

Negative Emotion Management Using a Smart Shirt and a Robot Assistant

Minh Pham¹, Ha Manh Do¹, Zhidong Su², Alex Bishop, and Weihua Sheng³

Abstract—Negative affects such as anger, fear, nervousness, depression, etc., may increase human's susceptibility to illness. In this letter, we propose a negative emotion management system that is able to recognize negative emotions through ECG signal and perform emotion regulation through a robot assistant, which has a potential for reducing health risks. A smart shirt is developed to collect the ECG signal from the human body. The robot assistant has the ability to engage in verbal conversations with humans. Recurrence Quantitative Analysis (RQA) is used to extract ECG features for emotion classification purpose. Along with our own dataset, two other public datasets, RECOLA and DECAF, are also used to evaluate our methodology. The detection of negative emotion can trigger the robot assistant to help the user get out of such situations through interactive conversations. We tested and evaluated the proposed framework through experiments. We also assessed the effectiveness of the interactions with the robot on the emotional well-being of older adults.

Index Terms—Human-centered robotics, robot companions, social human-robot interaction.

I. INTRODUCTION

A. Motivation

MENTAL health refers to behavioral, cognitive, and emotional well-being which affects how we feel, think, and act [1]. Positive emotional states correlate with healthy patterns of responding in cardiovascular activity and the immune system [2] while negative emotional states can cause unhealthy patterns of physiological functioning and lowered immune activity. Negative affects such as anger, fear, nervousness, depression,

anxiety, or hostility may increase human's susceptibility to illness [3]. Emotion regulation deficits contribute to the development of depression and anxiety, panic, and stress disorders [4]. To deal with such mental disorders, a common technique is to use cognitive behavioral therapy (CBT), or conversations between a therapist and a patient. Recently virtual therapists or robot assistants such as Woebot [5] and PARO [6] have been developed to deliver CBT. However, it will be even better if robots can prevent people from developing mental illness.

Therefore it is crucial to prevent humans from being in negative emotions for a long time. To this end, we need to solve two problems: 1) how to detect negative emotions, and 2) how to help the human get out of such emotions when detected. Human emotions can be recognized based on facial expression, voice, or body gestures. However, these signals can be consciously controlled by the human and may not reflect a real emotion. Another approach is to recognize human emotion using physiological signals such as heart activity (Electrocardiogram - ECG) [7], brain signal (Electroencephalography - EEG) [8], or galvanic skin response [9], which are more reliable and objective sources of information about human emotion. A person can regulate his or her own emotions, which refers to an intrinsic emotion regulation, by punching a pillow, texting or talking to a friend, taking a nap, or listening to music. On the other hand, a person's emotions can be altered through conversations with other persons, or so-called extrinsic emotion regulation [10]. Social robots have proved their capabilities in providing companionship and improving the users' mood [11]. Therefore we believe that social robots can offer a solution to extrinsic emotion regulation.

This letter presents an Emotion Management System (EMS) which consists of a smart shirt and a robot assistant. The smart shirt is used to collect ECG signal in order to classify emotional states. Once a negative emotion is detected, the robot assistant starts an interactive conversation to distract and help the human out of such negative emotions. The main contributions of this letter are three-fold. First, we proposed a closed loop system for health monitoring and healthcare delivery which proactively intervenes when health situation emerges, while most existing systems focus on health monitoring only. Second, we realized the major components of the EMS using wearable computing and social robotics technologies, which is the first of its kind. Third, we conducted experimental evaluation of the major components of the EMS with different age groups in a lab environment and tested the robot assistant with local seniors. The preliminary results are promising and establish a solid foundation for us to

Manuscript received October 14, 2020; accepted February 28, 2021. Date of publication March 23, 2021; date of current version April 6, 2021. This letter was recommended for publication by Associate Editor W. Johal and Editor G. Venture upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation (NSF) under Grants CISE/IIS 1427345, CISE/IIS 1910993, EHR/DUE 1928711, in part by NSF CISE/IIS 1838808, NSF OIA 1849213, in part by Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (under Grant ICT2021B14), this research is approved by the Oklahoma State University IRB Committee under IRB number EG-133. (Corresponding author: Weihua Sheng.)

Minh Pham is with the Baptist Health South Florida, Coral Gables, FL 33143 USA (e-mail: minhpham@okstate.edu).

Ha Manh Do is with the Louisville Automation & Robotics Research Institute (LARRI), University of Louisville, Louisville, KY 40292 USA (e-mail: ha.do@okstate.edu).

Zhidong Su and Weihua Sheng are with the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: zhidong.su@okstate.edu; weihua.sheng@okstate.edu).

Alex Bishop is with the Department of Human Development and Family Science, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: alex.bishop@okstate.edu).

Digital Object Identifier 10.1109/LRA.2021.3067867

pursue further research in this direction. The letter is organized as follows: the rest of Section I gives a review of previous works on emotion recognition and emotion regulation. Section II presents an overview of the EMS. Section III describes our methodology of negative emotion recognition and emotion regulation. Section IV presents the results of negative emotion recognition and emotion regulation by the robot assistant. Section V gives the conclusion and future work.

