Explorations of A Real-Time VR Emotion Prediction System Using Wearable Brain-Computer Interfacing

Nazmi Sofian Suhaimi¹, James Mountstephens¹ and Jason Teo¹

¹Faculty of Computer and Informatics, Universiti Malaysia Sabah, 88400, Kota Kinabalu, Sabah, Malaysia

Email: jwteo@ums.edu.my

Abstract. The following research describes the potential of using a four-class emotion classification using a four-channel wearable EEG headset combined with VR for evoking emotions from each individual. Multiple researchers have conducted and established emotion recognition by using a 2-D monitor screen for stimulus responses but this introduces artifacts such as the lack of concentration on-screen or external noise disturbance and the bulky and cumbersome wires on an EEG device were difficult and time-consuming to set up thus restricting to only the trained professionals to operate this complex and sensitive medical equipment. Therefore, using a small and portable EEG headset where it was accessible for consumers was used for the brainwave signal collection. The wearable EEG headset collects the brainwave samples at 256Hz at specific locations of the brain (Tp9, Tp10, AF7, AF8) and samples were transformed via FFT to obtain the five bands (Delta, Theta, Alpha, Beta, Gamma) and were classified using random forest classifier. An emotion prediction system was then developed and the trained model was used to benchmark the 4-class emotion prediction accuracy from each individual using a 4-channel low-cost EEG headset. Subsequently, a real-time prediction system was implemented and tested. Early findings showed that it could achieve predictions as high as 76.50% for intra-subject classification results.

1. Introduction
Human emotions are one of the components that govern the social interactions of society. As humans evolve, so does the level of social complexity in understanding human speeches and body languages which influences the decision-making and behavior of an individual [1]. Recent advancements in technology such as machine learning were successfully integrated into the system where it can interpret human physiological signals [2]. Furthermore, the developments in affective sciences have shown that human brain signals could likely be the key that could unlock potentials in human recognition system. The investigation into emotion prediction has been vastly explored with many of the researchers providing datasets such as DEAP [3], ASCERTAIN [4], and SEED [5] sharing valuable information such as the type of neurophysiological signal equipment used for data collection and the features used for emotion classification. Neurophysiological devices includes; electrocardiogram (ECG), electromyogram (EMG), electrodermal activity (EDA), electrooculogram (EOG) and electroencephalography (EEG) [6-10]. The datasets also shared videos and picture contents used for evoking specific emotions that were acknowledged by medical professionals. However, the use of a 2D monitor disrupts the attention of the individual, especially when viewed beyond the borders of the screen. This causes artifacts that were introduced into the dataset and affect the accuracy of the classification. Therefore, the use of virtual reality (VR) was considered as it had been adopted to treat conditions such
as phobias, anxieties, and other mental illnesses [11-13]. However, it was rarely seen as a tool for human recognition even though VR provides full immersion experiences that evoke stronger emotional responses [14].

The deployment of sensitive medical equipment for neurophysiological signal collection is costly and requires trained medical staff to operate such complex machinery thus causing accessibility an issue. Since the expression of emotion was triggered primarily from the central nervous system [15-17], the brain produces the impulses and the body reacts to the actions of these impulses. A consumer-grade EEG device would be feasible to collect the information directly from the brain. Finally, a live emotion prediction system will be developed to produce real-time results of the emotions with the integration of machine learning.

2. Methodology

2.1 Stimuli Preparation

The stimulus was prepared by selecting from various open-source VR video content. The VR videos have to satisfy the four quadrants of emotions by following the Arousal-Valence Space (AVS) [18], [19] model to represent each emotion in the respective quadrants, the model was shown in Figure 1. Arousal was represented in the x-axis where it ranges from engaged to non-engaged whereas valence was represented in the y-axis and ranges between negative and positive emotions [20]. Due to the extreme complexity of emotion selection, this research approached using a general representation of emotion where it is easy to understand and is embedded within human nature [21]. Each quadrant of the AVS model was represented with one emotion, happy, (quadrant 1), scared (quadrant 2), bored (quadrant 3), and calm (quadrant 4).

![Figure 1. Arousal-Valence Space (AVS) Model](image)

A total of 16 VR videos were selected to evoke emotional responses from each individual. Each VR video consists of 20-seconds clips representing a single emotion, four VR videos of continuous stimulation were presented in each quadrant. During the transition, there was a 10-second blank VR video (non-stimulating) scene for each individual to rest. The VR videos were presented in a single continuous clip with a total length of 350-seconds.

2.2 Dataset Collection

The EEG dataset used for training the system was collected from 31 healthy individuals (8 Females, 23 Males) aged between 23-35 where some were wearing glasses while others had normal visions. Each individual was presented with 2 headsets and a pair of earphones. The EEG headset was a product produced from InteraXon known as “Muse 2016” which has four-channels and capable of sampling at 256Hz and connects via Bluetooth 4.0 BLE communication protocol with electrodes positioned at Tp9, Tp10, AF7, and AF8 with reference at Fpz. The VR box headset was used to present the VR videos to
each individual using a smartphone attached and inserted into the box, Figure 2 shows the experiment setup when preparing for dataset collection.

![Figure 2](image.png)

**Figure 2.** An Individual wearing the experimental setup

### 2.3 Predictive System Development

The system was developed using a python environment where a combination of libraries was involved. Important libraries used to facilitate the system’s developments were “muselsl”, “numpy” “sklearn” and “flask”. **Muselsl** was used to establish the connection between **Muse 2016 EEG headset** with an external Bluetooth 4.0 BLE module. **Numpy** was used to transform the RAW input signals from the EEG headset and transformed to obtain the specific brainwave bands where; Delta ranges between 1Hz < 4Hz, Theta 4Hz < 8Hz, Alpha 8Hz < 12Hz, Beta 12Hz < 30Hz, Gamma 30Hz < 44Hz. The transformed brainwave signals were then fed into the random forest classifier using default parameter setup (n_estimators = 10, criterion = 'entropy', random_state = 42). Lastly, a localhost website was developed using **flask** module to show the predicted model for the audience.

