An automatic social engagement measurement during human-robot interaction

Wael Hasan Ali Almohammed[1], Sinan Adnan Muhisn[2], Zahraa Abed Aljasim Muhisn[3]
[1]Department of Computer Science, College of Computer Science and Information Technology, University of Kerbala, Karbala, Iraq
[2]College of Biotechnology, Al-Qasim Green University, Babylon, Iraq
[3]Computer Center, Al-Qasim Green University, Babylon, Iraq

Article Info

Article history:

Received Oct 7, 2024
Revised Feb 17, 2025
Accepted Mar 15, 2025

Keywords:

Human-robot interaction
Neural network
Social engagement
Social robotics
User engagement

ABSTRACT

Social engagement refers the expressions of existing interpersonal relationships during the interaction which represents the actual interesting of human in the interaction. However, social engagement measurement is a significant concern in social human-robot interaction (HRI) because of its role in understanding the interaction's trend and adapt robot's behavior accordingly. Hence, we achieved the two main objectives of this study. Firstly, enrichment the theoretical literature and related concepts. Secondly, proposed a robust neural network model which is multilayer perceptron (MLP) classifier to measure social engagement state during interaction. PInSoRo dataset was used for training and testing purpose. In particular, the parameters of MLP model were meticulously crafted to recognize the social engagement accurately. We evaluated the model's performance by several metrics and the result showed an interesting accuracy reached 94.85%. Given that, it supports the robot to has adaptive and responsive behavior in real time applications which is improving HRI eventually.

Corresponding Author:

Wael Hasan Ali Almohammed
Department of Computer Science, College of Computer Science and Information Technology
University of Kerbala
Karbala, Iraq
Email: wael.h@uokerbala.edu.iq

1. INTRODUCTION

Recently, social services robots as an assistant or a companion have begun integrated to our services environments. They are pervasively turning into part of everyday tasks in education, work, and healthcare. Human-robot interaction (HRI) and social robotics study how robots support human through social interaction with an insight on developing an interaction with individuals in different contexts effectively [1], [2]. Generally, social robots are designed as user-friendly even for users without technological background such as children. Researches in these fields have focused on the factors that influence individuals' behavior and perception toward robots [3]. Definitely, child-robot interaction is an essential and critical research field as social robots are significantly employed to work with. Children are interacting with robots in different way since they have different immature cognitive development and daily living skills as well as they have high ability to adapt and learn new technology [3], [4]. Normally, children do not interact with robot as a mechatronic device with a computer program, but the characteristics of robot these are usually expected to be similar to any living system. Furthermore, the perspectives of children toward robots are far different from those of adults. Hence, expanding this knowledge to children's behavior is crucial to positively engage with robot. Generally, robot's attitude effected directly on engagement of child with the robot.

Productivity and quality of interaction are vastly correlated with increasing of engagement level [5], [6]. Therefore, we use collision risk index (CRI) as a case to be studied in this paper.

Engagement concept is broadly studied in HRI as a core issue in the interaction. Although, the meaning of this term does not have an explicit definition yet. Generally, it refers to being involved in formal or informal social activities. However, some researchers have defined it as “the process by which interactors start, maintain and end their perceived connection to each other during an interaction” [7], [8]. Typically, engagement level expresses how the interaction between human and robot is successful. Indeed, the key goal is to sustain the human engaged during the interaction. Furthermore, engagement level can influence the interaction strategy whereas if the fluctuation in user engagement is able to be detected, the interaction strategy could be formulated to enforce the users experience and keep them engaged. In addition, realizing user’s engagement is significant to provide personalized support and avoid dropouts. Therefore, measurement of engagement’s level is a pivotal function in HRI [9]–[11]. Accordingly, tracing human’s engagement has been a promising research area. However, there are two main methods to measure the engagement which are automatic and manual [12]. In traditional way, a third party can recognize the engagement level by direct observation checklist and rate scale. On the other hand, using a learning system for automatic engagement measurement [9].

There are several forms of engagement such as affective engagement, social engagement, and cognitive engagement. Since the concept of engagement itself is yet unclear, thus there is no plain explanation for each form with its features and there is an overlapping in the definition of each form. However, most of existing work concentrated on cognitive and emotional engagement since these forms are more defined and understood to some extent so they are easier to be recognized while the social engagement form has gotten less attention [11], [13]. On other hand, by perusing the literature, researchers used diverse methods to measure the state of engagement forms. Machine learning and deep leering models have been employed in the most due to the fact that they have been proven their efficiency in field of pattern recognition especially with the quick and massive advancement in the computational software and hardware [14]–[16]. Nevertheless, the vast majority of the previous studies have been measuring the engagement state, regardless its form, by using binary classification of two classes which are engaged or not engaged while there are a range of engagement states in between, each one could be improved differently.

