

The Development of Attention Detection Model from Child Behavior for Robot-Assisted Autism Therapy

Penpitcha Wanglavan¹, Wisanu Jutharee², Thavida Maneewarn³, and
Boonserm Kaewkamnerdpong^{4*}

¹ Institute of Field roBOTics, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
(w.penpitcha@gmail.com)

² Institute of Field roBOTics, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
(wisanu.jutharee@gmail.com)

³ Institute of Field roBOTics, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
(praew@fib0.kmutt.ac.th)

⁴ Biological Engineering Program, Faculty of Engineering, King Mongkut's University of Technology Thonburi,
Bangkok, Thailand (boonserm.kae@kmutt.ac.th) * Corresponding author

Abstract: Using a robot as a mediator in therapeutic activity for autism therapy is a trending solution to reduce the burden for therapists. However, to promote continuous therapy, the use of robots could be extended for assisting parents or caregivers as well. Our group developed a robot called BLISS, which is a mobile robot with friendly toy-like appearance and sensors, to interact with children. The BLISS robot has been used to assist in learning interventions for autistic children through games. Previous studies showed that children paid attention and responded to the robot differently. Inexperienced parents often struggle to provide therapy to their autistic child. If the robot could learn to detect child attention and interact with the child appropriately, could the robot better help maintain child attention to the activity in therapy session? In this study, we developed an attention detection model based on child behavior during activity with the robot. We collected the child behavior data and the child psychologist evaluation for the behavior. We built the attention detection model for 2 levels of attention (high and low) by using k-nearest neighbor classifier, support vector machine, decision tree, and random forest. From the experiment, the random forest model returned the highest accuracy at 70%. We employed the model in the adaptive interaction system where the robot selects action based on child attention. In the pilot experiment, all four participants could stay engaged with the activity for equal to or more than the standard attention span. The promising results showed the potential of using robots to assist parents at home.

Keywords: Robot, Autism Spectrum Disorder, Autism Therapy, K-Nearest Neighbor Classifier, Support Vector Machine, Decision Tree, Random Forest

1. INTRODUCTION

The number of autistic patients entering the treatment process has dramatically increased. The number of specialists is not sufficient to thoroughly handle autistic patients. Parents and caregivers, therefore, take an important role in effective treatment. However, it is difficult for parents and caregivers who are not experienced in therapy because autistic children have problems in interacting and communicating with others; many autistic children do not cooperate in the therapy. Adopting robot and automation technology to the therapy may help lighten the workload from the therapists or parents who work with autism. Robots can do repetitive activities without developing negative emotions and can attract children during the therapeutic activities.

At present, there are research studies about using robotic technology to support therapeutic activities. For example, Milo [1] is a robot designed to look like humans that is used for practicing speech and facial expression. The NAO robot [2] has joints that can imitate human movement making it possible to teach the body movements and gestures to autistic patients. The Paro robot [3], which is a robot that is designed to look like a seal, has been used to promote animal therapy for

autistic patients so that patients feel relaxed and can practice both fine motor skills and gross motor skills. In Thailand, Kumdee *et al.* [4] developed the Fah Sai robot, which is a humanoid robot used for supporting teaching and learning of autistic children. It focuses on mimicking the basic gesturing and speaking for the children.

Many research studies focused on developing robots as assistants to the therapists. Our group aims toward develop robots to assist parents who provide therapy for their child at home. Because parents or caregivers are the ones who are closed to children, if robots could support them to provide their own therapy easier, they would provide continuous therapy to their child; the results of continuous therapy could lead to effective development. Our group developed the BLISS robot [5-6], which is a mobile robot that looks like a toy, to function as an assistant for parents to promote child development and learning through games.

Attawibulkul *et al.* [6] investigated the use of the BLISS robot in storytelling activity in autistic children. Comparing with the usual storytelling where a parent is a storyteller, it was found that using BLISS as assistant for parents in storytelling could better engage the child in the activity and increase positive child-parent interaction. With pre-programmed robotic activities,

parents confirmed during the interview that the assistant robot could reduce their burden during therapy sessions. Nevertheless, when the child loses attention to the activity, the parent still need to stimulate the child to attend to the activity. However, many parents may not yet be experts in dealing with children with autism. The effectiveness of each therapy session also depends on parent experience. If the assistant robot could learn to detect child attention and interact with the child appropriately, the robot could better help maintain child attention to the activity in therapy session.

