

SHORT REPORT

Open Access



Multiclass emotion prediction using heart rate and virtual reality stimuli

Aaron Frederick Bulagang, James Mountstephens and Jason Teo* 

*Correspondence:
jtwteo@ums.edu.my
Universiti Malaysia Sabah,
Jalan UMS, Kota Kinabalu,
88400 Sabah, Malaysia

Abstract

Background: Emotion prediction is a method that recognizes the human emotion derived from the subject's psychological data. The problem in question is the limited use of heart rate (HR) as the prediction feature through the use of common classifiers such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Random Forest (RF) in emotion prediction. This paper aims to investigate whether HR signals can be utilized to classify four-class emotions using the emotion model from Russell's in a virtual reality (VR) environment using machine learning.

Method: An experiment was conducted using the Empatica E4 wristband to acquire the participant's HR, a VR headset as the display device for participants to view the 360° emotional videos, and the Empatica E4 real-time application was used during the experiment to extract and process the participant's recorded heart rate.

Findings: For intra-subject classification, all three classifiers SVM, KNN, and RF achieved 100% as the highest accuracy while inter-subject classification achieved 46.7% for SVM, 42.9% for KNN and 43.3% for RF.

Conclusion: The results demonstrate the potential of SVM, KNN and RF classifiers to classify HR as a feature to be used in emotion prediction in four distinct emotion classes in a virtual reality environment. The potential applications include interactive gaming, affective entertainment, and VR health rehabilitation.

Keywords: Four-class emotion classification, Electrocardiography, Machine learning, Virtual reality, HR, Empatica E4, KNN, RF, SVM

Introduction

In machine learning, emotion classification is a method used in recognizing human emotions from subject physiological data. There are two levels of emotions, primary and secondary emotions. Emotions such as surprise, disgust, fear, anger, sadness and joy are considered as primary emotions. While reacting to the primary emotions experiences are considered as secondary emotions.

A significant aspect of everyday life is emotion: it influences decision-making, perception, human interaction, and human intelligence. Physiologically and mentally, feelings control the status of humans [6]. Positive emotions and negative emotions exist, positive emotions are more linked to the enhancement of human wellbeing and job performance, while negative emotions can cause health problems. A factor in clinical depression is the

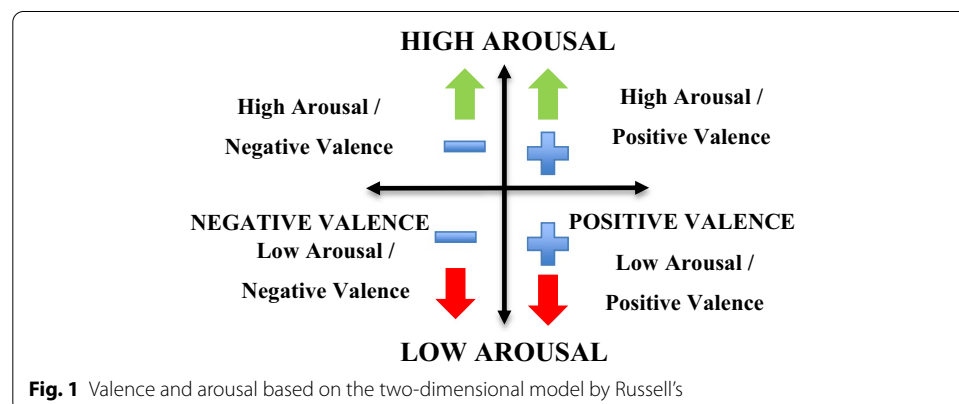
long-term accumulation of acquired negative emotions [2]. Physiological impulses in the human brain can be derived from the autonomous nervous system (ANS) and are not knowingly or deliberately triggered [8]. Baig and Kavakli [3], discuss and examine the classification of emotions using electrocardiography (ECG) and signals for electrodermography (EDG). “Introduction” section offers an overview of the model of emotion and Method of emotion classification researchers used in their studies. Although a previous study using electroencephalography (EEG) signals showed high classification accuracies of over 80%, emotion classification studies by EDG and ECG have also shown competitively high classification accuracies.

Valence and arousal can be classified and considered as emotions [1]. The quality of emotions is presented as valence, which ranges from negative (unpleasant) to positive (pleasant). While the level of quantitative activation is presented as arousal, which ranges from high (aroused) to low (not aroused). 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). An example of positive valence is happiness, negative valence is disgust, while low arousal is related to boredom, and high arousal is induced by surprise [7].