Therefore, this paper indicates two research questions to be answered: what is the definition of engagement in HRI and its components’ characteristics and how to develop an improved automatic engagement measurement model comparing to existing studies focusing on social engagement particularly. In order to answer these questions, we set two objectives for this study which are to implicitly define the concept of engagement and understand the characteristics of each form. As well as, extend the research by developing an efficient model to automatically measure the social engagement form specifically. This work proposed a neural network model to measure social engagement level during CRI. This work is tightly relevant to the automatic user activities recognition. It has used multimodal dataset composing of visual and audio modalities. Additionally, the proposed model classifies engagement state into multiple classes.

The rest of paper is structured as follows: section 2 discusses the definitions of engagement concept, types, measurement approaches, and a snapshot of related work. The method details including research design, dataset and analysis process, and experimental examines systematically in section 3. Finally, the result, conclusion and future directions evaluates and discusses in sections 4 and 5 respectively.

2. BACKGROUND AND RELATED STUDIES

This section covers the main concepts and context which is necessary to understand the research problem. It begins with theoretical background such as the engagement concept in HRI, it forms, measurement approach. Eventually, it highlights the related studies, identifying the main methods, findings, and existing problems:

2.1. Engagement in human-robot interaction

In the development of social intelligent technology such as (robot, computer, or virtual agent), there are different issues shall be considered in order to personalize the interaction. Indeed, engagement is one of main these issues that broadly utilized as a key social phenomenon in the HRI field [17]. The research filed of engagement robots with people (users) is obtaining an intensive attention and interest among researchers [18]. Regardless the common use of engagement, there is no explicit meaning or interpretation concept. Conversely, the definition of engagement is yet characterized by ambiguity and big variation [19], [20]. However, some studies define the concept of engagement with the technologies and its role in particular contexts. To demonstrate, Sidner *et al.* [7] was from the earlier to define engagement concept, in general form, as “the process by which two (or more) participants establish, maintain and end their perceived connection during interactions they jointly undertake”.

Later on, Poggi [21] define by using deeper terms as “the value that a participant in an interaction attributes to the goal of being together with the other participant(s) and continuing interaction”. In HRI context, engagement is a concept of the greatest significance due to its ability of shaping the design of, developing a more advanced, and adaptable interfaces for users as well as contribution to better interaction outcome. However, engagement has a dynamic nature which means it is changing over time and between interactions. With the nature of engagement in mind and referring to definition of Poggi [21], engagement is considered a quality measure of the interaction. Considering that, O'Brien and Toms defined the engagement as “a quality of user experience characterized by attributes of challenge, positive affect, endurability, aesthetic and sensory appeal, attention, feedback, variety/novelty, interactivity, and perceived user control” [22].

Definitely, the ultimate goal of HRI is to establish a high level of engagement during interaction, consequently, achieve the interaction’s task successfully. Hence, reinforcement of engagement enhances the quality of interaction which reflected eventually on increasing the possibility of achieving interaction’s goal [19], [20]. So that, measuring user’s engagement can give insight for developing the user interaction whereas literatures amply concluded the positive relationship between user engagement and task achievement. Robots may formulate interaction strategy to sustain the users engaged or improve the engagement level, if they got the ability to measure the state of user engagement during interaction. An accurate engagement measurement can support robots to adapt their behavior in order to increase the success of interaction’s task and enhance user experience [23], [24].

2.1.1. Engagement components

Along with the difficulty of stating an explicit and comprehensive definition of engagement term, many studies have been confirmed the point of view the engagement is a complicated concept and forms of multiple components which are relevant among themselves tightly but they are still detected by particular indicator for each behavior independently. Accordingly, engagement is divided into different components of engagement by different work such as cognitive, affective, behavioral, social, and task. Also, some studies considered a hybrid engagement component like social-emotional, social-task, and social-cognitive [9], [24]–[28]. In this study, we discussed all known individual components as follows.

a) Cognitive engagement

This component of engagement has been typically involved conscious components like investment, attention, and effort for instance when users invest their cognitive resources during the interaction away from emotional, physical, or social resources to reinforce the role of performance (e.g. I have to work hard) [27], [29]. On the whole, cognitive engagement concerns of how the users build their connection during interaction, thinking actively, answering the questions, and resolving the problems [30]. It can be defined as the efforts to understand and analyze the interaction concept including meta-cognitive behaviors such as how the user set’s goal, plans, and organize their effort to achieve the task. It was also defined as an intensity of engrossment, concentration, and focus to achieve the task during interaction [31], [32].

b) Behavioral engagement

Generally, it refers to user attention towards tasks completion during the interaction. Behavioral engagement has been defined as a proactive predisposition of user to adopt with the changes and experiences during the interaction, in addition, the desire to be enhanced toward these changes. It is considered the encouragement that motives the participation in the task [33]. Behavioral engagement is addressed at the task level when there are a goal-oriented tasks for establishing the engagement. Therefore, as long as behavioral engagement increases, the more positive impact it has on task achievement [34]. This component of engagement has been found in nature, purpose, lack of difficulty, and familiarity of the task, while it misses emotional and social factors. The key feature of this type is that the human can resume the behavioral engagement and completing the task after any interruption [27], [34].

c) Affective engagement (emotional)

