Journal of Engineering and Applied Sciences 13 (8): 2137-2144, 2018
ISSN: 1816-949X

Empirical Analysis of Intra vs. Inter-Subject Variability in VR EEG-Based Emotion Modelling

[1]N.S. Suhaimi, [1]J. Teo and [2]J. Mountstephens
[1]Evolutionary Computing Laboratory, Faculty of Computing and Informatics,
Universiti Malaysia Sabah (UMS), Jalan, 88400 Kota Kinabalu, Sabah, Malaysia
[2]Faculty of Computing and Informatics, Universiti Malaysia Sabah (UMS),
Jalan, 88400, Kota Kinabalu, Sabah, Malaysia

Abstract: This study presents the classification of emotions on EEG signals using commercial BCI headsets known as wearable EEG. One of the key issues in this research is the lack of mental classification using VR as the medium to stimulate emotion. Moreover, we endeavor to present the first comprehensive and systematic analysis of intra-versus inter-subject variability in EEG-based emotion classification using VR and wearable EEG. The approach towards this research is by using K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) as the machine learning classifiers. Firstly, each of the participants will be required to wear the EEG headset to record their brain waves when they are immersed inside the VR environment. The data points are then marked if they showed any physical signs of emotion or by observing the brain wave pattern. Secondly, the data will then be tested and trained with KNN and SVM algorithms. We conduct subject-dependent as well as subject-independent classifications in order to compare intra-against inter-subject variability, respectively in VR EEG-based emotion modeling. The highest subject-dependent classification accuracy achieved was 97.9% while the highest subject-independent classification accuracy obtained was 91.4% throughout the brain wave spectrum (α, β, γ, δ, θ). These methods showed highly promising results and will be further enhanced using other machine learning approaches such as deep learning in VR stimulus.

Key words: Neuroinformatics, emotion recognition, k-Nearest Neighbor (kNN), Support Vector Machine (SVM), Electroencephalography (EEG), Virtual Reality (VR)

INTRODUCTION

Virtual Reality (VR) has been on the rise for every industry especially in the smartphone industry where they want to capitalize this feature. Oculus, HTC Vive and Samsung Gear are some of the notable industries competing in the VR field. Some of the notable features of using VR from applications are such as gaming and 3360 videos that allows people to see and observe what other people have recorded throughout the journey and their surrounding environments.

In many of the demos from the companies that have conducted with VR, the participants would be immersed into the VR and demonstrated body movements and emotions that showed signs of fear such as riding a roller coaster and losing their sense of ground due to cognitive manipulation. However, there were not any application that have the adaptivity to capitalize the mental state of the user that could help the user which would then improve the user's experience inside the VR world.

The mental state of the user which is an ongoing electrical activity of the brain, particularly known as Electroencephalography (EEG) which shows the ongoing sensory such as motor movements, emotions and memories. This would provide an indication of the level of engagement for the user immersing into the VR which could be measured and quantified.

There are numerous researchers which have conducted in neuroinformatic to obtain the brainwave signals from humans. However, none has ever approached with the method of classifying mental state of users using VR as the stimulus. Technology has evolved throughout the years and EEG equipment have scaled down and are easily accessible by public to obtain.

In this research, we will be using the Muse headband which was developed by Interaxon for recording of the brainwave activity happening within the user's brain when they are immersed into the VR. The approach for this research will require the user to not have any motion sickness and impaired eyes. With the data obtained from

Corresponding Author: N.S. Suhaimi, Evolutionary Computing Laboratory, Faculty of Computing and Informatics, Universiti Malaysia Sabah (UMS), Jalan, 88400 Kota Kinabalu, Sabah, Malaysia

the brainwave, the data will then be filtered with KNN and SVM obtained from the brainwave such as muscle movements and eye twitches using R language data analysis. They will then be tested and trained with machine learning language to identify emotional brainwave patterns to allow adaptive interaction between VR and the current user's mental state.

