



Article

Fostering Emotion Recognition in Children with Autism Spectrum Disorder

Vinícius Silva ¹, Filomena Soares ^{1,*} , João Sena Esteves ¹ , Cristina P. Santos ²  and Ana Paula Pereira ³ 

¹ Centro Algoritmi, Campus of Azurém, University of Minho, 4800-058 Guimaraes, Portugal; a65312@alunos.uminho.pt (V.S.); sena@dei.uminho.pt (J.S.E.)

² CMEMS R&D, Center for Microelectromechanical Systems, University of Minho, 4800-058 Guimaraes, Portugal; cristina@dei.uminho.pt

³ Education Research Center, Institute of Education, University of Minho, 4710-057 Braga, Portugal; appereira@ie.uminho.pt

* Correspondence: fsoares@dei.uminho.pt

Abstract: Facial expressions are of utmost importance in social interactions, allowing communicative prompts for a speaking turn and feedback. Nevertheless, not all have the ability to express themselves socially and emotionally in verbal and non-verbal communication. In particular, individuals with Autism Spectrum Disorder (ASD) are characterized by impairments in social communication, repetitive patterns of behaviour, and restricted activities or interests. In the literature, the use of robotic tools is reported to promote social interaction with children with ASD. The main goal of this work is to develop a system capable of automatic detecting emotions through facial expressions and interfacing them with a robotic platform (Zeno R50 Robokind[®] robotic platform, named ZECA) in order to allow social interaction with children with ASD. ZECA was used as a mediator in social communication activities. The experimental setup and methodology for a real-time facial expression (happiness, sadness, anger, surprise, fear, and neutral) recognition system was based on the Intel[®] RealSense[™] 3D sensor and on facial features extraction and multiclass Support Vector Machine classifier. The results obtained allowed to infer that the proposed system is adequate in support sessions with children with ASD, giving a strong indication that it may be used in fostering emotion recognition and imitation skills.

Keywords: Human Computer Interaction; Zeno R50 Robot; Support Vector Machine; facial expressions; emotions recognition; Autism Spectrum Disorder



Citation: Silva, V.; Soares, F.; Esteves, J.S.; Santos, C.P.; Pereira, A.P. Fostering Emotion Recognition in Children with Autism Spectrum Disorder. *Multimodal Technol. Interact.* **2021**, *5*, 57. <https://doi.org/10.3390/mti5100057>

Academic Editor: Jerry Alan Fails

Received: 31 August 2021

Accepted: 17 September 2021

Published: 22 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Emotional states are usually reflected in facial expressions. Expressing and reading facial expressions is an effective way of social interaction and communication. However, some individuals have severe difficulties in recognizing facial communication clues. It is the case of individuals with Autism Spectrum Disorder (ASD) [1].

Nowadays, assistive robotics is focused on helping users with special needs in their daily activities. Assistive robots are designed to identify, measure, and react to social behaviours [2]. They can be a social support to motivate children and socially educate them. According to studies [3], it was observed that children with ASD can exhibit certain positive social behaviours when interacting with robots in contrast to what is perceived when interacting with social partners (peers, caregivers, and professionals). Furthermore, few worldwide projects pursue to include robots as part of the intervention program for individuals with ASD [4,5]. These studies have demonstrated that robots can promote a high degree of motivation and engagement in subjects, even in those who are reluctant to socially interact with professionals [5,6]. Studies also point out that individuals with ASD tend to show a preference for robot-like features over

non-robotics toys [7,8], and in some situations, respond faster when prompted by a robotic movement than by a human movement.

A humanoid robot can then be a useful tool to develop social–emotional skills of children with ASD, due to the engagement and positive learning outcome [9]. Additionally, robotic systems offer certain advantages when compared with virtual agents, likely due to the capability of robotic systems to use physical motion in a manner not possible in screen technologies [10].

As previously referred, successful human–human communication relies on the ability to read affective and emotional signals, but robotic systems are emotionally blind. The research that has been conducted in the field of Affective Computing tries to endow robotic systems with the capability of reading emotional signals, dynamically adapting the robot behaviour during interaction [11,12]. Different technological strategies have been used to try to mitigate the emotion recognition impairments that usually individuals with ASD present, mainly through the use of assistive robots capable of synthesising “affective behaviours” [9,13,14].

Following this idea, in the present work a humanoid robotic platform capable of expressing emotions is used as a mediator in social communication with children with ASD. Thus, the main goals of this work are (a) the development of a system capable of automatically detect five emotions (happiness, sadness, anger, surprise, fear) and neutral through facial cues and (b) interfacing the developed system with a robotic platform, allowing social communication with children with ASD. In order to achieve the mentioned goals, the proposed experimental setup uses the Intel[®] RealSense[™] 3D sensor and the Zeno R50 RoboKind[®] robotic platform. This layout uses a Support Vector Machine (SVM) technique to automatically classify the five emotions plus neutral expressed by the user in real time.

In the literature, few works have been devoted to the recognition of emotional expressions in games with children with ASD using robotic tools that are capable of synthesizing facial expressions. Furthermore, some systems proposed in the literature are controlled using the Wizard-of-Oz (WOZ) setup, meaning that the robot does not prompt an autonomous behaviour accordingly to the children’s actions [8,15,16]. The research presented in this article tries to tackle this gap. The robot is used as a mediator in social communication activities, endorsing emotion recognitions and imitation skills in an autonomous way.

In order to assess the system in a support context, an exploratory study was conducted involving 6 children with ASD during 7 sessions. The results obtained allowed to infer that the proposed system is able to interact with children with ASD in a comfortable and natural way, giving a strong indication that this system may be a suitable and relevant tool in the context of emotion recognitions and imitation skills with this target group.

This article is organized as follows: in Section 2 research regarding affective computing and emotion recognition impairment of children with ASD is addressed, and an overview of the topics on affective sensing and humanoid robots that are capable of expressing facial expressions are presented. Section 3 presents the overall system and the experimental methodology followed. The results and their discussion are presented in Section 4. Section 5 finalizes the article, with conclusions and future work.

2. Background

The affective behaviour displayed by humans is multi-modal, subtle, and complex. Humans use affective information such as facial expressions, eye gaze, various hand gestures, head motion, and posture to deduce the emotional state of each other [17]. The research on facial expression conducted by Paul Ekman [18] demonstrated the universality and discreteness of emotions by enumerating the basic emotions (happiness, sadness, anger, surprise, disgust, and fear), and creating the Facial Action Coding System (FACS) [19]. Head pose and eye gaze, together with facial expressions, are very important to convey emotional states. Head nods help emphasizing an idea during a conversation. In addition,

it also helps to agree or disagree with a point of view through nodding gestures such as ‘yes’ or ‘no’ head movements, synchronizing the interactional rhythm of the conversation [20]. Eye gaze is important for analysing attentiveness, competence, as well as intensity of emotions. They are processed together, when analysing human emotional states, from a computational point of view [21].

Although the process of imitating, recognizing, and displaying emotions can be an easy task for the majority of humans, it is a very difficult task for individuals with ASD [22]. Individuals with ASD are characterized by displaying repetitive patterns of behaviour, for having restricted activities or interests, and impairments in social communication. Furthermore, these individuals have difficulties in recognizing body language, making eye contact, understanding other people’s emotions, and lack of social or emotional reciprocity [23]. These difficulties in interpreting social situations in general cause children with ASD to lose or miss information on what is happening or happened during the social exchange [24].

Technological tools, such as assistive robots [5,6], have been employed in support sessions with children with ASD. Some of the robots used have a humanoid appearance [15]. Although researchers [25–27] have used a variety of facially expressive robots in their works, few have devoted their attention to the recognition of emotional expressions in games with children with ASD in an autonomously way. In fact, one of the fields of application of HRI (Human Robot Interaction) is ASD research, where social robots help users with special needs in their daily activities [2]. The following paragraphs summarize some of the relevant developed works involving a humanoid robot capable of displaying facial expression interacting with children with ASD.

FACE [28,29] is a female android built to allow children with ASD to deal with expressive and emotional information. The system was tested with five children with ASD and fifteen typically developing children. The evaluated emotions were the six basic emotions (happiness, sadness, anger, fear, disgust, and surprise). The results demonstrated that happiness, sadness, and anger were correctly labelled with high accuracy for both children with ASD and typically developing children. Conversely, fear, disgust, and surprise were not labelled correctly, particularly by participants with ASD. The overall recognition rate for FACE with children with ASD was 60.0%, and the recognition results for each emotion were the following: anger—100%, disgust—20%, fear—0%, happiness—100%, sadness 100%, surprise—40%. The results for FACE recognition rates with typically developing children were: anger—93%, disgust—20%, fear—46.7%, happiness—93.3%, sadness—86.7%, surprise—40%, and the average of all emotions was 61.1%.