B. Literature Review

1) *Emotion Recognition*: Recognizing emotion based on physiological signals is more reliable than external behaviors like facial expression or speech [12]. Signals such as electrical activities of human brain and heartbeat have been used in the research of emotion recognition [8]. Compared to brain signal, the heartbeat is easier to collect in practice. In [13], support vector machine was used to recognize emotions like amusement, sadness, and neutral from physiological signals. The heart rate and galvanic skin response were utilized as the emotional features. In [14], several convolutional and recurrent neural networks were used to predict affect from electrocardiogram and electrodermal activities. A regularized deep fusion framework based on multimodal physiological signals was used for emotion recognition in [15]. Emotion recognition based on physiological signals using convolution neural networks were also proposed in [16] and [17]. ECG-based emotion recognition has recently been adopted in human health research and human-computer interaction (HCI). ECG signal can be collected and processed in many different ways. For example, in [7], the authors proposed a method to perform ECG synthesis, signal decomposition, and feature extraction based on the Empirical Mode Decomposition (EMD) method. They used the International Affective Picture System (IAPS) dataset and their own dataset for emotion recognition. In [18], time and frequency domain analysis is applied for feature extraction. Several nonlinear indices are extracted based on Approximate Entropy, Lagged Pointcare Plot and Detrended Fluctuation Analysis (DFA). Their experimental results show a recognition accuracy of 84.72% on the valence dimension, and 84.26% on the arousal dimension. In [19], the non-linear feature Hurst is computed using Rescaled Range Statistics (RRS) and Finite Variance Scaling (FVS) methods. The accuracy of classifying six emotional states (happiness, sadness, fear, disgust, surprise and neutral) is 92.87%, and 76.45% using random and subject independent validation respectively.

2) *Emotion Regulation*: Emotion self-regulation is considered human's ability to control emotional reactions. However, regulation failure is common, especially when people are in negative emotions [20]. The failure of emotion self-regulation may lead to failures of self-regulation in other behaviors [21]. It is desirable to get the human out of negative emotion once it occurs. One of the solutions is to distract them. In [22], the laboratory studies show that inducing depressed people to focus on positive distractions can reduce their negative affect. Those studies also show that inducing depressed people to be distracted from negative thoughts leads to relatively more positive appraisals of situations, better problem solving, and less

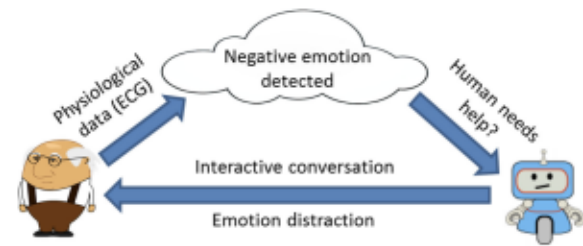


Fig. 1. The emotion management system.

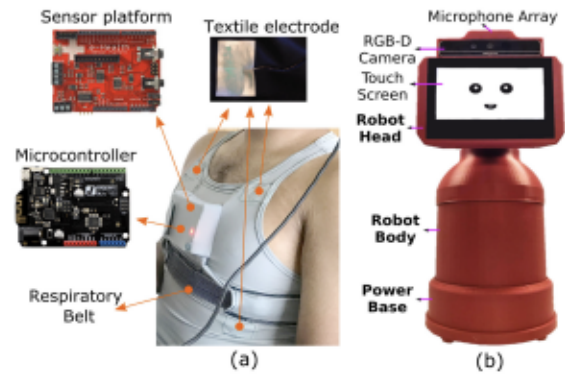


Fig. 2. The smart shirt and the robot assistant.

distress. According to [23], through experiments, the authors claim that people can be distracted from negative emotions by loading their working memory. The more working memory is being used by a distracting activity, the less room will remain for negative emotions to persist. Experimental results from [23] demonstrate that solving a math problem is an efficient way to load working memory. The participants in those experiments reported that their negative emotion was reduced after solving math problems.

II. SYSTEM OVERVIEW

A. Overview

As shown in Fig. 1 the proposed Emotion Management System (EMS) works as follows. Through a wearable unit called smart shirt, human's ECG signal is collected and processed to detect negative emotions. This detection result is shared with the robot assistant. If a negative emotion is detected, the robot assistant engages the human in interactive conversations, therefore forming a closed-loop health monitoring and healthcare delivery system.

B. Smart Shirt and Robot Assistant

The smart shirt consists of textile ECG electrodes, a respiration belt, a sensor platform and a microcontroller as shown in Fig. 2(a). To make it more comfortable, we use textile electrodes and conductive threads to collect physiological signals. Conductive metalized nylon fabric is chosen because it has great conductivity and is easy to sew by using a fusing method. A single lead ECG signal is collected from the garment using the

Lead-II configuration. The signal is amplified and sampled at a rate of 200 Hz, which is sufficient for detecting heart rate, studying heart rate variability and arrhythmias [24]. The respiration belt is an inductive transducer which measures the changes in thoracic or abdominal circumference during respiration. In this work, we only used the ECG signal for emotion recognition because: 1) The respiration data was not provided by the public data sets RECOLA and DECAF; 2) Using only ECG to build models will allow this research to extend to many commercial wearable devices such smartwatches, wristbands that can only collect ECG signals.

As shown in Fig. 2(b), the robot assistant is a 3D printed desktop robot. Its software features a conversational interface that combines speech recognition, spoken language understanding and speech synthesis, which makes it capable of verbal conversations with humans. It consists of three parts: a robot head, a body, and a base. The robot's head has two degrees of freedom and can turn to the direction where the sound is coming from. The robot also features a vision system that consists of an RGB-D camera used for face detection and an auditory system of four microphones allowing speech recognition and sound localization. Details of the robot design can be found in our previous work [25].

