

Article

Not peer-reviewed version

---

# In the wild Affect Analysis of Children with ASD using Heart rate

---

[Kamran Ali](#)\*, [Sachin Shah](#), [Charles E Hughes](#)

Posted Date: 21 April 2023

doi: 10.20944/preprints202304.0692.v1

Keywords: Emotion Recognition; smart bracelet; heart rate; wearable; Autism Spectrum Disorder (ASD)



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Article*

# In the Wild Affect Analysis of Children with ASD using Heart Rate

Kamran Ali <sup>1,\*</sup>, Sachin Shah <sup>1,2</sup> and Charles Hughes <sup>1</sup>

<sup>1</sup> Synthetic Reality Lab, Department of Computer Science, University of Central Florida Orlando, Florida; sachin.shah@knights.ucf.edu, charles.hughes@ucf.edu

<sup>2</sup> Department of Computer Science, University of Maryland

\* Correspondence: kamran.ali@ucf.edu

**Abstract:** This paper presents a framework to recognize the affective state of children with Autism Spectrum Disorder (ASD) in an in-the-wild setting using Heart Rate (HR) information. Our algorithm classifies a child's emotion into positive, negative, or neutral states by analyzing the heart rate signal. The HR signal is obtained from a smartwatch in real-time using our smartwatch application. The heart rate data is acquired when the child learns to code a robot while interacting with an avatar that assists the child in communications skills and programming the robot. In this paper, we also present a comparison of using HR data for the classification of emotions with classification based on features extracted from HR signals using Discrete Wavelet Transform (DWT). Our experimental results show that the proposed method produces a comparable performance with the state-of-the-art HR-based emotion recognition techniques, despite the fact that our experiments are performed in an uncontrolled setting as opposed to a lab environment. This work contributes to real-world affect analysis of children with ASD using HR information.

**Keywords:** Emotion Recognition; smart bracelet; heart rate; wearable; Autism Spectrum Disorder (ASD)

## 1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that limits social and emotional skills, and as a result, the ability of children suffering from ASD to interact and communicate is negatively influenced. The Centers for Disease Control and Prevention (CDC) reports that 1 in every 45 children in the US is diagnosed with ASD [1]. It can be difficult to recognize the emotions of individuals with ASD, therefore making it hard to infer their affective state during an interaction. However, new technological advancements have proven to be effective in understanding the emotional state of children with ASD.

In recent years, wearable devices have been used to recognize emotions, detect stress levels, and prevent accidents using behavioral parameters or physiological signals [2]. The low cost and wide availability of wearable devices such as smartwatches have introduced tremendous possibilities for research in affect analysis using physiological signals. An advantage of a wearable device, such as a smartwatch, is its ease of use in real-time emotion recognition systems. There are many physiological signals that can be used for emotion recognition, but heart rate is relatively easy to collect using wearable devices such as a smartwatch, bracelet, chest belt, and headset. Nowadays, many manufacturers have marketed smartwatches that can monitor heart rate employing photoplethysmography (PPG) sensors or electrocardiograph (ECG) electrodes. Heart rate sensors in devices like the Samsung Galaxy Watch, Apple Watch, Polar, Fitbit, and Xiaomi provide a reliable instrument for heart rate-based emotion recognition. Another significant aspect of using heart rate signals for affect recognition is its direct linkage with the human endocrine system and the autonomic nervous system. Thus, a more objective and accurate affective state of an individual can be acquired by using heart rate information. In this work, we employ Samsung Galaxy Watch 3 to acquire the heart

rate signal, as it is more comfortable to wear for participants having ASD compared to wearing a chest belt or a headset.

Previous research shows that heart rate changes with emotions. In [3], Ekman et al. showed that heart rate had unique responses to different affective states. It was found that heart rate increased during the affective state of anger and fear, and decreased in a state of disgust. Britton et al. revealed that heart rate during a happy state is lower than heart rate during neutral emotion [4]. Similarly, Valderas et al. found unique heart rate responses when subjects were experiencing relaxed and fearful emotions [5]. Valdera's experiments showed that the average heart rate is lower in a happy mood as compared to a sad mood.

Similarly, the field of robotics is opening many doors to innovate the treatment of individuals with ASD. Motivated by their deficiencies in social and emotional skills, some methods have employed social robots in interaction with children with ASD [6–8]. Promising results have been reported in the development of social and emotional traits of children with ASD while supported by social robots [9]. Similarly, in [10,11] Taylor et al. taught children with intellectual disabilities coding skills using the Dash robot developed by Wonder Workshop. In this paper, we employ an avatar, in a virtual learning environment, that assists children with autism (ASD) in improving communication skills while learning science, technology, engineering, and mathematics (STEM) skills. In particular, the child is given a challenge to program a robot, Dash™. Based on the progress and behavior of the child, the avatar provides varying levels of support so the student is successful in programming the robot.

One of the primary objectives and applications of this paper is to help a human puppeteer or inform an automation system by automatically recognizing the child's emotions during interactions. Emotion recognition has been applied by many research works to improve the interaction of children with ASD and social robots [1,12–14]. Different modalities, such as facial expressions [1,15–24], body posture [15,25], gestures [26], skin conductance [27,28], respiration [29], and temperature [30] have been used to perform emotion analysis of children with ASD. Heart rate is also used to recognize the emotion of children with ASD, but these methods use HR as an auxiliary signal that is combined with other modalities such as skin conductance [27] or with body posture [15]. In this paper, we propose an emotion classification technique that employs HR as a primary signal.

Most emotion analysis studies of children with ASD use various stimuli to evoke emotions in a lab-controlled environment. The majority of these studies have employed pictures and videos to evoke emotions. However, in [28], Fadhil et al. report that pictures are not the proper stimuli for evoking emotions in children with ASD. Although most studies have used video stimuli, other research works have employed serious games [31] and computer-based intervention tools [32]. In this work, we use a human-avatar interaction in a natural environment where children with ASD learn to code a robot with the assistance of an avatar using an iPad.

The compilation of ground truth labels from the captured data is challenging, laborious, and prone to human error. To tag the HR data according to emotions, many techniques have been used in the literature. For instance, in [33], the study participants use an android application to record their emotions by self-reporting them in their free time. Similarly, in [34,35], the HR signals are labeled by synchronizing the HR data with the stimuli videos. Since the emotion label of the stimuli is known, the HR signals aligned with those stimuli are tagged accordingly. The problem with this tagging process is that it assumes the participants experience the emotion of the stimuli and it is constant for all the participants. However, in [36], Lei et al. reveal that individuals experienced varied emotions to different stimuli.

