PUSAT PENGURUSAN PENYELIDIKAN

PENERBIT KUIS
KOLEJ UNIVERSITI ISLAM ANTARABANGSA SELANGOR

e-JURNAL PENYELIDIKAN DAN INOVASI
rmc.kuis.edu.my/jpi/
e-ISSN 2289-7909
KOLEJ UNIVERSITI ISLAM ANTARABANGSA SELANGOR

VOL. 6 NO.1 (APRIL 2019): pp 43-53

CLASSIFYING AFFECTIVE STATES IN VR ACCORDING TO RUSSELL'S 4-QUADRANT CIRCUMPLEX MODEL OF EMOTIONS VIA WEARABLE EEG AND MACHINE LEARNING

Nazmi Sofian Suhaimi, Jason Teo, James Mountstephens, Chrystalle Tan Bih Yuan, Krishna Dilip Murthy
nazmisofian@gmail.com, jtwteo@ums.edu.my, james@ums.edu.my
tan.chrystalle@gmail.com & dkmurthyy@ums.edu.my
Universiti Malaysia Sabah, UMS.

ABSTRACT

This study attempts to produce a novel database for classification on emotional analysis using virtual reality (VR) obtained from third party sources such as YouTube, Discovery VR, Jaunt VR, NYT VR, Veer VR and Google Cardboard, as the visual stimuli for evoking emotions from a commercial-off-the shelf (COTS) wearable Electroencephalography (EEG) headset together with a virtual reality headset and a pair of earphones for immersive experience obtained from 24 participants. While there are numerous dataset that can be used to stimulate for emotional analysis obtained using database for emotional analysis using physiological signal (DEAP) and database for emotional analysis in music. Neither had approached from a VR point of view and this would serve as the novelty of this study. The contents obtained from third party sources were stitched together to fullfil one of each of the Arousal-Valence Model (AVS) to enable proper classification and analysis of the emotion elicited by the subjects. The classifiers performed on the EEG datasets were K-nearest neighbor (KNN) and Support Vector Machines (SVM) and the datasets were split into Intra-Subject and Inter-Subject variability respectively to obtain an overview of the highest achievable accuracy. From Intra-Subject approach, the highest accuracy obtained from the classifiers were from SVM Class Weight Kernel above 97% on each quadrant. While Inter-Subject variability were obtained at 95.29%, 93.04%, 88.62% by KNN on quadrant 1,2 and 3 respectively while 92.86% were obtained from SVM-Polynomial Kernel on quadrant 4.

Keywords: Machine Learning Language, KNN, SVM, EEG, VR.

Received: 9 July 2018
Revised: 15 October 2018
Accepted: 13 Mac 2019

1.0 INTRODUCTION

Emotional analysis studies come from a variety of methods that could identify the state of a person is in. Such methods that are applied are facial expressions (Chanthaphan, 2015; Suja, 2016; Turan, 2015) and human speeches such as gestures or motions (Chenchah, 2015; Koolagudi, 2012). However, could these be the only method in assessing the state of a person is in or how truthful they can be with their emotions as a person could fake their tones when speaking or fake an expression to mask their facade (Wioleta, 2013). Some might not be capable of grasping the emotional state they are in and have difficulties in describing their feelings verbally (Suhaimi, 2018). Thanks to the evolution of technology, there are now new devices that could detect these unclear emotional states through physiological signals produced within the central nervous system that controls a person's reactions towards stimulus that is subjected onto them. These physiological signals are Electroencephalography (EEG), Electrocardiogram (ECG), Heart Rate Variability (HRV), Galvanic Skin Response (GSR), Muscle Activity or Electromyogram (EMG), Skin Temperature (SKT), Blood Volume Pulse (BVP) and Respiratory Volume (RESP).

Emotional states can be difficult to identify as there are numerous methods to evoke these states such as watching affective movies, video clips (Russel, 1980), video games (van de Laar, 2013), or listening to music (Lin, 2010) or music videos (Koelstra, 2012). Emotions can be defined into few models. Eckman's describes that a human would be embedded with the six discrete basic emotions; Happiness, Sadness, Surprise, Fear, Anger and Disgust (Ekman, 1999). Plutchik's model presents eight fundamentals of emotions; Joy, Trust, Fear, Surprise, Sadness, Disgust, Anger and Anticipation (Picard, 2001). Lastly, Russell's model focuses on two-dimensional evalution namely the Arousal-Valence Model (Plutchik, 2001).