ZECA [30], Zeno Engaging Children with Autism, is a humanoid robot from Robokind® (Zeno R50) that is used in the *Robótica-Autismo* research project at University of Minho (roboticaautismo.com). It seeks to use robotic platforms to improve social skills of individuals with ASD. ZECA was employed in a study with the purpose of analysing the use of a humanoid robot as a tool to teach emotions recognition and labelling. In order to evaluate the designed facial expressions, two experiments were conducted. In the first one, the system was tested by forty-two typically developing children aged between 8 and 10 years old (group A) that watched videos of ZECA performing the following facial expressions: neutral, surprise, sadness, happiness, fear, and anger. Then, sixty-one adults aged between 18 and 59 years old (group B) watched the same videos. Both groups completed a questionnaire that consisted in selecting the most appropriate correspondence for each video. The recognition rates of the facial expressions for group A were the following: anger—26.2%, fear—45.2%, happiness—83.3%, neutral—85.7%, sadness—97.6%, surprise—76.2%, and the average of all emotions was 69.0%. The recognition rates of the facial expressions for group B were the following: anger—24.6%, fear—77.0%, happiness—91.8%, neutral—90.2%, sadness—91.8%, surprise—86.6%, and the average of all emotions was 77.0%. The second experiment consisted of showing similar videos of ZECA performing the same facial expression, but now with gestures. The recognition rates of the facial expressions

improved in general, but with more impact for these two emotions: fear (73.8%) and anger (47.6%). Similar to group A, the recognition rates in group B, in general, also improved. The recognition rates of the facial expressions, adding gestures, for group B were the following: anger—70.5%, fear—93.4%, happiness—98.4%, neutral—91.8%, sadness—88.5%, surprise—83.6%, and the average of all emotions was 77.0%.

More recently, there has been a concern in developing more autonomous approaches [16,31,32] to interact with children with ASD.

Leo et al. [31] developed a system that automatically detects and tracks the child's face and then recognizes emotions on the basis of a machine learning pipeline based on Histogram of Oriented Gradients (HOG) descriptor and SVM. They used the Zeno R25 robot from Robokind[®] as mediator in activities concerning imitation of facial expressions. The system was evaluated by conducting two different experimental sessions: the first one tested the system using the CK + dataset; the second one involved 3 children with ASD in a preliminary exploratory session where 4 different expressions were investigated (happiness, sadness, anger, and fear). Considering the results obtained from the first experimental session, the following average accuracies for each facial expression were obtained: anger—88.6%, disgust—89.0%, fear—100%, happiness—100%, sadness—100%, and surprise—97.4%; achieving an average accuracy of 95.8%. From the results concerning the second session with the children with ASD, the authors concluded that the system can be effectively used to monitor the children's behaviours.

Chevalier et al. [16] developed an activity for facial expression imitation whereby the robot imitates the child's face to encourage the child to notice facial expressions in a play-based game. The proposed game consisted of using the Zeno R25 robot from Robokind[®], which is capable of displaying facial expressions, in a mirroring game where initially the robot imitates some of the child's facial cues and then the child gradually imitates the robot's facial expression. A usability study was conducted with 15 typically developing children aged between 4 and 6 years old. The authors concluded that, in general, the last step of the activity where the child imitates the robot facial expression was challenging, i.e., some children had difficulties focusing on the robot's face. Overall, the authors considered that the outcomes from the usability study were positive and believe that the target group, children with ASD, may benefit from it.

A multimodal and multilevel approach is proposed by Palestra et al. [32] where the robot acts as a social mediator, trying to elicit specific behaviours in children, taking into account their multimodal signals. The social robot used in this research was the Robokind[®] Zeno R25 humanoid robot that is capable of expressing human-like facial expressions. The system is composed by four software modules: head pose, body pose, eye contact, and facial expression. At the present stage of their research, the authors evaluated only the facial expression module in a preliminary study involving three high functioning children with ASD within the ages of 8–13 years old during two sessions. The facial expressions tested were anger, fear, happiness, and sadness. Each facial expression was consecutively imitated 4 times by the children. The authors evaluated the number of facial expressions that have been correctly imitated, the time needed to have eye contact, and the time needed to imitate the facial expression. The results obtained showed that the time of eye contact between session 1 and 2 decreased as well as the time needed to imitate the facial expression. Additionally, the success imitation rate for each facial expression increased in general from 20% in session 1 to 51.7% in session 2. Thus, the authors concluded that the robot can successfully play a mediator role in support sessions with children with ASD.

These works propose approaches that focus on how to increase the robot's autonomy. However, these systems do not take into account the head motion and the eye gaze as features for the classifier, but both play an important role in expressing affect and communicating social signals [21]. Thus, besides assessing the system performance in terms of metrics (e.g., accuracy), the present work presents an exploratory study involving six children with ASD during seven sessions with the goal of fostering facial expression

recognition skills, providing a more natural interaction by introducing some autonomy to the system.

3. Materials and Methods

This section presents the developed system to recognize the six facial expressions considered: ‘Happiness’, ‘Sadness’, ‘Anger’, ‘Surprise’, ‘Fear’, and ‘Neutral’. The experimental procedure and the game scenarios, the method to extract the facial features, as well as the database construction are also detailed.

3.1. Proposed System

The system implemented in this work consists of an Intel® RealSense™ sensor model F200, a computer, and the ZECA robot (Figure 1).



Figure 1. Experimental setup. Starting from the left: Intel® RealSense™ (model F200), computer and ZECA robot.

Intel® RealSense™ is a device for implementing gesture-based Human Computer Interaction (HCI) techniques manufactured by Intel® USA [33]. It contains a conventional RGB camera, an infrared laser projector, an infrared camera, and a microphone array. A grid is projected onto the scene by the infrared projector and the infrared camera records it, computing the depth information. The microphone array allows localizing sound sources in space and performing background noise cancellation. This device, along with the required software, the Intel® RealSense™ Software Development Kit (SDK), was used to obtain the face data from the user. This sensor was chosen mainly because of its small size, which is an advantage when conducting the final experiments in a school setting.

Zeno R50 a humanoid child-like robot manufactured by RoboKind® Texas USA (Figures 1 and 2) was used in the present work. ZECA, a common Portuguese name, is the acronym of *Zeno Engaging Children with Autism*. This robotic platform has 34 degrees of freedom: 4 are located in each arm, 6 in each leg, 11 in the head, and 1 in the waist [34]. The major feature that distinguishes Zeno R50 from other robots is the ability to express emotions thanks to servo motors mounted on its face and a special material, *Frubber*, which looks and feels like human skin.

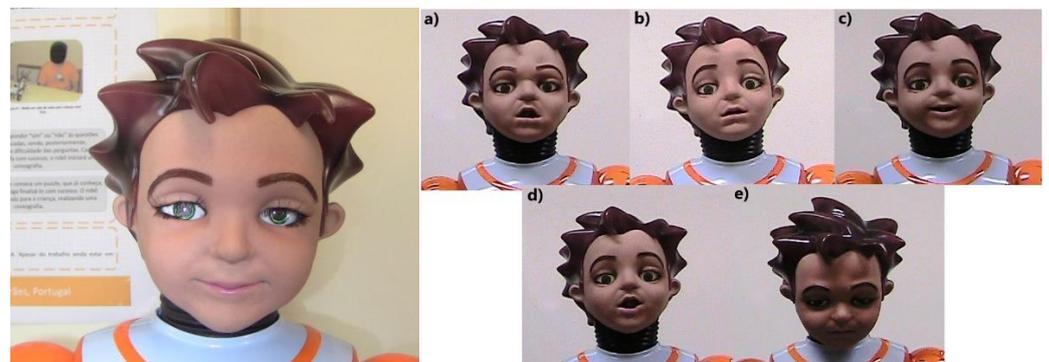


Figure 2. On the left: ZECA, a Zeno R50 RoboKind® humanoid robot. On the right: Facial expressions displayed by ZECA: (a) anger; (b) fear; (c) happiness; (d) surprise; (e) sadness [9].

3.2. Experimental Procedure and Game Scenarios

Children with ASD have difficulty in recognizing, imitating, and understanding emotional states [35]. In order to tackle these impairments, two activities were developed, and the experimental procedure was defined.