From the literature on tools for autism therapy, it is found that most robots could not yet create various interactions with users. They only respond as they are pre-programmed such as greetings, questioning, answering, and rewarding after activities. At present, the robot systems are being developed to be more responsive to the users. For example, a robot can perform predetermined activities based on the correctness of answers from the users. When completing each level of activities, robots may act in a different manner to the users, such as using voice, pointing and staring [7]. Some research studies used physiological information to better understand how children with autism emotionally respond to the tasks. Liu *et al.* [8] used the Biopac system to capture the electrophysiological signals based on cardiac activity and electromyogram (EMG) in order to analyze the affective states of the user. The affective model based on physiological information was built and used to adapt robot behaviors toward children with autism.

In therapeutic activities of children with autism who often have emotional and behavioral varieties, robotic systems have not been able to naturally attract the attention of children during activities. Although having wearable sensors attaching to children's body could allow us to obtain more information from the users, it may trouble the user during the activities, especially the autistic children who have problem accepting new things. In addition, a high cost equipment may not be affordable for most parents. Our research question is that could the robot learn to detect child attention from behavior observation? In this study, we developed an attention detection model on the BLISS robot system and tested the pilot use of model in real application with 4 child participants.

2. MATERIALS AND METHOD

2.1 The BLISS Robot

The BLISS robot is a semi-automatic robot with toy-like appearance as shown in Fig. 1. It has a friendly look and can express a variety of emotional expressions. It is an improved version from [5]. Apart from a RFID reader, a camera was added to receive input and behavior from children while doing activities with the robot. The BLISS robot can exhibit light, sound, locomotion (with two wheels) and ear movement to

attract children attention and to create robotic emotional patterns to interact with children. We designed the response system of the robot so that it does not show negative behaviors or emotions that scare children or make them too afraid to interact with the robot.



Fig. 1 The BLISS robot

The BLISS robot was designed to be a medium between a child and a therapist or parent in order to promote child development through triadic interaction and games. During activities with the BLISS robot, children can learn content knowledge and develop skills while enjoying the gaming lessons. The children can play games on the robot by using touchscreen and placing RFID cards on the reader located on the top of the robot. All activity data are collected for further analysis. Autism specialists could review the data and use them to plan their therapy sessions for children with autism.

2.2 System Overview

In order to create an attention detection model, we observed usual therapy sessions and identified behaviors that could be collected from the BLISS robot and might be features for detecting attention from children. Ten features were identified as follows:

- 1) The time spent from answering a question,
- 2) The distance between the face and the reference point,
- 3) The size of the face
- 4) The gazing status: closed or distanced
- 5) The direction status of gazing: looking straight at the robot or looking in other directions
- 6) The percentage of accuracy when answering a question.
- 7) The status of consecutive correct answering (more than 3 times),
- 8) The status of consecutive incorrect answering (more than 3 times)
- 9) The status of frequent screen touching (more than 5 times in less than 15 seconds)
- 10) User emotion status (happy = 1 / unhappy = 0)

Then, we developed the system to collect all ten features. We conducted an experiment with the BLISS robot to collect a dataset. In the experiment, while children played game activities with the BLISS robot,

the system collected data on activities and pictures of the children faces for analyzing behavior data. For each activity, a child psychologist gave an opinion whether the child behavior was considered high or low attention with the robot. Both the user behavior data and therapist's view were used to build classification models to predict the level of attention. The model can be used in the robot to detect children attention and choose a suitable action for the child. The system to obtain an attention detection model in this study can be illustrated in Fig. 2.

2.3 The Attention Detection Model

2.3.1 Dataset

There were 25 children participated in the data collection experiment. The experiment was conducted under child psychologist's supervision at Play & Learn by Kru Nong, Bangkok, Thailand. The user's behavior data were used as inputs to build a classification model. The user's attention levels during the activities evaluated by an expert child psychologist were used as outputs. There are two attention levels: high-level attention and low-level attention. From this experiment, the total number of 519 data were collected. There are 383 data for high-level attention and 136 data for low-level attention.

2.3.2 Data Preprocessing

The obtained behavior data were analyzed to determine the ten features identified earlier. Because at each time step the data could be in different values, we computed the moving average of each feature within a predefined period of time. In this study, we used the period of 10 seconds in order to balance between the avoidance of too much computation and the need of frequent data for timely detection.