Literature review: Numerous researchers had conducted emotion classification using various stimulus to obtain a certain emotion state of a person. One such is the use of music to classify the emotion a person felt from listening to them (Lin *et al*., 2010). Another use of emotional classification is by implicit tagging of a video and music videos (Koelstra and Patras, 2013; Koelstra *et al*., 2010) whereby the neurophysiological signals. facial expressions and the EEG signals of the participants were recorded. Each of the participants were then required to do a Self-Assessment Mannikin (SAM) in order to classify their emotions such as arousal, valence and domination. Games such as "Tetris" (Chanel *et al*., 2011) were used to navigate the blocks placement by using physiological signals and EEG signals to measure their stress level while playing the game itself. Emotion classification through text and speech based (Sreeja and Mahalakshmi, 2016; Dong-Mei and Fang, 2007) had also been conducted to test the model performance.

Virtual reality: Recently, the use of virtual reality as the stimulus to assess in emotion classification gain the attention to fully understand how the human brain research. The study of VR using footsteps to navigate through the environment (Pugnetti *et al*., 1996; Kober and Neuper, 2012) and to understand whether these footsteps would deepen the immersion of the presence inside the VR. A treatment used to heal subjects from affective disorders, brain injury or neurocognitive deficits by assessing through VRCPAT simulation under stress conditions (Wu *et al*., 2010). Neurophysiological signals and EEG signals were obtained from the experiment and it was found that the theta brain wave signal was found to be reliable in the activation of arousal. Simulation of a plane's movement inside the simulation (Li *et al*., 2015) was done according to the mental state of the user by relaxing or concentrating. Some of the classification such as KNN and SVM were widely used in classifying emotion and have recorded an average accuracy of 70% and above which suggests that the methods employed are possible in properly classifying each of the emotional state the user are feeling (AlZoubi *et al*., 2009; Yazdani *et al*., 2009; Hosseinifard *et al*., 2013; Subasi and Gursoy, 2010; Zulkifli *et al*., 2015).

The difficulty in obtaining EEG equipment: EEG equipment are difficult to obtain and peruse as the equipment itself are large and are not portable to be brought out for external use. Additionally, the cost of the equipment are high and are usually stored in the laboratory for safety purposes. Therefore, the accessibility to use the EEG equipment makes it difficult for any experimentation to be conducted. However, modern day technology has improved dramatically and scalable for consumers to obtain their EEG equipment. One such company is Interaxon which provided accessibility for consumers to be able to purchase and obtain the equipment with little to no maintenance required for meditation as a normal usage or for researchers (Li *et al*., 2015; Cassani *et al*., 2015; Galway *et al*., 2015; Karydis *et al*., 2015; Mann *et al*., 2015) to conduct their experimentation with the SDK tools provided by the company itself.

MATERIALS AND METHODS

About 13 (2 Female, 11 Male) participants were recruited from different background and culture aged between 22 through 65 who are residing within Sabah, Malaysia. Before the experiment began, each participant had been examined thoroughly to not have any prior brain damage, vision impairment, deafness, motion sickness and mental disorder. A third-party application known as "Muse Monitor" developed by James Clutterback was installed onto a smartphone to assist in observing the EEG signals from each participant and recording their EEG signals and uploading onto a cloud storage for further analysis using classification techniques such as KNN and SVM. Each of the participant were then asked to wear the Muse headband version onto the forehead and above their earlobes as shown in Fig. 1 (TP9, AF7, AF8, TP10 and the reference point at FPz) while making sure the connection between the device and the skin are properly receiving the signals from the brain.

Next, the participant were asked to wear VR box headset along with a 5.5" smartphone attached onto the VR headset with a pair of earphones connected to the smartphone. The smartphone attached inside the VR headset is installed with an application known as "Sisters" from the Google playstore as shown in Fig. 2 and 3, the application shown to provide a stimulus which would incite fear onto the participant. Each participant was then required to explore the surroundings of the VR environment and observe the objects within. The participants are allowed to move freely however they not allowed to move excessively as the connection between the muse headset and their skin contact might be lost and