The activity starts with ZECA greeting the child and the experimenter and prompts the experimenter to select, in the developed interface, the activity that is going to be performed: IMITATION or EMOTIONS. Then, the activity starts and ZECA gives the instruction to the chosen game.

In the IMITATION activity, the robot first displays one of the five facial expressions. Then, the child is prompted to identify the emotion associated with the facial expression ('Happiness', 'Sadness', 'Anger', 'Surprise', 'Fear'). The child answers by exhibiting the same facial expression that was prompted by the robot.

In the activity EMOTIONS, the robot starts by asking the child to perform a facial expression. The child answers by mimicking the facial expression that was asked by ZECA.

In both game scenarios, ZECA verifies if the answer is correct and prompts a reinforcement accordingly to the correctness of the answer. The type of reinforcement that is given to the child is based on a previous study [9,15] and consists of a combination of verbal, movement, and sound reinforcements (as an example of reinforcement, the robot would say "Congratulations!" while waving its arms in the air). When the time is up, ZECA asks if the experimenter wants to continue. The experimenter can extend or stop the activity. If the experimenter decides to stop the activity, the session ends, with a robot farewell.

3.3. Facial Features Extraction

Typically, emotions can be characterized as negative (sadness, anger, or fear), positive (happiness or surprise), or neutral. FACS (Facial Action Coding System) is the earliest method for characterizing the physical expression of emotions. Facial muscles contract and stretch while mimicking emotions through facial expressions. FACS defines the movements of these individual facial muscles, called Action Units (AU) [19].

The Intel® RealSense™ 3D sensor was used to extract facial AUs as well as detecting up to 78 facial landmarks using the depth information. The landmarks position in the image space can be used in different ways to extract the shape of facial features and the movements of facial features, also called geometric features. The geometric features can be extracted on the variation in shape of the triangles, or ellipses (eccentricity features) [36]. Additionally, the Intel® RealSense™ SDK can return the user head angles, Euler angles (pitch, roll, and yaw) [37], allowing to obtain the user's head motion, which is an important feature in the emotion communication process.

Table 1 lists the significance of the selected 10 facial landmarks. Table 2 lists the facial AUs from Intel® RealSense™ that were used in this work, differentiating those provided by the Intel® RealSense™ SDK from the ones obtained through facial landmarks (geometric features).

Table 1. Facial landmarks significance.

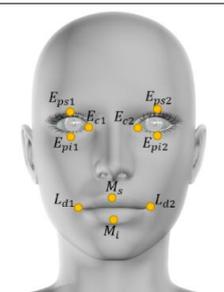
Facial Landmarks Location	Facial LandMark	Significance
	E_{c1}	Eye corner
	E_{c2}	Eye corner
	L_{d1}	Lip depressor
	L_{d2}	Lip depressor
	E_{ps1}	Eye palpebrale superius
	E_{ps2}	Eye palpebrale superius
	E_{pi1}	Eye palpebrale inferius
	E_{pi2}	Eye palpebrale inferius
	M_s	Mouth superius
	M_i	Mouth inferius

Table 2. Selected Facial AUs used in this work, provided by the Intel® RealSense™ SDK and the ones obtained through facial landmarks (geometric features).

Facial AUs	IRS ¹	FL ²	Geometric Feature Equation
Eye brow raiser (left and right)	✓		
Eye brow lower (left and right)	✓		
Eyelids		✓	$E_{ye_{right}} = E_{ps1y} - E_{pi1y}$ $E_{ye_{left}} = E_{ps2y} - E_{pi2y}$
Eyes up	✓		
Eyes down	✓		
Eyes left	✓		
Eyes right	✓		
Mouth open		✓	$MouthOpen = M_{sy} - M_{iy}$
Lip stretcher	✓		
Lip depressor (left and right)		✓	$Lip\ depressor_{right} = \sqrt{(E_{c1x} - L_{d1x})^2 + (E_{c1y} - L_{d1y})^2}$ $Lip\ depressor_{left} = \sqrt{(E_{c2x} - L_{d2x})^2 + (E_{c2y} - L_{d2y})^2}$

¹ Intel® RealSense™; ² Facial Landmarks.

The database used in this work was built using the 16 head features (face and neck), Table 2, acquired from the Intel® RealSense™ 3D sensor, and corresponding to the five emotions plus neutral.

A total of 43 participants (11 adults and 32 typically developing children) were considered for the database construction. The acquired features were normalized in a 0 to 100 intensity scale. Details of the implemented procedure can be found in [38].

4. Results and Discussion

The following section presents the results obtained with the proposed system in the recognition of the five emotions considered in this work, ‘Happiness’, ‘Sadness’, ‘Anger’, ‘Surprise’, ‘Fear’, and ‘Neutral’. In order to access the performance of the developed system different experimental evaluations were conducted.

Firstly, two SVM classifiers using the Linear and the non-linear Radial Basis Function (RBF) kernel were trained in order to recognize the six facial expressions: ‘Happiness’, ‘Sadness’, ‘Anger’, ‘Surprise’, ‘Fear’, and ‘Neutral’. The k-Fold Cross Validation method (k-Fold CV), where k = 10, was used to evaluate the classifier, as it contributes for the generalization of the classifier, avoiding overfitting. The following metrics—accuracy, sensitivity, specificity, Area Under the Curve (AUC), and the Mathews Correlation Coefficient (MCC)—were employed to evaluate each classifier performance. Then, an experimental study was conducted in a school environment with typically developing children. Finally, an exploratory study was performed with children with ASD.

It is worth mentioning that this study has the approval of the Ethics Committee of the University and Informed Consents from children’s parents or those responsible were obtained prior to the experiments.

4.1. Model Evaluation—Offline and Real-Time

The system was firstly evaluated offline, in a simulation environment using *MATLAB*, with the database created. Two multiclass SVM models were tested—the linear and the RBF kernels. Tables 3 and 4 show the comparison between both models in terms of accuracy, sensitivity, specificity, AUC, and MCC. By comparing the results, it is possible to conclude that the SVM model with the RBF kernel presents an overall superior performance when compared to the SVM model with the linear kernel. The accuracies per class increased in the RBF SVM model, especially the accuracy of the class ‘Fear’ (66% to 89%), Table 3. In

consequence, the overall accuracy of the SVM model with the RBF kernel also increased from 88.15% to 93.63%, which may indicate that the relation between class labels and attributes is nonlinear. The RBF model also outperformed the linear model in the other metrics, Table 4. Unlike the linear kernel, the RBF kernel can handle the case when the relation between class labels and attributes is nonlinear. Moreover, RBF has less hyper-parameters than other nonlinear kernels (e.g., polynomial kernel), which may decrease the complexity of model selection [39]. Additionally, RBF usually has lower computational complexity, which in turn improves real-time computational performance [40]. It is worth notice that the RBF kernel is widely used as a kernel function in emotion classification [41].

Table 3. Comparison of the accuracies per class between the SVM with a linear kernel and RBF kernel.

Class	Classifier	
	Linear	RBF
Happiness	99%	98%
Sadness	89%	94%
Anger	89%	95%
Surprise	94%	98%
Fear	66%	89%
Neutral	91%	90%

Table 4. Comparison of the overall performance between the SVM with a linear kernel and an RBF kernel.

Metric	Classifier	
	Linear	RBF
Accuracy	88.6%	93.6%
Sensitivity	87.2%	93.5%
Specificity	97.3%	98.5%
AUC	97.2%	99%
MCC	84.3%	92.2%

The work of Leo et al. [31], an approach that uses a conventional RGB camera, and the CK + database (a public dataset without children data) achieved an average accuracy of 94.3% for the six facial expressions. Despite using different experimental configurations, the present work achieved similar results of the state of art.

For real-time assessment, the proposed system was implemented and evaluated in a laboratorial environment with 14 adults (18–49 years old). The SVM model with RBF kernel implemented in the system was trained using the Accord Machine Learning C# library [42]. The participant sat in front of the sensor, looked at the Intel® RealSense™, and performed the emotion requested by the researcher.

Table 5 shows the recognition accuracy confusion matrix for the five emotions and neutral, with an overall accuracy of 88.3%. In general, the on-line system yields comparable results to that obtained in the off-line evaluation. ‘Happiness’ and ‘Sadness’ emotional states have accuracies over 90% and the other four facial expressions are consistently beyond 85%.