The ground truth labeling process becomes even more challenging when the data is collected from participants having ASD [12]. For these participants, it is very difficult to accurately determine their internal affective state, and due to the deficits in communication skills in children with autism, the conventional methods for emotion labeling are difficult to apply [12,27]. In this paper, we propose a semi-automatic emotion labeling technique that leverages the full context of the environment in the form of videos captured during the interaction of the participant with the avatar. We use an

off-the-shelf Facial Expression Recognition (FER) algorithm, TER-GAN [37] to produce an initial label recommendation by applying FER on the video frames. Based on the emotion prediction confidence of the FER algorithm, a human, having the knowledge of the full context of the situation decides the final ground truth label. The FER algorithm classifies a video frame into seven classes, i.e., six basic expressions of fear, anger, sadness, disgust, surprise, and happiness, and a neutral state. Similar to [33], we cluster these emotions into three classes: neutral (neutral), negative (fear, anger, sadness, disgust), and positive (happiness). After tagging the children-avatar interaction videos, we use the classical HR and video synchronization labeling technique to produce the ground truth emotion annotation of the HR signal.

After compiling our training and testing dataset, we then extract optimal features from the heart rate signal and input those features to the classifier for emotion recognition. We also provide a comparison between two different feature extraction techniques and perform experiments for intra-subject emotion categorization and inter-subject emotion recognition.

Table 1 summarizes the studies of emotion recognition of participants with ASD. Many of these techniques are complex and not suitable for real applications. In this paper, we propose an emotion classification technique that employs HR as a primary signal leveraging a wearable smartwatch.

**Table 1.** Summary of research on emotion recognition of ASD participants using various sensors.

Ref.	Related Work	Signal Type	Subject Number	Stimulation Materials	Performance
[17]	Grossard et al.	Video	36	Imitation of facial expressions of an avatar presented on the screen	Accuracy: 66.43 % (neutral, happy, sad, anger)
[18]	Coco et al.	Video	5	Video	Entropy score: (happiness: 1776, fear: 1574, sadness: 1644)
[25]	Marinoiu et al.	Body posture videos	7	Robot-assisted therapy sessions	RMSE: (valence: 0.099, arousal: 0.107)
[26]	Kumar et al.	Gesture videos	10	Unknown	F-Measure: (angry: 95.1%, fear: 99.1%, happy: 95.1%, neutral: 99.5%, sad: 93.7%)
[14]	Liu et al.	Skin conductance	4	Computer tasks	Accuracy: 82%
[29]	Sarabadani et al.	Respiration	15	Images	Accuracy: (low/positive vs. low/negative: 84.5% and high/positive vs. high negative: 78.1%)
[30]	Rusli et al.	Temperature (thermal imaging)	23	Video	Accuracy: 88%

The following presents the overall structure of the paper. The introduction section talks about the background, motivation, and main contributions of this paper. The second section presents the experimental details such as the demography of participants, the in-the-wild real-time learning environment, and the interaction of a child with an avatar. The third section discusses the semi-automatic emotion labeling process. The fourth section presents the feature extraction techniques.

The fifth section discusses the emotion recognition step. The sixth section discusses the experimental results. The seventh section compares our results with state-of-the-art HR-based emotion recognition techniques. Finally, the eighth section presents conclusions drawn from this research.

2. Experiment

2.1. Subject Information

A total of nine children (6 male and 3 female) who met the criteria for ASD were recruited for this project. Written parental consent and student assent were acquired from each child prior to participation. All of the procedures were approved, and we obtained both university and school district Institutional Review Board (IRB) approval for the study.

2.2. Interaction of Children with Avatar

Each child is given a task to program a robot, and the avatar interacts with the child to assist in completing this task. During this process, the children will not only learn STEM skills but will also develop a relationship with the virtual avatar through communication. The avatar interacts with the child using an iPad. The time duration of these sessions varies depending on the speed of completion of tasks by the children. Table 2 shows the time of interaction of each child with the avatar to complete the given task.

Table 2. Duration of the completion of the given task by the participants (seconds).

P1	P2	P3	P4	P5	P6	P7	P8	P9
828 s	846 s	786 s	540 s	660 s	480 s	583 s	611 s	779 s

Due to the nature of the task and the interaction with the avatar, the children experience different emotions at various stages of the session. For instance, the child becomes happy or surprised when the steps to program the robot are completed. Similarly, the child feels sad or angry when the robot Dash fails to move based on the child’s intent. We record the videos of these sessions, where each frame contains the face of the participant, the window containing the avatar, the window showing the robot Dash, and the audio of the interaction between the child and the avatar, as shown in Figure 1. Each child wears a smartwatch that collects heart rate information in the form of beats per minute (bpm). To transmit the heart rate data in raw form in real-time, we have developed a smartwatch application using the Tizen OS. Figure 2 shows the Samsung Galaxy watch having our Tizen-based real-time heart rate transmitting app, and our desktop app to receive the heart rate information.

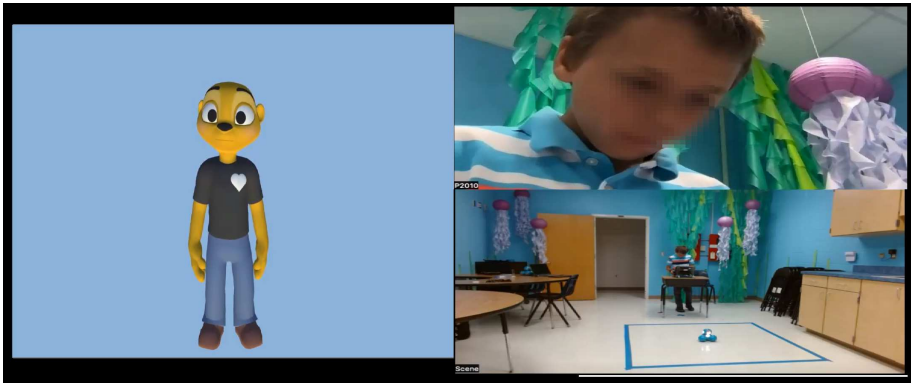
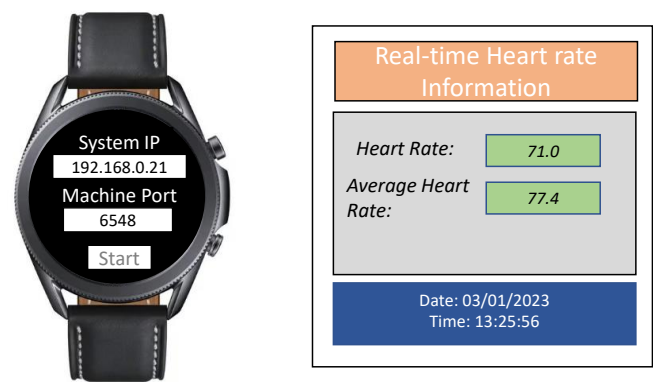


