Type of the Paper (Article, Review, Communication, etc.)

An Intelligent Toy Car for Autism Screening using Multi-Modal Features

Bijan Mehralizadeh 1, Bahar Baradaran 2, Shahab Nikkho,3, Pegah Soleiman2, and Hadi Moradi 2,\*

|  |
| --- |
| **Citation:** Lastname, F.; Lastname, F.; Lastname, F. Title. *Sustainability* **2021**, *13*, x. https://doi.org/10.3390/xxxxx  Academic Editor: Firstname Lastname  Received: date  Accepted: date  Published: date  **Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.    **Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). |

1 Affiliation 1; e-mail@e-mail.com

2 School of ECE, University of Tehran, Tehran, Iran; [e-mail@e-mail.com](mailto:e-mail@e-mail.com)

3 School of …., University of Oregon, Oregon, USA

**\*** Correspondence: moradih@ut.ac.ir; Tel.:+98-21-82084960,

**Early screening, which can lead to early diagnosis and intervention of children with autism (ASD), can significantly improve the life quality of children with autism. The observational process of ASD diagnosis and lack of experts make the technology-based ASD screening methods more necessary. Early ASD screening based on behaviors is one of the most reliable methods that could be done by analyzing children’s playing patterns. In this paper, we present an extension of our initial intelligent toy car functionalities by adding shaft encoders to detect attention into details and interest in rotating objects in children with ASD. , Using the two modalities to detect different ASD symptoms improved our screening accuracy by 10%.**

**Keywords:** keyword 1; keyword 2; keyword 3 (List three to ten pertinent keywords specific to the article yet reasonably common within the subject discipline.)

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that causes social communication and interaction problems [1]. Unfortunately, ASD is becoming more prevalent in the last two decades [2]. On the other hand, studies show that early diagnosis resulting in early intervention can effectively reduce the disorder’s impacts. Diagnosing autism requires experts to observe children and interview parents to determine the severity of the symptoms. Unfortunately, this is a time-consuming and challenging process. Furthermore, many cases remain undiagnosed due to the lack of widespread experts, especially in low and middle-income societies. Consequently, it becomes essential to develop technology-based screening methods to make screening services cheap and widely available to overcome these challenges.

Many technology-based methods are originated from the CHAT family[3] questionnaires, in paper-based format or mobile apps or web applications[4]. These methods employ machine learning algorithms to improve the screening accuracy [5]. Although these methods are proved effective, they rely on licensed clinicians and observers, making them time-consuming and exhausting [3, 6]. To overcome the shortcomings of the questionnaire-based systems, many researchers have focused on biological markers of ASD[7] using brain imaging techniques like fMRI[8] or EEG methods to find ASD symptoms[9].

Although these methods are proved effective, they require costly equipment. Furthermore, putting a child in an fMRI or putting on an EEG cap may cause many discomforts limiting the usage of these approaches. On the other hand, wearable devices like smart glasses or sensors are used for ASD screening [10, 11]. Despite the lower cost of these systems compared to fMRI and EEG, they still need to be conducted at dedicated centers. Intelligent observation of behaviors is a method to overcome the challenges of biometric and wearable methods. For instance, Moghaddas et al. developed a vision-based method to screen children with ASD based on the interaction between children with ASD and a parrot-like robot [12]. Although this approach reduces mentioned difficulties in wearable methods, but still depends on dedicated centers to conduct the screening.

That is why home-based IoT devices such as the intelligent toy car[13] were designed to perform screening in children’s natural settings at a very low cost. Along with this trend, in this study, we improved our intelligent toy car by incorporating two shaft encoders on the wheels of our car to investigate if children with autism focus on the wheels instead of the whole car more than Typically Developed (TD) children. This new modality, combined with its initial modality, which was based on capturing the acceleration features in playing with the car, helped improve screening accuracy by 10%. In other words, our contribution is in incorporating two modalities to better screen ASD.

2. Related work

There have been several studies focused on using technology for ASD screening. These methods try to observe the ASD symptoms automatically using biomarkers or behavioral markers.