Concerning the real-time performance, the emotion recognition system was tested in a frame rate of 30 fps on i5 quad-core Central Processing Units (CPUs) with 16 GB RAM. The required time for the system to perform efficiently facial expression recognition is 1–3 ms, which means that the working frequencies achievable for sampling and processing are very high and do not compromise the real-time feature of the interaction process. The training computational cost of the system is approximately 1–2 s for the multi-class SVM classifier.

Table 5. Confusion matrix—Real-time system.

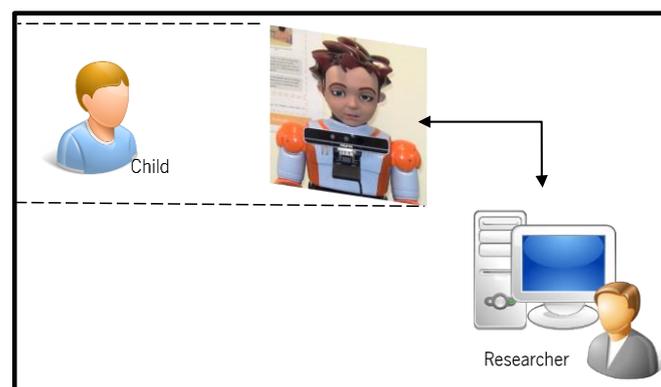
	Happiness	Sadness	Anger	Surprise	Fear	Neutral
Happiness	93%	0%	0%	0%	7%	0%
Sadness	0%	93%	0%	0%	0%	7%
Anger	7%	7%	86%	0%	0%	0%
Surprise	0%	7%	0%	86%	7%	0%
Fear	0%	14%	0%	0%	86%	0%
Neutral	0%	0%	7%	0%	7%	86%

The performance of the proposed system was compared to the results presented in [40]. This system was based on a *Kinect* sensor and used the *Bosphorus* database and SVM for facial expressions classification. The overall accuracy of the proposed system is 88% compared to 84% in [40]. Regarding the required time for the proposed system to perform facial emotion, recognition is 1–3 ms compared to 3–5 ms in [40].

4.2. Experimental Study with Typically Developing Children

This experimental phase was performed with typically developing children in a school environment. This study had two main goals: to test the system, with the two game scenarios, to detect the system constraints; to tune the conditions of the experimental scheme.

Following this trend, a set of preliminary experiments with the two game scenarios were carried out, involving 31 typically developing children aged between 6 and 9 years old. The facial expressions were asked (EMOTIONS scenario) or performed (IMITATE scenario) randomly by the robot. The experiments were performed individually in a triadic setup, i.e., child–ZECA–researcher. The robot had the role of mediator in the process of recognition and imitation of facial expressions. Figure 3 shows the experimental configuration used. Each child performed a two-to-three-minute session where the child had to perform five facial expressions—anger, fear, happiness, sadness, and surprise (one trial for each facial expression in each activity). The researcher oversaw the progress of the activity and monitored the system.

**Figure 3.** Experimental configuration: Child–ZECA–Researcher.

The quantitative behaviours analysed were the following: number of right and wrong answers and the children’s response time. The response time was counted from the time the robot gave the prompt to the time the child performed the correspondent facial expression.

Figure 4 shows the results of the IMITATE and EMOTIONS game scenarios obtained with 31 typically developing children. The results show that, in general, the system performed well in both activities. Both activities had similar high recognition rates in the classes ‘Happy’, ‘Sad’, and ‘Surprise’—87% vs. 88%, 90% vs. 97%, and 81% vs. 95%, respectively for each class and activity. However, ‘Fear’ and ‘Anger’ had the lowest recognition rates in the activity IMITATE (52% and 19%, respectively), when comparing to

the recognition rates of the same facial expressions in the activity EMOTIONS (81% and 58%, respectively).

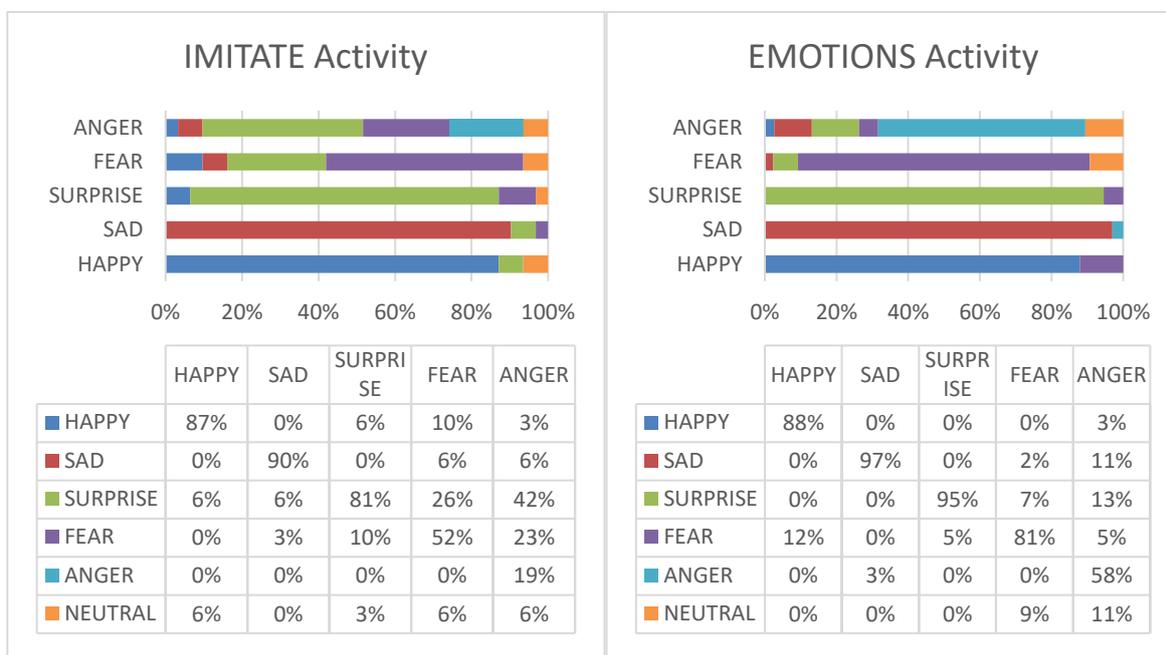


Figure 4. Results of the IMITATE and EMOTIONS game scenarios obtained with thirty-one typically developing children.

Table 6 presents the children’s mean response time, and standard deviation (SD) in each activity for each facial expression. In general, the children presented similar response times in both activities for the ‘Happiness’ and ‘Sadness’ facial expressions, which means that these emotions are the easiest recognizable facial expressions. The children response time corresponding to the ‘Surprise’ and ‘Fear’ facial expressions slightly decreased in the activity EMOTIONS, since the children only had to express the emotion requested by ZECA, without the need to recognize it. Concerning the response time of the children when displaying the ‘Anger’ expression, the response time decreased in the EMOTIONS activity. Additionally, the children answered faster to the prompt in the EMOTIONS activity in comparison to the response time in the IMITATE activity.

Table 6. Children’s mean response time in seconds per facial expression (SD) in the IMITATE and EMOTIONS activities.

Class	Activities	
	IMITATE	EMOTIONS
Happiness	19.47 (3.28)	19.46 (3.25)
Sadness	18.07 (2.63)	17.52 (2.34)
Anger	25.17 (1.08)	21.74 (3.54)
Surprised	19.72 (3.81)	17.56 (2.47)
Fear	21.57 (3.87)	19.84 (3.41)

By analysing the results, the low recognition rates of ‘Anger’ and ‘Fear’ in the activity IMITATE could probably be due to the fact that the children had to interpret the facial expression displayed by ZECA, perhaps meaning that they did not interpret well the facial expression, or the set of features contributing for the facial expression synthesized by ZECA was not well marked enough for the children to recognize. Moreover, these same facial expressions presented a higher recognition rate and were faster displayed by the children in the EMOTIONS activity. Additionally, in general,

the children took more time performing the ‘Anger’ expression in the activity IMITATE compared to the other facial expressions.

4.3. Exploratory Study with Children with ASD

This experimental phase (performed also in a school environment) had a twofold goal: to verify if the system can implement a procedure that makes the children able to interact in a comfortable and natural way and, on the other side, to evaluate the appropriateness of the system in a real environment with children with ASD. The main research question was: can the proposed system be used as an eligible tool in emotions recognition activities with children with ASD? This experimental study is crucial for the next steps in the research. In fact, only after concluding the study presented in this paper is it possible to proceed with further tests to infer the suitability of the proposed system as a complement to the traditional support sessions. The implementation of the proposed system in clinical support sessions must be performed with a larger sample, with a quantified recognition success rate and a quantified children evolution in terms of predefined behaviour indicators.