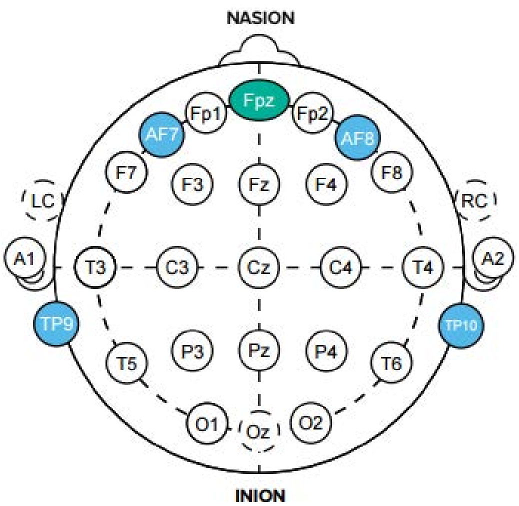

Fig. 1: The position of the head where the participants EEG signals was recorded

Fig. 2: The application used to stimulate the participant for evoking fear

Fig. 3: The VR environment inside “Sisters”

a large gap in between the EEG signals would not be usable for classification. Any external stimulus that would divert the attention of the participant were taken with great care to reduce any unwanted noise while recording the EEG signal. The collected data will then be analyzed using both KNN and SVM methods to classify the mental state of the participants and train the models to be able to identify and recognize the brainwave patterns and evaluate its performance on the analysis.

RESULTS AND DISCUSSION

The experiments were recorded from the third party application known as “Muse Monitor” and each of the data recordings were recorded at an interval of 0.5 sec to avoid any huge gap between the EEG signals. The recordings were then uploaded using the Comma-Separated-Value (CSV) format to be able to accommodate with the R language used for classification. Each of the recorded data contains information from different bands with each of the electrodes obtained from the Muse EEG headset performing the Fast-Fourier Transform (FFT) and their raw EEG signals. The horseshoe indicator shows the connectivity between the human skin and the electrode provided with marker buttons. Figure 4 provides an example of the information recorded from the Muse Monitor application.

The data is then rearranged to feed the data classifier and was saved in a different filename to avoid any changes to the original data which will be required for future references. The band of interest: α, β, Δ, δ and θ. A length of 20 data points (approximately 1 data points for every 0.5 sec) was selected before and after the marker was pressed to highlight the significant events that displayed physical emotional changes such as the doll that made a jump scare or any creepy motions that was displayed within the room. Resting state data points was selected from the data points before the subject was subjected to any stimulations (Fig. 5 and 6).

During the experimentation, there were a couple of participants whom did not portray any physical emotion such as jumpy body motions from the jump scare or shouts due to some scary scenes happening within the VR environment which made it difficult to put a timestamp marker on the data points to portray they were scared. In this case, we then marked according to the highest peak value of the EEG signal based from the Y-axis of the raw signal obtained from the recordings and the approximated time of the event which stimulates fear for the participants. One of the datasets could not be used due to a large gap of EEG signals and has to be removed from the analysis.

The datasets that were selected to be trained was 90% and the remaining 10% was to be trained and evaluate the performance of the classification. Additionally, a 10-fold cross validation was implemented into the classification process to automatically segregate the datasets into 9:1 parts for training and testing for KNN and multiple SVM classifiers such as SVM Class