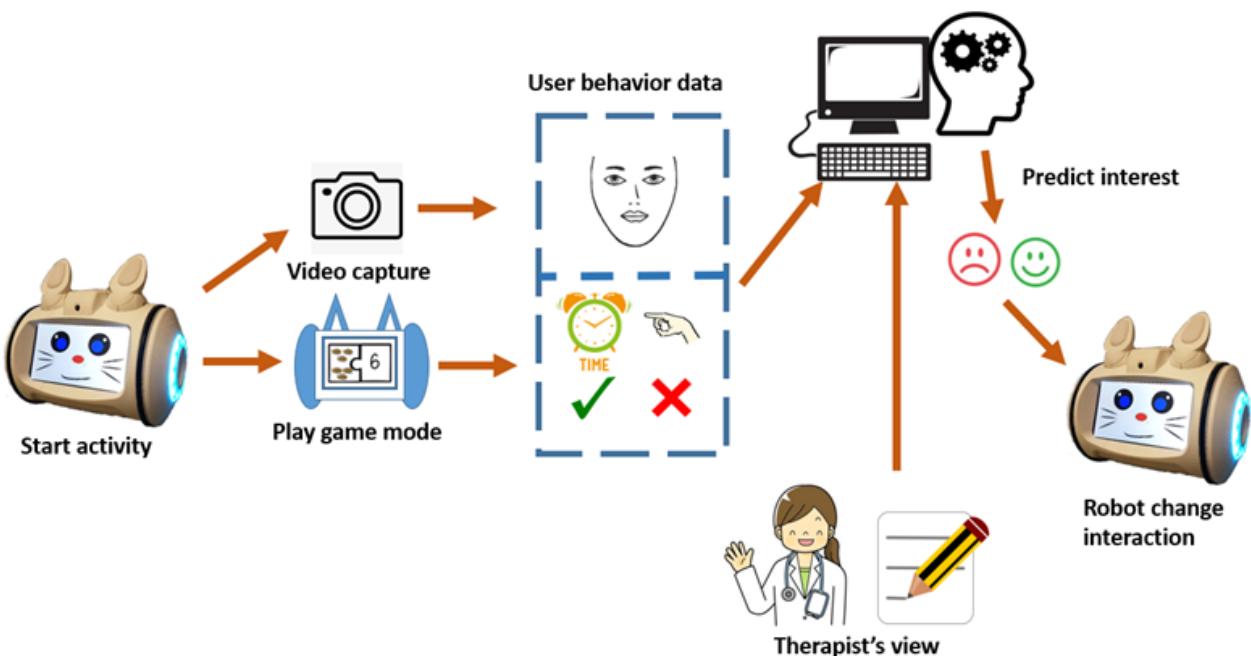
2.3.3 Feature Selection

Among ten features, we analyzed the existing data for the contribution score each feature has on the target attention classes and selected a set of most suitable features for building a classification model in the next step. By using the SelectKBest function in scikit-learn [9], we determined the 5 features with highest score obtained from the calculation of Chi-square statistic, which is commonly used for testing relationships between categorical or nominal variables. For feature selection, the feature that has high relationship with the target variable is considered an important feature; the feature is dependent to the target. The higher score value indicates the higher impact of that feature on the target.

2.3.4 Classification Model Construction

Based on the feature Selection, the selected 5 features with highest score were used in building a classification model to classify 2 levels of attention (high and low). Because each method has its own strengths and weaknesses, we developed classification model by using 4 methods in order to compare their results and select the best model with the highest classification accuracy. The five methods used in this study include k-Nearest Neighbor Classifier (kNN), Support Vector Machine (SVM), Decision Tree, and Random Forest

From the whole dataset, 25 percent were kept as a testing dataset to determine the classification accuracy; the remaining 75 percent of the dataset were used for building classification models. The data were randomly chosen into each set. From this dataset, 5-fold cross validation were used for model construction with GridSearch approach [9]. The classification accuracy of each result model was tested with the testing dataset (unseen by the methods).



3. RESULTS AND DISCUSSION

Based on the identified features from observing child behavior during doing activity with the BLISS robot, the scores from SelectKBest that measure the relationships between the features and the targets can be shown in Table 1. Five features with highest scores are shown in bold. The most important feature is the distance between the face and the reference point; this feature is also related to the size of the face. These two features indicate that how far the child is from the robot could reflect their attention level. In addition, the performance of the child during the activity could reflect their attention level; the high-score features involving the child performance include the time spent, the percentage of accuracy, and the status of consecutive correct answering.

Table 1 The scores from the calculation of Chi-square statistic for each feature.

