ROSbag-based Multimodal Affective Dataset for Emotional and Cognitive States

Wonse Jo, Shyam Sundar Kannan, Go-Eum Cha, Ahreum Lee, and Byung-Cheol Min

Abstract—This paper introduces a new ROSbag-based multimodal affective dataset for emotional and cognitive states generated using Robot Operating System (ROS). We utilized images and sounds from the International Affective Pictures System (IAPS) and the International Affective Digitized Sounds (IADS) to stimulate targeted emotions (happiness, sadness, anger, fear, surprise, disgust, and neutral), and a dual N-back game to stimulate different levels of cognitive workload. 30 human subjects participated in the user study; their physiological data was collected using the latest commercial wearable sensors, behavioral data was collected using hardware devices such as cameras, and subjective assessments were carried out through questionnaires. All data was stored in single ROSbag files rather than in conventional Comma-separated values (CSV) files. This not only ensures synchronization of signals and videos in a data set, but also allows researchers to easily analyze and verify their algorithms by connecting directly to this dataset through ROS. The generated affective dataset consists of 1,602 ROSbag files, and size of the dataset is about 787GB. The dataset is made publicly available. We expect that our dataset can be great resource for many researchers in the fields of affective computing, HCI, and HRI.

I. INTRODUCTION

The recent advancements in wearable devices have increased the attention to affective computing and Human-Computer Interaction (HCI). The easy availability of the wearable sensors has allowed for its integration with affective computing and has given rise to intelligent computing devices that can interpret the affective state of users and provide adaptive feedback to them accordingly. For instance, in an autonomous car, the level of autonomy could be dynamically adjusted based on the affective state of the human operator [1]. In addition to the field of HCI, the affective computing has been deeply influencing the field of robotics too, especially Human-Robot Interaction (HRI). For example, in the social robot interaction system, physical conditions of users extracted from cameras (e.g., facial expression and body gestures) and/or physiological states of users collected from sensors used to flexibly change communication methods to reduce human’s antipathy toward the robotics system [2], [3]. The development of these affective state prediction algorithms and estimation methods using machine learning and neural networks has boosted the availability of publically available annotated affective datasets [4]. The datasets have focused on recording the physiological responses of the participants using various stimuli. However, in most of the existing datasets, the data were recorded using laboratory type monitoring devices which are using wired [5], [6].

With the advent of wireless wearable sensors and other commercially available devices like Apple Watch, there has been an increasing interest in monitoring physiological vitals and estimating human’s state based on that. Lately, stress detection was done completely using wearable sensors [7]. To this current trend of monitoring human state using wearable sensors, it has been becoming important to build more physiological datasets based on wearable sensors.

In addition to the physiological sensor dataset, external behavioral information of the human is also useful in the estimation of the affective state [8]. Video recording of the face is one of the most commonly used modalities in affective computing and many datasets include facial data alongside the physiological sensory data [9], [10]. Another external modality that is widely used is the body gesture data [11], and it is a well-studied topic. The use of both physiological and behavioural data together enables the better estimation of the humans affective state [12], [13]. However, the use of behavioral data along physiological data for affective state estimation is not a well studied one. One reason for this is that there is not many datasets or affective state estimation method that combines both the physiological and behavioural data. Therefore, it is important to build multimodal datasets that consists of both physiological and behavioral data.

Estimation of humans affective state for effective HRI has been gaining increased interest in the recent days. The emergence of new robotics middleware (such as Robot Operating System (ROS) [14]) has also played a larger role in growing the variety of HRI research to integrate the robotics system with the affective computing. In ROS the data collected are usually stored as a ROSbag. The ROSbag format has more...
benefits than the CSV format for collecting and analyzing the dataset. Since the ROS ensures to synchronize the recording signals and videos, it is available to easily and directly analyze the dataset by replaying both using a single ROSbag file. Also, the ROS supports various program languages and operating systems, so that users can validate the developing algorithm and programs by connecting the dataset as like in real-time experiments. Plus, the dataset is available to convert to CVS format or others via additional ROS packages. Hence, it is important to build a dataset combining both physiological and behavioural data which is based on ROS.