For example, William J. Bosl et al. focused on early screening of ASD by a data-driven method based on the EEG’s data. They collected EEG measurements of 99 infants with an older sibling that received an ASD diagnosis and 89 low-risk controls. They screen ASD in children as early as three months of age with 95% sensitivity and PPV at some ages. They suggest EEG signals might be a valuable biomarker for ASD screening[9]. Also, MladenRakić et al. presented a method to improve ASD detection by combining structural and functional MRI data. They applied machine learning techniques on imaging data of 817 cases and successfully classified them with an accuracy of 85%[14]. Integrating biomarkers with other modalities has also proved effective; JiannanKang et al. identified ASD in children from 3 to 6 by inputting a combination of EEG and eye-tracking features collected with power spectrum analysis and areas of interest methods to an SVM classifier. They tested on a total number of 97 children and reached the maximum accuracy of 85%.[15].

Stereotypical Motor Movements (SMM) is one of the ASD symptoms that multiple methods have been developed to detect. Rad, N. M et al. proposed a Convolutional Neural Network that uses accelerometer sensor data worn on multiple body points to detect SMM. They applied feature learning and transfer learning approaches to improve their deep neural network performance[10].

Detecting and analyzing gaze is also a method in ASD screening; Anish Nag et al. compared the gaze data of 16 children with ASD, and 17 typically developed children were collected using google glass and gaze tracker. Although smart glasses perform promisingly, they do not outperform other automatic classifiers significantly[11]

Although wearable devices are a helpful method for ASD screening, it is always challenging to persuade a young toddler to wear such devices, especially children with special needs; besides, wearing such devices is usually a major distraction that affects the procedure. Robots are an excellent option for evaluating social interactions, but they are costly and usually require operators to handle the process.

One of the major symptoms of ASD is repetitive and stereotypical behaviors that are considered an essential indication in ASD’s diagnosing.[16] In recent years, many technology-based screening systems have been developed, many methods focused on vision-based approaches, behavioral analysis methods that use machine vision to detect and recognize movements and motor function patterns. R. Oberleitner et al.[17] developed a recognition system for detecting abnormal behaviors that can be used in screening, assessment, or rehabilitation. Rasool Taban et al. [18] record walking patterns by Kinect and then analyze them using central pattern generator parameters as their classifier features. They accurately distinguished between tip-toe walking and regular walking pattern. Guillermo Sapiro and et al. [19] developed a low-cost mobile app that uses machine learning and machine vision methods to detect movement patterns and assess eye tracking patterns.

Vision-based methods also used for studying the subject attention; Kathleen Campbell and et al. [20] developed an app that record and analyze the reaction of the toddlers to video stimuli that designed to engage child’s attention; their algorithm classifies by automatically detecting and tracking multiple facial landmarks and analyzing their patterns.

One of the best ways to study the behavior in children is through their play with toys and pet animals. Since children spend a considerable amount of time playing with toys at a young age, the repetitive patterns could easily be recognized. Studying playing patterns does not have challenges like the discomforting feeling of brain imaging or EEG analyzing methods, and unlike wearable devices, they do not affect child attention and are considerably more cost-effective than robots.

Sensorized toys are valuable tools in ASD screening, embed different sensors inside toys to capture playing patterns, and are classified based on proven effective, i.e., Lanini M. and et al. combined accelerometer, gyroscope, and magnetometers data.[21] Also, Moradi et al. introduced a platform for autism screening based on acceleration data of a toy car that, in their first version, a Wii remote controller perform as a sensor hub and a Matlab program developed to interact with the system to collect accelerometer data of x, y, and z axes to investigate distinctive playing patterns and implement an SVM classifier with 85% accuracy[13]

In this research, the intelligent toy car 2.0 is introduced. It extends the previous version functionality by adding shaft encoders to the wheels, introducing new features, and optimizing the feature selection method. It enables us to study the ASD symptoms with a multi modalities approach and simultaneously analyze the repetitive behaviors and the obsessive attention to the details.

3. System design

The intelligent toy car is designed to capture the signs of two major symptoms in children with ASD, i.e., obsessive attention to detail and repetitive behaviors. i.e., Thus, in the first design Intelligent Toy Car 1.0, a Wii Mute handle which includes an accelerometer, was placed in the car. Our new design, the intelligent toy car 2.0 (Fig. 1(a)), has had multiple upgrades with respect to its first design. The new system has an inexpensive IoT board ESP8266 NodeMCU to read sensor data and send them wirelessly through Wi-Fi via UDP protocol to ensure maximum data collection rate. Also, the cheap MEMS accelerometer ADXL345 is placed inside the car, and two magnetic shaft encoders are installed on the front and back axles of the car. The whole system runs on a battery, and all electronic parts are embedded inside the car deliberately to avoid any distraction. The diagram of the system is shown in Fig. 1(b).