III. METHODOLOGY

A. Emotion Recognition

We use the smart shirt to collect the ECG data from the human. The raw ECG signal is filtered using a bandpass filter within the frequency band at 5-15 Hz to remove unwanted noise, for example, electrical and muscle artifacts. Then, features are extracted using a sliding window of 4-second width and 2-second step size. The ECG signal ready for analysis has a sampling rate of 200 Hz. From the ECG data, the heart rate and heart rate variability can be derived by detecting R peaks. Features in time and frequency domain, as well as nonlinear features, are extracted as the input to the classification models. In this work we focus on recurrence quantitative analysis (RQA), which is a method of nonlinear data analysis, to investigate the dynamics of heart's electronic reactivity to emotion changes. There are only a few studies applying RQA on ECG signal to recognize human emotion. Recently, Goshvartpour *et al.* [26] indicated that the RQA is one of the most significant features to differentiate two groups, men or women, based on ECG responses while watching sad images. It has also been shown that nonlinear features have the advantage in physiological signals processing to deal with negative affect.

1) *Recurrence Plot (RP)*: The RQA measures are calculated based on recurrence plots, which is a graph visualizing the times that a phase space trajectory travels to the same place again [27]. The definition of RP is shown below

$$R_{i,j} = \Theta(\epsilon_t - \|\vec{x}_i - \vec{x}_j\|), \quad \vec{x}_i \in \mathbb{R}^m, \quad i, j = 1, \dots, N \quad (1)$$

where N is the number of considered states x_t , ϵ_t is a threshold distance, $\|\cdot\|$ a norm and $\Theta(\cdot)$ the Heaviside function. In our case, only one time series of ECG signal is available, therefore the phase space can be reconstructed by embedding dimension

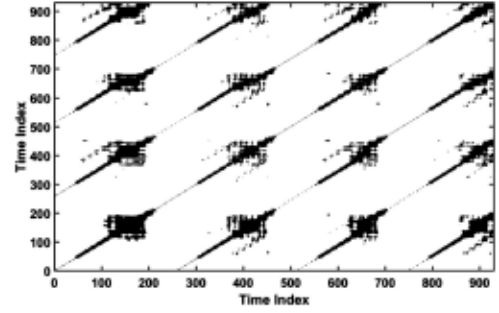


Fig. 3. Recurrence plot of ECG signal with non-negative emotion.

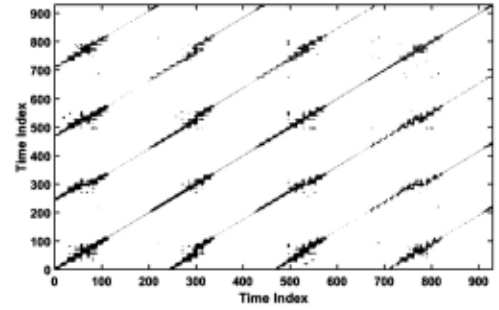


Fig. 4. Recurrence plot of ECG signal with negative emotion.

and time delay.

$$\vec{x}_i = (u(i), u(i + \tau), \dots, u(i + \tau(m - 1))), \quad (2)$$

where $u(i)$ is the time series, τ is the embedding time delay and m is the embedding dimension. To estimate the time delay, we use mutual information analysis [28], and the time delay τ which minimizes the mutual information is selected. The delay time $\tau = 14$ is chosen. The embedding dimension m of the ECG signal is determined by using the False Nearest Neighbors (FNN) method [29]. An appropriate dimension is selected in such a way that most of the nearest neighbors do not move apart significantly in the next higher dimension. The embedding dimension $m = 5$ is selected. Each point of the phase space trajectory x_i is evaluated if it is close enough to another point of the trajectory x_j (less than a specified threshold ϵ). In the 2-D recurrence plot, the states at time i and time j are denoted by black points if $R_{i,j} \equiv 1$ and white points if $R_{i,j} \equiv 0$. Fig. 3 and Fig. 4 show the recurrence plots of 4-seconds ECG data corresponding to non-negative and negative emotion label, respectively, with $\tau = 14$, $m = 5$ and $\epsilon = 0.07$. Based on the recurrence plot, the RQA measures are extracted and used to build the classification models.

2) *RQA Measures*: Based on the RP, the RQA measures such as RR, DET, LMAX, VMAX, ENT, TND, LAM, and TT are extracted [30].

- Recurrence rate RR: The percentage of recurrence points in an RP: $RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j}$
- Determinism DET: The percentage of recurrence points which form diagonal lines:

$$DET = \frac{\sum_{l=1}^N lP(l)}{\sum_{i=1}^N lP(l)}$$

- Longest diagonal line LMAX: The length of the longest diagonal line:

$$LMAX = \max(\{l_i; i = 1, \dots, N_l\})$$

- Longest vertical line VMAX: The length of the longest vertical line:

$$VMAX = \max(\{v_i; i = 1, \dots, N_v\})$$

- Entropy ENT: The Shannon entropy of the probability distribution of the diagonal line lengths $p(l)$:

$$ENT = - \sum_{l=l_{\min}}^N p(l) \ln p(l)$$

- Trend TND: The paling of the RP towards its edges:

$$TREND = \frac{\sum_{i=1}^N (t - \tilde{N}/2)(RR_i - (RR_i))}{\sum_{i=1}^N (t - \tilde{N}/2)^2}$$

- Trapping time TT: The average length of the vertical lines:

$$TT = TT = \frac{\sum_{v=v_{\min}}^N vP(v)}{\sum_{v=v_{\min}}^N P(v)}$$

- Laminarity LAM: The percentage of recurrence points which form vertical lines:

$$LAM = \frac{\sum_{v=v_{\min}}^N vP(v)}{\sum_{v=1}^N vP(v)}$$

Where

N ? the number of points on the phase space trajectory.