In this work, we present a ROSbag-based multimodal dataset comprising physiological data measured using wearable devices and behavioral data recorded using external devices. The data was collected from participants through a user study where various stimuli such as images, audio, and workload tasks were used. Fig. 1 outlines how the dataset was created and organized. During the user study physiological responses such as Photoplethysmography (PPG), Blood Volume Pulse (BVP), Heart Rate (HR), Interbeat Interval (IBI), Galvanic Skin Response (GSR), Electro-dermal Activity (EDA), Skin Temperature (ST), and Electromyography (EMG) were measured using commercially available wearable devices. In addition to the physiological sensors, a 3D frontal camera and a side-view camera were used to record face and body gestures, respectively. To investigate implicit behaviors of users, the variations in the keyboard typing and the mouse motion patterns were also recorded. During the study, the participants performed self-assessment of their affective level using questionnaires at the end of each experiment. These self-assessments can be used later for the training of the classifiers.

The main contributions of this paper are two-fold:

- A multimodal dataset comprising of physiological data such as EMG, EDA, BVP, ST, GSR, and PPG and behavioural data from cameras.
- The dataset enhances the currently available datasets with more sensory data. The dataset also presents self-reported values from the participants about their perceived affective state based on which classifiers could be trained.

II. RELATED WORKS

Human affects shape a huge part of the human experience such as attention, learning, memory, and even decision-making which are required to complete tasks. Therefore, understanding and measuring human affects in real time is vital to construct adaptive and context-aware interfaces that could enrich the user experience. To do so, affective computing research investigates how affect sensing and elicitation techniques can build the understanding of affect and contribute to the design of technologies [15]. Two main methods have been used to estimate human emotion and cognition states [16]. The first is to analyze internal human changes by monitoring physiological signals such as ECG, GSR, EMG, and so on. The other method involves human physical signals such as facial expression, gesture, voice, and so on. As human affects are too complex to present with a single signal, many researchers have applied multiple sensors to improve accuracy and reliability of the system [16], [17].

Most affective computing applications use annotated datasets to train machine learning models that recognize human psychological states [18], [17]. The majority of the dataset includes multimodal stimuli which were designed to elicit a particular human affect and sensor data that were collected when a subject was exposed to the stimuli. Depending on how the researchers defined the human affects and what types of sensors they used, characteristics of the annotated datasets are different. Although the independence between emotion and cognition is still a controversial topic [19], the researchers mainly focused on emotion recognition by providing different dimensions of emotion, so the affective dataset are getting increasingly diversified (such as, DECAF [6], DEAP [5], AMIGO [9], WESAD [7], and so on). Most of the existing dataset particularly focused on emotion recognition but did not design a deliberate experimental setting to detect one’s cognitive state which could affect one’s emotional states.

In this regard, we present a dataset for detecting emotional and cognitive states which is collected from various wearable devices that can monitor and collect human physiological and behavioral data in an unobtrusive manner.

III. DESIGN OF USER STUDY

We designed a user study to build a new affective dataset including physiological and behavioral data based on participants’ emotional and cognitive states. All participants were asked to perform two tasks, an emotion elicitation task and a cognitive workload task. This study was approved by the Purdue Universitys Institutional Review Board (Purdue IRB Protocol: #1812021453).

A. Experimental Setup

The user study was conducted in a closed indoor setup as shown in Fig. 2. The participants were seated in front of a screen with the various wearable sensors and other external sensors connected to a developed ROS-based monitoring system. Fig. 1 depicts a diagram of the monitoring system for reading physiological and behavioral data, as well as rating self-assessment. The main laptop behind the screen is to connect all sensors and devices, as well as to execute Graphical user interface (GUI) programs for displaying emotion stimulus sets and a memory test game on the screen. The programs are connected with the ROS to synchronize and save the data to a ROSbag file that is to track and record all rostopic messages communicated within the ROS.