A bi-dimensional perspective is formed when valence and arousal are combined, the bipolar model is a widely used classification model from Russell, known with the moniker Circumplex model of emotions [1]. Russell’s model of two-dimensional that is based on arousal and valence is presented in Fig. 1.

The scale of valence-arousal by Russell’s is arguably the most widely used model in researches related with emotion, in particular research, the dataset from Database for Emotion Analysis (DEAP) using Physiological Signals) is used in emotion classification [5]. The observation Russell’s model is good at recognizing emotions that is positive or negative within a 2D space model of emotion, together with states of arousal that are low or high, the problem arises when distinguishing emotions within a similar quadrant is attempted. When distinguishing the emotions such as anger and fear is attempted, it is both within the same quadrant of high arousal with negative valence according to the 2D model, that is where the emotion model with higher-dimension was proposed [4].

The signal provided by HR is valuable in researching the changes physiologically of the heart in different scenarios. This has been used widely in research involving treating heart disease, epilepsy, and arrhythmia. The application of HR has been used in



the evaluation of psychological and mental conditions. Furthermore, HR signal information can recognize the emotional stress of humans. HR can play an important part when attempting to study human emotions when certain types of emotional stimuli are presented to them. The reliable recognition of emotions has multiple applications within neuromarketing, personalized entertainment, affective computing fields, virtual rehabilitation, and virtual learning. Hence, this paper's objective is investigating whether the use of HR signals alone with Virtual Reality (VR) as the stimuli is sufficient as a machine learning feature in classifying emotions in four different classes. To the best of our knowledge, there has been no attempt to investigate HR for emotion prediction in four distinct emotion classes. Next, we briefly present two related studies that utilized HR for emotion prediction but both papers only predicted two distinct emotion classes with a neutral label.

Table 1 above shows the comparison of HR signals used for emotion prediction in two related studies. In [11], the participants were 25 subjects while the stimuli used was the CEVS video dataset showing a various range of emotions to the participants. A number of features were extracted from the original HR signals including amplitude, slope and information entropy among others. Using a Gradient Boosting Decision Tree (GBDT) classifier, the best accuracy of 84% was obtained. However, this prediction was conducted for two emotions only, which were happy and sad that covers only the valence axis, with one neutral label. In [9], the study included 5 participants only who were asked to record their emotions via an Android application when they experienced an emotion. It was shown from their experiments that the best accuracy of 79% can be obtained by using the Discrete Wavelet Transform (DWT) to extract features from the original HR signals features with a time window of length 180 and using SVM as the classifier. Again, the study only conducted prediction along the valence axis, predicting emotions clustered in either positive or negative emotions only (with one neutral label). In both of these related studies, no results were presented for four-class emotion predictions, which is the focus of this study. Moreover, it was not stated if the studies conducted were for intra-subject classification or inter-subject classification, which is clearly differentiated in this study, where the implications between these two types of experimental setup have very significant outcomes on the classification accuracy, as will be demonstrated in the experimental section's discussion of results.

Table 1 Comparison of related HR-based emotion prediction studies

Author	Signals	Participants	Stimuli	Feature extraction	Classifier	Accuracy
[11]	HR	25	China Emotional Video Stimuli (CEVS)	Amplitude Slope Information Entropy and others	Gradient Boosting Decision Tree	84% (2 emotions with 1 neutral label)
[9]	HR	5	Android Application	Discrete Wavelet Transform (DWT)	SVM	79% (2 emotions with 1 neutral label)

The following presents the outline of the paper. The first section introduces the aims and background of this research brief, the second section presents the methodology including the stimuli for emotion, the test group demography, the hardware of the experiment, and the setup of the experiment. The third section presents the analysis and results of the experiment, where the experimental findings using HR as the feature for prediction was reviewed. Finally the fourth section concludes this research brief.

Methodology

Overall experimental flow

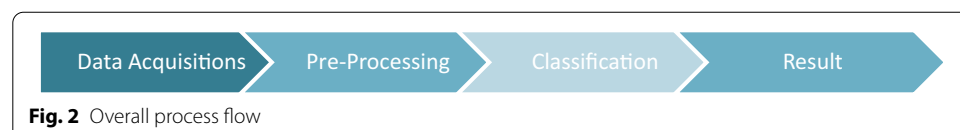
The experimental methodology reflects the setup procedures for the in-lab experiment. The experiment framework was prepared based on the four general steps in the signal processing activity as shown in Fig. 2.