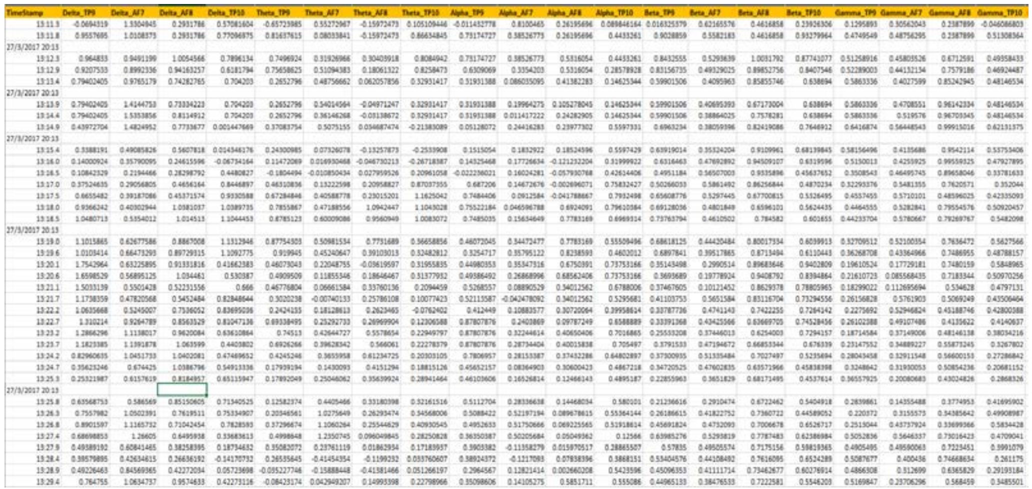

Fig. 4: The data that was recorded from muse monitor application

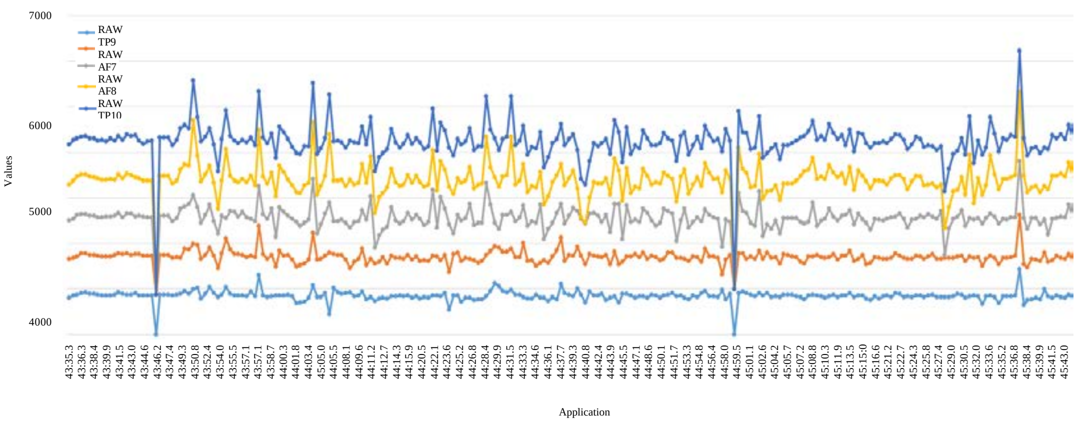

Fig. 5: RAW brainwave pattern obtained from the recording of the muse monitor application

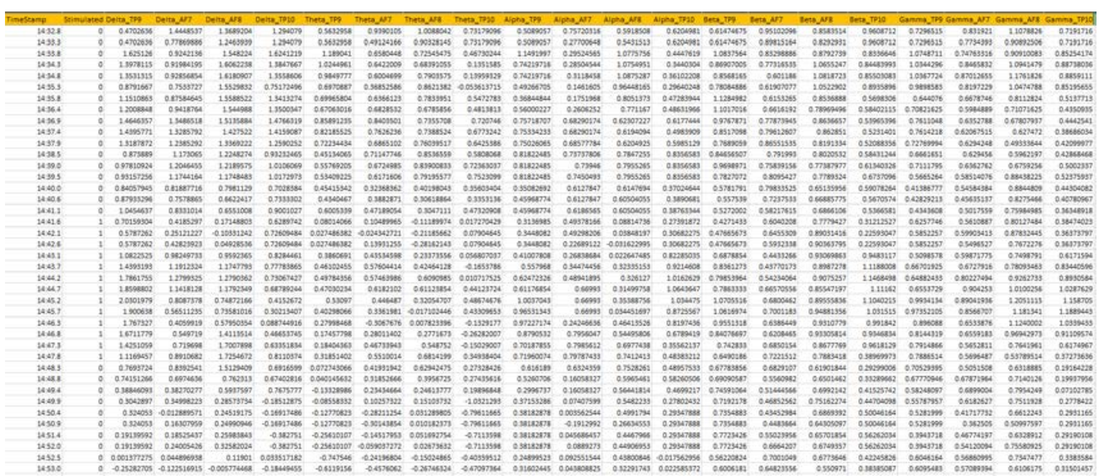

Fig. 6: Adjustment of data to be fed into the classifier

Table 1: The Accuracy of individual participant (intra-subject variability) using 10-fold CV with KNN classifier to train and test the observed brainwave patterns

