Iran University of Medical Sciences (#IR.IUMS.REC.1395.95301469).

References

1. Scassellati B, Admoni H, Matarić M (2012) Robots for use in autism research. Annu Rev Biomed Eng 14:275–294
2. Taheri AR, Alemi M, Meghdari A, PourEtemad HR, Basiri NM (2014) Social robots as assistants for autism therapy in Iran: research in progress. In: 2014 second RSI/ISM international conference on robotics and mechatronics (ICRoM). IEEE, pp 760–766
3. Ingersoll B (2008) The social role of imitation in autism: implications for the treatment of imitation deficits. Infants Young Child 21(2):107–119
4. Carpenter M, Pennington BF, Rogers SJ (2002) Interrelations among social-cognitive skills in young children with autism. J Autism Dev Disord 32(2):91–106
5. Uzgiris I (1999) Imitation as activity: its developmental aspects. In: Nadel J, Butterworth G (eds) Imitation in infancy, ch. 6. Cambridge University, Cambridge, England, pp 186–206, ISBN: ISBN-0-521-58033-1. https://eric.ed.gov/?id=ED433967
6. Cabibihan JJ, Javed H, Ang M, Aljunied SM (2013) Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. Int J Soc Robot 5(4):593–618
7. Taheri A, Meghdari A, Alemi M, Pouretemad H (2018) Human–robot interaction in autism treatment: a case study on three pairs of autistic children as twins, siblings, and classmates. Int J Social Robot 10(1):93–113
8. Kajopoulos J, Wong AHY, Yuen AWC, Dung TA, Kee TY, Wykowska A (2015) Robot-assisted training of joint attention skills in children diagnosed with autism. In: International conference on social robotics. Springer, Cham, pp 296–305
9. Shukla J, Cristiano J, Oliver J, Puig D (2019) Robot assisted interventions for individuals with intellectual disabilities: impact on users and caregivers. Int J Soc Robot 11(4):631–649
10. Ismail LI, Verhoeven T, Dambre J, Wyffels F (2019) Leveraging robotics research for children with autism: a review. Int J Soc Robot 11(3):389–410
11. Diehl JJ, Schmitt LM, Villano M, Crowell CR (2012) The clinical use of robots for individuals with autism spectrum disorders: a critical review. Res Autism Spectr Disord 6(1):249–262
12. Pennisi P, Tonacci A, Tartarisco G, Billeci L, Ruta L, Gangemi S, Pioggia G (2016) Autism and social robotics: a systematic review. Autism Res 9(2):165–183

13. Pour AG, Taheri A, Alemi M, Meghdari A (2018) Human–robot facial expression reciprocal interaction platform: case studies on children with autism. Int J Soc Robot 10(2):179–198

14. Costa AP, Charpiot L, Lera FR, Ziafati P, Nazarikhorram A, van der Torre L, Steffgen G (2018) A comparison between a person and a robot in the attention, imitation, and repetitive and stereotypical behaviors of children with autism spectrum disorder. In: Proceedings workshop on social human–robot interaction of human-care service robots at HRI2018

15. So WC, Wong MY, Cabibihan JJ, Lam CY, Chan RY, Qian HH (2016) Using robot animation to promote gestural skills in children with autism spectrum disorders. J Comput Assist Learn 32(6):632–646

16. Ali S, Mehmood F, Dancey D, Ayaz Y, Khan MJ, Naseer N et al (2019) An adaptive multi-robot therapy for improving joint attention and imitation of ASD children. IEEE Access 7:81808–81825

17. Aly A (2020) Human posture recognition and gesture imitation with a humanoid robot. arXiv:2002.01779

18. Fujimoto I, Matsumoto T, De Silva PRS, Kobayashi M, Higashi M (2011) Mimicking and evaluating human motion to improve the imitation skill of children with autism through a robot. Int J Soc Robot 3(4):349–357

19. Pirsiavash H, Vondrick C, Torralba A (2014) Assessing the quality of actions. In: European conference on computer vision. Springer, Cham, pp 556–571

20. Zariffa J, Kapadia N, Kramer JL, Taylor P, Alizadeh-Meghrazi M, Zivanovic V et al (2011) Relationship between clinical assessments of function and measurements from an upper-limb robotic rehabilitation device in cervical spinal cord injury. IEEE Trans Neural Syst Rehabil Eng 20(3):341–350

21. Daunoravičienė K, Žižienė J, Pauk J, Idzkowski A, Raudonytė I, Juocevičius A et al (2017) Stroke-affected upper extremity movement assessment via continuous relative phase analysis. Measurement 110:84–89

22. Bosecker C, Dipietro L, Volpe B, Igo Krebs H (2010) Kinematic robot-based evaluation scales and clinical counterparts to measure upper limb motor performance in patients with chronic stroke. Neurorehabil Neural Repair 24(1):62–69

23. Butler EE, Ladd AL, LaMont LE, Rose J (2010) Temporal–spatial parameters of the upper limb during a reach & grasp cycle for children. Gait Posture 32(3):301–306

24. Zhou L, Bai S, Hansen MR, Rasmussen J (2011) Modeling of human arm energy expenditure for predicting energy optimal trajectories. Model Identif Control 31(3):91–101. https://doi.org/10.4173/mic.2011.3.1

