Tuning Support Vector Machines for Improving Four-Class Emotion Classification in Virtual Reality (VR) using Heart Rate Features

Aaron Frederick Bulagang¹, James Mountstephens², Jason Teo Tze Wi³

¹Graduate Researcher, Evolutionary Computing Laboratory, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.
²Senior Lecturer, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.
³Professor, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.

jtwteo@ums.edu.my

Abstract. The main objective of this paper is to conduct three experiments using Support Vector Machine (SVM) with different parameter settings to find and compare the accuracy of each SVM setting. Heart rate (HR) signals were collected with a medical-grade wearable heart rate monitor from Empatica (E4 Wristband) and processed using the Empatica Realtime Monitor application during this investigation. HR was employed as the method to capture the test subjects’ physiological signals via plethysmography. The three experiments were conducted using a wrist-worn monitor to gain HR signal, and a VR Headset for subjects to view 360 degrees video stimuli. A total of 10 subjects participated in this experiment. Data from the 10 subjects were then processed with Python with SVM. The data was classified for four distinct emotion classes using both inter-subject classification and intra-subject classification testing approaches, with inter-subject classification yielding an accuracy of 53.39% while intra-subject classification ranges from 56.92% to 86.15%. These results demonstrate the potential of achieving higher accuracy results using different parameter settings via the use of HR as the input feature to the machine learning classifier, which appears to be a promising sensor modality for four-class emotion classification in virtual reality using wearable technology.

1. Introduction

Emotion classification is the branch of machine learning which attempts to recognize human emotions from physiological data. Emotion can be characterized at two levels: primary emotions that include joy, sadness, anger, fear, disgust, surprise; and secondary emotions that are exhibited as a reaction to the primary emotion being experienced. Most commonly, emotions are classified based on arousal and valence (Alarcao & Fonseca, 2017). Valence represents the quality of emotion, ranging from unpleasant (negative) to pleasant (positive) whereas arousal denotes the quantitative activation level, from aroused (high) to not aroused (low). For example, happiness has a positive valence, disgust has a negative valence, boredom relates with low arousal, while surprise evokes high arousal (Menard et al., 2015). When combing arousal and valence to form a bi-dimensional perspective, the classification
widely used is the bipolar model advocated by Russell, known as the Circumplex Model of Emotions (Alarcao & Fonseca, 2017). Figure 1 shows Russell’s two-dimensional model based on valence and arousal.

![Figure 1: Russell’s two-dimensional model based on valence and arousal](image)

Russell’s valence-arousal scale is a broadly utilized model in emotion studies, specifically in studies that focus on using the DEAP (Database for Emotion Analysis using Physiological Signals) dataset (Koelstra, 2011) for emotion classification. Nonetheless, it has been seen that Russell’s 2D emotion space model is excellent in perceiving positive or negative emotions, along with high or low arousal states, the issue emerges when attempting to recognize emotions that are classified within the same quadrant, e.g. attempting to distinguish between fear and anger, which both in the same zone with negative valence and high arousal according to this 2D model, which is why proposition of a higher-dimensional emotion models have been suggested (Koelsch et al., 2015).

Heart rate (HR) gives a signal that is significant in studying the physiological changes of the heart in various situations. It has been utilized broadly in the investigation of few ailments including coronary illness, arrhythmia, and epilepsy. The utilization of HR has additionally been applied to accessing mental and psychological conditions. Similarly, HR signal can produce relevant information in understanding human emotional stress. As such, HR plays a significant role during the study of human emotion when presented with certain emotional stimuli. The dependable detection of emotions enables numerous applications as such in the fields of affective computing, personalized entertainment, neuromarketing, virtual learning, and virtual rehabilitation, to name a few. Hence, this paper aims to investigate whether HR alone can be used as a feature for machine learning to classify emotions by achieving higher accuracy within the four distinct classes. To the best of our knowledge, such an attempt has yet to be accounted for in the time of this writing.

This paper is presented as follows. Section 1 briefly introduces the background and goals of this study. Section 2 presents in-depth the methodology along with the emotional stimuli, the demography of the test group, hardware utilized, and experimental setup. Section 3 presents the results and analysis section, where the results of the experiment with the use of HR monitor were discussed thoroughly. Section 4 concludes the paper by showing a summarized finding of the experiment and what research work is planned for future investigation.

2. Methodology

2.1 VR Content Selection

The arrangement for this experiment utilizes Virtual Reality (VR) to exhibit the emotional stimuli’s content to the participants, which 360° videos are presented to evoke emotional responses from participants. The content was chosen to evoke driven emotions from each of the four quadrants as presented in Figure 1. Four short videos are presented together for each quadrant to have a high and
sustained response from the participant for the focused emotion. Towards the completion of each quadrant, a 10-second rest period was given where no visual stimuli are presented to enable the participant to return to a baseline emotional state to prepare for the following quadrant’s emotional presentation stimuli. Hence, 16 videos were compiled and stitched together for the entire experiment over the four emotional quadrants. Figure 2 shows the flow of the experiment videos presented to the participants. The entire sequence of stitched content presentation to the participants spans a total of 6 minutes and 5 seconds including the initial baseline period.