Participant	K = 5 (%)	K = 7 (%)	K = 9 (%)	K = 11 (%)	K = 13 (%)	K = 15 (%)	K = 17 (%)	K = 19 (%)	K = 21 (%)	K = 23 (%)
A	90.83	90.83	90.00	88.96	88.54	88.33	88.13	87.92	87.71	87.71
B	87.59	86.72	85.32	85.34	85.81	85.81	85.83	85.38	85.81	85.81
C	81.71	83.20	81.97	82.22	83.21	82.97	82.73	82.23	82.23	82.23
D	94.07	93.13	93.13	92.51	92.04	92.04	90.48	88.92	87.84	87.21
E	82.72	81.12	78.67	77.31	78.13	77.08	77.10	76.04	74.94	75.76
F	88.40	87.10	86.68	83.38	82.55	82.55	79.63	78.38	77.98	77.15
G	97.48	96.93	96.93	96.93	96.93	96.93	96.93	96.93	96.93	96.93
H	94.32	93.84	93.84	93.37	94.06	93.83	93.83	94.55	94.31	94.31
I	94.06	93.72	92.33	92.68	92.34	90.25	89.90	89.90	89.90	89.90
J	84.35	84.84	84.10	85.10	83.85	84.85	83.59	84.34	84.34	85.10
K	90.67	88.88	90.34	90.68	90.68	90.68	90.68	90.68	90.68	90.68
L	92.05	91.03	90.33	91.71	91.71	91.02	91.02	91.02	91.02	91.02
Average accuracy	89.86	82.41	81.82	81.55	81.53	81.26	80.76	80.48	80.28	80.29

Table 2: The accuracy of individual participant (intra-subject variability) using 10-fold CV with multiple SVM classifiers to train and test the observed brainwave patterns

Participant	Class Weights (CW) (%)	Linear Kernel (LK) (%)	Linear Kernel 2 (LK2) (%)	Polynomial Kernel (PK) (%)	Radial Basis Function Kernel (RBFK) (%)	Radial Basis Function Kernel (RBFK radial cost) (%)	Radial Basis Function Kernel (RBFK radial sigma) (%)
A	95.43	90.43	90.01	94.59	92.71	92.50	94.16
B	90.35	86.21	85.02	88.45	85.30	88.91	88.48
C	88.38	82.23	82.23	87.91	82.97	82.98	84.20
D	96.88	87.99	87.99	96.26	90.65	90.80	92.19
E	85.45	78.17	79.48	84.12	80.59	79.53	80.05
F	93.80	83.42	84.65	93.35	86.30	88.75	89.63
G	98.33	96.63	96.93	97.75	96.92	96.93	96.92
H	97.14	94.52	94.79	95.76	94.32	94.31	95.02
I	97.92	89.91	93.02	97.88	93.03	93.02	94.41
J	87.38	84.36	84.61	87.61	85.36	85.36	85.36
K	93.86	91.76	91.38	94.33	90.67	90.67	90.67
L	96.24	92.09	92.75	95.16	91.75	91.02	92.41
Average accuracy	86.24	81.36	81.76	85.63	82.35	82.68	83.35

Weights (SVM-CW), SVM Linear Kernel (SVM-LK), SVM Polynomial Kernel (SVM-PK), SVM Radial Basis Function Kernel (SVM-RBFK), Radial Cost (SVM-RBFK-RC) and Radial Sigma (SVM-RBFK-RS). The evaluation was conducted using intra-subject and inter-subject method. Intra-subject is conducted by training and testing specifically configured for an individual while inter-subject is conducted by combining all of the individual samples and testing the whole datasets performance. Figure 7 shows the selected datasets for training and Fig. 8 shows the selected datasets to be tested for the performance evaluation.

Figure 7 is an example of the selected datasets used for training and Fig. 8 is an example of the selected data used for testing to evaluate the performance for KNN and SVM running at 10-fold cross validation. The results for intra-subject variability are shown in Table 1-4 shows the inter-subject variability results. Figure 9 provides a visual comparison of the best performance obtained from both KNN with “k” value of 5 and SVM-CW. Figure 10 provides a visual comparison of the average accuracy performance obtained from KNN with “k” value of 5 and the multiple

Table 3: Overall accuracy of the classifier for KNN using 10-fold CV (inter-subject variability)

Classifier	“k” value	Accuracy (%)
KNN	5	89.57
	7	88.89
	9	88.25
	11	87.90
	13	87.72
	15	87.61
	17	87.29
	19	87.63
	21	87.45
	23	87.08

Table 4: Overall accuracy of the multiple classifiers for SVM using 10-fold CV (inter-subject variability)