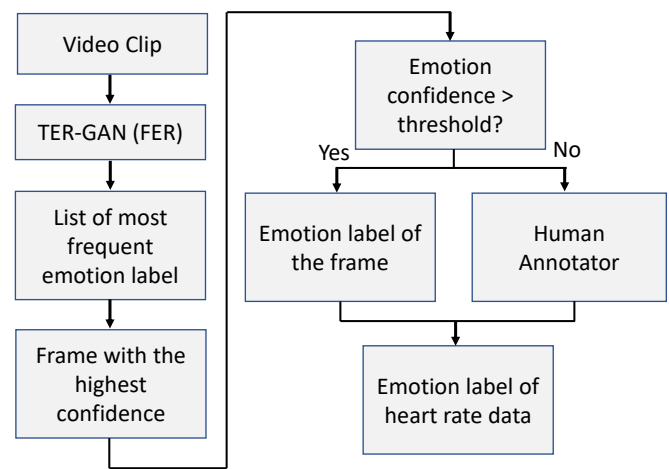
Figure 1. The experimental set-up for the child-avatar interaction.



**Figure 2.** Tizen-based real-time heart rate transmitting app and our desktop app to receive the heart rate information.

**3. Semi-Automated Emotion Annotation Process**

The heart rate data is aligned with the videos of the children working with the avatar to complete the given task. Then, to label the heart rate data, we employ a semi-automated facial expression recognition-based emotion classification algorithm. Figure 3 shows the flow diagram of the semi-automated emotion annotation process. During this annotation process, we leverage the off-the-shelf TER-GAN [37] FER modal, and add some parameters to classify two more emotions, i.e., neutral and contempt, and fine-tune it on the in-the-wild AffectNet dataset. We divide each video into clips containing  $n$  video frames. Then, the video clip is inputted to the FER modal frame by frame to obtain a representative frame for the entire video clip. A representative frame is chosen from the video clip, and the length of the video clip is synchronized with the transmission frequency of the heart rate sensor of the smartwatch. The representative frame is chosen based on two criteria, 1) the label of the representative frame should be the most frequent label, and 2) the prediction confidence of the frame should be the highest in the most frequent list. After automatically obtaining the representative frame and its emotion label, the algorithm decides whether or not to employ a human annotator based on the confidence of the model predicting the emotion label. If the confidence value is lower than a threshold, then the human annotator steps in, and after analyzing the full context of the situation, the final label of the representative frame is assigned. Since the heart rate is aligned with the video data, the emotion label of the video is assigned to the corresponding heart rate data. In this paper, we are interested in classifying three emotions: neutral, positive, and negative. Therefore, the heart rate data is clustered into these three groups.



**Figure 3.** The flow diagram of the semi-automated emotion annotation process.



#### 4. Feature Extraction

The performance of our emotion classifier depends on the quality of the features extracted from the heart rate signal. As mentioned above, one of the main goals of this project is to provide real-time support to a human puppeteer or automated system by classifying the emotion of a child based on heart rate information. Given this goal, we were motivated to avoid delays in the real-time processing of the heart rate signal. As such, we experimented with three different time windows (five seconds, three seconds, and two seconds). Therefore, the heart rate data is in the form of vectors:

$$V = (h_{(t-(n-1))}, h_{(t-(n-2))}, \dots, h_t) \quad (1)$$

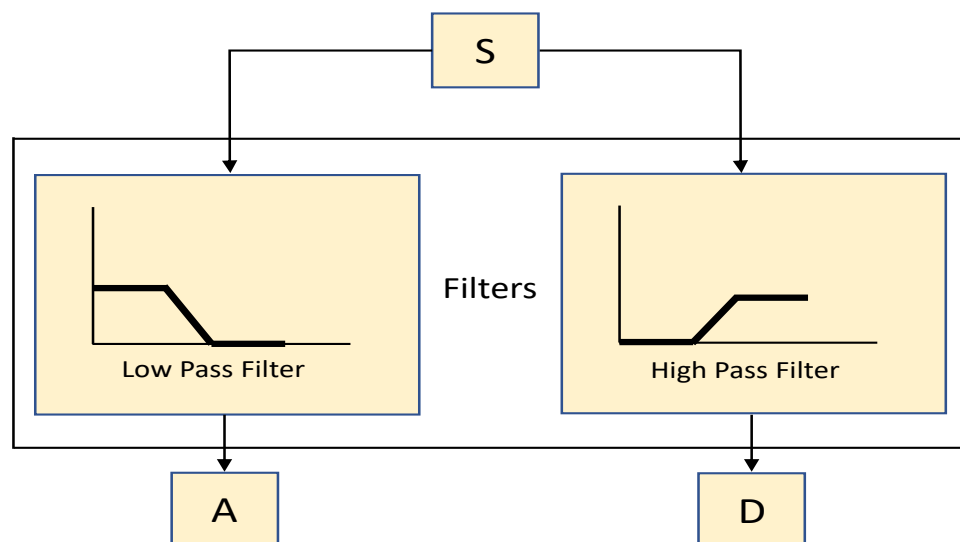
- $h_t$  represents heart rate at time  $t$ .
- $n$  corresponds to the length of the time window.

We then extract features from the heart rate signal for each time interval.

##### 4.1. Discrete Wavelet Transform

Wavelet transform is widely used in signal processing applications to analyze a signal in the time-frequency domain. This mathematical tool is also used by many research works to analyze heart rate data [9,38,39]. The discrete wavelet transform (DWT) is preferred over conventional signal analysis techniques in decomposing the waves into an optimal resolution, both in time and frequency. Thus, there is no requirement that the signal be stationary. Due to these desirable properties, DWT is frequently used in many research works to perform time-scale analysis, signal compression, and signal decomposition.

DWT filters decompose a signal into two bands at any particular level, i.e., approximations and detail bands of a signal. The approximations ( $A$ ) correspond to the low-frequency components of the signal at a high resolution. The details ( $D$ ) are high-frequency components of the signal at a lower resolution. During the sub-sampling process, the components of a signal are divided by 2 for multi-resolution analysis, as shown in Figure 4. The pre-processed heart rate data is inputted to the DWT decompositions. This multi-scale DWT decomposition is also called sub-band coding. The sub-sampling at every scale decomposes a signal into half the number of samples. Figure 5 shows the multiscale decomposition of a signal into sub-bands at various levels. In this paper, we extract DWT features for emotion recognition by decomposing the heart rate signal using the Haar wavelets.