### 2.4 Real-time Emotion Prediction System

Five healthy individuals (1 Female, 4 Males) aged between 30-63, were gathered to collect additional data for benchmarks with the live prediction system, three individuals had prescription glasses and two others had normal visions. Each individual was presented with the same 350-sec VR video used from the earlier dataset collected. Each individual had no prior knowledge of the said VR video and therefore removes the possibility of affecting the outcome results, Figure 3 shows the live emotion prediction system working on a localhost webserver.
3. Results and Discussion

Based on figure 4, it was observed that bored emotion obtained the highest accuracy compared with the rest of the emotion obtaining at 39.75% followed by scared emotion at 35.50%, happy emotion at 25.50%, and lastly calm emotion at only 9.50%. This shows that the system’s capability of predicting emotional responses on a four-class emotion using a consumer four-channel EEG headset was feasible.

**Overall Accuracy of Each Emotion Prediction**

| Emotion | Accuracy |
|---------|----------|
| Calm    | 9.50%    |
| Scared  | 35.50%   |
| Happy   | 25.50%   |
| Bored   | 39.75%   |

**Figure 3.** Live Four-Class Emotion Prediction System

**Figure 4.** Overall Accuracy of Each Emotion
In Figure 5, each individual’s emotion prediction accuracy was compared to analyze which emotion had better performance over the other. It was observed that each individual had mixed results with some emotions having better prediction capability than the other three emotions. P1, P2, and P3 had better emotion predictions on bored emotion achieving 41.25%, 72.50%, and 62.50% respectively while only P5 had the highest single emotion prediction at 76.25% for bored. However, P2 and P5 had very low accuracy on three other emotions.

![Figure 5. Prediction Accuracy of Each Individual](image)

Presently, there was a lot of research conducts on emotion prediction using various neurophysiological signals using 2D screens such as DEAP [3], ASCERTAIN [4], SEED [5], and many others but little was known on purely VR stimulus and using a consumer-grade EEG headset. Furthermore, this system is probably the world’s first live-prediction four-class emotion system using VR.

4. Conclusion and Future Work
In this research, the live-prediction system was capable of predicting emotions using a four-channel wearable EEG headset combined with a VR headset to evoke emotions. For future work, integration with using an eye-tracker system may improve the accuracy of the emotion classification and be able to draw heatmaps to map the user’s focus attention on certain objects displayed on the screen. Combinations of different electrode positions such as frontal lobes only or temporal lobes only may yield some evidence as to where emotional experiences are stored or the activation of emotion might be located elsewhere in the brain region. Classifiers used to train the model will need to be explored further to compare the trained models.

References
[1] Miceli M and C. Castelfranchi 2019 New Ideas Psychol.55, 42–49
[2] Delić V 2019 Comput. Intell. Neurosci. 1155 4368036
[3] Koelstra S 2012 IEEE Trans. Affect. Comput. 3 18–31
[4] Subramanian R, Wache J, Abadi M K, Vieriu R L, Winkler S and Sebe N 2018 IEEE Trans. Affect. Comput. 9 147–160
[5] Song T, Zheng W, Song P and Cui Z 2018 IEEE Trans. Affect. Comput. 3045 1–10
[6] Chen Y H 2014 Sensors (Switzerland) 14 23758–23780
[7] Demirel C, Kandemir H and Kose H 2018 *IEEE Signal Process. Commun. Appl. Conf.* pp. 1–4
[8] Dey A, Piumsomboon T, Lee Y and Billinghurst M 2017 *Conf. Hum. Factors Comput. Syst. - Proc.* 2017 4045–4056
[9] Li T H, Liu W, Zheng W L and Lu B L 2019 *Int. IEEE/EMBS Conf. Neural Eng. NER* 2019 607–610
[10] Ménard M, Richard P, Hamdi H, Daucé B and Yamaguchi T 2015 *Int. Conf. Physiol. Comput. Syst. Proc.* pp. 26–32
[11] Freina L and Ott M 2015 *Proc. eLearning Softw. Educ.* p 8
[12] Lee C and Wong G K C 2019 *J. Clin. Neurosci.* 62 14–20
[13] Wang S, Mao Z, Zeng C, Gong H, Li S and Chen B 2010 *Int. Conf. Geoinformatics* pp 1–5
[14] Horvat M, DoBrinic M, Novosel M and Jeric P 2018 *Int. Conv. Inf. Commun. Technol. Electron. Microelectron.* pp 1006–1010
[15] Bhattacharjee S 2019 *Adv. Intell. Syst. Comput.* 814 377–382
[16] Koelstra S and Patras I 2013 *Image Vis. Comput.* 31 164–174
[17] Mann S, Janzen R, Wu H, Lu M H and Guleria N 2016 *IEEE Games Entertain. Media Conf. GEM* 2015 7377239
[18] Aguinaga A R, Lopéz Ramírez M A and del Rosaria Baltazar Flores M 2015 *Int. J. Adv. Comput. Sci. Appl.* 6 230–238
[19] Bai J 2017 *Int. Conf. Cogn. Informatics Cogn. Comput.* pp. 121–129
[20] Verma G K and Tiwary U S 2017 *Multimed. Tools Appl.* 76 2159–2183
[21] Ekman P 1987 *J. Pers. Soc. Psychol.* 53 712–717

**Acknowledgments**

This work is supported by the research grant ref: IF0318M1003 (ICF0001-2018) received from the Ministry of Science, Technology, and Innovation (MOSTI), Malaysia.