Classifier Kernel Function	Average accuracy (%)
SVM	
Class Weights (CW)	91.40
Linear Kernel (LK)	86.81
Linear Kernel 2 (LK2)	86.81
Polynomial Kernel (PK)	89.64
Radial Basis Function Kernel (RBFK)	87.27
Radial Basis Function Kernel (RBFK Radial Cost)	87.08
Radial Basis Function Kernel (RBFK Radial Sigma)	87.77

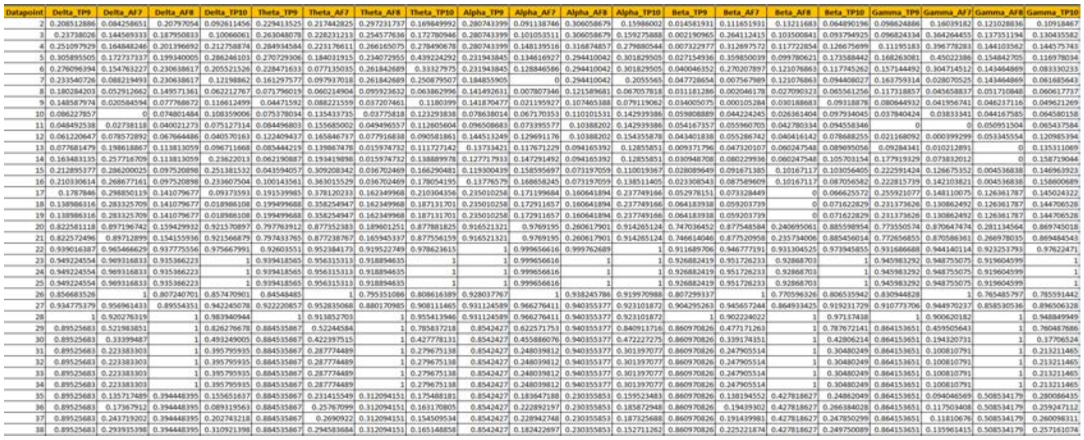

Fig. 7: Datasets which are selected to be trained

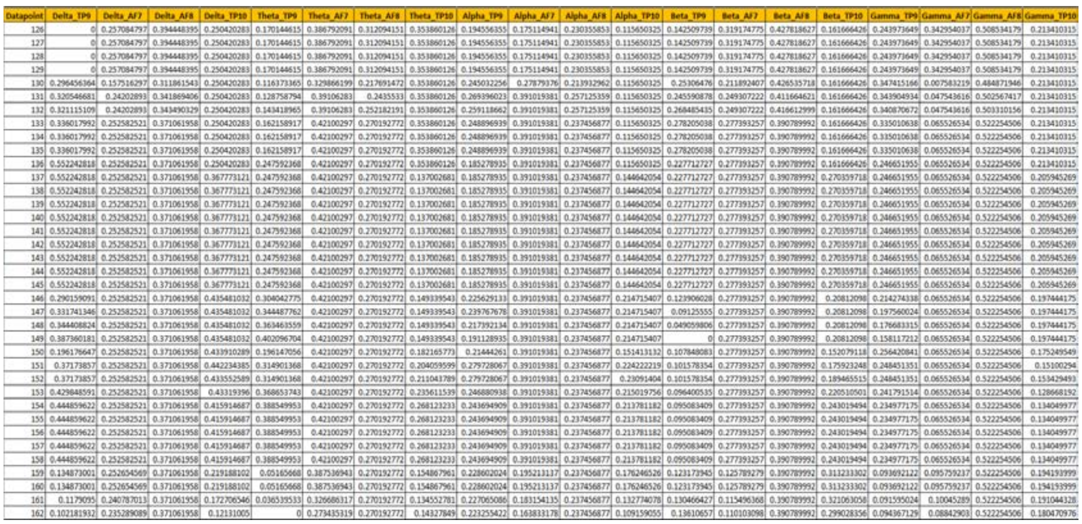

Fig. 8: Dataset which are selected to be tested for the accuracy prediction

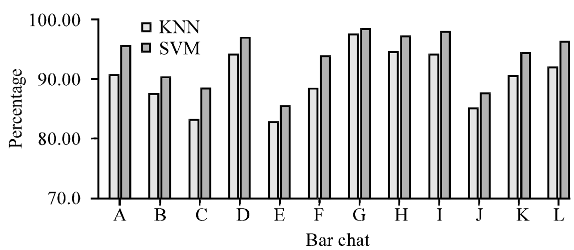

Fig. 9: Bar chart accuracy representation for intra-subject variability comparison from Table 4 and 5 with KNN and SVM classifiers

kernel functions from SVM. Figure 11 provides a visual comparison of the best performance obtained from both KNN and SVM.