Figure 2: Overall flow of the VR presentation for evoking the emotional responses of participants

Before the commencement of each experiment, participants were clarified the progression of the experiment, and after thoroughly understood, a consent form was presented for them to sign. Next, the VR headset was attached to their head while adjusting the straps to their comfort level. The earphone is then attached for a further immersion experience.

2.2 Demography, Hardware, and Setup
An aggregate of 10 participants (9 males) volunteered in this experiment with the age of 20-28, comprising of individuals who were studying or working. Before starting the experiment, all participants were advised about the potential side effect experience they may have such as nausea, headache, motion sickness, and dizziness due to the use of VR as the stimuli presentation device.

The hardware that was used to record the heart rate is the Empatica E4 wristband, a medical-grade wearable device for obtaining physiological data. It is a non-invasive device and utilizes photoplethysmography sensor to obtain HR data by recognizing changes in the blood volume through the skin optically as an indication of the user’s pulse, which measures at a frequency of 64Hz. The device can be worn on the wrist of either the left or right hand. After attaching the wristband, the VR headset and earphones, the participant is in a seated position, the 360° video is played at the same time with the heart rate reading being capture wirelessly through a smartphone application.
Figure 3 shows the wearable device Empatica E4 on the participant’s wrist, to record the participant’s HR during the experiment, the Empatica E4 Realtime application was utilized as shown in Figure 4. Figure 5 shows the VR headset attached to the participant’s head to view the 360 degrees videos. The headset from HTC enables the participant to be completely immersed in the environment they are in following the video being played at that time. Participants can move in 360 degrees to view the surrounding environments of the video. The headset is attached to the participant’s head by adjusting to their viewing focus and comfort. During the data collection with the heart rate monitor, the participant’s HR signal will be collected by the smartphone application as displayed in Figure 6.

Figure 6: E4 Realtime Heart Rate Visualization.

3. Results and Discussion

3.1 Min, Max and Average BPM
After the experiment was conducted, the data retrieved from the Empatica E4 Realtime application was then uploaded to the cloud where the CSV file can be downloaded to be further processed. The following are the results for the 10 participants who participated in the experiment. Based on the timestamp of the data file, each instance was mapped to each of the four corresponding quadrants' class labels.
Figure 7: Mean, Max and Average Heart Beat of Participants

Figure 7 above shows the data collected from the 10 participants. The figure shows the minimum, maximum and average heart rate for the intra-subject during the experiment. The heart rate ranges from a minimum of 54.58 bpm from participant 3, a maximum of 92.88 bpm from participant 9 and a total average of 70.96 bpm over the entire 10 participants.

Figure 8: Min, Max and Average HR over the entire cohort of participants

Figure 8 presents the min, max and average HR for the 10 participants with 54.58 being the lowest, and 92.88 the highest, and 70.96 being the average.

3.2 Intra- and Inter-subject classification using SVM with parameter tuning

The individual data collected which is used for training and testing within the same participant is referred to as intra-subject classification whereas the entire collection of data from all 10 participants which is used for training and testing across all participants is referred to as inter-subject classification. Intra-subject classification data were classified individually using R while inter-subject classification data were collectively classified by grouping all 10 participants’ data into one CSV file. Ten-fold cross-validation was used in all experiments.

The HR data acquired is used as the sole input feature to machine learning classifiers in an attempt to classify the emotions into four distinct classes based on HR data alone. The machine learning approaches used were Support Vector Machine (SVM) with different parameter tuning used in three experiments to identify and classify the emotions based on the heart rate data gathered from the
participants in response to the stimuli presented during the experiment. The three-experiment conducted used the following parameter tuning, first experiment (0.1, 1, 10, 100), second experiment (0.01, 0.1, 1, 10) and the third experiment with (0.001, 0.01, 0.1, 1). The following is the classification accuracy results for inter-subject classification and intra-subject classification for the experiment conducted.

3.3 Intra-Subject Classification using SVM with Parameter Tuning Results

Figure 9: Experiment 1 Intra-Subject Classification Accuracy Results

Figure 9 shows the accuracy results from Experiment 1, with the accuracy ranging from 75.38% as the highest accuracy obtained from subjects 2 and 3, to 61.53% as the lowest from subjects 4 and 8.

Figure 10: Experiment 2 Intra-Subject Classification Accuracy Results
Figure 10 shows the accuracy results from Experiment 2, with the accuracy varying from 86.15% as the highest accuracy obtained from subject 2, to the lowest being 63.07% obtained from subjects 1, 4, and 7.