**Figure 4.** The low pass and high pass filtering of the DWT.

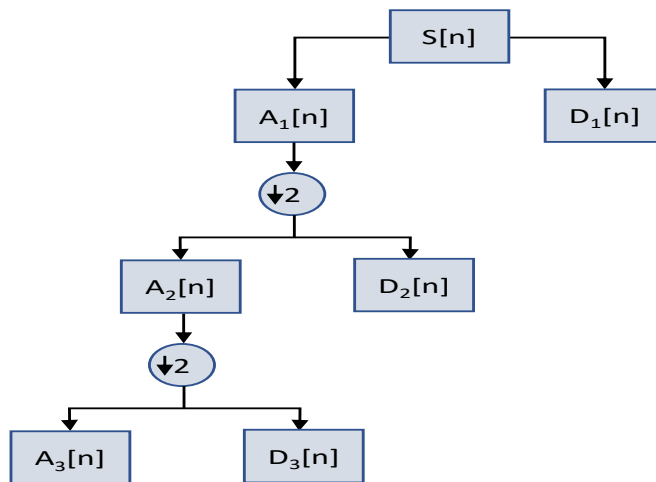


Figure 5. Discrete Wavelet Transform Sub-band Coding.

## 5. Emotion Recognition

After extracting DWT features from the heart rate signal, we use three different classifiers to recognize emotions. We employ SVM, KNN, and Random Forest classifiers for intra-subject and inter-subject classification. Intra-subject emotion classification is performed when the heart rate data from the subject is acquired individually, and then the classifier is trained and tested on the same data. Whereas, in inter-subject emotion recognition, the heart rate data is collected from all participants rather than individually, and is used for the training and testing of the recognition modal. All the experiments were performed using a ten-fold cross-validation method. For comparison purposes, similar to [35], we also run experiments using the heart rate data as the input feature to the classifiers. Figure 6 shows the heart rate signals of three different participants in the negative, positive, and neutral states.

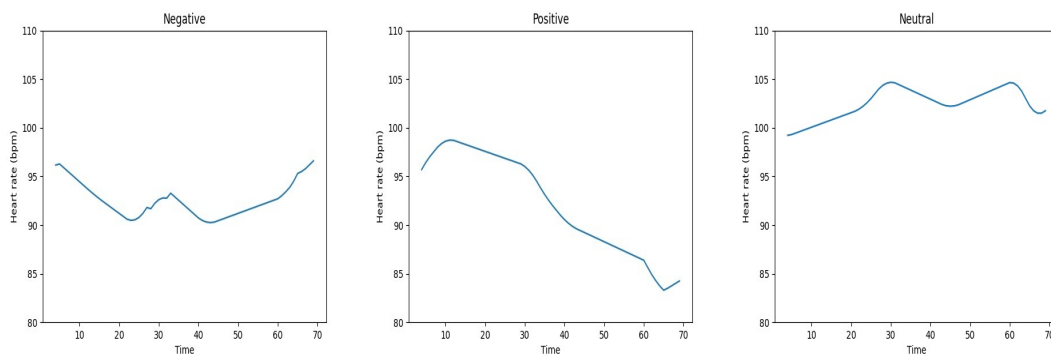


Figure 6. The heart rate signals of three different participants in the negative, positive, and neutral states.

We calculate the classification accuracy of the three classifiers using the following formula:

$$\text{Accuracy} = \frac{TP/TN}{TP + TN + FP + FN} \quad (2)$$

Where TP represents true positive, TN denotes true negative, FP stands for false positive, and FN is the abbreviation of false negative [35].

## 6. Experimental Results and Discussion

Summary statistics of the heart rate data acquired during the interaction of children with the avatar is shown in Figure 7. Figure 7 shows the average heart rate, the minimum heart rate, and the



maximum heart rate of all participants. As it can be seen that during the completion of tasks and the interaction with the avatar, the participants go through a range of heart rate activities. The lowest heart rate of 62 BPM is from Participant 6, and the highest heart rate of 124 BPM is from Participant 1.

Similarly, Figure 8 shows the maximum, minimum, and average beats per minute of all nine participants. The average heart rate of all nine participants is 96.8 BPM, while the maximum heart rate is 124 BPM and the minimum heart rate is 62 BPM.

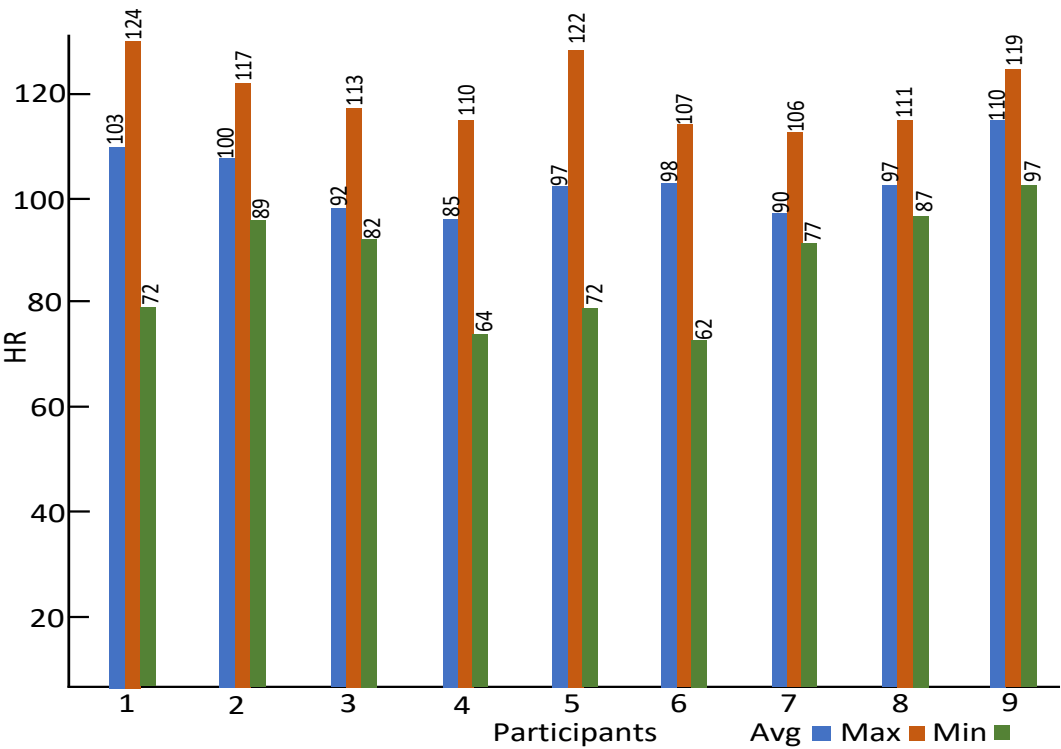


Figure 7. Average, minimum, and maximum heart rate of all participants.