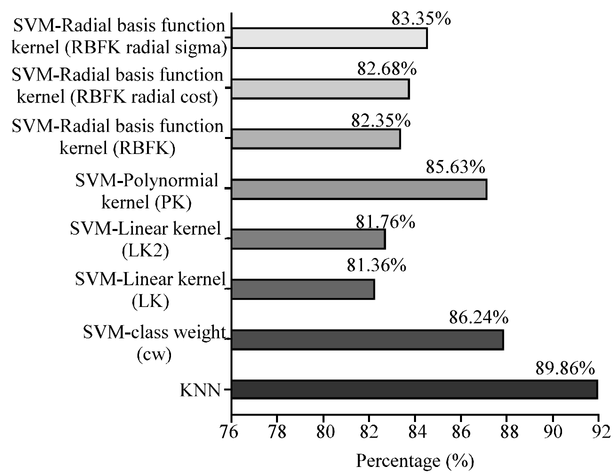

Fig. 10: Average accuracy of each classifier with 10-fold CV in intra-subject variability

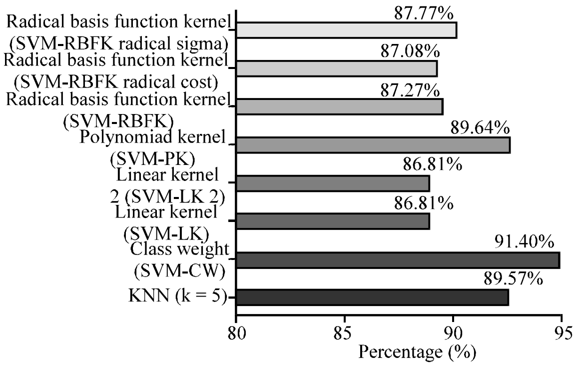

Fig. 11: Accuracy of each highest performance classifiers with 10-fold CV in inter-subject variability

CONCLUSION

From the data obtained, there were differences between the intra-subject and inter-subject variability performance. The highest accuracy obtained from individual participants was at 97.48% in KNN whilst SVM obtained the highest accuracy form individual participants at 97.92%. There seems to be a small drop in accuracy performance for KNN in inter-subject variability even when the optimal "k" value found was at a value of 5 between intra-subject (89.86% on average) and inter subject variability (89.57% on average). SVM performance shows the highest accuracy performance over KNN. Inter subject variability does show a larger improvement of accuracy over the intra-subject variability in SVM classifiers with the kernel function of class weights from inter-subject variability achieving the highest accuracy at 91.40% compared to the intra-subject variability at 86.24%. SVM classification shows a greater opportunity in classifying emotion under VR and inter-subject variability over KNN classifier.

For future recommendations, the number of participants will be increased up to 40 participants to be able to obtain a clearer and convincing data on the accuracy of the mental classification. Other suggestions of mental classifications would be to conduct a different type of emotions to have a better grasp on the EEG signals.

Some of the participants were not able to fit into the Muse headband most likely due to the small circumference of the head as the Muse headband tends to fit into a more circular shaped head. This would also make recordings of the EEG signals difficult as the conductive band does not have proper skin contact.

Lastly, there is the difficulty of setting up a proper environment for the participant to fully immerse into the VR as not all participant were available to travel to the vicinity of the experiment that was conducted. Instead, we have to arrange logistics to travel to their premise and or using an empty ground area to conduct the experiment while the surrounding environment was dark and quiet. All necessary steps to avoid external stimulus were reassured before the conduct of the research.

ACKNOWLEDGEMENT

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

REFERENCES

AlZoubi, O., R.A. Calvo and R.H. Stevens, 2009. Classification of EEG for Affect Recognition: An Adaptive Approach. In: AI 2009: Advances in Artificial Intelligence, Nicholson A. and X. Li (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-10438-1, pp: 52-61.