Features	Score
The time spent from answering a question	80.943
The distance between the face and the reference point	154.613
The size of the face	130.111
The gazing status: closed or distanced	1.856
The direction status of gazing	0.006
The percentage of accuracy when answering a question	136.241
The status of consecutive correct answering	8.314
The status of consecutive incorrect answering	2.439
The status of frequent screen touching	2.439
User emotion status	0.871

The selected features were used to create classification models for classifying 2 levels of attention

Table 2 The comparison of classification performance for resulting attention detection models from 4 methods

Model		Accuracy	Precision		Recall		F1 Score	
			High	Low	High	Low	High	Low
KNN	Train	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Test	0.67	0.74	0.52	0.76	0.50	0.75	0.51
SVM	Train	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Test	0.65	0.65	0.00	1.00	0.00	0.79	0.00
Decision Tree	Train	0.99	0.97	1.00	1.00	0.98	0.98	0.99
	Test	0.67	0.72	0.52	0.80	0.40	0.76	0.46
Random Forest	Train	0.96	1.00	0.95	0.89	1.00	0.94	0.97
	Test	0.70	0.84	0.54	0.66	0.77	0.74	0.64

(high and low). By using GridSearch approach on 4 classification methods, the best parameter set for each method can be identified and used to create the classification model for each method. The classification performance in terms of accuracy, precision, recall and F-1 score for all four methods are shown in Table 2. After training, all models become over-fitted to the training data; the performance of all models become closed to 1. When performing prediction on the testing dataset, we found that the model from Random Forest resulted with the highest accuracy (at 70%).

In order to test the attention detection model, we could not perform the same process for collecting the dataset as the model detects and generates classification results every 10 seconds. For human experts, a longer time is needed to assess the child attention. We decided to conduct a pilot experiment to investigate whether a robot with changing actions based on child attention could extend the child engagement in the activity. We set two actions: stimulation (exhibited when the attention level is low) and reward (exhibited when the attention level is high). Every 10 seconds, the robot used the random forest model to classify the child attention level. If it is appropriate for exhibiting an action, the stimulation or reward action would be performed. However, if the robot is in the course of its tasks for hosting the activity, the actions would be refrained. We collected the time that the child stayed and played with the robot. We expected that children would attend to the activity with the robot longer or at least longer than their own attention span based on their age.

Four children participated in the adaptive interaction experiment. Table 3 shows the age of the participants, the standard attention span according to their age, and the time participants stayed engaged in the experiment. The engaged time in the experiment is equal to or more than the participants' standard attention span. The Participant B and C could stay engaged longer than the others. Fig 3 and 4 show the example of detailed results from the Participant B who stayed engaged for the longest time, more than twice of the participant's attention span.

Table 3 The comparison of engagement time and standard attention span of participants in the adaptive interaction experiment.

Participants	Age (Years)	Standard Attention Span (Minutes)	Engaged Time in Experiment (Minutes)
A	6	14	15
B	8	18	40
C	5	12	21
D	8	18	18

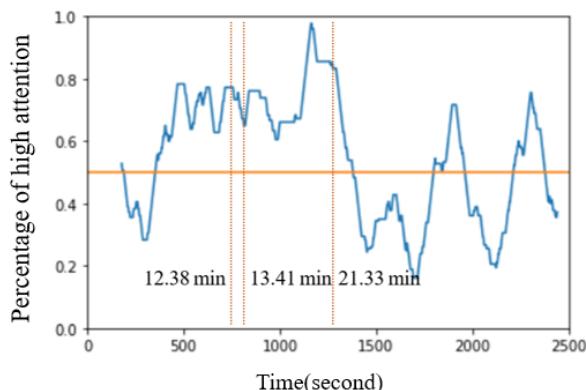


Fig. 3 The dynamical change in the child attention detected by the random forest model

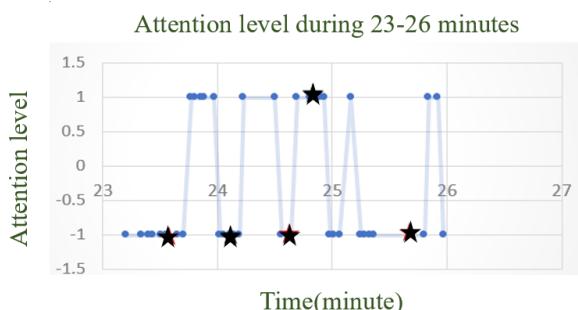


Fig. 4 An example of generated interaction based on the attention level between 23 and 26 minutes of the experiment

Fig. 3 illustrates the example of the dynamical change in the child attention detected by the random forest model in one of the participants. The attention dynamics is shown in term of the moving average of percentage of high-level attention over time with 3-minute window. This participant spent 40 minutes engage in the activity with the BLISS robot. The participant is 8 years old. The standard attention span is 18 minutes. It is evidenced that the adaptive interaction based on attention could extend the child attention to more than twice of the child's attention span. Fig. 4 illustrates the detail of generated interaction based on the attention level between 23 and 26 minutes of the experiment; the star symbol denotes the action of stimulation or reward that the robot exhibited. It can be seen that when the robot stimulated the child, the

attention level later became high, but the robot reward action may not quite motivate the child to keep high-level attention.