Following this trend, a set of preliminary experiments were carried out involving six children with ASD (high-functioning autism or Asperger’s syndrome) aged between 8 and 9, four boys and two girls. Based on the children’s skills and as recommended by the professionals, the original group of six was uniformly divided into two subsets of three children each (subset one, with children A, B, and C; subset two, with children D, E, and F). In subset one, three different facial expressions were investigated (anger, happiness, and sadness). On the other hand, in the subset two, five facial expressions were investigated (anger, fear, happiness, sadness, and surprised). The facial expressions were asked or performed randomly by the robot (EMOTIONS and IMITATE game scenarios, respectively). The experiments were performed individually in activities involving the professional, the robot, and the child. The robot had the role of mediator in the process of imitation and recognition of facial expressions. The professional only intervened if necessary to “regulate” a child’s behaviour. The researcher supervises the progress of the activity and monitors the system. Figure 5 shows the experimental configuration used where each child was placed in front of the robot. Seven sessions of two to three minutes each were performed.

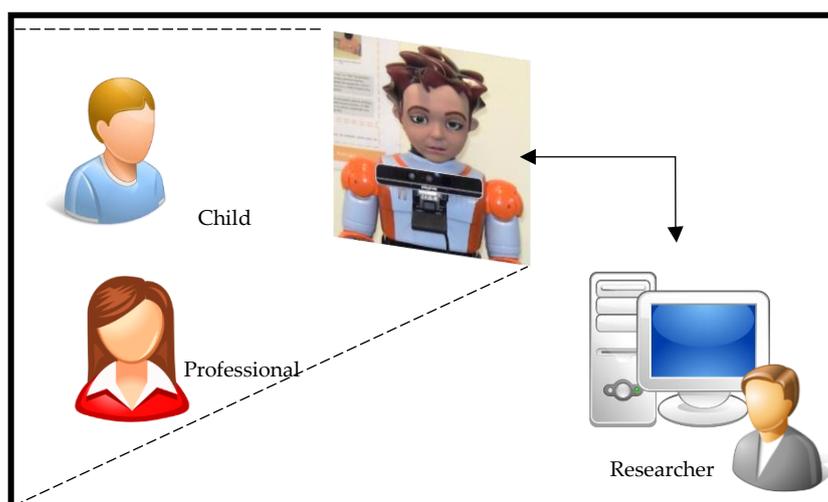


Figure 5. Experimental configuration: Child–ZECA–Professional–Researcher.

The quantitative behaviours analysed were the following: number of right and wrong answers and the child’s response time per session. The response time was counted from the time the robot gave the prompt to the time the child performed the correspondent facial expression.

4.3.1. Results from Subset One

Figures 6 and 7 show the results of the two game scenarios obtained with two of the three children (A and B) from the subset one. The results of child C were inconclusive as most sessions were unsuccessfully ended. The child did not perform as expected, since he was more attracted by the robot components or he was tired/annoyed, and consequently he was not focused on the activity.

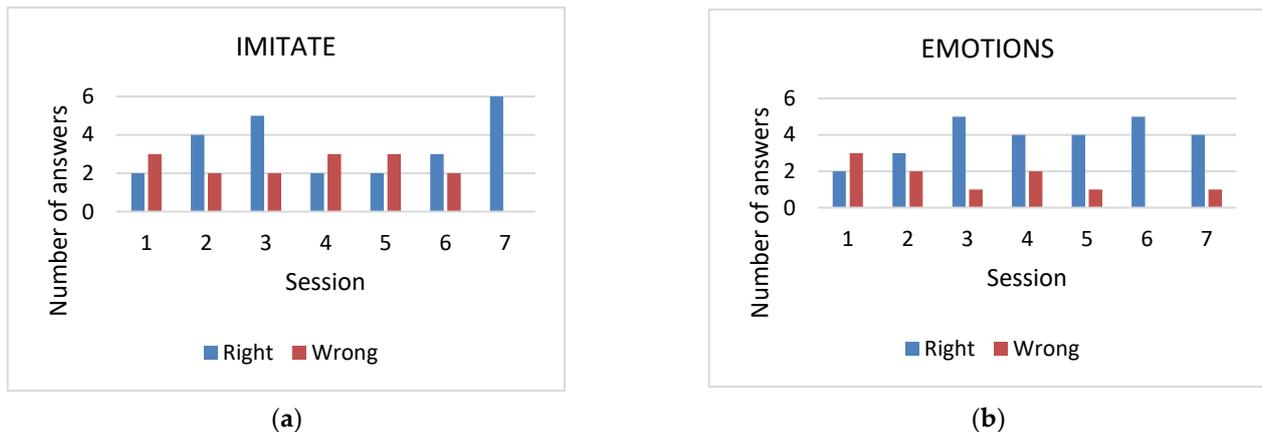


Figure 6. Child A results—On (a), the progress over seven sessions with the IMITATE activity. On (b), the progress over seven sessions with the EMOTIONS activity.

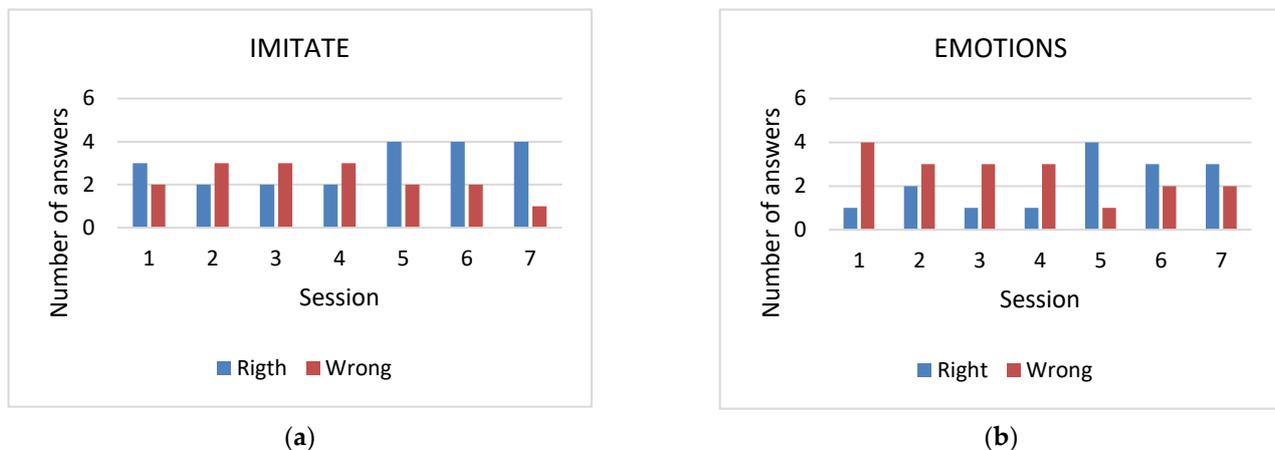


Figure 7. Child B results—On (a), the progress over seven sessions with the IMITATE activity. On (b), the progress over seven sessions with the EMOTIONS activity.

The results in Figure 6, on the left, show that in the first session child A gave more incorrect answers than correct ones, whereas child B gave slightly more correct answers in the same session (Figure 7 on the left). Then, in the following sessions the performance of child A slightly improved by having more correct answers than incorrect. However, the performance of child B slightly worsened, improving only in the last three sessions. Conversely, in session 4 and 5, the progress of the child A slightly worsened, by giving more incorrect answers. In the last session, both children had a good performance. It is possible to conclude that the overall performance of the child A in the IMITATE activity fluctuated with a good performance in the last session, whereas by analysing the results of the same activity for the child B there was a positive evolution. In the EMOTIONS activity, Figures 6 and 7 on the right, child A had a distinctly better performance than child B in the first sessions. Considering the last three sessions, the performance of child B improved, equalling up to the performance of the child A. It is possible to infer that both children had a positive evolution in the EMOTIONS activity. Some difference in performance is

consistent with the fact that the effects of ASD and the severity of symptoms differ from person to person [1].

Tables 7 and 8 present children's mean response time, and standard deviation (SD) of successful answers in the activities IMITATE and EMOTIONS, given in each session. Both participants took more time answering to the prompt in the last session, Session 7. Child A was usually faster to answer the prompt from the robot than child B. Additionally, it is possible to see that in the last three sessions on the EMOTIONS activity, where the participants took more time at performing the facial expression asked, the performance improved in both cases, having more correct answers than incorrect ones (Figures 6 and 7).