This experiment starts with the data acquisition process, a total of 20 healthy participants with no history of heart cases volunteered. During the data acquisition process, each of the participants wears a wearable device called the Empatica E4 which is used to detect and record their HR signals, which will be further explained in Subsection 2.3. It is then paired with a virtual reality (VR) headset for the participant to view 360° videos, where the videos are shown to the participants according to the four classes of emotions from Russell's model as explained earlier in Fig. 1. The setup, hardware, and demography are further discussed in "Experimental setup, hardware and demography" section. After the data acquisition phase is completed, the recorded data were then uploaded to the cloud through the Empatica Connect smartphone application. The smartphone application is used to view, record and process the HR reading in real-time.

In the pre-processing phase, the data recorded is downloaded from the cloud, where each of the participant's data is then labeled synchronously with the time of the videos to match the emotional responses to the HR signals. This is done to all the 20 participant's data set. Next is the classification phase, where three classifiers were used to compare which achieved the best performance in terms of accuracy. Two types of classification processes were executed during the experiment, an intra-subject classification, which involves classification within a specific individual and inter-subject classification, which classifies the combined data across all participants. The next subsection will discuss further the VR content selection in this experiment.

The selection content for VR

VR was used in this experimental setup as the device to present the stimuli emotional content to the subjects, the emotional content is in the form of 360° videos which are used in evoking responses that are emotional from the subjects. The four quadrants previously shown in Fig. 1 evoke a specific emotion in the content chosen. Short videos of four were presented to the subjects for every quadrant to evoke a high and responses that are sustained by the subjects for the emotions focused. In each completed quadrant, subjects are



given a rest period of 10-s where visual stimuli are not present, this is to allow the subjects to reset their emotional state to a baseline in preparation for the following quadrant's. Hence, 16 videos in total were stitched and compiled together for the overall experiment across each emotional quadrant of four. The experiment flow presented in Fig. 3 shows the video time and quadrant shown to the subjects. The total duration for the overall content that was compiled and shown to the subject is 6 min and 5 s that includes the baseline period to reset the subject emotional state.

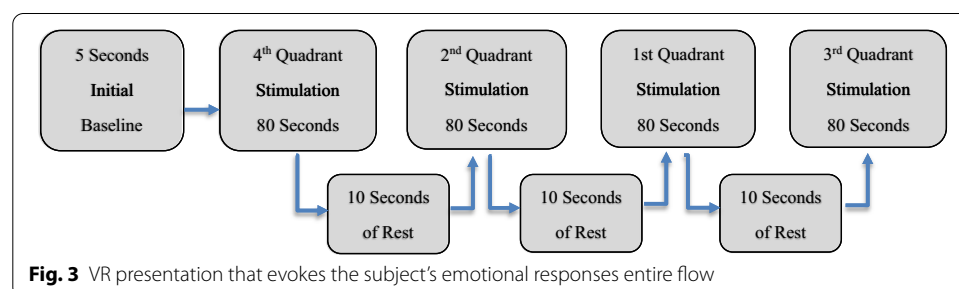
Prior to starting an experiment, first, the subject was given a thorough explanation of the experiment flow, then after they have understood the flow, they were then asked to sign a consent form. Next, the VR headset is mounted to the subject head while an adjustment of the straps was done to ensure the subject level of comfort. Then, the earphone is plugged in the subject ear to provide a deeper immersive experience.

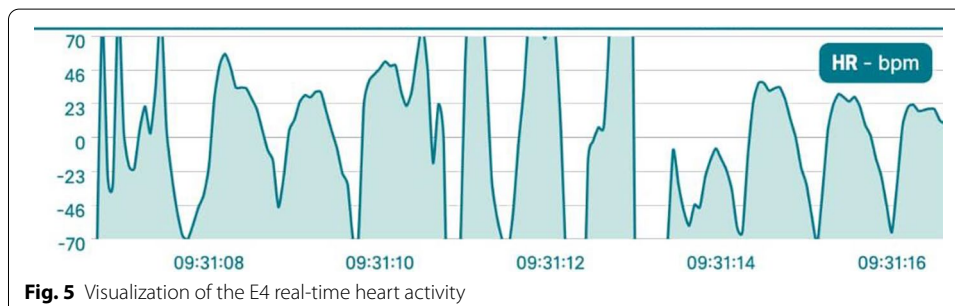
Experimental setup, hardware and demography

This experiment has a total of 20 subjects (12 males and 8 females) that participated in ranging from 20 to 28 years old, consisting of individuals who are working or studying. Before the experiment started, all subjects were briefed about the side effect that they may potentially experience such as motion sickness, headache, dizziness, and nausea. These are the side effect that VR usage is known to have.