Obviously, the emotional engagement is defined as the mirror of affections and reaction among users (humans) and robots who are the parts of interactions which might be an internal and an external. In particular, the emotional engagement comprises of several affective states, to name few, enjoyment, mood, the feelings, and attitudes of the users who are joining the interaction. Nevertheless, the enthusiastic feeling and the enjoyment are the dominant affective states which have been investigated in the vast majority of the done studies [29], [35]. The theory, that says “positive emotions give a signal of purpose and excitement to the brain, accelerating learning and enhancing motivation”, ensures the tight association between the positive emotions and engagement level. Hence, the affective engagement amounts of user’s enjoyment in the interaction environment. Yet, it does not consider as indicator of the ultimate interaction effect, regardless as positive as it could be [24], [36].

d) Social engagement

Generally speaking, social engagement is the way of interaction between the human and its environment (other human, technology or task) in an adequate contextually approach and shows complicated internal dynamics which indicates the occupation of interaction state. it is a main metric for measure the human's cognitive and socio-emotional state collectively. Also, it is defined as the quantity and quality of verbal and non-verbal social interaction with robot [37]. In HRI term, social engagement refers to the involvement of the human with robots which have a friendly and sociable interaction capability. Additionally, it is added to the other engagement components due to its reference to the human dynamics and consider the engagement as an expression of existing interpersonal relationships during the interaction [38]. Furthermore, it differs from other engagement components because of having different conscious concentricity through the interaction. Whereas the other components disregard a significant factor to assess the engagement during interaction which is the actual interesting and readiness of human to begin the interaction [39].

2.1.2. Engagement measurement approaches

To begin with, a valid, reliable, and sturdy engagement measurement is a significant factor for developing an interactive robot from human's perspective since the nature of engagement is challenging to be measured. There are different approaches to measure engagement state during the HRI that have been fairly studied. Thereafter, each category has divided into sub-categories considering the data modalities and techniques used. Here, this study highlights a general overview and key points of each category:

a) Manual measurement

It is a traditional and ubiquitous approach to measure the engagement state of user during HRI. The predominant techniques in this approach are observational techniques and self-report and questionnaire. This approach has been wildly employed in various fields; HRI included. In case of observation methods, the interaction's administrator relies on the observation to measure the level of user's engagement. To name few of techniques that have employed in this case, ethograms and observational rating scales [40]. An example of Ethograms, video coding incorporating observed emotions that indicates to analyzing and labelling the video recording to categorize the emotion state for the individuals in the video [41].

On the other hand, the example of observational rating scales is observational measurement of engagement which uses an observation checklist to measure level of engagement. On other hand, self-report and questionnaire involved the interaction's user such as user engagement scale [42]. This approach endures some drawbacks such as the subjectivity of the administrator, time-discrepancy issue as the engagement is measured after the interaction, and lack of adaptability for robot during the interaction.

b) Automatic measurement

In order to overcome the limitations of manual engagement measurement, several studies began with development of an automatic measurement methods. In fact, the idea of automatic measurement of engagement in HRI is relatively recent then it has earned more attention lately [6]. Mostly, the studies utilize video and audio modalities of data as well as the neurological and physiological data for measurement such as heart rate, relative motion index (RMI), and electroencephalogram (EEG). However, the diversity of social robotics' applications has drawn more attention toward the visual data since each social robot has a built-in camera. Accordingly, the vast majority of latest studies use a cue-based approaches to recognize the social cues which could measure the social engagement level as well as other types of engagements [33].

A development of intelligent robots, that socially interacted with human and autonomously adapted its behavior during the interaction, requires an ability to measure the engagement state in proper and continuous way which is consider as a key challenge for social robotics researchers. The transition to automatic measurement offered an ability to determine if the user already engaged to the robot and waiting for its response within the interaction time. Thereupon, the robot can use the engagement state to adapt its behavior conveniently toward enhancing the interaction outcome. Additionally, the advancement of machine learning and deep learning models have led to expand the improvement of automatic engagement measurement in term of accuracy and computational time [43], [44].

For the purpose of automatic measurement rule-based, machine learning, and deep learning are used. Firstly, the rule-based techniques that choose various rules among the presence of the main social signals. Then, each rule is measured by a state machine that calculate the final engagement level. Also, it can adopt a threshold-based rule for measurement purpose. Secondly, machine learning models are largely used in engagement measurement since it enriches the development of HRI and affective computing studies that automatically characterize the human behavior. Finally, deep learning that is somewhat late in engagement measurement studies. However, both machine learning and deep learning use by mapping the features of raw data to get the target level of engagement. Deep learning sparked by the weakness of machine learning models to deal with high-dimensional features and large variations raw data. In addition, deep learning minimizes the complicated mapping into a group of sub-mappings [14].

2.2. Related work

Indeed, an adequate works have been proposed for engagement measurement in HRI among different scenarios. These works have obvious diversity in computational model used, data modality, feature sets, and the number of engagement classes. In this section, we showcase of selected work that employed different machine learning and deep learning models for automatic engagement measurement in HRI.