N_l - the number of diagonal lines in the recurrence plot.

N_v ? the number of vertical lines in the recurrence plot.

$P(l), P(v)$ - the histogram of the line lengths of diagonal/vertical lines.

$\tilde{N}$ - the maximal number of diagonals parallel to the LOI which will be considered for the calculation of TND.

3) *Classification*: The emotional state labels include negative (1) and non-negative (0). In the study, the negative and non-negative labels are categorized based on the valence values of individual emotions provided in the public data set RECOLA and DECAF. RECOLA dataset [31] was obtained from 23 participants, in which their audiovisual and physiological data were collected while performing experiments. The video and audio data were recorded at 25fps and 44.1 kHz, respectively. The bio-signals include ECG and EDA (Electrodermal Activity). In addition, the participants' personal information, such as age, gender and mother tongue, were also recorded. The annotations were performed on the ANNEMO web-based annotation tool by six people, three females and three males. The label of emotions, presented as valence and arousal, was marked with a framerate of 40 ms. DECAF dataset [32] was obtained from 30 participants (14 females and 16 males). The recorded biosignals include brain signal (Magnetoencephalogram -MEG), ECG, Bi-polar horizontal Electro-oculogram (hEOG) and Bi-polar trapezius Electro-myogram (tEMG). The Near-Infra-Red (NIR) facial data were also recorded. The emotional status was elicited by letting the participants watch different music videos and movie video clips. We classified the emotions with valence less than 0 as negative emotions, which include anger, disgust, fear and sadness. Emotions with valence greater than or equal to 0 are grouped in non-negative category, which includes calm, happiness, and excitement.

We apply the same methodology of data preprocessing and feature extraction to the RECOLA and DECAF dataset for comparison purpose. However, we need to do some extra data processing on emotion labels. In the RECOLA dataset, an emotion in terms of valence and arousal is labeled by six annotators every

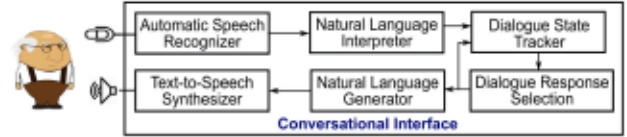


Fig. 5. The conversational interface of the robot assistant.

40 ms. Our models used RQA measures extracted from 4-second width windows of ECG signal. Thus, we had to determine a label of emotion for every 4-second window, and we used different thresholds (70% 80% and 90%) to decide to keep or reject that data. We would reject it if, in that 4-second window, there is no label with the outcome exceeding the thresholds. The valence value fluctuates significantly in small amounts of time, which is not feasible for the heart signal to react properly to the emotion changes. To deal with this issue, first, we took the average of all valence values evaluated from six annotators. The inter-rater agreement has been evaluated by the authors of the RECOLA and DECAF dataset. A median filter is applied to smoothen the valence, where we used the MATLAB function MEDFILT1 with a window size of 200 ms. The valence values are then converted to a binary format with 0 as the threshold point. A sliding window with 4-second width and 2-second step size is used to loop through the labeled data. A window threshold σ is used, which determines the unique label of the whole window size of the data, i.e. 1000 data points. If there are more than $\sigma\%$ of 1000 data points with the same label, then that label is assigned to the window. Otherwise, the window of data is not used for training. The σ value is chosen at 70%, 80% and 90%, respectively. In the DECAF dataset, an emotion label is assigned to a whole session of the experiment in which a person watches a movie clip. In this case, the label is kept unchanged and the preprocessing method is applied the same way as we do on our dataset.

4) *Emotion Recognition Method*: We use the following machine learning algorithms for emotion recognition: neural network, decision tree, gradient boosting and SVM RBF. Those respective models are built by using the SAS software. A 5-fold cross validation strategy is used to train the models.

B. Emotion Regulation

The robot initiates interactive conversations with the human in order to help the human get out of negative emotions if detected. A conversational interface (CI) is developed in the robot, which provides functions such as converting voice to text, analyzing intents and actions, generating responses and converting text to speech. As shown in Fig. 5, the CI features a speech recognizer, a natural language interpreter, a dialogue state tracker, dialogue response selection, a natural language generator, and a text-to-speech synthesizer. Automatic Speech Recognizer converts a speech signal to a sequence of words using Google Cloud Speech APIs. Natural Language Interpreter is responsible for extracting the structure and meaning of input text, which is enabled by Google Cloud Natural Language APIs. Dialogue State Tracker and Dialogue Response Selection are

supported by the Google API AI. We combine the API AI with the expression evaluation in our emotion regulation system. This allows the conversations to be tracked properly. Natural Language Generator is used to generate textual summaries from the input data set. Generally, it consists of data analysis and text generator. In our case, the input data is only from what the human says during the conversation. Text-to-Speech Synthesizer is used to convert normal language text into speech with proper tones of voice. Google text-to-speech is used in the robot assistant.

According to the literature [23], an efficient way to distract humans is to make them actively use their working memory area, for example solving a simple mathematics problem. Therefore, in this study, our idea is to enable the robot to engage the human in interactive conversations. We develop several types of conversation with different complexity levels to adapt to different intensities of the negative emotions the human may have. The conversations between the human and the robot can be used for analysis and improving the interactive conversations in the future. A sample conversation between the robot and the human runs as follows.