The Empatica E4 wearable is the wristband used as the hardware in recording the subject heart rate, it is a wearable device that is medical-grade used in data acquisition related to physiological. It is a wearable device, hence it is non-invasive towards the subject when obtaining the HR via a photoplethysmography sensor which works by detecting the activity within the blood volume, optically, the skin indicates the pulse of the user, which measures the HR at a fixed frequency of 64 Hz by default of the wearable device and is non-customizable. The device is usually placed on the right or left wrist of the subject. After the wristband is placed, the earphones, VR Headset, the subject is then placed in a seating position, when the 360° video is started, the heart rate is simultaneously captured wirelessly via the E4 real-time application.

In Fig. 4, the wearable device shown is the Empatica E4 placed on the subject wrist in Fig. 4a, and in Fig. 4b is the Empatica E4 Real-time application used during the experiment to acquire the participant's heart activity. The Empatica E4 real-time application enables a real-time view of the HR and EDG signals via smartphone. There are four available data to stream via the application, including, heart rate, skin activity, acceleration and temperature. The application uploads the data recorded automatically to the





server which can be downloaded from the E4 Connect page. Other than that, it is used to view the session and battery level of the E4 wearable device.

The subject HR data will be acquired using the application on the smartphone shown in Fig. 5 during the data acquisition with the Empatica E4.

The VR headset shown in Fig. 6 was used for the participants to watch the video content for each quadrant in 360 degrees. The HTC Vive VR headset which allows the subject to be immersed completely into their surrounding, according to the content at that time being played. Within the 360° video environment, subjects can freely move around by changing the trajectory of their focus.

Discussion and results

Average BPM, max, and min

After the completion of the experiment, the data acquired through the real-time application from the Empatica E4 wearable was transferred to the cloud, the transferred data can then be downloaded in Common Separated Value (CSV) format to be



Fig. 6 Subject with the HTV vive VR headset mounted

processed. Next are the results of the 20 subjects that participated in this experiment. Each row of the data was synchronized to the timestamp that reflects on each of the classes in the four quadrants.

The 20 subject's data acquired from the experiment are shown in Fig. 7. This figure presents the heart rate average, maximum, and minimum for intra-subject (individual). The range of HR activity with the lowest Beat Per Minute (BPM) is from the third subject with 54.6 bpm, subject 18 had the highest bpm with 110.6 bpm, while the 20 subject's entire total average bpm was 76.1 bpm.