B. Participants

For this user study, we recruited 30 participants from University; the 11 females and 19 males had an age range of 18 to 37 years (mean: 25.1; s.d: 4.497). It was ensured that none of the participants had any skin allergies to metal or plastic, medical history of brain, mental, or heart diseases and vision or muscle impairment, so that all the wearable
Fig. 2: A user study setting. Commercial wearable sensors including Empatica E4, Shimmer3 GSR and PPG, Polar H10, and Myo armband are utilized. Behavioral sensors including a USB camera (for side view), Intel RealSense (frontal depth and RGB images), a microphone, and a mouse & keyboard are also utilized.

devices could be used. The participants were compensated with $10 for their participation.

C. Equipment

As shown in Fig. 2 physiological and behavioral sensors used in the monitoring system are wearable and commercial devices, so it is not limited the participant’s native body movements which are essential for monitoring.

The physiological sensors connected to the monitoring system are as follows:

- **Empatica E4** is a wristband with an array of sensors for physiological monitoring: EDA, BVP, IBI HR, and ST [20].
- **Myo** is an armband that measures the 8-channel EMG signals. It includes the 8 electrodes placed inside the band to measure the 8-channels EMG signals [21].
- **Polar H10** is worn-chest strap wearable measuring the HR via electrodes attached on a participant’s chest [22].
- **Shimmer3 GSR+** measures GSR and the PPG using electrodes that are attached to the fingers [23].

The behavioral sensors included in the monitoring system are as follows:

- **Intel RealSense** is to record 3D-depth and 2D color videos, and mounted on the top of the TV screen for capturing participant’s face [24].
- **USB camera** is a basic camera to monitor the side view of the participants.
- **Mouse & Keyboard** is used to track mouse cursor and monitor pushed keys.
- **Microphone** is to record the participant’s voice.

D. Stimulus

For the emotion elicitation task, the images and the audio clips were taken from the IAPS and IADS which are widely used and validated in the physiology field for provoking specific emotions [25], [26]. We particularly exploited 21 pictures of the International Affective Picture System (IAPS) [27] and 21 audio clips of the International Affective Digitized Sound System (IADS) [28]. We used these visual and auditory stimuli to elicit targeted seven-emotions (e.g., happiness, sadness, anger, fear, surprise, and neutral). Table I shows the finally selected stimulus data for this user study.

The used images and the number of IAPS and IADS are included on the dataset. For the cognitive workload task, we employed dual N-back games [29]. To provoke different levels of cognitive workload (e.g., low, medium, and high), we controlled the number of back steps (N) of games from 1-back to 3-back to adjust the difficulty of the games.

E. Experimental Protocol

In the user study, participants were given two tasks as illustrated Fig. 3. In both tasks, there were 42 rounds (21 rounds using IAPS and 21 rounds using IADS) in the emotion elicitation task using the IAPS and IADS sets, and three rounds in the cognitive workload tasks. Each round lasted for 60 seconds. The first was the emotion elicitation task which was composed of 21 rounds. The participants were asked to look at a white cross on the screen for 10 seconds (called fixation cross), then interact with images for 6 seconds or listen to short audio clips for 6 seconds, and rate their perceived emotion with a 9 point Self-Assessment Manikin (SAM) scale [30]. The images and the audios

TABLE I: Selected stimulus data for basic emotions.

| Type of emotion | IAPS images       | IADS: audios       |
|-----------------|-------------------|--------------------|
| Happiness       | #110, #2070, #2550| #10, #226, #820    |
| Sadness         | #2800, #3230, #3350| #105, #278, #812   |
| Anger           | #4621, #6560, #6840| #106, #290, #420   |
| Fear            | #1120, #1201, #1930| #276, #286, #712   |
| Surprise        | #1616, #3022, #8180| #114, #360, #425   |
| Disgust         | #7380, #9300, #9320| #210, #255, #700   |
| Neutral         | #7080, #7175, #7217| #262, #319, #723   |
TABLE II: Summary of the Dataset.