Table 7. Subset one: Children's mean response time in seconds for successful answers (SD) in the IMITATE activity. In general, the response time increased in the last session.

Session Number	Child A	Child B
1	16.54 (0.02)	19.64 (1.5)
2	18.16 (1.95)	18.09 (4.23)
3	17.88 (1.27)	17.45 (0.44)
4	19.66 (0.99)	16.54 (0.44)
5	18.53 (1.98)	17.06 (0.18)
6	18.65 (1.21)	18.79 (0.54)
7	19.63 (2.08)	18.99 (3.08)

Table 8. Subset one: Children's mean response time in seconds for successful answers (SD) in the EMOTIONS activity. In general, the response time increased in the last sessions.

Session Number	Child A	Child B
1	16.02 (0)	16.90 (0)
2	16.57 (0.64)	18.21 (2.17)
3	17.47 (1.62)	18.26 (0)
4	16.82 (1.0)	24.82 (0)
5	20.49 (4.27)	20.68 (2.67)
6	21.63 (1.68)	21.07 (2.06)
7	20.77 (1.99)	20.00 (1.33)

4.3.2. Results from Subset Two

Figures 8–10 show the results of the two game scenarios obtained with the three children (D, E, and F) from the subset two.

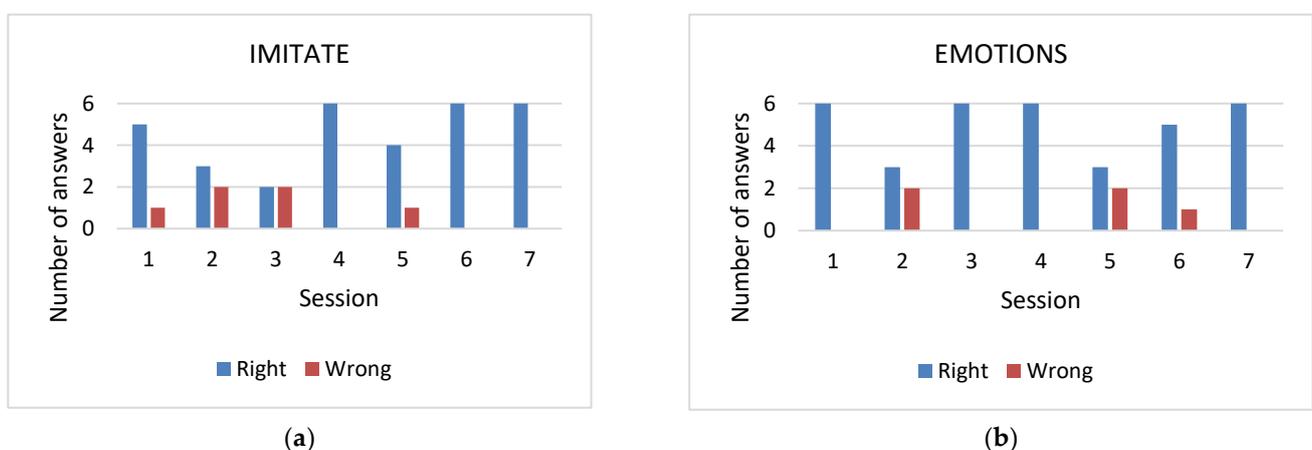


Figure 8. Child D results—On (a), the progress over seven sessions with the IMITATE activity. On (b), the progress over seven sessions with the EMOTIONS activity.

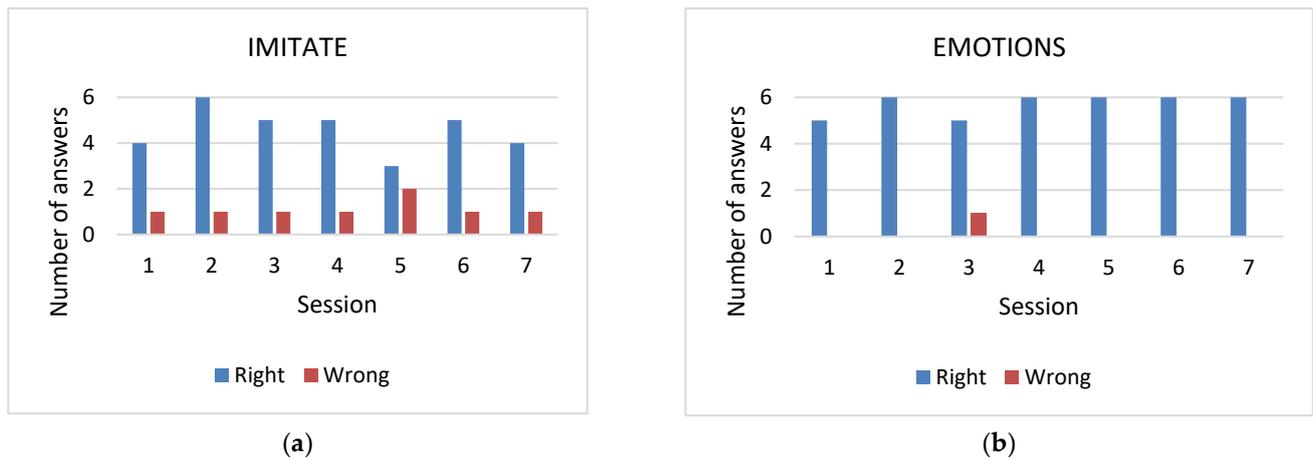


Figure 9. Child E results—On (a), the progress over seven sessions with the IMITATE activity. On (b), the progress over seven sessions with the EMOTIONS activity.

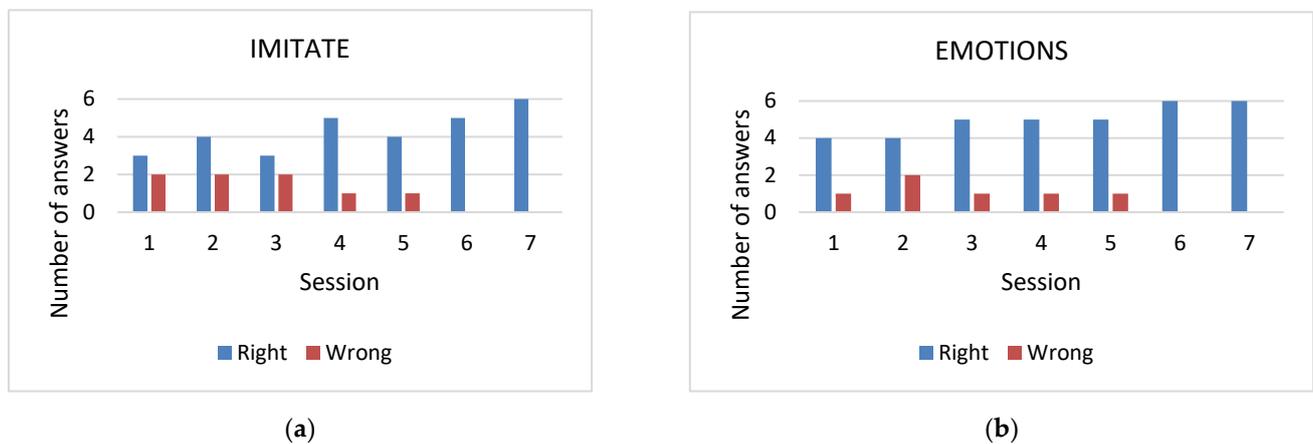


Figure 10. Child F results—On (a), the progress over seven sessions with the IMITATE activity. On (b), the progress over seven sessions with the EMOTIONS activity.

By analysing the performance of the three children in the activity IMITATE, children D and F had a positive evolution, whereas the performance of the child E fluctuated, having an overall good performance. In the EMOTIONS activity, the three children had, in general, a good performance over the sessions. In particular, child F had a more notable positive evolution. This child improved his performance in displaying the anger facial expression, since until the last two sessions he did not correctly perform the anger expression.

Tables 9 and 10 present the children's mean response time, and standard deviation (SD), in the activities IMITATE and EMOTIONS, of successful answers given in each session. All children took more time answering to the prompt in the last session, Session 7. Child D was usually faster to answer the prompt from the robot than his partners.