The intelligent toy car firmware is based on the Arduino ecosystem to make future R&D more effortless. Also, an ROS (Robotic Operating System) package is developed for interfacing with the system. It makes integrating the intelligent toy car in other systems more straightforward.

|  |  |
| --- | --- |
|  | Diagram  Description automatically generated |
| (**a**) | (**b**) |

**Figure 1**. (a) the intelligent toy car and (b) the schematic of the system

4. Experiments

The data collection process took place in the Dosste-Autism center (Autism friends center) in Tehran, Iran. The intelligent toy car was tested on 50 children ranging from 3 to 6 years old in three groups children with ASD, TD children, and other (CP and fragile X syndrome) shown in Table 1. Since it has been shown that the play complexity and toy engagement of children with ASD in both genders for the car-like toys are almost similar [17][18], we did not normalize the number of cases based on their gender. The subjects played with the intelligent toy car for about 3 to 5 minutes in a 3x4 meters room. The children could play in the test room alone or with their parents or therapists. The recorded data from each participant consists of time, acceleration in 3 dimensions, front and back wheel rotation counts saved in a database. A unique id in the database only identified each participant, and to preserve user anonymity and privacy, no personal data was recorded during the procedure. Furthermore, the parents’ consent was taken for all the participants. ASD group has 28 children that five of them did not seem interested in playing with the intelligent toy car and neglected it. All TD children were very interested in playing with the intelligent toy car, and children with fragile X syndrome and CP also played with the intelligent toy car but with less enthusiasm. Generally, the TD children playing was more energic, and they moved the car in the greater area than other groups.

Table 1. Details of the participants

|  |  |  |  |
| --- | --- | --- | --- |
|  | **autistic** | **TD** | **Other (CP and Fragile X)** |
| Number | 28 | 18 | 4 |
| Mean age | 4.63 | 4.61 | 5.5 |
| Median age | 4.0 | 4.0 | 5.5 |

5. Feature extraction

As mentioned earlier, the intelligent toy car is designed to capture the motion behaviors, focus on details, and interest in rotating items, which all are symptoms of children with ASD. In our previous study, we used movement patterns extracted from acceleration data for classification [13]. In this research, the same patterns are extracted, and the encoders' data are added into the model to be able to determine interest in rotating and rotation of items as well. To capture all three symptoms together, two steps are taken: a) extraction of features representing the pattern of the car movement. This is similar to the previous work done in [17]. Also, two other futures are extracted from acceleration signals representing the rolling and pitching of the car while playing with the wheels. To extract these features Short Term Fourier Transform is used. b) extraction of features representing focusing on details, interest in items’ rotation, and interest in rotating items, i.e. wheels’ rotation and rotating wheels. To extract these features the summation of two shaft encoder signals are used.

It should be noted that the whole children's playtime with the intelligent toy car can be divided into four states: 1) not playing, 2) playing only with wheels, 3) playing on the ground, and 4) playing in the air.

In the not playing section, the intelligent toy car is almost stationary and has no movements, and both encoders and acceleration signals are almost zero. The playing only with wheels section is when the intelligent toy car is almost motionless while rotating its wheels. In such a case, the shaft encoders show changes while there is no significant change in the acceleration. This state exactly represents the interest in rotating items and items’ rotation. The playing on the ground section describes those portions of playing where the intelligent toy car is moved which creates changes in both acceleration and shaft encoder signals. The movement patterns can be extracted in this state. Finally, the playing in the air section is when the intelligent toy car is moved in the air, and the acceleration is changing while its wheels are not rotating. Thus, the encoders’ signals show zero rotation. Based on the above state analysis, beside the original features proposed in [17], the following extra features (Table 2) were designed and extracted. The interactive play time means the summation of time spent in states 2 to 4.