Initially, a dynamic Bayesian network model has utilized to measure the engagement of children with autism spectrum disorder (ASD) interacting with the NAO robot. The evaluation data by the professional caregivers used as input to the model and the best performance of the model is reached 93.60% [45]. In the like manner, Papakostas *et al*. [15] conducted a multimodal machine learning approach for measuring binary engagement state for children with learning difficulties during educational scenario of interaction. A visual and audio data were collected and processed and the AdaBoost decision tree ensemble model has achieved 93.33%. Additionally, Engwall *et al*. [16] proposed a machine learning model of combined support vector machine (SVM) for engagement measurement during HRI in context of second language learning. The data collected from video record and the highest measurement accuracy has achieved is 79.00%.

On the other hand, some other studies used a deep learning models for this purpose. For instance, long short-term memory (LSTM)-based neural network has been employed during unrestricted child-robot collaboration for their engagement measurement. The study has been used the data child's poses and it achieved a competent accuracy 77.11% considering the difficulty of the problem and interaction scenario [4]. Also, Javed *et al* [25] proposed a multilayer and multichannel of convolutional neural network (CNN) for automatic measurement of engagement in children with ASD. The evaluation showed the best performance of proposed framework is 81.00% accuracy using collected data of video, audio and motion-tracking. In the same context, another study of proposed a deep learning models CNN and LSTM. It tested the model with several visual datasets of different contexts and the optimal performance reached is 89.00% accuracy [14]. Table 1 summarizes the mentioned-above studies by stated used model and best performance rate.

Regardless some intersection with other works, this study has an outstanding contribution by enrichment the theoretical literature of engagement concept and categorizes engagement to independent components with clear insight and characteristics. As well as, it proposed a neural network classifier for automatic measuring multi-class of social engagement particularly and it was achieved a remarkable accuracy rate. Therefore, such results would have practical implications for improvement the interaction's quality in different fields.

Table 1. Summary of previous work and their result

Reference	Year	Model	Accuracy (%)
[45]	2017	Dynamic Bayesian network	93.60
[15]	2021	AdaBoost decision tree ensemble	93.33
[16]	2022	SVM	79.00
[4]	2019	LSTM-based neural network	77.11
[25]	2020	CNN	81.00
[14]	2020	CNN and LSTM	89.00

3. EXPERIMENTAL METHOD

The key issue that has been addressed here is to measure, automatically, social engagement state of children utilizing visual and audio modalities. The experiments were carried out by using Python 3.10 in Google Colab environment. There are several libraries have been used during the experiment process such as mainly NumPy and Pandas for data preprocessing and handling, scikit-learn library for design and implement classification model, Matplotlib library for visualizing the results and others libraries for other particular tasks. A multiclass classification using MPL classifier whereas each class represents a different state of child's social engagement as detailed in next sections. Figure 1 visualize the general workflow of the study's experiment.

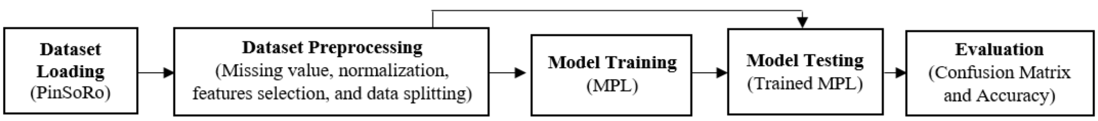

Figure 1. General research workflow

3.1. Dataset description

For sake of training and testing the proposed model, a publicly available dataset, named PInSoRo, has been used. This dataset was collected during a series of underspecified free-play child-child and child-robot interaction. However, it recorded an over 45 hours of social interactions among 45 child-child pairs and 30 child-robot pairs. Besides, it used a hand-coded recordings occurring in natural social interactions between children. It has a video recording, skeletal information, 3D recordings of the faces, and full audio records. The key strength of considering visual and audio data is that the setup of interaction environment will be relatively comfortable and close to reality. Eventually, the dataset is rich because of many characteristics such as covering a vast range of interaction situation, demonstrating complex social dynamics, natural and original behaviours due to the unspecificity, and wealth of multimodal interactions [46]. Particularly, the dataset specified five primary and distinct states of social engagement which are: firstly, solitary that indicates child disengagement. Secondly, onlooker signifies that the child is watching the interaction but does not really join. Thirdly, parallel means the child join the interaction's game but playing solely. Fourthly, associative refers that the child joins the game without coordination with others actively. Lastly, cooperative that signifies the child joined the game with an organized role and start sensing of team work.

3.2. Data pre-processing

Prior to feed the network by the dataset, several pre-processing techniques have been applied on the raw dataset toward enhancing network's performance. Initially, different action has been taken for different data type, for instance, cleaning the data, handling the missing values, and managing the categorical features. However, imbalance dataset is a key challenge in engagement measurement, then the dataset has passed through some steps to be balanced and normalized. Additionally, principal component analysis (PCA) technique has been utilized to reduce the high dimensional issue in the raw dataset. On the other hand, the irrelevant features have been eliminated from the dataset.