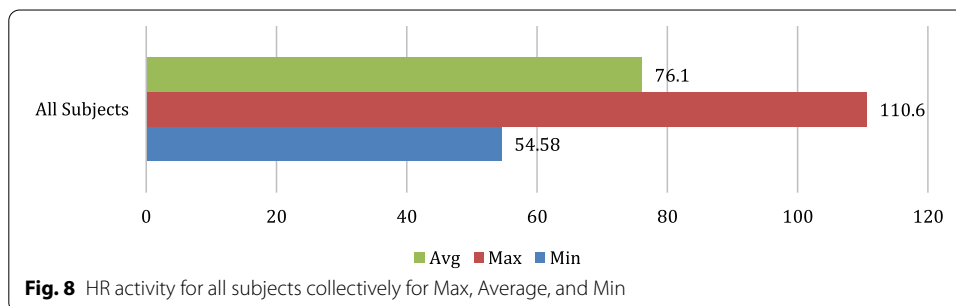
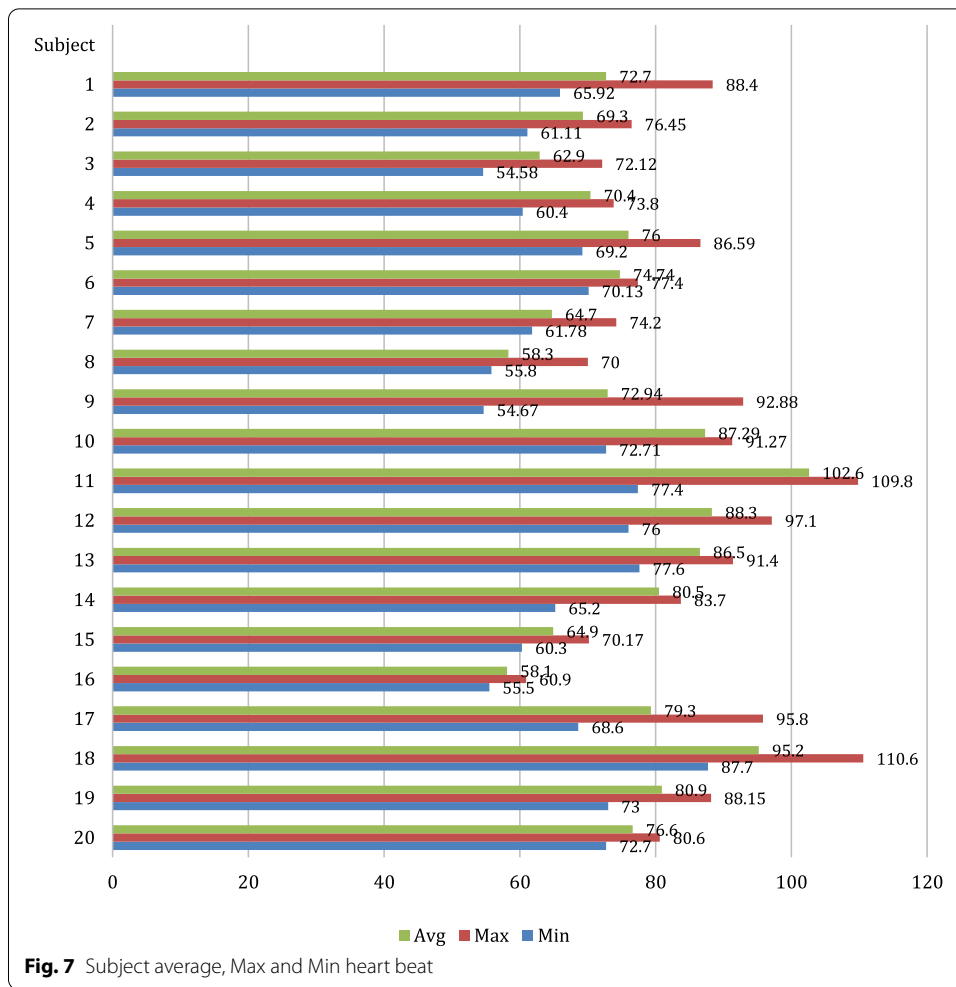
Figure 8 shows the max, min, and average BPM of all 20 subjects data collectively. The average for 20 subjects is 76.1 BPM, while the max is 110.6 BPM and min is 54.6 BPM.

SVM, KNN and Random Forest classifiers on individual and overall-subject classification

Intra-subject classification refers to the data collected from the subject individually, the same subject data is then used for testing and training, whereas inter-subject classification refers to the data collected from the 20 subjects overall rather than individually that is then used for testing and training over all subjects. Python was used in the classification of the intra-subject data individually, while the inter-subject data was grouped collectively from the 20 subjects into a single CSV file. The experiment used cross-validation of ten-fold.

The input feature used in the machine learning classifier is the HR data collected from the subjects, an attempt to use HR data solely as the feature in emotion classification of four different classes. SVM, RF and KNN were used as the approach for machine learning by using Python, this is to recognize the emotion classification that is based on the heart rate activity retrieved from the experiment with the subject that is the direct response to the 360° video stimuli shown in the experiment.