| Participants | 30 (Female: 11 and Male: 19) |
| Number of ROSbag files | 1,602 files (about 787 GB) |
| Emotion ratings | Arousal, Valence, Dominance, Word rating |
| Workload rating | Mental/Physical/Temporal demand, Performance, Effort, Frustration |
| Physiological signals | PPG from wrist and chest, EDA, IBI, ST, ECG, GSR, and EMG |
| Video types | Frontal face videos (RGB and depth), side view video |

were selected such that they can stimulate various human emotions. Fig. 3a and Fig. 3b explain the procedures of the emotion elicitation task using the images and sound stimulus, respectively.

The second task is the cognitive workload task which consisted of three rounds by presenting different levels of difficulty, low, medium, and high. The participants were asked to complete the Dual N-back games. During the experiment, the humans physiological and behavioral conditions were monitored using the proposed monitoring system in section IIIC. After they completed each session, they were asked to rate their perceived cognitive workload with NASA-Task Load Index (NASA-TLX) [31]. Fig. 3c shows the procedures of the cognitive workload tasks.

IV. DATASET CONSTRUCTION

In this section, we explain the details of the proposed dataset configuration: physiological and behavior sensor data. Table II presents the summary of the dataset.

A. PHYSIOLOGICAL SENSOR DATA

The dataset includes BVP, ST, EDA, and IBI from Empatica E4 sensor with 30Hz sampling time, BVP and GSR from Shimmer3 GSR unit with 30Hz sampling time, HR from Polar H10 with 1Hz, and 8-channel EMGs from Myo armband with 50 sampling time.

Fig. 4 shows an example of physiological data in the dataset (IAPS #1201, P13). The first plot from top is the BVP signals, the second plot is the average of the IBI data, the third plot is the average of the EDA, the fourth plot is the average of ST data. Those data are collected from the Empatica E4 sensor. The fifth and sixth plots are raw PPG and GSR data of the Shimmer3 sensor. The seventh plot is the result of HR data of the Polar H10. The last plot is raw data of 8-channel EMGs of the Myo armband.

In the figures, the gray area indicates the duration when the stimulus was exposed to the participants during the experiments. The left side of the gray area is a baseline section where the participant lies in the fixation section. The right side of the gray area is a self-assessment reporting section for participants to fill the subjective questionnaires out.

Table III summarizes the rostopic message information of the physiological data in the dataset.

TABLE III: List of rostopic messages for physiological sensors.

| Sensor | Name of rostopic messages | Type of rostopic messages |
|--------|--------------------------|---------------------------|
| Empatica E4 | /physiological_data | empatica_e4_msgs/DataArray 1 |
| Shimmer3 | /shimmer3/3ACMR | std_msgs/Float64 |
| | /shimmer3/PPG | std_msgs/Float64 |
| Polar H10 | /polar_2100hrs | std_msgs/Float32MultiArray |
| Myo Armband | /myo_raw/myo_emg | ros.myo/imuArray |
| | /myo_raw/myo_jmu | ros.myo/imuArray |

B. BEHAVIORAL SENSOR DATA

The dataset includes three different kinds of image sequences taken by two cameras. The Intel RealSense camera located at the front captured facial expressions and upper body gestures in 30 frames per second (fps). At the same time, depth camera results separately were recorded in 30 fps. The USB camera at the side of participants obtained induced behavioral responses in 10 fps. As well, the participant’s

1Empatica E4 ROS message: https://github.com/hyeonuukhin/empatica_e4_msgs
2Myo Armband ROS message: https://github.com/dzhu/myo-arom
speech was recorded via a microphone mounted on the participant’s neck for the user study.

The collected experimental data showed that the tasks elicited participants’ emotional and cognitive states. For example, a piece of the proposed dataset with the participant P13 and visual stimulus IAPS#1201 is shown in Fig. 5. Given the recorded stream of participants, as presented in Fig. 5b and 5c, the behavioral data include facial expressions and body movements, which imply emotional reactions.