3.3. Proposed model

A fully connected multilayer perceptron (MLP) has been used to our task. The features of pre-processed dataset were employed to train the selected deep learning neural network, MLP, to measure the social engagement state of user during child-robot interaction. MLP is one of the ubiquitous feedforward neural network to map set of input features and the corresponding classes. The general architecture of MPL consists of multilayers with nodes that fully connected to each other. The input and output layers are the first and last layers sequentially in addition to one, at very least, or multiple hidden layers in between. Moreover, the number of nodes is varying for every layer in accordance to the number of inputs and outputs.

The MLP has been selected for social engagement measurement task due to several reasons such as its notable efficiency in solving the non-liner decision boundary as well as complicated pattern recognition problems by using non-linier activation function which essential for real-world data such human-robot engagement, in addition to its robustness by dealing with high-dimensional data. MLP has generalization ability to untrained data which overcomes overfitting and maintain new examples effectively. Its capability Also, unlike the classic machine learning techniques, MLP overcomes the feature selection and feature extraction issues and deals with subtleties for capturing the social engagement state accurately. It has the ability to process the intricate tapestry of social dynamics during child-robot interaction like physical gestures and facial expressions for instance. However, the network has trained through the uniform sampling of dataset can minimize overfitting and time difficulties.

3.4. Measurement metrics

In order to evaluate the performance of proposed model for automatic social engagement measurement, several approaches have been applied. Firstly, comparing the classification result with the actual classes by calculating the accuracy, precision, recall, and F1-score as the following equation. Overall, the model achieved an impressive classification accuracy rate of 94.85%. In the same context, precision—which defines the model's performance by calculating the ratio of true positives (TP) to the total predicted positives (TP+false positives (FP)), as shown in (1)—reached 93.00%. Likewise, recall, which measures the ratio of TP to the total actual positives (TP+false negatives (FN)), as described in (2), reached 95.00%. Meanwhile, the F1-score, defined as the harmonic mean of precision and recall as shown in (3), the rate was 94.00%. Table 2 summarizes the results obtained by the proposed network.

$$Precision = \frac{TP}{(TP+FP)} \quad (1)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (2)$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

Table 2. Summary of performance measurement metrics for each class

Classes	Precision	Recall	F1-score
Cooperative	1.00	1.00	1.00
Associative	0.91	1.00	0.96
Parallel	0.93	0.91	0.92
Onlooker	0.91	0.89	0.92
Solitary	0.97	0.98	0.97

Secondly, the model's performance has been evaluated by one of the most used approach for evaluating performance of machine learning and deep learning models which is confusion matrix. Moreover, confusion matrix is a comprehensive presentation for evaluating multi-classes classifiers' performance. Also, it provides a visualized depiction that plainly reveal insight into the number of predicted classes to the number of actual classes. However, in respect to our model's performance the confusion matrix presents a breakdown in details of measuring the social engagement state of children. Lastly, we apply the receiver operating characteristics (ROC) to visualize the measurement performance considering the correct and incorrect measurement rate. ROC plotted the trade between the true positive rate and false positive rate. Figure 2 depicted the ROC of social engagement measurement model.

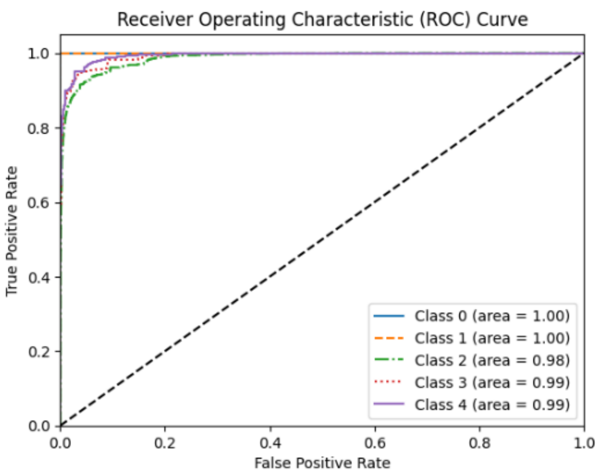

Figure 2. ROC of social engagement measurement model

4. RESULTS AND DISCUSSION

In the present paper, the proposed neural network met the expectation whereas it has shown a remarkable result which verified during the evaluation phases as detailed in the previous section. Also, the evaluation of our model's experimental demonstrated an outperformance in the overall classification accuracy comparing to the result of other work as we can state by seeing the previous works in Table 1. A notable matter in the result is that the classification of all classes is convergent, still, there is a differentiation in the cooperative class which may be attributed to the fact that the number of cooperative class samples in the dataset are the least.

the objectives of this study have been achieved whereas firstly, a well understanding and discussion for the core of social engagement, its features, and difference about other components of engagement are presented which reflected on the setting of model and the accuracy's improvement eventually. Secondly, developing a high accurate measurement model for social engagement state during HRI. Yet, the results highlight a limitation in measurement of onlooker state that has a slight decline as shown in the recall (89.00%). It could be caused by the features' overlap of this class with other or the selected hyper parameters have not reached the optimal and affecting the model performance. Overall, the proposed model holds promises for social engagement measurement in HRI.