The formula for calculating the classification accuracy using the three classifiers, namely SVM, KNN, and RF, is as follows:



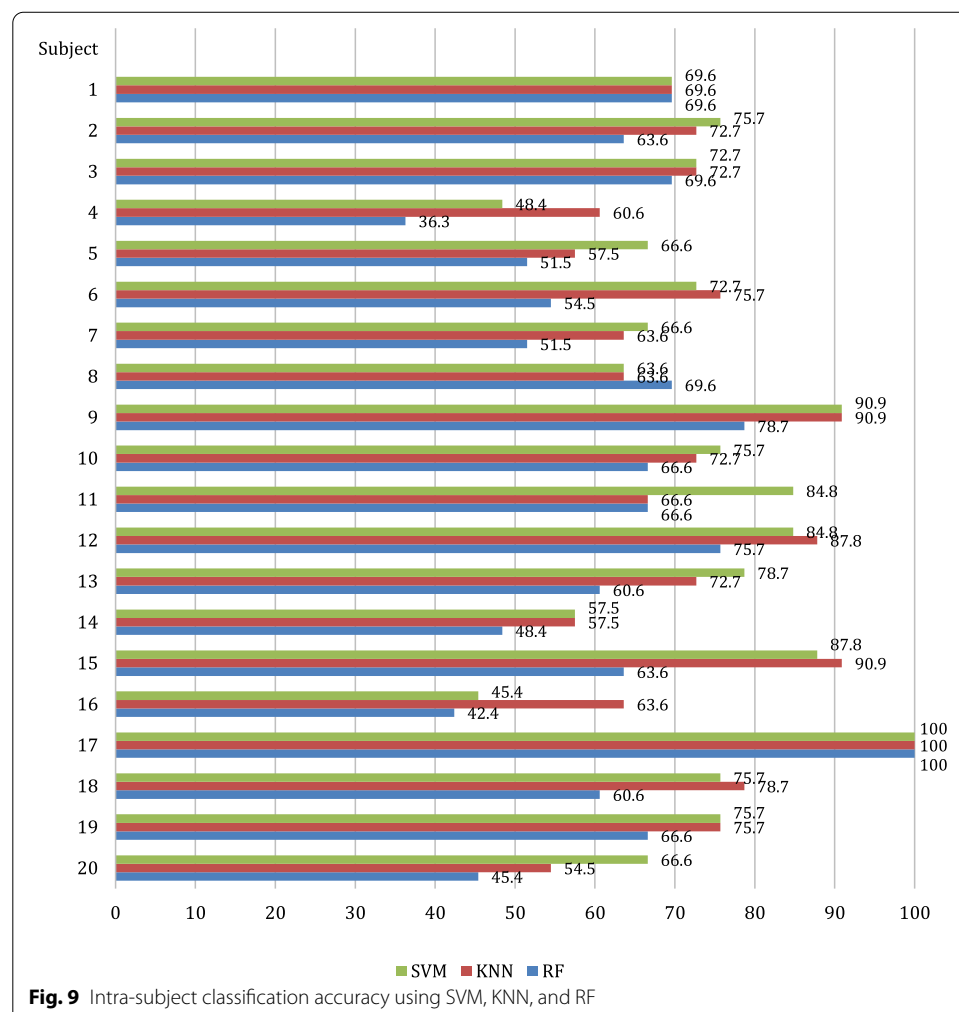
$$Accuracy = \frac{TP/TN}{TP + TN + FP + FN}$$

where TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative [10].

The classification for both inter-subject and intra-subject accuracy results of the experiments using SVM, KNN, and RF classifier is presented in the following figures.

Figure 9 shows the accuracy results for SVM, KNN, and RF classifiers for intra-subject classification. SVM's accuracy results range from the lowest accuracy of 45.4% from participant 16 and to the highest accuracy of 100% from participant 17; KNN's results range from the lowest accuracy of 54.5% from participant 20 and to the highest accuracy of 100% from participant 17; while RF's results ranged from the lowest accuracy of 36.3% from participant 4 to the highest accuracy of 100% from participant 17. All three classifier has one subject that achieved 100% from participant 17, while the lowest-performing subject is participant 4 with 36.3% with the RF classifier.

The results for inter-subject classification results using SVM, KNN, and Random Forest are shown in Fig. 10. The inter-subject classification generated accuracies of 46.7 for SVM, 42.9% for KNN and 43.3% for Random Forest, which shows that inter-subject classification of emotions using HR is significantly harder compared to intra-subject classification, which is to be expected since the HR varies considerably across different individuals when exposed to emotional stimuli.



Confusion matrix analysis for inter-subject classification

Table 2 below shows the confusion matrix for 20 inter-subject classifications across the four-class emotional quadrants. The principal diagonal represents the percentage of the successful recognition of each class.