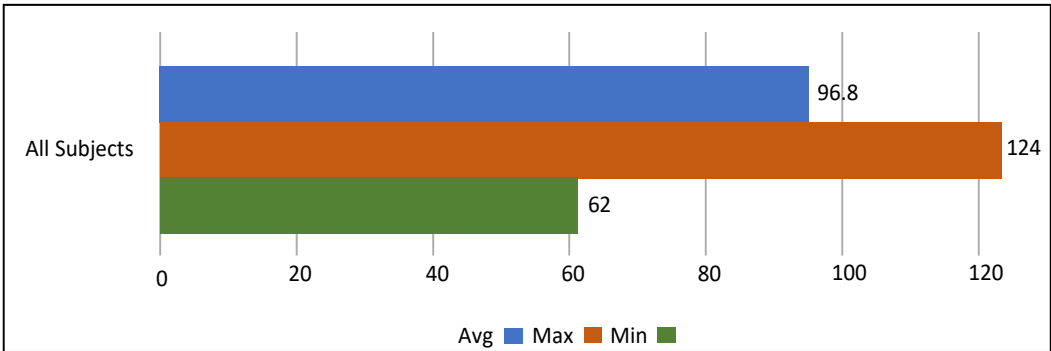


Figure 8. The average, minimum, and maximum heart rate of all participants collectively.

The emotion recognition results of both the intra-subject and the inter-subject data using DWT and heart rate features employing SVM, KNN, and RF classifiers are discussed in the following paragraphs. We performed experiments using three different window sizes and found that the window size of two enhances the performance of our algorithm both in terms of accuracy and speed, which facilitates the real-time application of our emotion recognition technique.

The intra-subject classification accuracies of the three classifiers using DWT features are shown in Figure 9. In the case of SVM, the highest accuracy of 100% is obtained from Participant 6, and the lowest recognition accuracy of 40.1 % is obtained from Participant 4. For KNN, Participant 6 obtained

the highest accuracy of 100 %, and Participant 3's emotion recognition accuracy of 51.4 % is the lowest. In the case of RF, the highest accuracy of 99.5 % is obtained from Participant 6, and the lowest accuracy of 39.2 % is obtained from Participant 9.

Similarly, the intra-subject classification accuracies of the three classifiers using HR data are shown in Figure 10. In the case of SVM, the highest accuracy of 100 % is obtained from Participant 6, and the lowest recognition accuracy of 29.7 % is obtained from Participant 1. In the case of KNN, Participant 5 obtained the highest accuracy of 100 %, and Participant 2's emotion recognition accuracy of 32.1 % is the lowest of all the accuracies. For RF, the highest accuracy of 99.2 % is obtained from Participant 6, and the lowest accuracy of 35.6 % is obtained from Participant 1.

The emotion recognition accuracies using the DWT features for inter-subject classification employing SVM, KNN, and RF are shown in Figure 11. The highest emotion recognition accuracy of 39.8 % is obtained using SVM. While, the emotion classification accuracies produced by KNN and RF are 33.4 % and 35.7 %, respectively. Similarly, the highest classification accuracy of 38.1 % is obtained using SVM with the heart rate data, while RF produced the lowest recognition accuracy of 31.9 % employing heart rate data as an input feature, as shown in Figure 12. Hence, this comparison shows that a slightly better performance in emotion recognition can be achieved by using DWT-based features. Similarly, comparing the intra-subject and inter-subject recognition accuracies, we can see that the inter-subject emotion detection task is much more difficult than the intra-subject emotion classification due to the variation present in heart rate data for each individual.

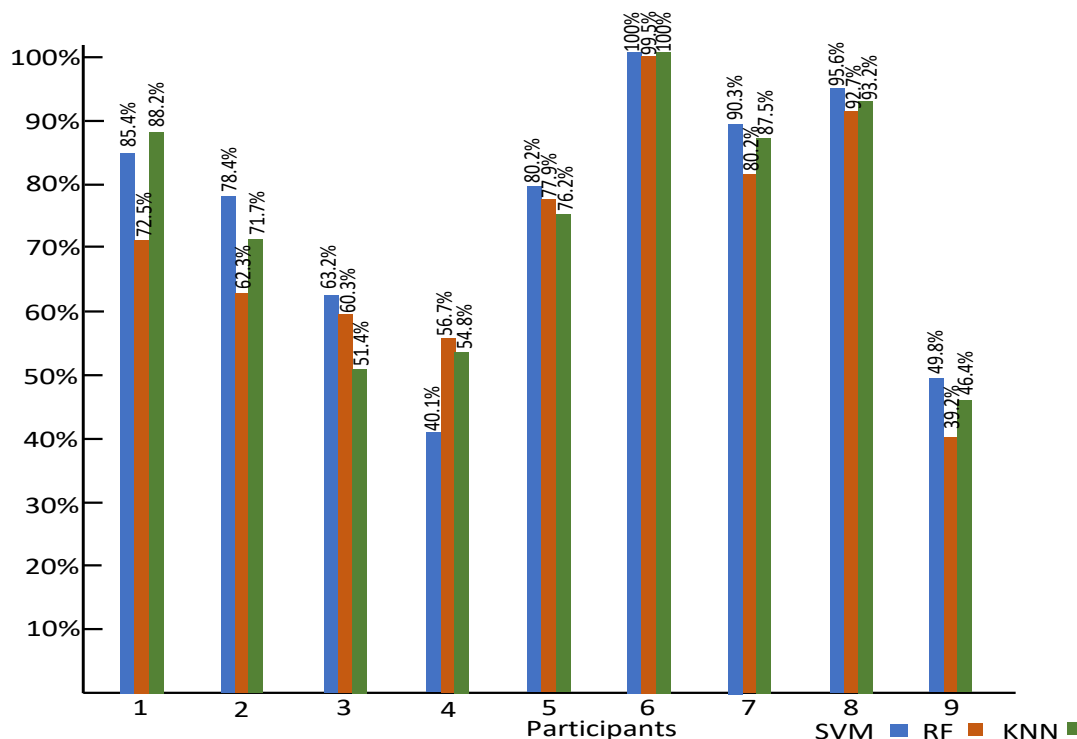


Figure 9. The intra-subject classification accuracies using DWT features.