To evoke the emotions, the subject would be subjected to audible and visual stimulation. Hence, using Virtual Reality (VR), researchers can stimulate intricate real-life situations to investigate the complex human behaviors inside a controller environmental setting while providing a sense of immersivity or "being there" in an interactive environment that would override all other sensory information a user receives (Soleymani, 2012). As technology progresses, so does the capability of the VR immersion experience provide. Some notable entertainment industries such as Facebook, Sony and HTC are now looking into VR for video games, news, and even filming Cellan-Jones, 2016). Though there were only little advancements in EEG based emotional classification using VR as stimulus.

While there is a large dataset for emotional analysis such as the DEAP dataset, no available dataset was prepared for the perusal of emotional analysis in VR. Hence, this project will focus on preparing dataset for VR based contents and the classification of EEG signals according to the Arousal-Valence Model on each quadrant.

2.0 METHODS

2.1 Selection of VR Contents

The experimental setup for this project requires the need to stimulate responses from the participants using lists of Virtual Reality (VR) contents are those available online. The contents that was found to be closely relatable to stimulate each quadrant's responses according to Figure 1. Each quadrant was provided with 4 videos to achieve the highest impact possible of the emotion that was targeted. A total number of 16 videos were compiled and stitched. Figure 2 illustrates the flow of the video that was presented to the participants during the experimentation. The total time for the stitched videos lasted for 6 minutes and 1 second according to the finalized video.

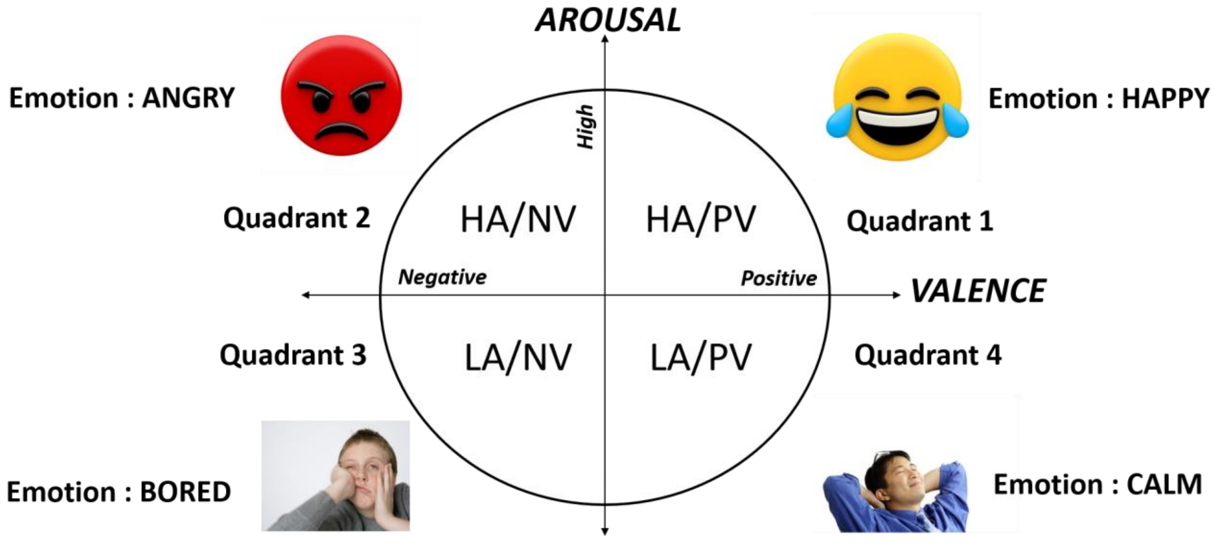

Figure 1: Arousal Valence Model (AVS).

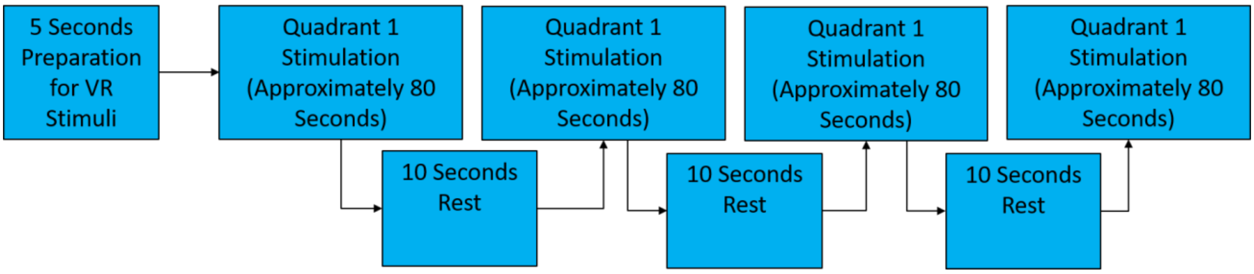

Figure 2: Timestamp of the Stitched Video.