The result shows the prediction of the classification for SVM for where the emotion is classified in four-classes starting with low arousal/positive valence (LA/PV), low arousal/negative valence (LA/NV), high arousal/negative valence (HA/NV) and (HA/PV). Table 2 shows across participants, the most difficult quadrant to predict was LA/PV at only 40% while the most successful prediction came from the HA/NV quadrant at 96%. Overall, it also appears that classifying negative valence emotions were easier than positive valence emotions.

Results comparison against related studies

Table 3 presents the result comparison of using HR signals.

This table shows the result comparison of other results besides the result obtained from this experiment, where the highest accuracy was 84% while this experiment achieved between 46.7% and 100% depending on whether it was predicted as an intra-subject classification problem or an inter-subject classification problem. As previously explained, neither of the other two published studies explained their experimental protocol in terms of whether it was focused on solving the intra-subject or inter-subject classification. Furthermore, both of these studies only classified two distinct emotions focusing solely on the valence axis only, which is clearly a far simpler classification task compared to our more challenging and arguably more useful approach of classifying into four distinct classes across both the valence and arousal dimensions of emotions according to Russell's model of emotions.

Conclusion

This paper prime objective was investigating the use of HR signals could be the sole feature used in the classification of human emotions into four distinct quadrants according to Russell's Circumplex Model of Emotions when emotional stimuli were presented to participants via a VR environment. From the experimental results, it was found that four-class intra-subject emotion classification yielded accuracies ranging from 45.4% to 100% for SVM, 54.5% to 100% for KNN, and 36.3% to 100% for Random Forest while four-class inter-subject emotion classification yielded accuracies of 46.7% for SVM, 42.9% for KNN and 43.3% for RF. These results show that HR

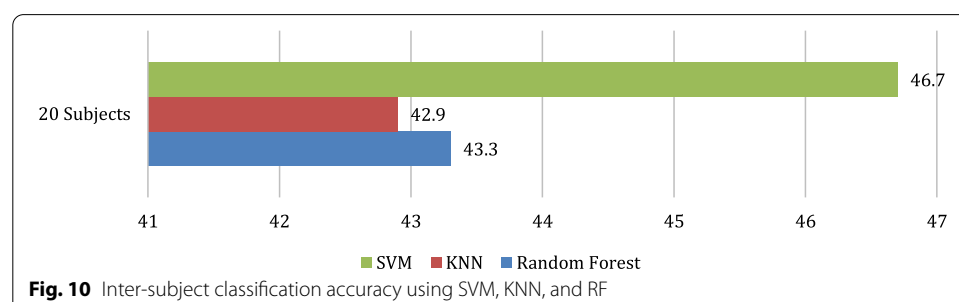


Table 2 Confusion matrix for 20 inter-subject

Prediction	Low arousal/positive valence (↓ +) (%)	Low arousal/negative valence (↓ −) (%)	High arousal/negative valence (↑ −) (%)	High arousal/positive valence (↑ +) (%)
LA/PV (↓ +)	40	49	43	21
LA/NV (↓ −)	16	88	35	7.4
HA/NV (↑ −)	18	35	96	23.5
HA/HV (↑ +)	21	24	25	79

Table 3 Result comparison with other researchers

Author	Signals	Participants	Stimuli	Classifier	Accuracy
[11]	HR	25	China Emotional Video Stimuli (CEVS)	Gradient boosting decision tree	84% (2 emotion classes only)
[9]	HR	5	Android application	SVM	79% (2 emotion classes only)
Our results	HR	20	Virtual reality (VR) 360° videos	SVM, KNN, RF	100% for intra-subject and 46.7% for inter-subject (4 emotion classes)

is an excellent potential approach in emotion classification for four-class emotion using VR, in particular for predicting emotions in four classes for data specific to the participants since intra-subject classification produced accuracies of more than 80% using SVM and KNN (Tables 2 and 3).

In our next immediate work, we will be expanding the study to 30 participants. Furthermore, the use of skin conductance data as a different and possibly added modality for emotion classification will be investigated as well since electrodermography (EDG) data is also acquired by the Empatica E4 wristband. A comparison between the HR and EDG as well as using both signals for emotion classification will be explored. We also intend to apply deep learning and feature extraction approaches to compare the classification performance against SVM, KNN, and RF.