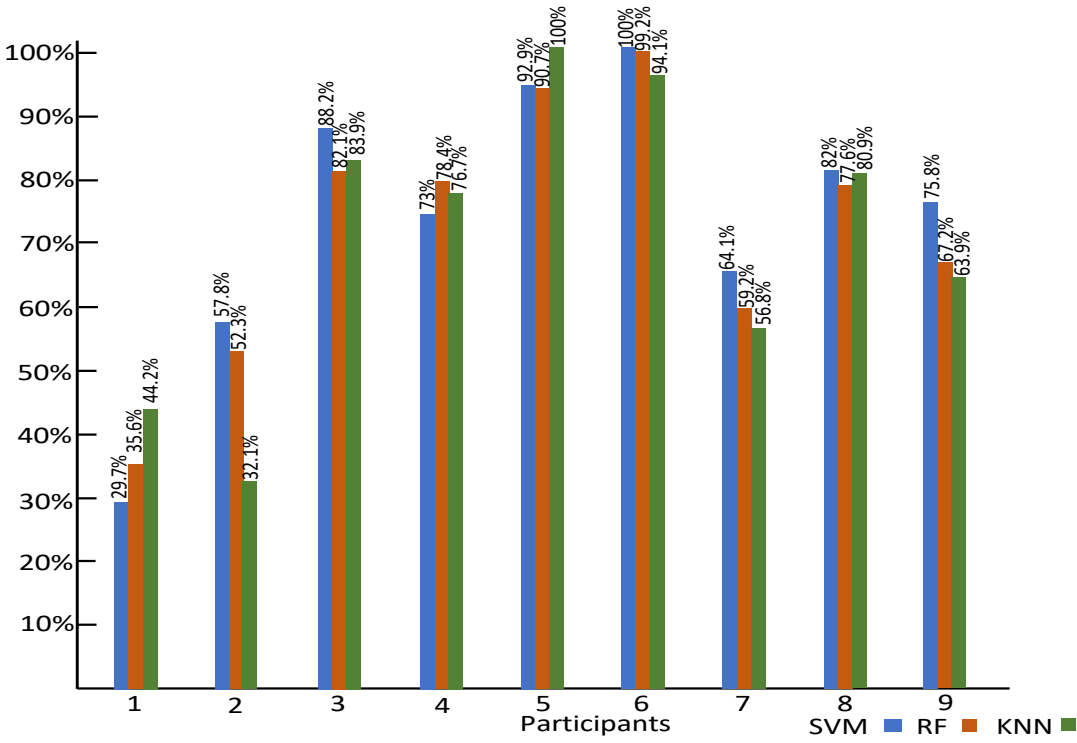


Figure 10. The intra-subject classification accuracies using HR.

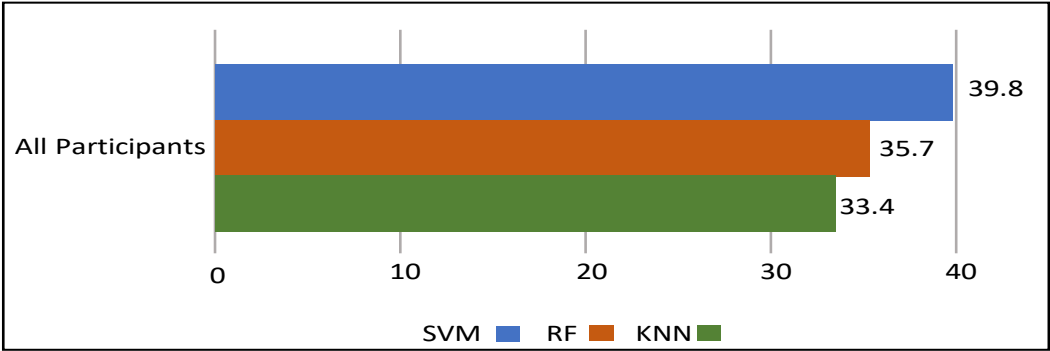


Figure 11. The inter-subject classification accuracies using DWT features.

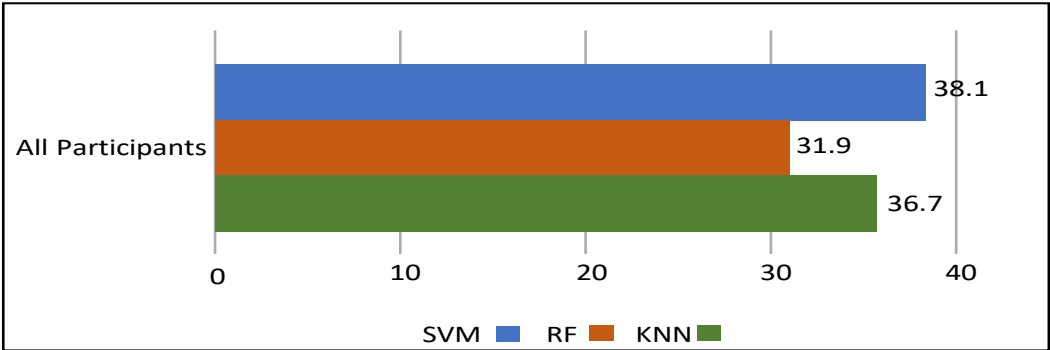


Figure 12. The inter-subject classification accuracies using HR.

7. Comparison with Related Studies

We compare our emotion recognition results with the state-of-the-art HR-based emotion classification techniques in Table 3. As shown in the table, the highest recognition accuracy of

84 % is obtained by [34], while the second highest accuracy of 79 % is produced by the emotion recognition technique proposed in [33]. As reported in [35], the experimental protocol and details of the heart rate data used for the validation of these techniques are not explained in their papers. It is not known whether their emotion recognition algorithm is validated using the intra-subject HR data or the inter-subject HR data. Therefore, a better comparison of our classification accuracy can be performed by comparing our intra-subject recognition accuracy of 100 % with the intra-subject classification accuracy in [35] which is also 100 %. Similarly, the inter-subject recognition accuracy of our technique is comparable with the inter-subject classification accuracy of the technique proposed in [35]. Despite the fact that, in our case, we train and validate our algorithm with the in-the-wild HR dataset obtained during the real-time interaction of the participant with an avatar without well-defined external stimuli and a constant lab environment. Another reason for a little lower recognition accuracy of our technique is that the number of participants in our case is nine, while the emotion recognition algorithm in [35] is trained and tested using a dataset of twenty participants. We note that it is common to have small samples when working with children diagnosed to be on the spectrum versus other populations.