2.2 Demography Target, Hardware Selection and Application setup

A total number of 24 participants (3 females and 21 males) volunteered to participate in the experiment aged between 20-31 and is either studying or working in their respective fields. All individuals were screened before the experiment and were healthy with good senses (Sight, Hearing, Smell, Touch, and Taste). Before the experimentation began, all participants were thoroughly briefed on the potential impacts that may occur such as nausea, headache, motion sickness and dizziness.

Hardwares that were used to record the brainwave data were of the Commercially-off-The-Shelf (COTS) EEG headset device known as Muse, developed by Interaxon. The wearable EEG headset is a non-invasive device that does not penetrate the human skin but has electrodes that detects the 10-20 EEG system on the human scalps. An illustration of the 10-20 EEG system is shown in figure 3.

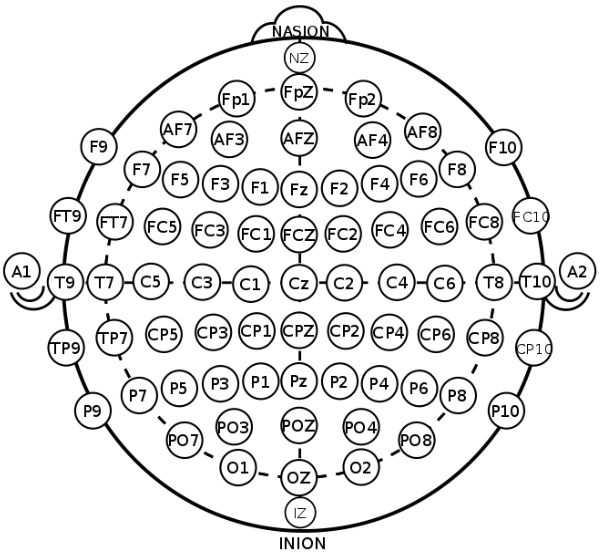

Figure 3: 10-20 EEG System of the position for the electrode to positon on the human scalp.

The wearable EEG has 5 channels that can detect the human brainwave signal. These 5 channels are placed according to figure 4. TP9 and TP10 are placed behind and above the earlobes while AF7, AF8 and Fpz are placed at the front of the scalp.

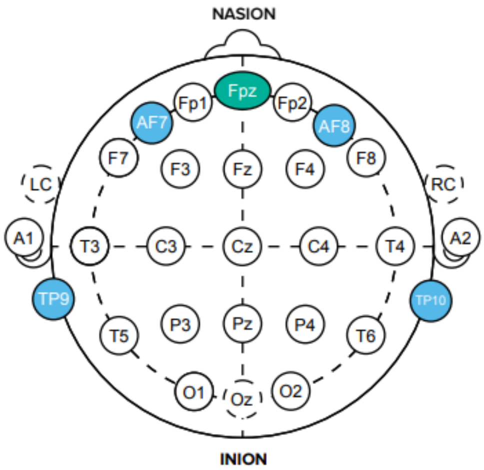

Figure 4: 5-Channel positions Muse headband places on the human scalp.

To allow the partipants to achieve an immersive experience of the 360 videos, another headset was required to accommodate such experience which is the VR headset that places a smartphone within the headset and plugged in with a pair of earphones. This immersive experience allows users to have the freedom of movement and watching their surroundings changing. Indirectly, this would also provide larger impact on the stimuli of the video. Figure 5 illustrates the hardware acquisition and figure 6 demonstrates the hardware setup for the participants. Because the participant were given freedom of movements, external environments were taken into account to avoid any unwanted harm to the participants such as slamming into walls, bumping onto chairs or tables and also not to be disturbed with any external artefacts that may alter the immersion level of the participant which degrades the accuracy results.

Figure 5: Hardwares that were used to facilitate the 360-Video Experiment and EEG recordings.

Figure 6: A demonstration of the hardware setup while being shown with the 360-video contents.