5. CONCLUSION AND FUTURE WORK

This paper studied the social engagement state measurement task for human, children exclusively, interacting with a social robot in order to set up an adaptive, responsive, and intelligent interaction in real time application which enhances the HRI at last place. It presented the definition of engagement in HRI field and dived deeper to each component's characteristics. However, the measurement process has been considered as multi class classification issue. MPL model was used to tackle this problem and was achieved a distinguish results. The proposed model utilized a multimodal dataset which consists of visual and audio data for training and testing purpose. The overall accuracy is 94.85% that appeared an improvement comparing to other done studies. The result is promising toward building a more sociable and adaptable robot and leverage the interaction. In future, we will work to measure the social engagement state in integrated way which means measuring the states in between the distinct states such as (onlooker and parallel) or (associative and cooperative) at the same time and test the model with real application, then simulate the proposed model to the virtual robotics environment such as ROS. As well as, keep working on improving the accuracy with more data and model parameters.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the use of service and facilities of the Artificial Intelligence Research Lab at University of Kerbala.

FUNDING INFORMATION

Authors acknowledge limited funding provided by Department of Computer Science, College of Computer Science and Information Technology, University of Kerbala.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Wael Hasan Ali Almohammed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Sinan Adnan Muhisn		✓		✓	✓	✓		✓		✓	✓			
Zahraa Abed Aljasim Muhisn	✓			✓		✓	✓			✓	✓		✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author on request.

REFERENCES

[1] A. Alabdulkareem, N. Alhakbani, and A. Al-Nafjan, "A systematic review of research on robot-assisted therapy for children with autism," *Sensors*, vol. 22, no. 3, Jan. 2022, doi: 10.3390/s22030944.
[2] N. Gasteiger, M. Hellou, and H. S. Ahn, "Factors for personalization and localization to optimize human-robot interaction: a literature review," *International Journal of Social Robotics*, vol. 15, no. 4, pp. 689–701, 2023, doi: 10.1007/s12369-021-00811-8.
[3] C. L. v. Straten, J. Peter, and R. Kühne, "Child-robot relationship formation: a narrative review of empirical research," *International Journal of Social Robotics*, vol. 12, no. 2, pp. 325–344, 2020, doi: 10.1007/s12369-019-00569-0.

[4] J. Hadfield, G. Chalvatzaki, P. Koutras, M. Khamassi, C. S. Tzafestas, and P. Maragos, "A deep learning approach for multi-view engagement estimation of children in a child-robot joint attention task," *IEEE International Conference on Intelligent Robots and Systems*, pp. 1251–1256, 2019, doi: 10.1109/IROS40897.2019.8968443.

[5] C. Filippini and A. Merla, "Systematic review of affective computing techniques for infant robot interaction," *International Journal of Social Robotics*, vol. 15, no. 3, pp. 393–409, 2023, doi: 10.1007/s12369-023-00985-3.

[6] J. Nasir, B. Bruno, M. Chetouani, and P. Dillenbourg, "What if social robots look for productive engagement?: automated assessment of goal-centric engagement in learning applications," *International Journal of Social Robotics*, vol. 14, no. 1, pp. 55–71, 2022, doi: 10.1007/s12369-021-00766-w.

[7] C. L. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich, "Explorations in engagement for humans and robots," *Artificial Intelligence*, vol. 166, no. 1–2, pp. 140–164, 2005, doi: 10.1016/j.artint.2005.03.005.

[8] M. Paliga, "Human–cobot interaction fluency and cobot operators' job performance. The mediating role of work engagement: a survey," *Robotics and Autonomous Systems*, vol. 155, 2022, doi: 10.1016/j.robot.2022.104191.

[9] S. N. Karimah and S. Hasegawa, "Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods," *Smart Learning Environments*, vol. 9, no. 1, 2022, doi: 10.1186/s40561-022-00212-y.

[10] D. Lala, K. Inoue, P. Milhorat, and T. Kawahara, "Detection of social signals for recognizing engagement in human-robot interaction," *arXiv-Computer Science*, pp. 1-8, 2017.

[11] C. Lytridis, C. Bazinas, G. A. Papakostas, and V. Kaburlasos, "On measuring engagement level during child-robot interaction in education," *Advances in Intelligent Systems and Computing*, vol. 1023, pp. 3–13, 2020, doi: 10.1007/978-3-030-26945-6_1.

[12] M. A. A. Dewan, M. Murshed, and F. Lin, "Engagement detection in online learning: a review," *Smart Learning Environments*, vol. 6, no. 1, 2019, doi: 10.1186/s40561-018-0080-z.

[13] M. R. Lima, M. Wairagkar, N. Natarajan, S. Vaitheswaran, and R. Vaidyanathan, "Robotic telemedicine for mental health: a multimodal approach to improve human-robot engagement," *Frontiers in Robotics and AI*, vol. 8, 2021, doi: 10.3389/frobt.2021.618866.