- *Robot: Hi Minh, how are you doing now?*
Human: I am a little bit sad.
- *Robot: I am sorry that you feel sad. Can you join me playing something? Can you give me two numbers, please?*
Human: It's OK. Number 7 and number 3.
- *Robot: What is the summation of 7 and 3?*
Human: It is 10.
- *Robot: Awesome, it's correct. What are you feeling now, Minh?*
Human: I am still sad.
- *Robot: I am sorry that you are still sad. Can you do multiplication of 7 and 3 and plus 5.*
Human: It is 24, correct?
- *Robot: No, it's wrong. Can you try one more time?*
Human: Oh, it is 26.
- *Robot: Awesome, it's correct. How do you feel now, Minh?*
Human: I feel better now.
- *Robot: I am glad that you feel better now. Do you want to chat with me for a while?*
Human: No, I have to go out.
- *Robot: Ok, bye for now.*

IV. EXPERIMENTAL EVALUATION

A. Emotion Recognition Experiment

1) *Experimental Setup:* In order to elicit emotion and collect data from the participants, a set of hardware components including a smart shirt, an Oculus Rift, and a keypad are used, as shown in Fig. 6. Recent development of Virtual Reality (VR) technology provides a 3D immersive environment and offers a better solution to inducing user emotion reactions [33]. In our research, we use the Oculus Rift headset and a set of videos to conduct the experiment. By wearing the VR goggle to watch videos, the wearer is isolated from the surrounding environment, so it can minimize unwanted interference from outside affecting his/her emotion. This equipment works with 3D videos, 360° videos, as well as normal 2D videos. Compared to other studies



Fig. 6. Emotion elicitation using Oculus Rift.

using traditional methods of watching 2D videos, our method makes it easier to induce the participant's emotion. Videos are manually classified into different categories corresponding to emotional states such as anger, disgust, fear, sadness, neutral, calm, and happiness.

2) *Data Collection:* The experiment is conducted in a dark room so that the participants can be immersed in the virtual environment as much as possible. There are totally ten male participants taking part in this data collection activity, who are graduate students of age between 26 and 40. The current version of the garment prototype we developed did not work well with female bodies in terms of data acquisition quality. Therefore, only male participants were recruited. As part of the future work, we will redesign the garment to make it fit for female bodies. Each participant wears the smart shirt with textile electrodes integrated to collect the ECG signal while watching the videos. We searched 3D and 360-degree videos on YouTube using emotional power words in each emotion category. The total length of the collected videos is 132 minutes 36 seconds. We had 3 people who watched the videos and assessed them. We collected 43 videos with different categories. The lengths of the videos are around 2 to 4 minutes. Finally, 43 videos were selected for the experiment, including: anger:6, disgust:8, fear:8, sadness:7, calm:8, happiness:6, excitement:6. During the experiments, each participant was asked to watch a neutral video before watching a video for emotion elicitation, then he was asked a few questions such as in which parts of the video that he felt sad/disgusted/etc., and the corresponding intensity with a scale of 1 to 10. The number of videos shown to participants were also different because different participants had different reactions to the same video.

To make sure they really experience the desired emotion, on one hand, we use different categories of video. On the other hand, a post-interview is conducted after each video to ask about their feeling. In addition, we equip them with a wireless keypad so that they can press a button to mark the moment when they feel they are experiencing a certain emotion during experiments, as shown in Figure 6. Each video in our collection was used to elicit a single emotion, i.e. anger, disgust, fear, sadness, neutral, calm, or happiness. Thus, only one key was used to mark the emotion. The ECG data at that time is marked with the label of that emotion. All data from the shirt and the keypad are synchronized and sent to a computer wirelessly for

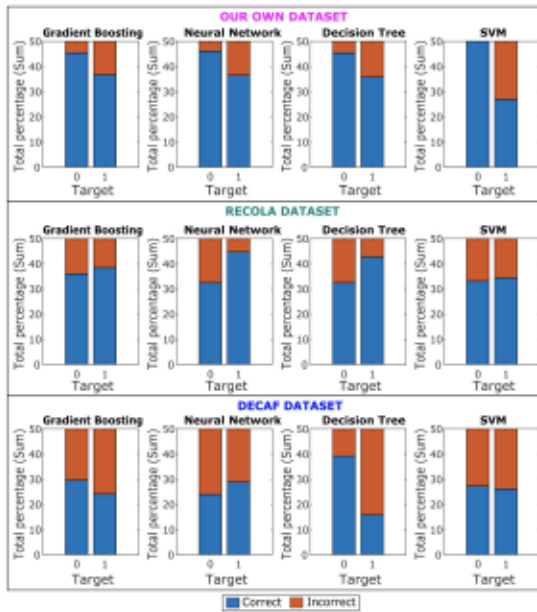


Fig. 7. Emotion recognition results on our own dataset, RECOLA dataset, and DECAF dataset.

TABLE I
RESULT OF NEGATIVE EMOTION CLASSIFICATION

	Our Dataset	RECOLA			DECAF
		70%	80%	90%	
Neural Network	0.821	0.771	0.65	0.599	0.53
Decision Tree	0.813	0.747	0.638	0.63	0.60
Gradient Boosting	0.828	0.741	0.7	0.598	0.533
SVM RBF	0.769	0.671	0.639	0.609	0.529

storing and data analysis purposes. Ten participants were asked to watch videos in different categories. We selected 43 videos and a normal movie had a 3-minute length. Each participant watched 1 or 2 videos of each category (anger, disgust, fear, and sadness, calm, happiness, and excitement). Therefore, for each of the 43 videos, there are 15 observations on average. The total watching time is approximately 310 minutes, but only 120 minutes were marked with labels, and the corresponding ECG signals were used for data processing.