Table IV summarizes the rostopic messages information of the behavioral data in the dataset.

### TABLE IV: List of rostopic messages for behavioral sensors.

| Devices        | Name of rostopic message | Type of rostopic message |
|----------------|--------------------------|--------------------------|
| Intel RealSense| /camera/color/image_raw  | sensor_msgs/Image         |
|                | /camera/depth/image_raw  | sensor_msgs/Image         |
| CSR camera     | /image_rect_raw          | sensor_msgs/Image         |
| Microphone     | /audio/audio             | audio_common_msgs/AudioData|
| Mouse          | /mouse/position          | std_msgs/Int32Array       |
| Keyboard       | /keyboard/position       | std_msgs/Int32Array       |

V. SUBJECTIVE RATING ANALYSIS

A. SAM Rating in the Emotion Elicitation Task

All participants’ SAM subjective measures (e.g., arousal, valence, and dominance) in each emotion elicitation task are compared to the reference values published in [32], [33]. The results were plotted on a grid map image like Fig. 6(a) where we used Root-Mean Square Error (RMSE) to compare with them. Fig. 6a shows the result of the comparison analysis in the emotion elicitation task using IAPS. Fig. 6b shows the result of the comparison analysis in the emotion elicitation task using IADS. In both figures, the x-axis is the participant’s number from P1 to P30 and the y-axis is the number of the dataset. In order to show the overall results of the comparison analysis of the self-assessments, we displayed the results using gradual colors from blue to red. The closer the index value to 0 (blue) means that the more similar it is to the reference value. On the other hand, the closer the index value to 45 (red) means that the more difference is from the reference value.

For the results of the SAM scales in the emotion elicitation task using IAPS, the lowest similarity of the dataset is #3350 of P3 with RMSE 42.69, and the highest similarity of the dataset is IAPS#3022 of P26 with RMSE 0.04. P25 produced the highest similarity with mean RMSE 1.66, and P28 produced the lowest similarity with mean RMSE 6.81. The overall average of RMSE is 4.26 with a standard deviation (SD) 3.90.

For the SAM scales in the emotion elicitation task using IADS, the lowest quality of the dataset is IADS#286 of P2 with RMSE 44.28, and the highest quality of the dataset is #820 of P26 with RMSE 0.01. P15 produced the highest similarity with mean RMSE 1.69, and P14 produced the lowest similarity with mean RMSE 10.02. The overall average of
Fig. 7: The results of the Dual N-Back game score and NASA-TLX questionnaire according to the level of the workload.

**A. Examples of replaying the dataset**

Since the dataset is encapsulated into the ROSbag files, the dataset can be easily played back in any ROS-compatible robot system, such as ROS system in Linux system, and Matlab.

For using the ROS system, users should install the ROS on Linux, then decompress the compressed dataset. An example of reading a ROSbag file on Linux system is below:

```
$ rosbag decompress [rosbag_name].bag
$ rqt_bag or rosbag play [rosbag_name].bag
```

For using Matlab, user should install ROS toolbox\(^3\) that is capable of accessing the ROS and exchanging data. An example of reading a ROSbag file in Matlab is below:

```
%% Read a rosbag file
input_bag = rosbag('[:rosbag_name].bag');
%% Display available topics included in the rosbag
input_bag.AvailableTopics
%% Display all message data along with time stamps.
input_bag.MessageList;
%% Select topics from the all message list
selected_topic = select(input_bag, 'Topic', '[:rosbag_name]');
selected_topic_msgStructs = readMessages(
    selected_topic, 'DataFormat', 'struct');
```

**B. Applications: Emotion Analysis**

Emotion analysis was conducted from the recorded videos at the front and side to elicit the emotions. Open-source libraries can extract the feature of posture and facial expressions. When it comes to facial expressions, Face Emotion Recognition (FER) [34], which is one of the python language-based libraries, performed emotion prediction.