Cassani, R., H. Banville and T.H. Falk, 2015. MuLES: An open source EEG acquisition and streaming server for quick and simple prototyping and recording. Proceedings of the 20th International Conference on Intelligent User Interfaces Companion, March 29-April 01, 2015, ACM, Atlanta, Georgia, ISBN:978-1-4503-3308-5, pp: 9-12.

Chanel, G., C. Rebetez, M. Betrancourt and T. Pun, 2011. Emotion assessment from physiological signals for adaptation of game difficulty. IEEE. Trans. Syst. Man Cybern. Part A. Syst. Hum., 41: 1052-1063.

Dong-Mei, Y.U. and J.A. Fang, 2007. Research on a methodology to model speech emotion. J. Eng. Appl. Sci., 2: 1262-1267.

Galway, L., P. McCullagh, G. Lightbody, C. Brennan and D. Trainor, 2015. The potential of the brain-computer interface for learning: A technology review. Proceedings of the 2015 IEEE International Conference on Computer and Information Technology Ubiquitous Computing and Communications: Dependable, Autonomic and Secure Computing Pervasive Intelligence and Computing (CIT-IUCC-DASC-PICOM'15), October 26-28, 2015, IEEE, Liverpool, UK., ISBN: 978-1-5090-0154-5, pp: 1554-1559.

Hosseinifard, B., M.H. Moradi and R. Rostami, 2013. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. Comput. Methods Programs Biomed., 109: 339-345.

Kober, S.E. and C. Neuper, 2012. Using auditory event-related EEG potentials to assess presence in virtual reality. Intl. J. Hum. Comput. Stud., 70: 577-587.

Koelstra, S. and I. Patras, 2013. Fusion of facial expressions and EEG for implicit affective tagging. Image Vision Comput., 31: 164-174.

Koelstra, S., A. Yazdani, M. Soleymani, C. Muhl and J.S. Lee *et al*., 2010. Single Trial Classification of EEG and Peripheral Physiological Signals for Recognition of Emotions Induced by Music Videos. In: Brain Informatics, Yao, Y., R. Sun, T. Poggio, J. Liu and N. Zhong, *et al*., (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-15313-6, pp: 89-100.

Li, Z., J. Xu and T. Zhu, 2015. Prediction of brain states of concentration and relaxation in real time with portable electroencephalographs. Hum. Comput. Interact., 1: 1-18.

Lin, Y.P., C.H. Wang, T.P. Jung, T.L. Wu and S.K. Jeng *et al*., 2010. EEG-based emotion recognition in music listening. IEEE. Trans. Biomed. Eng., 57: 1798-1806.

Mann, S., R., Janzen, H. Wu, M.H. Lu and N. Guleria, 2015. Bright ideas: A wearable interactive Inventometer (brainwave-based idea display). Proceedings of the 2015 IEEE Conference on Games Entertainment Media (GEM'15), October 14-16, 2015, IEEE, Toronto, Ontario, ISBN:978-1-4673-7453-8, pp: 1-8.

Pugnetti, L., L. Mendozzi, E. Barberi, F.D. Rose and E.A. Attree, 1996. Nervous system correlates of virtual reality experience. Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technology, July 8-10, 1996, University of Reading, Reading, England, ISBN:0-7049-1140-X, pp: 239-246.

Sreeja, P.S. and G.S. Mahalakshmi, 2016. Comparison of probabilistic corpus-based method and vector space model for emotion recognition from poems. Asian J. Inf. Technol., 15: 908-915.

Subasi, A. and M.I. Gursoy, 2010. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst. Appl., 37: 8659-8666.

Wang, X.W., D. Nie and B.L. Lu, 2014. Emotional state classification from EEG data using machine learning approach. Neurocomput., 129: 94-106.

Yazdani, A., T. Ebrahimi and U. Hoffmann, 2009. Classification of EEG signals using Dempster Shafer theory and a K-nearest neighbor classifier. Proceedings of the 4th International IEEE-EMBS Conference on Neural Engineering (NER'09), April 29-May 2, 2009, IEEE, Antalya, Turkey, ISBN: 978-1-4244-2072-8, pp: 327-330.

Zulkifli, N.A.A., S.H. Sawal, S.A. Ahmad and M.S. Islam, 2015. Review on Support Vector Machine (SVM) classifier for human emotion pattern recognition from EEG signals. Asian J. Inf. Technol., 14: 135-146.