3) *Results*: A summary of classification results on our dataset, RECOLA dataset, and DECAF dataset is shown in Fig. 7. Fig. 7 is a visualization of the confusion matrices when applying different models on different data sets. Class 0 means Non-Negative emotion and Class 1 means Negative emotion. Before applying the models, to avoid bias in classification, we performed under-sampling to make the number of observations of each class similar to each other, which means 50% negative and 50% non-negative. Blue portions are true positive and true negative. Red portions are false positive, and false negative. As shown in Table I, for our dataset, the Gradient Boosting model has the highest accuracy of classification, which is 82.8%. The Neural Network results in the highest accuracy when it is applied

to the RECOLA dataset with the window threshold $\sigma = 70\%$. When applying classification models on the DECAF dataset, however, the highest accuracy is only 60%.

The difference in classification accuracy among the three datasets may be caused by the difference in the way of eliciting and labeling emotions. In our data collection process, using the Oculus Rift offers an immersive VR environment to elicit expected emotions. An emotion label is marked by the participant while watching movie clips, and the corresponding one-minute ECG signal is marked too. That one-minute ECG reflects the emotion changes, and the training data built up from such ECG signals results in higher accuracy. This result also shows that our method offers a better way to collect data for emotion recognition using ECG data.

B. Emotion Regulation Experiment

1) *Experimental Setup*: Another experiment was conducted with the same ten subjects in the data collection phase to evaluate the interactive conversation. The purpose of this experiment is to make the participants have negative emotion by watching the video through the Oculus Rift. The binary classification model which was built in the previous phase was applied to the real-time ECG data to classify emotion. The robot engaged the participant in a conversation if a negative emotion was detected. Finally, a survey was conducted in which they were asked if they really experienced a certain emotion with a level of intensity (from 1 to 10), and if they felt distracted or not after the conversation with the following choices: 1) Not distracted at all or 2) Distracted (or somewhat distracted).

2) *Results*: The survey result shows that out of ten people, six felt sad and their mean arousal is 5.3 out of 10. Four felt disgusted and their mean arousal is 7.3 out of 10. Our experiment shows that 90% of people got somewhat distracted or distracted and that people who had higher intensity were more likely to be distracted. The result is shown in Fig. 8. In order to test the whole system, we asked each participant to wear the smart shirt and watch a sadness video played by the robot. The ECG signal was collected for emotion recognition. As demonstrated in Fig. 9, when a sad emotion was detected, the robot started an interactive conversation to ask him or her to solve simple mathematics problems until she or he got out of the sad emotion. The experiments show that the proposed EMS can be used for negative emotion management.

C. Effectiveness Experiment of the Robot as Companions for Older Adults

1) *Experimental Setup*: The robot was used in experiments to further assess the impact of a brief social interaction between the robot and older adults on subjective ratings related to emotional well-being. In this study, 26 community-dwelling older adults in a mid-sized community in North-Central Oklahoma, aged 60 to 92, were recruited to take part in an experimental study. These 26 participants included 17 females and 9 males, who are primarily identified as White/ Caucasian (73.08%). On average, most participants reported having good overall health, vision, and hearing.

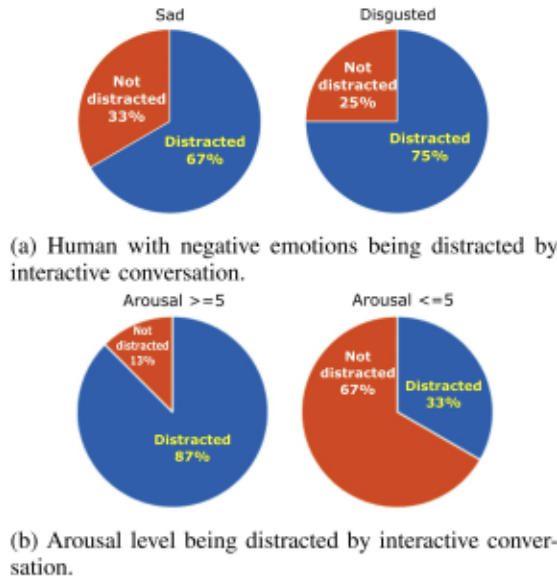


Fig. 8. Experimental result.

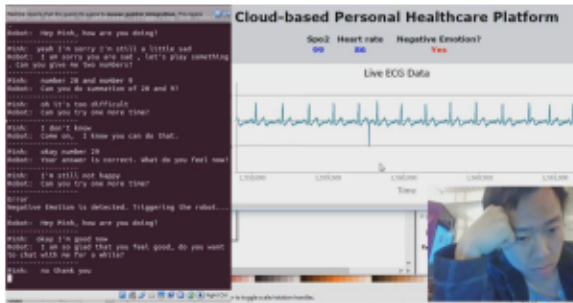


Fig. 9. A dashboard monitoring the real-time physiological signals (right) and an interactive conversation being triggered after the negative emotion is detected (left).

Prior to participation, all individuals read and signed a university approved IRB informed consent. Prior to the engagement phase, participants completed a pre-survey which included questions regarding personal demographics and self-reported loneliness, fatigue, and affect before completing condition-specific activities. Next, participants completed their activities during the engagement phase. Participants engaged in a one-on-one interaction with the robot and completed various activities with the robot, including conversational tasks (e.g., asking about the time, weather, and news), hearing an inspirational quote, telling knock-knock jokes, playing Rock-Letter-Scissors, doing a simple math problem, and listening to music. After the engagement phase concluded, all participants completed the post-survey, which was identical to the pre-survey.