Fig. 8b shows computed emotion based on the facial expressions (P3’s IAPS#1201). The gray area in Fig. 8b shows emotion recognition (FER) of the facial expressions (P3’s IAPS#1201).

---

3ROS Toolbox in Matlab: [https://www.mathworks.com/products/ros.html](https://www.mathworks.com/products/ros.html)
emotions and that other behavioral or physiological features indicate the exposure duration of visual or auditory stimuli. The left side of the gray area is the exposure time with 10 seconds fixation cross. The right side of the gray area indicates the period during the self-assessment. The participant P13 rated the emotion response as ‘Disgust’, pleasure level 2, and arousal level 6 about IAPS#1201. Compared to the highest emotion probability of ‘Happiness’ with 10 seconds fixation cross, not only is the calculated emotion different from self-assessed one, but the facial expressions are also not matched with the SAM scale assessment. This implies that only analyzing facial expressions may not be enough to fully understand human emotions and that other behavioral or physiological features and analyses may need to be combined.

VII. Conclusion and Future Works

In this paper, we introduced a new ROSbag-based affective dataset including physiological and behavioral data depending on the emotional and cognitive states. For building the affective dataset, we designed the user study to stimulate the targeted emotions using IAPS and IADS sets and levels of the cognitive workload using dual N-back games, and executed this study by recruiting 30 participants. In the user study, we recorded the status of the user study including physiological data from the commercial wearable devices and the behavioral data using hardware devices, as well as the results of the subjective questionnaires using SAM and NASA-TLX. All data were saved in single ROSbag files rather than CSV files. This not only ensures synchronization of signals and videos in a data set, but also allows researchers to easily analyze and verify their algorithms by connecting directly to this dataset through ROS. The generated dataset consists of 1,602 ROSbag files, and the size of the dataset is about 787GB. We expect that our dataset can be a great resource for many researchers in the fields of affective computing, HCI, and HRI.

In the future, we will utilize even more (latest) physiological sensors and hardware devices and various psychological experiments related to workload, in order to update the affective dataset. We will also analyze more details of the dataset by extracting features from the collected data and validate the dataset using advanced machine learning techniques to estimate human’s emotional and cognitive states.

ACKNOWLEDGMENT

This work was supported by NSF CAREER Award IIS-1846221.