4. CONCLUSION

In this study, we developed an attention detection model on the BLISS robot system. The results showed that the robot system could detect the attention level (high and low) from child behavior during the activity by using random forest model with 70% accuracy. The result can be used to further develop the model to change the robot's responsive behavior by which the robot can choose the suitable actions (stimulation or reward) to the user's attention state. We did a pilot test using the attention detection model for adaptive interaction. The BLISS robot choose the appropriate behaviors according to the attention level. It was found that the robot could extend the time that the child engaged in the activity more than the child attention span. The child psychologist observed the experiment gave a comment that a robot with attention detection model could really help experts evaluate child attention in therapy sessions. We planned to extend our pilot study to a greater number of children. Moreover, different autistic children could have different symptoms and could respond well to different robot actions. We planned to develop the robot system for learning from the user responses and generating suitable stimulation and reward actions to promote the personalized effective therapy for autistic children.

ACKNOWLEDGMENT

The study is financially supported by King Mongkut's University of Technology Thonburi (Fiscal Year of 2019). The authors are thankful for children with autism participated in this experiment. Finally, the authors are immensely thankful for information and assistance in this study from Ms. Tichagorn Boonsri, child psychologist, Play & Learn by Krunong.

REFERENCES

- [1] J. H. Yousif, H. A. Kazem, M. T. Chaichan, "Evaluation Implementation of Humanoid Robot for Autistic Children: A Review," *International Journal of Computation and Applied Sciences*, Vol. 6, No. 1, pp. 412-420, 2019.
- [2] S. Shamsuddin, L. I. Ismail, H. Yussof, N. I. Zahari, S. Bahari, H. Hashim and A. Jaffar, "Humanoid robot NAO: Review of control and motion exploration," *Proc. of IEEE International Conference on Control System, Computing and Engineering, Penang*, pp. 511-516, 2011.
- [3] C. J. Calo, N. Hunt-Bull, L. Lewis and T. Metzler, "Ethical Implications of Using the Paro Robot, with a Focus on Dementia Patient Care", *Human-Robot Interaction in Elder Care*, pp.20-24, 2011.
- [4] O. Kumdee, S. Chumpiya, C. Ruangjirakit, P. Ritthipravat, "Speech Training System for Thai

- Therapeutic Robot”, *Proc. of the IASTED International Conference on Biomedical Engineering*, 2013.
- [5] S. Santatiwongchai, W. Jutharee, K. Ounjai and B. Kaewkamnerpong, “BLISS: Using Robot in Learning Intervention to Promote Social Skills for Autism Therapy,” in *The 10th International Convention on Rehabilitation Engineering & Assistive Technology (i-CREATE)*, Bangkok, Thailand, July 25-28, 2016.
 - [6] S. Attawibulkul, N. Sornsuwonrangsee, W. Jutharee and B. Kaewkamnerpong, “Using Storytelling Robot for Supporting Autistic Children in Theory of Mind,” *International Journal of Bioscience, Biochemistry and Bioinformatics* vol. 9, no. 2, pp. 100-108, 2019.
 - [7] H.-L. Cao, P. G. Esteban, D.B. Albert, R. Simut, G. Van de Perre, D. Lefever and B. Vanderborght, “A Collaborative Homeostatic-Based Behavior Controller for Social Robots in Human–Robot Interaction Experiments”, *International Journal of Social Robotics*, Vol. 9, No. 5, pp. 675-690, 2017.
 - [8] C. Liu, K. Conn, N. Sarkar and W. Stone, “Online Affect Detection and Robot Behavior Adaptation for Intervention of Children with Autism”, *IEEE Transactions on Robotics*, Vol. 24, No. 4, pp. 883-896, 2008.
 - [9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos and D. Cournapeau, Scikit-learn: Machine Learning in Python, *JMLR* 12, pp. 2825-2830, 2011.