Regarding the qualitative analysis, the children's first reaction to the robot in the first session was positive: they were interested in the face of the robot, touching it repeatedly and always in a gentle way. None of the children abandoned the room. Moreover, with the exception of child C from the first subset, none of the participants got up out of the chair during the sessions, indicating, in general, that they were interested in the activity.

Comparing with other studies in the literature, the authors from the work [32] tested a facial expression module for a robotic platform (Zeno R25) in a preliminary study involving three high functioning children with ASD within the ages of 8–13 years old during two sessions. The facial expressions tested were anger, fear, happiness, and sadness. The results obtained from their study allowed to conclude that the success rate increased from the

first to the second session, which are similar results to the present work. Conversely, the children's response time to each robot prompt decreased between sessions, which may be due to the fact that the facial expressions were consecutively imitated 4 times by the children. In the present work, in order to mitigate the repetition factor and memorization, the facial expressions were randomly generated by the robot which may imply the increase in the children's response time between the first and the last session (sessions 1 and 7).

Table 9. Subset two: Children's mean response time in seconds for successful answers (SD) in the IMITATE activity. In general, the response time increased in the last session.

Session Number	Child D	Child E	Child F
1	17.68 (1.74)	19.14 (1.5)	20.37 (2.59)
2	16.54 (0.02)	21.41 (4.23)	19.05 (1.54)
3	17.64 (1.75)	16.88 (0.44)	22.20 (4.31)
4	17.89 (2.68)	17.18 (0.44)	16.54 (0.02)
5	18.43 (2.74)	17.31 (0.18)	17.89 (1.45)
6	18.17 (2.34)	17.04 (0.54)	18.73 (2.77)
7	20.47 (3.27)	18.45 (3.08)	20.01 (3.01)

Table 10. Subset two: Children's mean response time in seconds for successful answers (SD) in the EMOTIONS activity. In general, the response time increased in the last session.

Session Number	Child D	Child E	Child F
1	18.27 (2.6)	19.92 (4.42)	21.58 (2.78)
2	19.00 (3.57)	18.80 (3.21)	18.75 (2.05)
3	19.50 (3.33)	18.63 (2.88)	17.08 (0.74)
4	17.89 (0.72)	17.79 (1.83)	18.29 (1.2)
5	16.63 (0.68)	17.35 (1.7)	17.73 (1.68)
6	16.75 (1.0)	17.56 (1.75)	17.84 (1.49)
7	18.27 (2.65)	18.45 (3.16)	18.82 (3.11)

5. Conclusions

Facial expressions are a basic source of information about human emotional state. Failure of the emotion recognition skills might have consequences in a child's social development and learning [22]. In fact, individuals with ASD usually have difficulties in perceiving emotional expressions in their peers.

Assistive robots can be a useful tool to develop social–emotional skills in the support process of children with ASD. Currently, assistive robots are getting “more emotional intelligent” since affective computing has been employed, allowing to build a connection between the emotionally expressive human and the emotionally lacking computer.

The purpose of this study was to develop a system capable of automatically detecting facial expressions through facial cues and to interface the described system with a robotic platform in order to allow social interaction with children with ASD. To achieve the proposed goals, an experimental layout that uses the Intel[®] RealSense[™] 3D sensor and the ZECA robot was developed. This system uses SVM technique to automatically classify, in real-time, the emotion expressed by the user.

The developed system was tested in different configurations in order to assess its performance. The system was first tested in simulation using MATLAB and the performance of the two kernels was compared. RBF presented the best results, as the relation between class labels and attributes is nonlinear, with an average accuracy of 93.6%. Although using different experimental configurations, the present work achieved similar results of the state of art.

Then, the real-time subsystem was tested in a laboratorial environment with a set of 14 participants, obtaining an overall accuracy of 88%. The required time for the system to efficiently perform facial expression recognition is 1–3 ms at a frame rate of 30 fps on an i5 quad-core CPUs with 16 GB RAM. Then, the proposed subsystem was compared to other

state-of-the-art 3D facial expression recognition development in terms of overall accuracy, obtaining a performance of 88% against 84%, respectively.

An initial experimental study was conducted with typically developing children in a school environment with a main goal of testing the system, to detect the system constraints in a support session. The results obtained in this initial experimental phase showed that in the activity IMITATE all the facial expressions with exception of 'Anger' and 'Fear' had high recognition rates. The lower recognition rates of 'Anger' and 'Fear' could probably be due to the fact that the children had to interpret the facial expression displayed by ZECA, meaning perhaps that they did not interpret well the facial expression or the set of features contributing for the facial expression synthesized by ZECA was not well marked enough for the children to recognize. Moreover, these same facial expressions presented a higher recognition rate and were faster displayed by the children in the EMOTIONS activity.

Finally, an exploratory study, involving six children with ASD aged between eight and nine, was conducted in a school environment in order to evaluate the two game scenarios: the IMITATE, where the child has to mimic the ZECA's facial expression, and EMOTIONS, where the child has to perform the facial expression asked by ZECA. The original group of six was uniformly divided into two subsets of three children (one and two). In the subset one, three different facial expressions were investigated (anger, happiness, and sadness). On the other hand, in the subset two, five facial expressions were investigated (anger, fear, happiness, sadness, and surprised). The effects of ASD and the severity of symptoms differ from person to person, so it is expected that each child presents a unique pattern of progress throughout the sessions. Indeed, the results show that each child had a different learning progress. For example, one of the children (B) experienced more difficulties than the other children in the first sessions. However, in the last two sessions his/her performance improved and in fact, by analysing the results from both subsets, it is possible to infer that, in general, children had a positive evolution over the sessions, more expressed in the subset 2. In general, all children took more time answering to the prompt in the last session. The increase of the response time over the sessions might be related to the children thinking and considering all options they had available.

The results obtained allowed to conclude that the proposed system is able to interact with children with ASD in a comfortable and natural way, giving a positive indication about the use of this particular system in the context of emotions recognition and imitation skills. Although the sample is small (and further tests are mandatory), the results point out that the proposed system can be used as an eligible mediator in emotions recognition activities with children with ASD.

Therefore, future research should conduct more experiments to conclude the suitability of the proposed system to be used as a complement to the traditional interventions and use larger sample sizes in order to increase the reliability and replicability of data.