[14] F. D. Duchetto, P. Baxter, and M. Hanheide, "Are you still with me? Continuous engagement assessment from a robot's point of view," *Frontiers in Robotics and AI*, vol. 7, 2020, doi: 10.3389/frobt.2020.00116.

[15] G. A. Papakostas *et al.*, "Estimating children engagement interacting with robots in special education using machine learning," *Mathematical Problems in Engineering*, vol. 2021, 2021, doi: 10.1155/2021/9955212.

[16] O. Engwall, R. Cumbal, J. Lopes, M. Ljung, and L. Månsson, "Identification of low-engaged learners in robot-led second language conversations with adults," *ACM Transactions on Human-Robot Interaction*, vol. 11, no. 2, 2022, doi: 10.1145/3503799.

[17] A. Sorrentino, G. Mancioppi, L. Coviello, F. Cavallo, and L. Fiorini, "Feasibility study on the role of personality, emotion, and engagement in socially assistive robotics: a cognitive assessment scenario," *Informatics*, vol. 8, no. 2, 2021, doi: 10.3390/informatics8020023.

[18] J. Avelino, L. Garcia-Marques, R. Ventura, and A. Bernardino, "Break the ice: a survey on socially aware engagement for human-robot first encounters," *International Journal of Social Robotics*, vol. 13, no. 8, pp. 1851–1877, 2021, doi: 10.1007/s12369-020-00720-2.

[19] M. A. Trahan, J. Kuo, M. C. Carlson, and L. N. Gitlin, "A systematic review of strategies to foster activity engagement in persons with dementia," *Health Education & Behavior*, vol. 41, pp. 70S-83S, 2014, doi: 10.1177/1090198114531782.

[20] I. Neal, S. H. J. Du Toit, and M. Lovarini, "The use of technology to promote meaningful engagement for adults with dementia in residential aged care: a scoping review," *International Psychogeriatrics*, vol. 32, no. 8, pp. 913–935, 2020, doi: 10.1017/S1041610219001388.

[21] I. Poggi, *Mind, hands, face, and body: a goal and belief view of multimodal communication*, Berlin, Germany: Weidler Buchverlag, 2007.

[22] K. Doherty and G. Doherty, "Engagement in HCI: conception, theory and measurement," *ACM Computing Surveys*, vol. 51, no. 5, 2019, doi: 10.1145/3234149.

[23] Y. Feng, E. I. Barakova, S. Yu, J. Hu, and G. W. M. Rauterberg, "Effects of the level of interactivity of a social robot and the response of the augmented reality display in contextual interactions of people with dementia," *Sensors*, vol. 20, no. 13, pp. 1–12, 2020, doi: 10.3390/s20133771.

[24] Y. Feng, G. Perugia, S. Yu, E. I. Barakova, J. Hu, and G. W. M. Rauterberg, "Context-enhanced human-robot interaction: exploring the role of system interactivity and multimodal stimuli on the engagement of people with dementia," *International Journal of Social Robotics*, vol. 14, no. 3, pp. 807–826, 2022, doi: 10.1007/s12369-021-00823-4.

[25] H. Javed, W. H. Lee, and C. H. Park, "Corrigendum: toward an automated measure of social engagement for children with autism spectrum disorder-a personalized computational modeling approach," *Frontiers in Robotics and AI*, vol. 7, 2020, doi: 10.3389/frobt.2020.00067.

[26] M. Mihai, C. N. Albert, V. C. Mihai, and D. E. Dumitras, "Emotional and social engagement in the english language classroom for higher education students in the COVID-19 online context," *Sustainability*, vol. 14, no. 8, 2022, doi: 10.3390/su14084527.

[27] Y. Kim, S. Butail, M. Tscholl, L. Liu, and Y. Wang, "An exploratory approach to measuring collaborative engagement in child robot interaction," *ACM International Conference Proceeding Series*, pp. 209–217, 2020, doi: 10.1145/3375462.3375522.

[28] A. F. Paiva, C. Cunha, G. Voss, and A. D. Matos, "The interrelationship between social connectedness and social engagement and its relation with cognition: a study using SHARE data," *Ageing and Society*, vol. 43, no. 8, pp. 1735–1753, 2023, doi: 10.1017/S0144686X2100129X.

[29] H. Salam, O. Celiktutan, I. Hupont, H. Gunes, and M. Chetouani, "Fully automatic analysis of engagement and its relationship to personality in human-robot interactions," *IEEE Access*, vol. 5, pp. 705–721, 2017, doi: 10.1109/ACCESS.2016.2614525.

[30] W. L. Q. Oga-Baldwin and L. K. Fryer, "Engagement growth in language learning classrooms: a latent growth analysis of engagement in japanese elementary schools," *Student Engagement in the Language Classroom*, pp. 224–240, 2020.

[31] S. Y. B. Huang, C. H. Huang, and T. W. Chang, "A new concept of work engagement theory in cognitive engagement, emotional engagement, and physical engagement," *Frontiers in Psychology*, vol. 12, 2022, doi: 10.3389/fpsyg.2021.663440.