25. Balasubramanian S, Melendez-Calderon A, Roby-Brami A, Burdet E (2015) On the analysis of movement smoothness. J Neuroeng Rehabil 12(1):112

26. Feng Y, Jia Q, Wei W (2018) A control architecture of robot-assisted intervention for children with autism spectrum disorders. J Robot 2018:3246708. https://doi.org/10.1155/2018/3246708

27. Gordon AS (1995) Automated video assessment of human performance. In: Proceedings of AI-ED, pp 16–19

28. Perše M, Kristan M, Perš J, Kovačič S (2007) Automatic evaluation of organized basketball activity using bayesian networks. na

29. https://www.softbankrobotics.com/emea/en/nao (2020)

30. Dong Y (2018) An application of Deep Neural Networks to the in-flight parameter identification for detection and characterization of aircraft icing. Aerosp Sci Technol 77:34–49

31. Dong Y (2019) Implementing deep learning for comprehensive aircraft icing and actuator/sensor fault detection/identification. Eng Appl Artif Intell 83:28–44

32. Rozo L, Silverio J, Calinon S, Caldwell DG (2016) Learning controllers for reactive and proactive behaviors in human–robot collaboration. Front Robot AI 3:30

33. Calinon S, Guenter F, Billard A (2007) On learning, representing, and generalizing a task in a humanoid robot. IEEE Trans Syst Man Cybern Part B (Cybern) 37(2):286–298

34. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77(2):257–286

35. Duquette A, Michaud F, Mercier H (2008) Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. Auton Robots 24(2):147–157

36. Giannopulu I, Montreynaud V, Watanabe T (2014) PEKOPPA: a minimalistic toy robot to analyse a listener-speaker situation in neurotypical and autistic children aged 6 years. In: Proceedings of the second international conference on human-agent interaction, pp 9–16

37. Ricks DJ, Colton MB (2010) Trends and considerations in robot-assisted autism therapy. In: 2010 IEEE international conference on robotics and automation. IEEE, pp 4354–4359

38. Pioggia G, Igliozzi R, Sica ML, Ferro M, Muratori F, Ahluwalia A, De Rossi D (2008) Exploring emotional and imitational android-based interactions in autistic spectrum disorders. J Cyber Ther Rehabil 1(1):49–61

39. Ingersoll B (2010) Brief report: pilot randomized controlled trial of reciprocal imitation training for teaching elicited and spontaneous imitation to children with autism. J Autism Dev Disord 40(9):1154–1160

40. Taheri A, Meghdari A, Alemi M, Pouretemad H (2018) Human–robot interaction in autism treatment: a case study on three pairs of autistic children as twins, siblings, and classmates. Int J Soc Robot 10(1):93–113

41. Taheri A, Meghdari A, Alemi M, Pouretemad H (2019) Teaching music to children with autism: a social robotics challenge. Sci Iran 26(Special Issue on: Socio-Cognitive Engineering):40–58

Alireza Taheri is an Assistant Professor of Mechanical Engineering with an emphasis on Social and Cognitive Robotics at Sharif University of Technology, Tehran, Iran. He is a Research Associate at the Social and Cognitive Robotics Lab at Sharif University of Technology.

Ali Meghdari is a Professor of Mechanical Engineering and Robotics at Sharif University of Technology (SUT) in Tehran. Professor Meghdari has performed extensive research in various areas of robotics; social and cognitive robotics, mechatronics, bio-robotics, and modeling of biomechanical systems. He has been the recipient of various scholarships and awards, the latest being: the 2012 Allameh Tabatabaei distinguished professorship award by the National Elites Foundation of Iran (BMN), the 2001 Mechanical Engineering Distinguished Professorship Award from the Ministry of Science, Research & Technology (MSRT) in Iran, and the 1997 ISESCO Award in Technology from Morocco. He is currently the Director of the Centre of Excellence in Design, Robotics, and Automation (CEDRA), an affiliate member of the Iranian Academy of Sciences (IAS), a Fellow of the American Society of Mechanical Engineers (ASME), and the Founder and Chancellor of Islamic Azad University-Fereshtegaan International Branch (for students with special needs; primarily the Deaf and Blind).

Mohammad H. Mahoor received the B.Sc. degree in electrical engineering from Abadan Institute of Technology with the highest honor in 1996 and the M.Sc. degree in biomedical engineering from Sharif University of technology, Tehran in 1998. He joined the department of Electrical and Computer Engineering at the University of Miami, Florida in 2003 where he received his Ph.D. in Nov. 2007. As an outstanding graduate student, he was granted the Eliahu & Joyce Jury Award in 2007 at the University of Miami. He then worked as a Post-Doctoral fellow at the University of Miami where he received the Marino Autism Research Institute award for his research on facial expressions analysis of children with Autism. Dr. Mahoor joined the University of Denver, Colorado in Fall 2008 where he is currently Professor of Electrical and Computer Engineering.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;

2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;

3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;

4. use bots or other automated methods to access the content or redirect messages

5. override any security feature or exclusionary protocol; or

6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com