Abbreviations

2D: Two dimension (2D Space emotion model); 360: Three hundred sixty (Type of Stimuli use that allows subjects to move and view the video freely); BPM: Beats per minute (Heart rate of subjects recorded in number of beats per minute); CSV: Comma separated values (Excel file that stores the recorded data of subjects); DEAP: Database for Emotion Analysis Using Physiological Signals (A database containing physiological signals for emotion analysis); E4: Empatica E4 (A wearable by Empatica used in collecting heart rate and skin activity of the subject); EDG: Electrodermography (The electrical activity of the skin); HA/NV: High arousal/negative valence (Scale based on Russell's emotion model); HA/PV: High arousal/positive valence (Scale based on Russell's emotion model); HR: Heart Rate (The speed of the heart beating); Hz: Hertz (The frequency used in acquiring skin activity data from the subject); KNN: K-Nearest Neighbor (Type of classifier in machine learning); LA/NV: Low arousal/negative valence (Scale based on Russell's emotion model); LA/PV: Low arousal/positive valence (Scale based on Russell's emotion model); Python: Python Programming Language (Python is a high-level and general-purpose programming language); SVM: Support Vector Machine (Type of classifier in machine learning); VR: Virtual reality (The simulated experience that can be similar or different from the real world).

Acknowledgments

This work was supported by the Ministry of Higher Education, Malaysia [Grant Number FRGS0512-1/2019].

Authors' contributions

AFB—Writing—original draft. JM—Writing—Review and editing, supervision. JT—Writing—review and editing, supervision, funding acquisition. All authors read and approved the final manuscript.

Funding

This work was supported by the Ministry of Higher Education, Malaysia [Grant Number FRGS0512-1/2019].

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Received: 18 August 2020 Accepted: 14 December 2020

Published online: 07 January 2021

References

1. Alarcao SM, Fonseca MJ. Emotions recognition using EEG signals: a survey. *IEEE Trans Affect Comput*. 2017;10(3):374–93. <https://doi.org/10.1109/TAFFC.2017.2714671>.
2. Ali M, Machot AH, Al F, Kyamakya K. emotion recognition involving physiological and speech signals: a comprehensive review. In: *Studies in systems, decision and control*. vol. 109. Springer International Publishing; p. 287–302; 2018.
3. Baig MZ, Kavakli M. A survey on psycho-physiological analysis & measurement methods in multimodal systems. *Multimodal Technol Interact*. 2019;3(2):37.
4. Koelsch S, Jacobs AM, Menninghaus W, Liebal K. The quartet theory of human emotions: an integrative and neuro-functional model. *Phys Life Rev*. 2015;13:1–27. <https://doi.org/10.1016/j.plrev.2015.03.001>.
5. Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, Patras I. DEAP: a database for emotion analysis; using physiological signals. *IEEE Trans Affect Comput*. 2011;3(1):18–31.
6. Kumar J, Kumar J. A machine learning approach to classify emotions using GSR. 2015;2(12):72–76.
7. Ménard M, Richard P, Hamdi H, Daucé B, Yamaguchi T. Emotion recognition based on heart rate and skin conductance. In: *PhyCS 2015—2nd international conference on physiological computing systems, proceedings*, 26–32. 2015. <https://doi.org/10.5220/0005241100260032>.
8. Minhad, KN., Ali, S. HMD., and Reaz, MBI. (2017) "A Design Framework for Human Emotion Recognition Using Electrocardiogram and Skin Conductance Response Signals," *J. Eng. Sci. Technol.*, vol. 12, no. 11, pp. 3102–3119, 2017.
9. Nguyen NT, Nguyen NV, My Huynh T, Tran, Nguyen Binh T. A potential approach for emotion prediction using heart rate signals. In: *9th international conference on knowledge and systems engineering (KSE)*, Ho Chi Minh city, Vietnam. 2017.
10. Setyohadi DB, Kusrohmaniah S, Gunawan SB. Galvanic skin response data classification for emotion detection. *Int J Electr Comput Eng*. 2018. <https://doi.org/10.11591/ijece.v8i5.pp4004-4014>.
11. Shu L, Yu Y, Chen W, Hua H, Li Q, Jin J, Xu X. Wearable emotion recognition using heart rate data from a smart bracelet. *MDPI Sensors*. 2020;20:718. <https://doi.org/10.3390/s20030718>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