Perceptions of loneliness and social isolation were measured using the UCLA Loneliness scale [34]. To address current emotional well-being (affect), participants completed the PANAS [35]. Half of the questions address positive affect (PANAS-P) whereas the other half address negative affect (PANAS-N). Participants rated current positive or negative emotions (e.g., upset, hostile, inspired, attentive) on a Likert-type

scale from 1 (Never) to 5 (Always). Summary scores for each subscale were calculated with high scores indicating high levels of positive/negative affect. Fatigue was measured using the Iowa Fatigue Scale (IFS) [36] which is an 11-item measure adapted to address current levels of fatigue. The IFS assesses overall fatigue relating to cognition, drowsiness, energy levels, and productivity. Participants rated how much each statement represented how they were feeling using a 5-point Likert-type scale from 1 (Not at all) to 5 (Extremely). A summary score was calculated in which a high score reflected high fatigue and a low score represented low fatigue.

We utilized innovative statistical methods to adopt a more person-centered analytical approach. Data were analyzed with concatenate ordinal analyses using Observation Oriented Modeling (OOM) [37]. The resulting statistics from OOM analysis include a percent correct classification (PCC) index, indicative of the percentage of participants whose responses match the predicted pattern, as well as a c-value, or chance value. The chance value uses a series of randomized trials to determine the probability of obtaining the resulting PCC value.

2) **Results:** Using OOM we found 16 of the 24 experimental participants fit the hypothesized pattern and reported decreased feelings of loneliness after interacting with the robot (PCC = 66.67, c-value = .05). For positive affect, OOM analyses indicated that 16 experimental participants reported increased positive affect (PCC = 61.54, c-value = .03). When assessing negative affect, OOM analyses indicated that the experimental condition again matched the hypothesized pattern, such that 18 participants reported decreased negative affect (PCC = 69.23, c-value = .004). Therefore, it appears that the interaction with the robot reduced feelings of negative affect for both conditions fairly equally. Changes in fatigue were assessed using scores on the IFS. OOM analyses revealed that 21 experimental participants reported decreased feelings of fatigue after interacting with the robot (PCC = 80.77, c-value = .002).

V. CONCLUSION AND FUTURE WORK

In this letter, we developed an emotion management system that recognizes human negative emotions based on the ECG signal and regulates the emotion through a robot assistant. The RQA, a nonlinear data analysis method, is used for feature extraction and results in high accuracy of negative emotion detection. Our robot assistant can initiate interactive conversations to help the human get out of negative emotions. Based on our experiment, the accuracy of the negative emotion classification is around 82.8%. The major components of the closed loop system were validated in a lab environment and the companion robot was also used in a test with local seniors to evaluate its effectiveness in promoting emotional well-being. The preliminary results are promising with noticeable improvement on test subjects' mood after the interactions. Our method can be further improved. It is desirable to conduct a large scale test of the complete system in more realistic environments. Also there are situations where negative emotions can be beneficial which is not considered in this work. Another limitation of our work is that the suppression or distraction from negative mood is not a sustainable solution

especially as an intervention by a robotic assistant. The robot needs to be more intelligent to understand the context in which the negative emotion occurs. More reactions or functions should be implemented for the robot so they can be used accordingly. In the future, we plan to combine ECG and audiovisual signals such as facial expression and voice. Such a multi-modality fusion method will allow us to recognize negative emotions more accurately in more circumstances, especially when the ECG quality is not good.