[32] A. Henschel, R. Hortensius, and E. S. Cross, "Social cognition in the age of human-robot interaction," *Trends in Neurosciences*, vol. 43, no. 6, pp. 373–384, 2020, doi: 10.1016/j.tins.2020.03.013.

[33] A. Rossi, M. Raiano, and S. Rossi, "Affective, cognitive and behavioural engagement detection for human-robot interaction in a bartending scenario," *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021*, pp. 208–213, 2021, doi: 10.1109/RO-MAN50785.2021.9515435.

[34] P. Dao, "Effects of task goal orientation on learner engagement in task performance," *IRAL - International Review of Applied Linguistics in Language Teaching*, vol. 59, no. 3, pp. 315–334, 2021, doi: 10.1515/iral-2018-0188.

[35] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, "Context-sensitive affect recognition for a robotic game companion," *ACM Transactions on Interactive Intelligent Systems*, vol. 4, no. 2, 2014, doi: 10.1145/2622615.

[36] T. K. F. Chiu, "Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic," *Journal of Research on Technology in Education*, vol. 54, no. S1, pp. S14–S30, 2022, doi: 10.1080/15391523.2021.1891998.

[37] Y. Kishida and C. Kemp, "Measuring child engagement in inclusive early childhood settings: implications for practice," *Australasian Journal of Early Childhood*, vol. 31, no. 2, pp. 14–19, 2006, doi: 10.1177/183693910603100204.

[38] S. Lemaignan, F. Garcia, A. Jacq, and P. Dillenbourg, "From real-time attention assessment to 'with-me-ness' in human-robot interaction," *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 157–164, 2016, doi: 10.1109/HRI.2016.7451747.

[39] T. Fukuda, Y. Fukada, J. Falout, and T. Murphey, "How ideal classmates priming increases EFL classroom prosocial engagement," *Student Engagement in the Language Classroom*, pp. 182–201, 2021, doi: 10.21832/9781788923613-013.

[40] A. Kapoor and R. W. Picard, "Multimodal affect recognition in learning environments," *Proceedings of the 13th ACM International Conference on Multimedia, MM 2005*, pp. 677–682, 2005, doi: 10.1145/1101149.1101300.

[41] C. Jones, B. Sung, and W. Moyle, "Assessing engagement in people with dementia: a new approach to assessment using video analysis," *Archives of Psychiatric Nursing*, vol. 29, no. 6, pp. 377–382, 2015, doi: 10.1016/j.apnu.2015.06.019.

[42] J. Whitehill, Z. Serpell, Y. C. Lin, A. Foster, and J. R. Movellan, "The faces of engagement: automatic recognition of student engagement from facial expressions," *IEEE Transactions on Affective Computing*, vol. 5, no. 1, pp. 86–98, 2014, doi: 10.1109/TAFFC.2014.2316163.

[43] S. H. W. Chuah and J. Yu, "The future of service: the power of emotion in human-robot interaction," *Journal of Retailing and Consumer Services*, vol. 61, 2021, doi: 10.1016/j.jretconser.2021.102551.

[44] T. Hempel, L. Dinges, and A. Al-Hamadi, "Sentiment-based engagement strategies for intuitive human-robot interaction," *Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 4, pp. 680–686, 2023, doi: 10.5220/0011772900003417.

[45] Y. Feng, Q. Jia, M. Chu, and W. Wei, "Engagement evaluation for autism intervention by robots based on dynamic Bayesian network and expert elicitation," *IEEE Access*, vol. 5, pp. 19494–19504, 2017, doi: 10.1109/ACCESS.2017.2754291.

[46] S. Lemaignan, C. E. R. Edmunds, E. Senft, and T. Belpaeme, "The PInSoRo dataset: supporting the datadriven study of child-child and child-robot social dynamics," *PLoS ONE*, vol. 13, no. 10, 2018, doi: 10.1371/journal.pone.0205999.

BIOGRAPHIES OF AUTHORS

Wael Hasan Ali Almohammed holds a master degree in Information Technology from Universiti Utara Malaysia, Malaysia in 2015. He also received his B.Sc. (Computer Science) from University of Kerbala, Iraq in 2012. He is currently an assistant lecturer at Department of Computer Science, University of Kerbala, Karbala, Iraq. His research includes machine learning, affective computing, human-robot interaction, and network security. He can be contacted at email: wael.h@uokerbala.edu.iq.

Sinan Adnan Muhisn holds a master degree in Information Technology from Universiti Utara Malaysia, Malaysia in 2015. He is currently an assistant lecturer at the Faculty of Biotechnology, Al-Qasim Green University, Babylon, Iraq. His research includes enterprise resource planning (ERP), machine learning, and e-learning. He can be contacted at email: sinan@uoqasim.edu.iq.

Zahraa Abed Aljasim Muhisn holds a master degree in Information Technology from Universiti Utara Malaysia, Malaysia in 2015. She is currently an Assistant Professor at Computer Science, Al-Qasim Green University, Babylon, Iraq. Her research includes machine learning, software engineering, knowledge management, and e-learning. She can be contacted at email: zahraa.a@uoqasim.edu.iq.