![Experiment 3 with Parameter Tuning of (0.001, 0.01, 0.1, 1)](chart)

**Figure 11: Experiment 3 Intra-Subject Classification Accuracy Results**

Figure 11 shows the accuracy results from Experiment 3, with the accuracy results yielding 84.61% as the highest accuracy yielded from subject 9 to the lowest of 56.92% obtained from subject 4.

### 3.4 Inter-Subject Classification using SVM with Parameter Tuning Results

The following figures present the accuracy results for inter-subject classification using SVM based on the three experiments with the same parameter tuning set.

![Inter10](chart)

**Figure 12: Experiment 1 Inter-Subject Classification Accuracy Results**

Figure 12 shows the classification accuracy results from Experiment 1 for Inter-subject classification for across 10 subjects, 10 Inter-subject classifications yielding 50.92% accuracy.
Figure 13: Experiment 2 Inter-Subject Classification Accuracy Results

Figure 13 shows the accuracy results from Experiment 2, with the accuracy results for 10 subjects in Inter-subject classification yielding 53.39% accuracy.

Figure 14: Experiment 3 Inter-Subject Classification Accuracy Results

Figure 14 shows the accuracy results from Experiment 3, with the accuracy results for 10 subjects in Inter-subject classification yielding 52.00% accuracy.

4. Conclusion
The main objective of this paper is to conduct three experiments using SVM with parameter tuning to find and compare the accuracy results of the Inter- and intra-subject classification. With the Heart Rate data collected from the Empatica E4 wristband and VR environment as a stimulus, the classification was done based on Russell’s Circumplex Model of emotion. Based on the three experiments conducted, it was found that the highest accuracy for intra-subject classification results was yielded from experiment 2 from Subject 2 with 86.15%, where the parameters that were used in that experiment being (0.01, 0.1, 1, 10) while the Inter-subject classification shows highest accuracy results of 53.39% for Inter-subject classification across ten subjects from the second experiment. The next immediate work will focus on extending the study to a bigger cohort of 30 test participants. Furthermore, skin conductance will also be investigated as a potential sensor modality in addition to HR. An exploration and comparison between HR, EDG and the use of both for emotion classification are planned to be conducted, as well as applying deep learning approaches to compare the results with prior experiments conducted using conventional machine learning classifiers.

Acknowledgments
This work was supported by the Ministry of Education, Malaysia [ref: FRG0512-1/2019].
References

[1] Alarcao S M and Fonseca M J 2017 IEEE Transactions on Affective Computing, 10(3), 374-393.

[2] Candra H 2017 Emotion Recognition Using Facial Expression and Electroencephalography Features with Support Vector Machine Classifier Student.

[3] Koelsch S, Jacobs A M, Menninghaus W and Liebal K 2015 Physics of Life Reviews, 13, 1-27

[4] Koelstra S, Muhl C, Soleymani M, Lee J S, Yazdani A, Ebrahimi T and Patras I 2011 IEEE Transactions on affective computing, 3(1), 18-31

[5] Menard M, Richard P, Hamdi H, Dauce B, and Yamaguchi T 2015 2nd International Conference on Physiological Computing Systems, Proceedings, 26-32.

[6] Ekman P. Basic emotions. Handbook of cognition and emotion. 1999 Feb 25;98(45-60):16.

[7] Dissanayake T, Rajapaksha Y, Ragel R, Nawinne I. An ensemble learning approach for electrocardiogram sensor based human emotion recognition. Sensors (Switzerland). 2019;19(20):1-24.

[8] Marin-Morales J, Higuera-Trujillo JL, Greco A, Guixeres J, Llinares C, Scilingo EP, Alcañiz M, Valenza G. Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors. Scientific reports. 2018 Sep 12;8(1)

[9] Maskeliunas R. Anxiety Level Recognition for Virtual Reality Therapy System Using Physiological Signals. 2019.

[10] Pane E, Hendrawan M, Wibawa A, Purnomo M. Identifying Rules for Electroencephalograph (EEG) Emotion Recognition and Classification. Proc 2017 5th Int Conf Instrumentation, Commun Inf Technol Biomed Eng ICICI-BME 2017. 2018;(November):167-172.

[11] Labati R, Muñoz E, Piuri V, Sassi R, Scotti F. Deep-ECG: Convolutional Neural Networks for ECG biometric recognition. Pattern Recognit Lett. 2019;126:78-85.

[12] Rebollo-neira L, Cerná D. Biomedical Signal Processing and Control Wavelet based dictionaries for dimensionality reduction of ECG signals. 2019;54.

[13] Young B. New standards for ECG equipment. J Electrocardiol. 2019;

[14] Zhou S, Tan B. Electrocardiogram soft computing using hybrid deep learning. Appl Soft Comput J. 2019

[15] Abdullah S, Choudhury T. Sensing Technologies for Monitoring Serious Mental Illnesses. IEEE Multimed. 2018;25(1):61-75.

[16] Choi EJ, Kim DK. Arousal and Valence Classification Model Based on Long Short-Term Memory and DEAP Data for Mental Healthcare Management. 2018;24(4).