REFERENCES

- [1] MentalHealth.gov, What is Mental Health? Accessed: Mar. 31, 2021. [Online]. Available: <https://www.mentalhealth.gov/basics/what-is-mental-health>
- [2] P. Salovey, A. J. Rothman, J. B. Detweiler, and W. T. Steward, "Emotional states and physical health," *Amer. Psychol.*, vol. 55, no. 1, pp. 110–121, 2000.
- [3] S. Cohen, W. J. Doyle, D. P. Skoner, P. Fireman, J. M. Gwaltney Jr, and J. T. Newsom, "State and trait negative affect as Predictors objective subjective symptoms of respiratory viral infections," *J. of personality and Social Psychol.*, vol. 68, no. 1, pp. 159–169, 1995.
- [4] M. Berking and B. Whitley, "Emotion regulation: Definition and relevance for mental health," *Affect Regulation Training*. Berlin, Germany: Springer, 2014, pp. 5–17.
- [5] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial," *JMIR Ment. Health*, vol. 4, no. 2, p. e19, 2017.
- [6] R. Yu *et al.*, "Use of a therapeutic, socially assistive pet robot (paro) in improving mood and stimulating social interaction and communication for people with dementia: Study protocol for a randomized controlled trial," *JMIR Res. Protoc.*, vol. 4, no. 2, 2015, Art. no. e4189.
- [7] F. Agraftioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 102–115, Jan.–Mar. 2012.
- [8] S. Nasehi, H. Pourghassem, and I. Isfahan, "An optimal eeg-based emotion recognition algorithm using gabor," *WSEAS Trans. Signal Process.*, vol. 3, no. 8, pp. 87–99, 2012.
- [9] M. Liu, D. Fan, X. Zhang, and X. Gong, "Human emotion recognition based on galvanic skin response signal feature selection and svm," in *Proc. Int. Conf. Smart City Syst. Eng.*, 2016, pp. 157–160.
- [10] J. Zaki and W. C. Williams, "Interpersonal emotion regulation," *Emotion* (Washington, D.C.), vol. 13, no. 5, pp. 803–810, Oct. 2013. [Online]. Available: <https://doi.org/10.1037/a0033839>
- [11] S. Sabanovic, C. C. Bennett, W.-L. Chang, and L. Huber, "PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia," in *Proc. IEEE 13th Int. Conf. Rehabil. Robot.*, Jun. 2013, pp. 1–6.
- [12] L. Shu *et al.*, "A review of emotion recognition using physiological signals," *Sensors*, vol. 18, Jun. 2018, Art. no. 2074.
- [13] J. Domínguez-Jiménez, K. Campo-Landines, J. Martínez-Santos, E. Delahoz, and S. Contreras-Ortiz, "A machine learning model for emotion recognition from physiological signals," *Biomed. Signal Process. Control*, vol. 55, 2020, Art. no. 101646.
- [14] G. Keren, T. Kirschstein, E. Marchi, F. Ringeval, and B. Schuller, "End-to-end learning for dimensional emotion recognition from physiological signals," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2017, pp. 985–990.
- [15] X. Zhang *et al.*, "Emotion recognition from multimodal physiological signals using a regularized deep fusion of kernel machine," *IEEE Trans. Cybern.*, pp. 1–14, 2020.
- [16] T. Song, G. Lu, and J. Yan, "Emotion recognition based on physiological signals using convolution neural networks," in *Proc. 12th Int. Conf. Mach. Learn. Comput.*, New York, NY, USA: Assoc. Comput. Machinery, 2020, pp. 161–165. [Online]. Available: <https://doi.org/10.1145/3383972.3384003>
- [17] M. Lee, Y. K. Lee, M.-T. Orcid, Lim, and T.-K. Kang, "Emotion recognition using convolutional neural network with selected statistical photoplethysmogram features," *Appl. Sci.*, vol. 10, 2020, Art. no. 3501.
- [18] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, "Recognizing emotions induced by affective sounds through heart rate variability," *IEEE Trans. Affective Comput.*, vol. 6, no. 4, pp. 385–394, Oct.–Dec. 2015.
- [19] J. Selvaraj, M. Murugappan, K. Wan, and S. Yaacob, "Classification of emotional states from electrocardiogram signals: A non-linear approach based on hurst," *Biomed. Eng. Online*, vol. 12, no. 1, pp. 1–18, 2013.
- [20] T. F. Heatherton and D. D. Wagner, "Cognitive neuroscience of self-regulation failure," *Trends Cogn. Sci.*, vol. 15, no. 3, pp. 132–139, 2011.
- [21] E. J. Leehr, K. Krohmer, K. Schag, T. Dresler, S. Zipfel, and K. E. Giel, "Emotion regulation model in binge eating disorder and obesity—a systematic review," *Neurosci. Biobehavioral Rev.*, vol. 49, pp. 125–134, 2015.
- [22] S. Nolen-Hoeksema, B. E. Wisco, and S. Lyubomirsky, "Rethinking rumination," *Perspectives Psychol. Sci.*, vol. 3, no. 5, pp. 400–424, 2008.
- [23] L. F. Van Dillen and S. L. Koole, "Clearing the mind: A working memory model of distraction from negative mood," *Emotion*, vol. 7, no. 4, pp. 715–723, 2007.
- [24] A. J. Camm *et al.*, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, 1996.
- [25] F. Erivaldo Fernandes, H. M. Do, K. Muniraju, W. Sheng, and A. J. Bishop, "Cognitive orientation assessment for older adults using social robots," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2017, pp. 196–201.
- [26] A. Goshvarpour, A. Abbasi, and A. Goshvarpour, "Do men and women have different eeg responses to sad pictures?" *Biomed. Signal Process. Control*, vol. 38, pp. 67–73, 2017.
- [27] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *EPL (Europhysics Lett.)*, vol. 4, no. 9, pp. 973–977, 1987.
- [28] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, vol. 33, no. 2, 1986, Art. no. 1134.
- [29] M. B. Kennel, R. Brown, and H. D. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A*, vol. 45, no. 6, 1992, Art. no. 3403.
- [30] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, "Recurrence-plot-based measures of complexity and their application to heart-rate-variability data," *Phys. Rev. E*, vol. 66, no. 2, 2002, Art. no. 026702.
- [31] F. Ringeval, A. Sonderegger, J. Sauer, and D. Lalanne, "Introducing the recola multimodal corpus of remote collaborative and affective interactions," in *Proc. 10th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit.*, 2013, pp. 1–8.
- [32] M. K. Abadi, R. Subramanian, S. M. Kia, P. Avesani, I. Patras, and N. Sebe, "DECAF: MEG-based multimodal database for decoding affective physiological responses," *IEEE Trans. Affective Comput.*, vol. 6, no. 3, pp. 209–222, Jul.–Sep. 2015.
- [33] J. Marín-Morales *et al.*, "Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors," *Sci. Rep.*, vol. 8, no. 1, pp. 1–15, Sep. 2018.
- [34] D. W. Russell, "Ucla loneliness scale (version 3): Reliability, validity, and factor structure," *J. Pers. Assessment*, vol. 66, no. 1, pp. 20–40, 1996.
- [35] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The panas scales," *J. Pers. Social Psychol.*, vol. 54, no. 6, 1988, Art. no. 1063.
- [36] A. Hartz, S. Bentler, and D. Watson, "Measuring fatigue severity in primary care patients," *J. Psychosomatic Res.*, vol. 54, no. 6, pp. 515–521, 2003.
- [37] J. W. Grice, *Observation Oriented Modeling: Analysis of Cause in the Behavioral Sciences*. 1st ed. Oxford, UK: Academic Press, 2011.