2.3 Data Cleanup

At the end of each video trial, there were a total number of 39 data elements recorded from the wearable EEG Muse headset. Figure 7 shows the 39 data elements that were recorded and the of the 39 elements, only 20 were required to perform the classification for training and testing of the accuracy on emotion which were highlighted (Delta, Theta, Alpha, Beta and Gamma). There are two approaches towards performing the classification, one of them is through intra-subject variability whereby the dataset was specifically catered towards a single individual (approximately around 600 datapoint) while the other is inter-subject variability which combines all of the individual datasets into a large dataset (approximately 15,000 datapoints).

TimeStamp	Delta_TP9	Delta_AF7	Delta_AF8	Delta_TP10	Theta_TP9	Theta_AF7	Theta_AF8

Theta_TP10	Alpha_TP9	Alpha_AF7	Alpha_AF8	Alpha_TP10	Beta_TP9	Beta_AF7	Beta_AF8

Beta_TP10	Gamma_TP9	Gamma_AF7	Gamma_AF8	Gamma_TP10	RAW_TP9	RAW_AF7	RAW_AF8

RAW_TP10	AUX_RIGHT	Accelerometer_X	Accelerometer_Y	Accelerometer_Z	Gyro_X	Gyro_Y	Gyro_Z

HeadBandOn	HSI_TP9	HSI_AF7	HSI_AF8	HSI_TP10	Battery	Elements

Figure 7: Recorded Elements from Wearable EEG Muse Headset.

2.4 Classification Methods

In order to classify the emotion, a statistical tool known as RStudio was used to perform Machine Learning Language for KNN and SVM to perform normalization, pre-processing and any additional data cleanup to achieve the output required. Each quadrant on the AVS model were performed and compared and was shown in the results section. Further analysis will also be discussed in the discussion section.

3.0 RESULTS

In this section, the results obtained from performing KNN and SVM kernels are displayed here in both table and charts.