Table 2. details of features

|  |  |  |
| --- | --- | --- |
|  | features | description |
| 1 | not playing ratio | ratio of not playing to total play time |
| 2 | playing only with wheels ratio | ratio of playing only with wheels to total play time |
| 3 | playing on the ground ratio | ratio of playing on the ground to total play time |
| 4 | playing in the air ratio | ratio of playing in the air to total play time |
| 5 | interactive playing only with wheels ratio | ratio of playing only with wheels to interactive play time |
| 6 | interactive playing on the ground ratio | ratio of playing on the ground to interactive play time |
| 7 | interactive playing in the air ratio | ratio of playing in the air to interactive play time |
| 8 | total wheels turn rate | total number of wheels turns during the total play time |
| 9 | tilt about X axis | number of fast movements about X axis |
| 10 | tilt about Y axis | number of fast movements about Y axis |

6. Classification Structure

To train classifiers, the collection of 45 subjects' data was divided into two groups: the training set with 80% of samples and the test set with 20% of remaining samples. The training set is used to train the classifier, and the test set is used to measure the classifier's performance. The K-fold cross-validation method[22] separates test and training sets in k=5 different ways to generalize the result and make it more reliable. The average of the accuracy, sensitivity, and specificity of the trainings are reported. In this research, a Support Vector Machine (SVM)[22] is a suitable machine learning method that can effectively classify this kind of data. By testing three kinds of SVM, SVM with a linear kernel is selected for its considerably better performance.

7. Feature Selection

Since the size of the training set is small compared to the size of the feature vector, feature reduction is necessary before applying machine learning methods. In the first step to reduce the number of features, the correlation between features, in each modality, were examined. Then from each highly correlated feature set, the best feature was selected using forward selection and backward elimination methods [24] [25]. This method effectively reduces the number of feature vector by about 30%. After reducing the number of features in each modality, the best features were selected in another round of feature selection.

Then in the final step, all the selected features from the two modalities were combined in an early fusion fashion to select and train the best classifier. As seen in [fig2], the best result is obtained by integrating five acceleration features and two shaft encoder features.

7. Results

Classification result based on previous studies is used as a baseline to examine the performance of adding encoder features in the model. The same SVM method is performed on the selected acceleration features presented in [13] to classify the new data. The best accuracy based on these acceleration features is 71.11%, as shown in Table 3. Adding encoder features increases the accuracy to 78.61%. By adding the new acceleration feature, the accuracy is increased from 71.11% to 75.83%.

Moreover, by adding both encoders' features and the new acceleration feature, the accuracy reaches 85.55% (Table 3). Based on the optimized feature selection method used in this study, the most effective shaft encoders' features are playing only with wheels ratio, not playing ratio, and total wheels turns rate. The most effective acceleration features are jolt in the Y direction, tilt about Y axis, the energy of the signal in the X direction, Correlation of the acceleration signals between X and Y directions, the fourth highest frequency in X direction, the fifth highest frequency in Z direction, and the relative amplitude of the highest frequency in Y direction.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Classifier** | **Accuracy** | **Sensitivity** | **Specificity** | **precision** |
| Baseline features | 71.11 | 67.14 | 73.00 | 80.00 |
| Baseline and encoder features | 78.61 | 75.00 | 68.00 | 87.50 |
| Baseline and new acceleration feature | 75.83 | 65.48 | 77.00 | 64.00 |
| Baseline, encoder features and new acceleration feature | 85.56 | 81.67 | 81.00 | 87.67 |

8. Discussion

As shown in Table 3, the accuracy of the new proposed multi-modal approach is better than the previously suggested single modal approach. Furthermore, adding the extra acceleration feature, i.e., …, in the feature set improved the classification accuracy.

The novelty of this research is its multi-modality structure resulting in the examination of ASD through a wider variety of symptoms. In order to reach this purpose, encoders are added to the system, which increases number of features. Since having a smaller learning space enhances the model functionality, feature reduction algorithms are implemented on each modality separately. After applying two stage feature reduction-including high correlated features elimination and most important features selection- selected features are combined together. Ultimately, the data is classified based on early fusion method which leads to the best result.

9. Conclusion

In this paper, we introduced the intelligent toy car 2.0 by which multi-modal ASD screening is planned. The new design incorporates shaft encoders to capture the tendency of children with ASD into details and rotating items. Furthermore, we improved the feature selection strategy to increase the system accuracy by multi-modal analyzing ASD symptoms. The advantage of this system over other screening methods is in its low cost and limited need for expertise. It can be used at homes, daycares, or clinics for initial screening.

For the future work, we have to test the system on a wider population and test it on a variety of cognitive deficits to see if it can differentiate between different cognitive deficits or not. Consequently, at this stage, it can be used as a warning system to alarm the parents and care givers to perform further evaluation through experts.

Finally, the intelligent toy car can be used beside other screening devices to increase the accuracy by considering other modalities of children with ASD. We expect that having more modalities observed can help to better screening.

References