**Table 3.** Comparison with other HR-based emotion recognition techniques.

Author	Participants	Stimuli	Classifier	No. Classes	Accuracy
Shu et al. [34]	25	China Emotional Video Stimuli (CEVS)	Gradient boosting tree decision	3	84 %
Bulagang et al. [35]	20	Virtual reality (VR) 360° videos	SVM, KNN, RF	4	100 % for intra-subject and 46.7 % for inter-subject
Nguyen et al. [33]	5	Android application	SVM	3	79 %
Ours	9	Realtime interaction with avatar	SVM, KNN, RF	3	100 % for intra-subject and 39.8 % for inter-subject

## 8. Conclusions

The main objective of this paper is to develop a real-time heart rate-based emotion classification technique to recognize the affective state of children with ASD while they interact with an avatar in an in-the-wild setting as opposed to a lab-controlled environment. We present a semi-automated facial expression recognition-based emotion annotation technique for the heart rate signal. The emotion labels obtained from our proposed tagging method are then grouped into three clusters: positive, negative, and neutral emotions. To classify an affective mood of a child with ASD into three emotional states, we extract two sets of features from the HR signal using a window size of two seconds and evaluate the effectiveness of these two sets of features on three classifiers, namely, SVM, KNN, and RF. We also compile two types of HR datasets, i.e., an intra-subject dataset and an inter-subject dataset. Our experimental results show that a 100 % classification accuracy is obtained by extracting DWT and HR features from the intra-subject dataset of Participant 6 and inputting these features to SVM. The experiments performed using the inter-subject HR dataset show that the highest emotion recognition accuracy of 39.8 % is produced by using the DWT features with SVM, which is comparable to the state-of-the-art inter-subject HR-based emotion classification technique. The variation in heart rate due to individual differences present in the inter-subject dataset contributes to lower recognition accuracy, as observed by other HR-based emotion recognition techniques.

**Author Contributions:** Conceptualization, K.A. and C.H.; methodology, K.A.; software, K.A., S.S.; validation, K.A.; formal analysis, K.A.; investigation, K.A. and C.H.; resources, K.A. and C.H.; data curation, K.A.; writing—original draft preparation, K.A.; writing—review and editing, K.A., S.S. and C.H.; visualization, K.A.; supervision, C.H.; project administration, C.H.; funding acquisition, C.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported in part by grants from the National Science Foundation Grants 2114808, and from the U.S. Department of Education Grants H327S210005, H327S200009H. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

**Institutional Review Board Statement:** The studies involving human participants were reviewed and approved by the University of Central Florida Institutional Review Board. Further, the study was conducted in accordance with institutional guidelines and adhered to the principles of the Declaration of Helsinki and Title 45 of the US Code of Federal Regulations (Part 46, Protection of Human Subjects). The parents of the participants provided written consent and the children provided verbal assent with the option to stop participating at any time.

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

**Data Availability Statement:** The videos/frames showing the faces of participants will not be made public as the IRB requires that access to each individual's images have explicit parental permission.

**Acknowledgments:** The authors would like to thank Lisa Dieker, Rebecca Hines, Ilene Wilkins, Kate Ingraham, Caitlyn Bukaty, Karyn Scott, Eric Imperiale, Wilbert Padilla, and Maria Demesa for their assistance in data collection and thoughtful input throughout this project.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

FER	Facial Expression Recognition
HR	Heart rate
DWT	Discrete Wavelet Transform
KNN	K-nearest neighbors
RF	Random Forest
SVM	Support Vector Machine

## References

1. Javed, Hifza, Myounghoon Jeon, and Chung H. Park. Adaptive framework for emotional engagement in child-robot interactions for autism interventions. In Proceedings of the International Conference on Ubiquitous Robots, USA, 2018.
2. Fioriello, Francesca, et al. A wearable heart rate measurement device for children with autism spectrum disorder. *Sci. Rep.* **2020**, *10*, 1-7.
3. Ekman, Paul. An argument for basic emotions. *Cogn. Emot.* **1992**, *6*, 169-200.
4. Britton, A.; Shipley, M.; Malik, M.; Hnatkova, K.; Hemingway, H.; Marmot, M. Changes in Heart Rate and Heart Rate Variability Over Time in Middle-Aged Men and Women in the General Population (from the Whitehall II Cohort Study) *Am. J. Cardiol.* **2007**, *100*, 524-527.
5. Valderas, M.T.; Bolea, J.; Laguna, P.; Vallverdú, M.; Bailón, R. Human emotion recognition using heart rate variability analysis with spectral bands based on respiration In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, Italy, 2015.
6. Richardson, K., Coeckelbergh, M., Wakunuma, K., Billing, E., Ziemke, T., Gomez, P., Vanderborght, B., Belpaeme. Robot enhanced therapy for children with Autism (DREAM): A social model of autism. *IEEE Technol. Soc. Mag.* **2018**, *37*, 30-39.
7. Pennisi, P., Tonacci, A., Tartarisco, G., Billeci, L., Ruta, L., Gangemi, S., Pioggia, G. Autism and social robotics: A systematic review. *Autism Res.* **2016**, *9*, 165-183.
8. Scassellati, B., Admoni, H., Matarić, M. Robots for use in autism research *Ann. Rev. Biomed. Eng.* **2012**, *14*, 275-294.
9. Ferari, E., Robins, B., Dautenhahn, K. Robot as a social mediator - a play scenario implementation with children with autism. In Proceedings of the International Conference on Interaction Design and Children, Como, Italy, 2009.
10. Taylor, M. S. Computer programming with preK1st grade students with intellectual disabilities *J. Spec. Educ.* **2018**, *52*, 78-88.