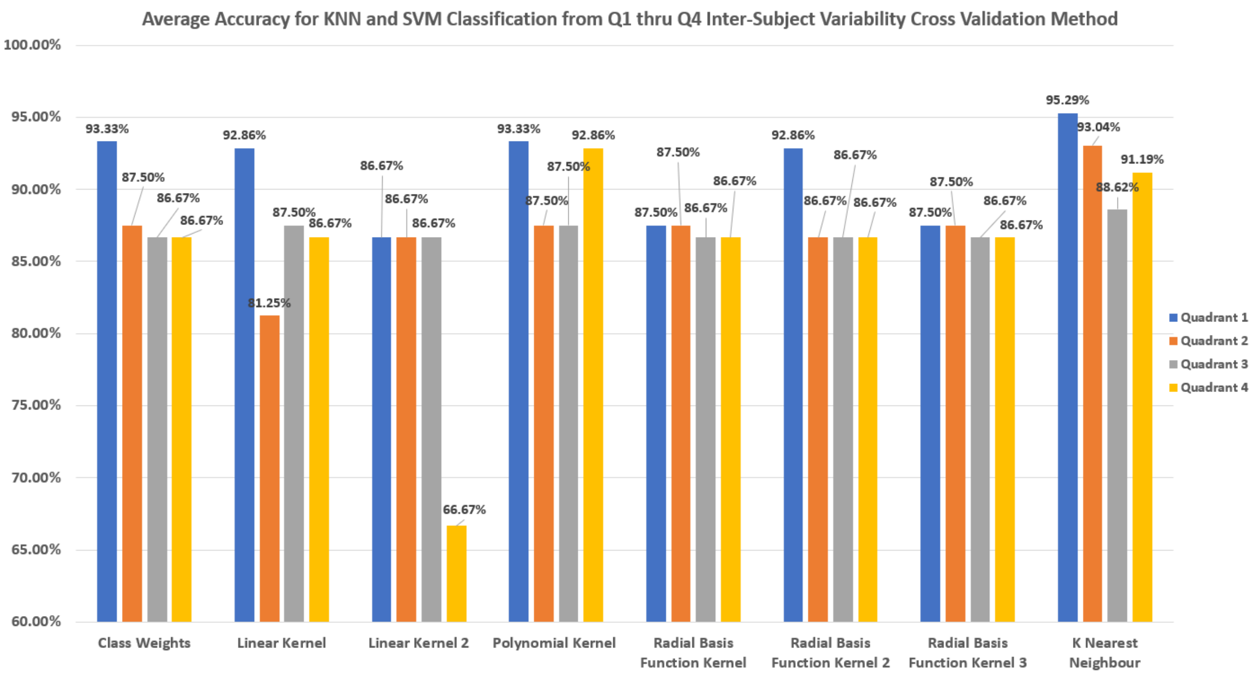

Figure 8: Tabulated results of KNN and SVM using Inter-Subject Variability

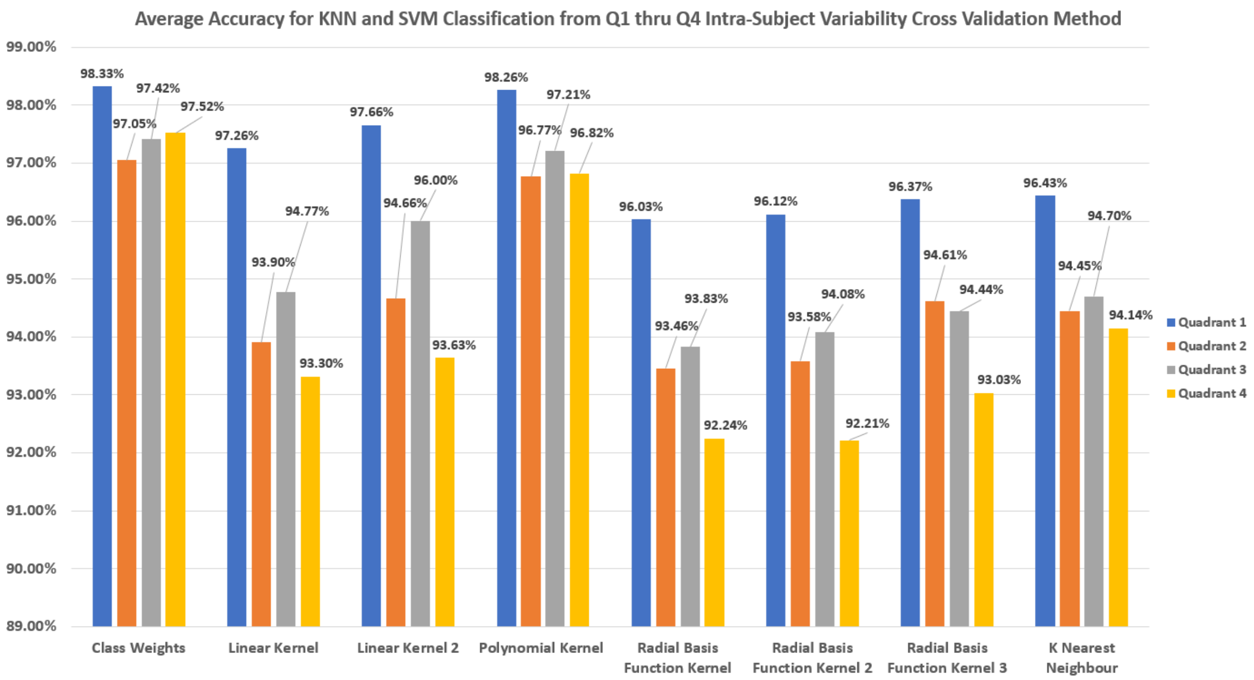

Figure 9: Tabulated results of KNN and SVM using Intra-Subject Variability

Table 1: Tabulated results of KNN and SVM from Inter-Subject Variability test.

Quadrant / Kernel	1 (HA/HV)	2 (HA/NV)	3 (LA/NV)	4 (LA/HV)
Class Weights	93.33%	87.50%	86.67%	86.67%
Linear Kernel	92.86%	81.25%	87.50%	86.67%
Linear Kernel 2	86.67%	86.67%	86.67%	66.67%
Polynomial Kernel	93.33%	87.50%	87.50%	92.86%
Radial Basis Function Kernel	87.50%	87.50%	86.67%	86.67%
Radial Basis Function Kernel 2	92.86%	86.67%	86.67%	86.67%
Radial Basis Function Kernel 3	87.50%	87.50%	86.67%	86.67%
K Nearest Neighbour	95.29%	93.04%	88.62%	91.19%

Table 2: Tabulated results of KNN and SVM from Intra-Subject Variability test.

Quadrant / Kernel	1 (HA/HV)	2 (HA/NV)	3 (LA/NV)	4 (LA/HV)
Class Weights	98.33%	97.05%	97.42%	97.52%
Linear Kernel	97.26%	93.90%	94.77%	93.30%
Linear Kernel 2	97.66%	94.66%	96.00%	93.63%
Polynomial Kernel	98.26%	96.77%	97.21%	96.82%
Radial Basis Function Kernel	96.03%	93.46%	93.83%	92.24%
Radial Basis Function Kernel 2	96.12%	93.58%	94.08%	92.21%
Radial Basis Function Kernel 3	96.37%	94.61%	94.44%	93.03%
K Nearest Neighbour	96.43%	94.45%	94.70%	94.14%

4.0 DISCUSSION

From the results obtained running classifiers with KNN and SVM, the overall highest accuracy achieved was from Intra-Subject variability where most of its classifiers performed the best compared to Inter-Subject Variability. This was to be expected from performing using Intra-Subject variability method as the classifiers were tuned to perform best on the individual subject with small numbers of test and training datasets.