Author Contributions: Writing—original draft preparation, V.S.; data curation, V.S.; formal analysis, V.S. and A.P.P.; writing—review and editing F.S., J.S.E., C.P.S. and A.P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by FCT—Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020. Vinicius Silva thanks FCT for the PhD scholarship SFRH/BD/SFRH/BD/133314/2017.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the University of Minho (internal protocol code SECVS 028/2014, April 2014).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: The authors are grateful to all participants, performers and supervisors of the tests for their voluntary cooperation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fridenson-hayo, S.; Berggren, S.; Lassalle, A.; Tal, S.; Pigat, D.; Bölte, S.; Baron-Cohen, S.; Golan, O. Basic and Complex Emotion Recognition in Children with Autism: Cross Cultural Findings. *Mol. Autism* **2016**, *7*, 52. [CrossRef]
2. Tapus, A.; Member, S.; Scassellati, B. The Grand Challenges in Socially Assistive Robotics. *IEEE Robot. Autom. Mag.* **2007**, *14*, 1–7. [CrossRef]
3. Ricks, D.J.; Colton, M.B. Trends and considerations in robot-assisted autism therapy. In Proceedings of the 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010; pp. 4354–4359. [CrossRef]
4. Dautenhahn, K. Design issues on interactive environments for children with autism. In Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT 2000), Serpa, Portugal, 8–10 September 2001; pp. 153–162. Available online: <https://uhra.herts.ac.uk/bitstream/handle/2299/1944/902099.pdf?sequence=1&isAllowed=y> (accessed on 22 September 2021).
5. Taheri, A.; Meghdari, A.; Alemi, M.; Pouretamad, H. Human-Robot Interaction in Autism Treatment: A Case Study on Three Pairs of Autistic Children as Twins, Siblings, and Classmates. *Int. J. Soc. Robot.* **2018**, *10*, 93–113. [CrossRef]
6. Scassellati, B. How social robots will help us to diagnose, treat, and understand autism. *Robot. Res.* **2007**, *28*, 552–563. [CrossRef]
7. Lund, H.H. Modular playware as a playful diagnosis tool for autistic children. In Proceedings of the 2009 IEEE International Conference on Rehabilitation Robotics 2009, Kyoto, Japan, 23–26 June 2009; pp. 899–904. [CrossRef]
8. Pennisi, P.; Tonacci, A.; Tartarisco, G.; Billeci, L.; Ruta, L.; Gangemi, S.; Pioggia, G. Autism and social robotics: A systematic review. *Autism Res.* **2016**, *9*, 165–183. [CrossRef]
9. Costa, S. *Affective Robotics for Socio-Emotional Skills Development in Children with Autism Spectrum Disorders*; University of Minho: Guimarães, Portugal, 2014.
10. Messinger, D.S.; Duvivier, L.L.; Warren, Z.E.; Mahoor, M.; Baker, J.; Warlaumont, A.; Ruvolo, P. Affective Computing, Emotional Development, and Autism. In *The Oxford Handbook of Affective Computing*; Oxford University Press: Oxford, UK, 2014; pp. 516–536.
11. Picard, R.; Klein, J. Computers that recognize and respond to user emotion: Theoretical and practical implications. *Interact. Comput.* **2002**, *14*, 141–169. [CrossRef]
12. Robinson, P.; el Kaliouby, R. Computation of emotions in man and machines. *Philos. Trans. R. Soc. B Biol. Sci.* **2009**, *364*, 3441–3447. [CrossRef] [PubMed]
13. Lee, C.-H.J.; Kim, K.; Breazeal, C.; Picard, R. Shybot: Friend-stranger interaction for children living with autism. In Proceedings of the ACM CHI 2008 Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; Volume 2, pp. 3375–3380.
14. Pop, C.A.; Simut, R.; Pintea, S.; Saldien, J.; Rusu, A.; David, D.; Vanderfaeillie, J.; Lefeber, D.; Vanderborght, B. Can the social robot probo help children with autism to identify situation-based emotions? A series of single case experiments. *Int. J. Hum. Robot.* **2013**, *10*, 1350025. [CrossRef]
15. Soares, F.O.; Costa, S.C.; Santos, C.P.; Pereira, A.P.S.; Hiolle, A.R.; Silva, V. Socio-emotional development in high functioning children with Autism Spectrum Disorders using a humanoid robot. *Interact. Stud.* **2019**, *20*, 205–233. [CrossRef]
16. Chevalier, P.; Li, J.J.; Ainger, E.; Alcorn, A.M.; Babovic, S.; Charisi, V.; Petrovic, S.; Schadenberg, B.R.; Pellicano, E.; Evers, V. Dialogue Design for a Robot-Based Face-Mirroring Game to Engage Autistic Children with Emotional Expressions. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2017; Volume 10652, pp. 546–555.
17. Ambady, N.; Rosenthal, R. Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis. *Psychol. Bull.* **1992**, *111*, 256. [CrossRef]
18. Ekman, P.; Rosenberg, E. *What the Face Reveals*; Oxford University Press: Oxford, UK, 2005; ISBN 0-19-510447-1.
19. Ekman, P.; Friesen, W.V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*; Consulting Psychologists Press: Palo Alto, CA, USA, 1978; ISBN 0931835011.
20. Hadar, U.; Steiner, T.J.; Rose, F.C. Head movement during listening turns in conversation. *J. Nonverbal Behav.* **1985**, *9*, 214–228. [CrossRef]
21. Baltrusaitis, T. *Automatic Facial Expression Analysis*; University of Cambridge: London, UK, 2014.
22. Uljarevic, M.; Hamilton, A. Recognition of emotions in autism: A formal meta-analysis. *J. Autism Dev. Disord.* **2013**, *43*, 1517–1526. [CrossRef] [PubMed]
23. Hopkins, I.M.; Gower, M.W.; Perez, T.A.; Smith, D.S.; Amthor, F.R.; Wimsatt, F.C.; Biasini, F.J. Avatar assistant: Improving social skills in students with an asd through a computer-based intervention. *J. Autism Dev. Disord.* **2011**, *41*, 1543–1555. [CrossRef] [PubMed]
24. Happé, F.; Briskman, J.; Frith, U.; Happé, F.; Frith, U. Exploring the cognitive phenotype of autism: Weak “central coherence” in parents and siblings of children with autism: II. Real-life skills and preferences. *J. Child Psychol. Psychiatry Allied Discip.* **2001**, *42*, 309–316. [CrossRef]
25. Sosnowski, S.; Kuehnlenz, K.; Buss, M. EDDIE—An Emotion Display with Dynamic Intuitive Expressions. In Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2006), Hatfield, UK, 6–8 September 2006; pp. 3113–3118. [CrossRef]
26. Breazeal, C. *Sociable Machines: Expressive Social Exchange between Humans and Robots*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2000.

27. Hashimoto, T.; Kobayashi, H.; Kato, N. Educational system with the android robot SAYA and field trial. In Proceedings of the 2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011), Taipei, Taiwan, 27–30 June 2011; Volume 8, pp. 766–771. [[CrossRef](#)]
28. Mazzei, D.; Lazzeri, N.; Billeci, L.; Iglizzo, R.; Mancini, A.; Ahluwalia, A.; Muratori, F.; De Rossi, D. Development and evaluation of a social robot platform for therapy in autism. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 4515–4518. [[CrossRef](#)]
29. Mazzei, D.; Lazzeri, N.; Hanson, D.; De Rossi, D. HEFES: An Hybrid Engine for Facial Expressions Synthesis to control human-like androids and avatars. In Proceedings of the 2012 4th IEEE RAS & EMBS International Conference on biomedical robotics and biomechatronics (BioRob), Rome, Italy, 24–27 June 2012; pp. 195–200. [[CrossRef](#)]
30. Costa, S.; Soares, F.; Santos, C. Facial expressions and gestures to convey emotions with a humanoid robot. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2013; pp. 542–551. [[CrossRef](#)]
31. Leo, M.; Del Coco, M.; Carcagnì, P.; Distanto, C.; Bernava, M.; Pioggia, G.; Palestra, G. Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment. In Proceedings of the 2015 IEEE International Conference on Computer Vision Workshops, Santiago, Chile, 7–13 December 2015; pp. 537–545.
32. Palestra, G.; Varni, G.; Chetouani, M.; Esposito, F. A multimodal and multilevel system for robotics treatment of autism in children. In Proceedings of the International Workshop on Social Learning and Multimodal Interaction for Designing Artificial Agents—DAA, Tokyo, Japan, 16 November 2016; pp. 1–6.
33. Intel® RealSense™ Technology. 2015. Available online: <http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html> (accessed on 15 January 2016).
34. R50 Documentation. 2014. Available online: <http://www.robkindrobots.com/support-documentation/r50/> (accessed on 12 July 2015).
35. Philip, R.C.; Whalley, H.C.; Stanfield, A.C.; Sprengelmeyer, R.; Santos, I.M.; Young, A.W.; Atkinson, A.P.; Calder, A.J.; Johnstone, E.C.; Lawrie, S.M.; et al. Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychol. Med.* **2010**, *40*, 1919–1929. [[CrossRef](#)] [[PubMed](#)]
36. Loconsole, C.; Miranda, C.R.; Augusto, G.; Frisoli, A. Real-Time Emotion Recognition: A Novel Method for Geometrical Facial Features Extraction. In Proceedings of the 2014 International Conference on Computer Vision Theory and Applications (VISAPP), Lisbon, Portugal, 5–8 January 2014; pp. 378–385.
37. Face Pose Data [F200, SR300]. 2015. Available online: https://software.intel.com/sites/landingpage/realsense/camera-sdk/v1.1/documentation/html/index.html?doc_face_face_pose_data.html (accessed on 16 January 2016).
38. Silva, V.; Soares, F.; Esteves, J.S.; Figueiredo, J.; Leão, C.P.; Santos, C.; Paula, A. Real-time Emotions Recognition System. In Proceedings of the 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Lisbon, Portugal, 18–20 October 2016; pp. 201–206.
39. Hsu, C.W.; Chang, C.C.; Lin, C.J. *A Practical Guide to Support Vector Classification*; National Taiwan University: Taipei, Taiwan, 2010.
40. Zhang, Y.; Zhang, L.; Hossain, M.A. Adaptive 3D facial action intensity estimation and emotion recognition. *Expert Syst. Appl.* **2015**, *42*, 1446–1464. [[CrossRef](#)]
41. Michel, P.; El Kaliouby, R. Real time facial expression recognition in video using support vector machines. In Proceedings of the 5th International Conference on Multimodal Interfaces, Vancouver, CB, Canada, 5–7 November 2003. [[CrossRef](#)]
42. Souza, C.R. The Accord.NET Framework. 2014. Available online: <http://accord-framework.net> (accessed on 22 May 2016).