11. Taylor, M. S., Vasquez, E., and Donehower, C. Computer programming with early elementary students with Down syndrome. *J. Spec. Educ. Technol.* **2017**, 32, 149–159..
12. Landowska, Agnieszka, et al. Automatic emotion recognition in children with autism: A systematic literature review. *Sensors* **2022**, 22.4, 1649.
13. Pollreisz, David, and Nima TaheriNejad. A simple algorithm for emotion recognition, using physiological signals of a smart watch. In Proceedings of the international conference of the IEEE engineering in medicine and biology society. South Korea, 2017.
14. Liu, Changchun, et al. Affect recognition in robot-assisted rehabilitation of children with autism spectrum disorder. In Proceedings of the International Conference on Robotics and Automation. Italy, 2007.
15. Rudovic, O.; Lee, J.; Dai, M.; Schuller, B.; Picard, R.W. Personalized machine learning for robot perception of affect and engagement in autism therapy. *Sci. Robot.* **2018**, 3, eaao6760.
16. Pour, A.G.; Taheri, A.; Alemi, M.; Meghdari, A. Human–Robot Facial Expression Reciprocal Interaction Platform: Case Studies on Children with Autism. *Soc. Robot.* **2018**, 10, 179–198.
17. Grossard, C. et al. Children with autism spectrum disorder produce more ambiguous and less socially meaningful facial expressions: An experimental study using random forest classifiers. *Mol. Autism* **2020**, 11, 5.
18. Del Coco, M.; Leo, M.; Carcagnì, P.; Spagnolo, P.; Mazzeo, P.L.; Bernava, M.; Marino, F.; Pioggia, G.; Distant, C. A Computer Vision Based Approach for Understanding Emotional Involvements in Children with Autism Spectrum Disorders. In Proceedings of the IEEE International Conference on Computer Vision Workshops. Italy, 2017.
19. Leo, M.; Del Coco, M.; Carcagnì, P.; Distant, C.; Bernava, M.; Pioggia, G.; Palestra, G. Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Chile, 2015.
20. Silva, V.; Soares, F.; Esteves, J. Mirroring and recognizing emotions through facial expressions for a RoboKind platform. In Proceedings of the IEEE 5th Portuguese Meeting on Bioengineering, Portugal, 2017.
21. Guo, C.; Zhang, K.; Chen, J.; Xu, R.; Gao, L. Design and application of facial expression analysis system in empathy ability of children with autism spectrum disorder. In Proceedings of the Conference on Computer Science and Intelligence Systems, Online, 2021.
22. Silva, V.; Soares, F.; Esteves, J.S.; Santos, C.P.; Pereira, A.P. Fostering Emotion Recognition in Children with Autism Spectrum Disorder. *Multimodal Technol. Interact.* **2021**, 5, 57.
23. Landowska, A.; Robins, B. Robot Eye Perspective in Perceiving Facial Expressions in Interaction with Children with Autism. In Web, Artificial Intelligence and Network Applications; Barolli, L., Amato, F., Moscato, F., Enokido, T., Takizawa, M., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 1287–1297.
24. Li, J.; Bhat, A.; Barmaki, R. A Two-stage Multi-Modal Affect Analysis Framework for Children with Autism Spectrum Disorder. *arXiv* **2021**, arXiv:2106.09199.
25. Marinoiu, E.; Zafir, M.; Olaru, V.; Sminchisescu, C. 3D Human Sensing, Action and Emotion Recognition in Robot Assisted Therapy of Children with Autism. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2158–2167.
26. Santhoshkumar, R.; Kalaiselvi Geetha, M. Emotion Recognition System for Autism Children using Non-verbal Communication. *Innov. Technol. Explor. Eng* **2019**, 8, 159–165.
27. Liu, C.; Conn, K.; Sarkar, N.; Stone, W. Online Affect Detection and Adaptation in Robot Assisted Rehabilitation for Children with Autism. In Proceedings of the RO-MAN 2007–The 16th IEEE International Symposium on Robot and Human Interactive Communication, Jeju, Korea, 26–29 August 2007; pp. 588–593.
28. Fadhil, T.Z.; Mandeel, A.R. Live Monitoring System for Recognizing Varied Emotions of Autistic Children. In Proceedings of the 2018 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Iraq, 9–11 October 2018; pp. 151–155.
29. Sarabadani, S.; Schudlo, L.C.; Samadani, A.; Kushki, A. Physiological Detection of Affective States in Children with Autism Spectrum Disorder. *IEEE Trans. Affect. Comput* **2018**, 11, 588–600.
30. Rusli, N.; Sidek, S.N.; Yusof, H.M.; Ishak, N.I.; Khalid, M.; Dzulkarnain, A.A.A. Implementation of Wavelet Analysis on Thermal Images for Affective States Recognition of Children With Autism Spectrum Disorder. *IEEE Access* **2020**, 8, 120818–120834.



31. Di Palma, S.; Tonacci, A.; Narzisi, A.; Domenici, C.; Pioggia, G.; Muratori, F.; Billeci, L. Monitoring of autonomic response to sociocognitive tasks during treatment in children with Autism Spectrum Disorders by wearable technologies: A feasibility study. *Comput. Biol. Med.* **2017**, *85*, 143–152.
32. Liu, C.; Conn, K.; Sarkar, N.; Stone, W. Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. *Int. J.-Hum. Stud.* **2008**, *66*, 662–677.
33. Nguyen NT, Nguyen NV, My Huynh T, Tran, Nguyen Binh T. A potential approach for emotion prediction using heart rate signals. In Proceedings of the international conference on knowledge and systems engineering, Vietnam, 2017.
34. Shu L, Yu Y, Chen W, Hua H, Li Q, Jin J, Xu X. Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors* **2020**, *20*, 718.
35. Bulagang, Aaron Frederick, James Mountstephens, and Jason Teo. Multiclass emotion prediction using heart rate and virtual reality stimuli *J. Big Data* **2021**, *8*, 1-12.
36. Lei, Jing, Johannan Sala, and Shashi K. Jasra. Identifying correlation between facial expression and heart rate and skin conductance with iMotions biometric platform. *J. Emerg. Forensic Sci. Res.* **2017**, *2*, 53-83.
37. Ali, Kamran, and Charles E. Hughes. Facial Expression Recognition By Using a Disentangled Identity-Invariant Expression Representation. In Proceedings of the International Conference on Pattern Recognition, Online, 2020.
38. Castellanos, Nazareth P., and Valeri A. Makarov. Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis *J. Neurosci. Methods* **2006**, *158*, 300-312.
39. Dimoulas, Charalampos, et al. Long-term signal detection, segmentation and summarization using wavelets and fractal dimension: A bioacoustics application in gastrointestinal-motility monitoring. *Comput. Biol. Med.* **2007**, *37*, 438-462.
40. Glavinovitch, Alexei, M. N. S. Swamy, and E. I. Plotkin. Wavelet-based segmentation techniques in the detection of microarousals in the sleep eeg. 48th Midwest Symposium on Circuits and Systems, 2005.

## Author Biographies

### KAMRAN ALI

Kamran Ali is a postdoctoral research associate in the computer science department at the University of Central Florida, Orlando, FL, 32816, USA. His research interests include computer vision and machine learning. He is the corresponding author of this article. Contact him at kamran.ali@ucf.edu.

### SACHIN SHAH

Sachin Shah is a graduate student in computer science at the University of Maryland, College Park, MD, USA. He is also a researcher with the University of Central Florida Synthetic Reality Lab. His research interests include computational imaging, machine learning, and computer graphics. Contact him at sachin.shah@knights.ucf.edu.

### CHARLES HUGHES

Charles Hughes is a pegasus professor of Computer Science at the University of Central Florida, Orlando, FL, 32816, USA. His research interests include virtual learning environments, computer graphics, machine learning, and visual programming systems. Contact him at charles.hughes@ucf.edu.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.