The classifiers in here were performed in reference with the datasets comparing only in their respective quadrants of being stimulated and non-stimulated and would explain the high number of accuracy obtained from the classifiers. However, comparing to previous work [Steuer; 1992] where the experiment was stimulated with a 3D game app and had only obtained an accuracy of 82.44% and 82.43% on both KNN and SVM classifiers. This suggests that the animated environment may not have been truly immersive and felt somewhat unreal as compared to the contents shown in this 360-video where the environment in realistic.

Apart from that, the selection of kernel to classify the emotion would also affect the accuracy of the classification as was seen in the results. Class weights kernel was seen to be performing well in Intra-Subject variabilty where the number of dataset was low. K-nearest neighbor was seen performing well under Inter-Subject variability with a larger number of datasets.

5.0 CONCLUSION

From this study, we were able classify each of the quadrants emotional analysis from the VR contents that was stitched to stimulate the subject responses towards them. This dataset will then be further analyzed by performing classifiers on all 4 quadrants simultaneously to obtain their accuracy performance. This study will also attempt to approach using Deep Learning Language to compare with the perforamance obtained from Machine Learning Language.

ACKNOWLEDGEMENT

This work was supported by an FRGS research grant ref: FRG0435 from the Ministry of Higher Education, Malaysia.

REFERENCES

Cellan-Jones, R. (2016). *The year when VR goes from virtual to reality*. BBC news. Technology section. Retrieved from http://www.bbc.com/news/technology- 35205783.

Chanthaphan, N., Uchimura, K., Satonaka, T., & Makioka, T. (2015, November). Facial emotion recognition based on facial motion stream generated by Kinect. In *Signal-Image Technology & Internet-Based Systems (SITIS), 2015 11th International Conference on* (pp. 117-124). IEEE.

Chenchah, F., & Lachiri, Z (2015). Acoustic Emotion Recognition Using Linear and Nonlinear Cepstral Coefficients. International Journal of Advanced Computer Science and Applications, 6(11), 135-138.

Ekman, Paul (1999), "Basic Emotions", in Dalgleish, T; Power, M, *Handbook of Cognition and Emotion* (PDF), Sussex, UK: John Wiley & Sons

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., & Patras, I. (2012). DEAP : A Database for Emotion Analysis using Physiological Signals. *IEEE Transactions on Affective Computing*, *3*(1), 18–31.

Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech: a review. *International journal of speech technology*, *15*(2), 99-117.

Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., & Chen, J. H. (2010). EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, *57*(7), 1798-1806.

Suhaimi, N. S., Teo, J., & Mountstephens, J. (2018, March). Emotional State Classification in Virtual Reality Using Wearable Electroencephalography. In *IOP Conference Series: Materials Science and Engineering* (Vol. 341, No. 1, p. 012016). IOP Publishing.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE transactions on pattern analysis and machine intelligence*, *23*(10), 1175-1191.

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, *89*(4), 344-350.

Russell, James (1980). "A circumplex model of affect". *Journal of Personality and Social Psychology*. **39**: 1161–1178. doi:10.1037/h0077714

Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, *3*(1), 42-55.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, *42*(4), 73-93.

Suja, P., & Tripathi, S. (2016, February). Real-time emotion recognition from facial images using Raspberry Pi II. In *Signal Processing and Integrated Networks (SPIN), 2016 3rd International Conference on* (pp. 666-670). IEEE.

Turan, C., Lam, K. M., & He, X. (2015, December). Facial expression recognition with emotion-based feature fusion. In *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2015 Asia-Pacific* (pp. 1-6). IEEE.

van de Laar, B., Gürkök, H., Bos, D. P. O., Poel, M., & Nijholt, A. (2013). Experiencing BCI control in a popular computer game. *IEEE Transactions on Computational Intelligence and AI in Games*, *5*(2), 176-184.

Wioleta, S. (2013, June). Using physiological signals for emotion recognition. In *Human System Interaction (HSI), 2013 The 6th International Conference on* (pp. 556-561). IEEE.