REFERENCES

[1] Z. Kowalczyk, M. Czubenko, and T. Merta, “Emotion monitoring system for drivers,” IFAC-PapersOnLine, vol. 52, no. 8, pp. 200–205, 2019.
[2] A. Bera, T. Randhavane, R. Prinja, and D. Manocha, “Sociosense: Robot navigation amongst pedestrians with social and psychological constraints,” in 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2017, pp. 7018–7025.
[3] A. Bera, T. Randhavane, R. Prinja, K. Kapsasakis, A. Wang, K. Gray, and D. Manocha, “The emotionally intelligent robot: Improving social navigation in crowded environments,” arXiv preprint arXiv:1903.03217, 2019.
[4] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, “End-to-end multimodal emotion recognition using deep neural networks,” IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1301–1309, 2017.
[5] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, “Deap: A database for emotion analysis; using physiological signals,” IEEE transactions on affective computing, vol. 3, no. 1, pp. 18–31, 2011.
[6] M. K. Abadi, R. Subramanian, S. M. Kiu, P. Avesani, I. Patras, and N. Sebe, “Decaf: Még-based multimodal database for decoding affective physiological responses,” IEEE Transactions on Affective Computing, vol. 6, no. 3, pp. 209–222, 2015.
[7] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, “Introducing wesad, a multimodal dataset for wearable stress and affect detection,” in Proceedings of the 26th ACM International Conference on Multimodal Interaction, 2018, pp. 400–408.
[8] C. Shan, Learning Human Emotion from Body Gesture. Boston, MA: Springer US, 2012, pp. 1887–1889. [Online]. Available: https://doi.org/10.1007/978-1-4419-1428-6_1905
[9] J. A. M. Correa, M. K. Abadi, N. Sebe, and I. Patras, “Amigos: A dataset for affect, personality and mood research on individuals and groups,” IEEE Transactions on Affective Computing, 2018.
[10] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, “Ascertain: Emotion and personality recognition using commercial sensors,” IEEE Transactions on Affective Computing, vol. 9, no. 2, pp. 147–160, 2016.
[11] F. Noroozi, D. Kaminska, C. Corneanu, T. Sapinski, S. Escalera, and G. Anbarjafari, “Survey on emotional body gesture recognition,” IEEE transactions on affective computing, 2018.
[12] O. Schleusing, P. Renevey, M. Bertschi, S. Dasen, J. . Koller, and R. Paradiso, “Monitoring physiological and behavioral signals to detect mood changes of bipolar patients,” in 2011 5th International Symposium on Medical Information and Communication Technology, 2011, pp. 130–134.
[13] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. Hutcherson, C. Nass, and O. John, “Real-time classification of evoked emotions using facial feature tracking and physiological responses,” International journal of human-computer studies, vol. 66, no. 5, pp. 303–317, 2008.
[14] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, “Ros: an open-source robot operating system,” in ICRA workshop on open source software, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.
[15] O. C. Santos, “Emotions and personality in adaptive e-learning systems: an affective computing perspective,” in Emotions and personality in personalized services. Springer, 2016, pp. 263–285.
[16] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang, “A review of emotion recognition using physiological signals,” *Sensors*, vol. 18, no. 7, p. 2074, 2018.

[17] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, “A review of affective computing: From unimodal analysis to multimodal fusion,” *Information Fusion*, vol. 37, pp. 98–125, 2017.

[18] A. Mollahosseini, B. Hasan, and M. H. Mahoor, “Affectnet: A database for facial expression, valence, and arousal computing in the wild,” *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 18–31, 2017.

[19] H. Leventhal and K. Scherer, “The relationship of emotion to cognition: A functional approach to a semantic controversy,” *Cognition and emotion*, vol. 1, no. 1, pp. 3–28, 1987.

[20] “Real-time physiological signals - e4 eda/gsr sensor.” [Online]. Available: https://e4.empatica.com/e4-wristband

[21] “Myo armband.” [Online]. Available: https://support.getmyo.com/hc/en-us

[22] “Polar h10: Heart rate monitor chest strap.” [Online]. Available: https://www.polar.com/us-en/products/accessories/h10_heart_rate_sensor

[23] “Shimmer3 gsr unit.” [Online]. Available: https://www.shimmersensing.com/products/shimmer3-wireless-gsr-sensor

[24] “Intel realsense depth and tracking cameras.” [Online]. Available: https://www.intelrealsense.com/

[25] R. Sánchez-Reolid, A. S. García, M. A. Vicente-Querol, L. Fernández-Aguilar, M. T. López, A. Fernández-Caballero, and P. González, “Artificial neural networks to assess emotional states from brain-computer interface,” *Electronics*, vol. 7, no. 12, p. 384, 2018.

[26] J.-L. Hsu, Y.-L. Zhen, T.-C. Lin, and Y.-S. Chiu, “Affective content analysis of music emotion through eeg,” *Multimedia Systems*, vol. 24, no. 2, pp. 195–210, 2018.

[27] P. J. Lang, M. M. Bradley, and B. N. Cuthbert *et al.*, “International affective picture system (iaps): Technical manual and affective ratings,” *NIMH Center for the Study of Emotion and Attention*, vol. 1, pp. 39–58, 1997.

[28] M. Bradley and P. J. Lang, *The International affective digitized sounds (IADS): stimuli, instruction manual and affective ratings*. NIMH Center for the Study of Emotion and Attention, 1999.

[29] M. Hampson, N. R. Driesen, P. Skudlarski, J. C. Gore, and R. T. Con-