

EEG-Based Analysis for Learning through Virtual Reality Environment

Sayed Ahmed Alwedaie, Habib Al Khabbaz, Sayed Redha Hadi and Riyadh Al-Hakim*

Department of Computer Engineering, College of Information Technology, University of Bahrain, Bahrain

Abstract

Recently, many researchers studied learning through VR environment in various fields. Their assessment tools were based on tests, quizzes and/or statistical analysis of questionnaires. This study is based on the analysis of EEG signals collected from the students' brains directly to capture their feelings and engagement during the lecture in both traditional and VR methods of teaching.

To recognize the emotions of the students, the fine K-Nearest Neighbor (KNN) algorithm is used. To calculate the engagement score for a student, a well-known engagement score formula issued.

The participants chosen are students of Anatomy and Physiology course. All participants were subject to three sessions of EEG signal acquisition for both Real Lecture and Virtual Reality, each session is five-minutes long. For better accuracy, EEG signals were captured three times for each student in each lecturing method. Based on the data-analyzing methods applied, which are Dependent Paired Samples T-Test and Independent Paired Samples T-Test, positive emotions in a real lecture are better than positive emotions in a VR-Lecture. However, the engagement score in both classes was approximately the same.

Keywords: Electroencephalography (EEG); Emotion classification; Engagement; Education comparison; Virtual reality (VR); Physiology and anatomy

Introduction

Recently, online learning has offered students a time-saving education plan, as they do not need to attend a classroom. Despite these advantages, online education lacks interactivity available with traditional learning.

Virtual reality (VR), which has been defined as I3, which stands for "Immersion-Interaction- Imagination" [1], is a technology that has become extremely popular in recent years. VR promises to deliver the best aspects of both traditional and online distance learning into a single platform. VR tools can generate realistic images, sounds, videos and other sensory inputs that replicate an environment in which the student will be immersed.

Over the past decade, many researchers compared VR with traditional learning in various fields. Most of their assessment tools were based on tests, quizzes and/or statistical analysis of questionnaires.

Crosier [2] investigated the potential of Virtual Reality (VR) for teaching radioactivity at secondary school level. Evaluation was carried out in a local school and compared directly to the traditional. Results indicated that both ability level and the order in which the conditions were completed significantly affected the attitude scores. High ability students reported higher attitude scores in general, and specifically in VR classes.

Cantwell [3] conducted a modified pre-test/post-test and attitude study to determine the effectiveness of virtual field trips, gain insight into how they are best used in an introductory Earth Science course. He concluded that the virtual field trip did not successfully mimic teaching observation and data evaluation learning goals; however, it was able to address question and hypothesis posing skills and establish an appreciation for the complexity of a scientific issue.

Moazami et al. [4] enrolled 35, fifth year undergraduate dental students from the School of Dentistry, Shiraz University of Medical Sciences. With the same lecturer, they used two different learning models (virtual and traditional) to plan a course on the topic "rotary

instrumentation of root canals". The study groups completed their courses over three consecutive weeks. Their improvement was assessed immediately and two months after completion of the course by a valid, reliable test. Despite the difficulties encountered in designing the virtual learning environment, the study was conducted successfully. Based on the findings of this study, the virtual learning was more effective than lecture-based training.

Nicholson et al. [5] demonstrated, in contrast to earlier studies, that a computer-based 3D anatomical model enhances medical students' learning of ear anatomy.

Ng CL et al. [6] innovated a 3D model to facilitate teaching of the complicated anatomic area anatomy of the epitympanum of middle ear. Their study demonstrated that it is efficacious in short-term recall. By allowing the learner to visualize relations of the epitympanum from all directions, the model aids the learner in appreciation of anatomy and identification of structures of this region.

Codd and Choudhury evaluated the use of 3D virtual reality when compared with traditional anatomy teaching methods. Three groups were identified from the University of Manchester second year Human Anatomy Research Skills Module class: a "control" group (no prior knowledge of forearm anatomy), a "traditional methods" group (taught using dissection and textbooks), and a "model" group (taught solely using e-resource). The groups were assessed on anatomy of the forearm by a ten-question practical examination. ANOVA analysis showed the model group mean test score to be significantly higher

*Corresponding author: Riyadh Al-Hakim, Department of Computer Engineering, College of Information Technology, University of Bahrain, Bahrain, Tel: +973 1743 7684; Fax: +973 1743 7334; E-mail: rkamal@uob.edu.bh

Received January 10, 2018; Accepted January 27, 2018; Published February 03, 2018

Citation: Alwedaie SA, Khabbaz HA, Hadi SR, Al-Hakim R (2018) EEG-Based Analysis for Learning through Virtual Reality Environment. J Biosens Bioelectron 9: 249. doi: 10.4172/2155-6210.1000249

Copyright: © 2018 Alwedaie SA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

than the control group and not significantly different to the traditional methods group. Feedback from all users of the e-resource was positive. Virtual reality anatomy learning can be used to compliment traditional teaching methods effectively.

In this study, a novel approach of comparison is implemented. The comparison is based on capturing and statistically analyzing raw EEG data (emotion and engagement) from participants, (who are Physiology and Anatomy students in Biology department in University of Bahrain), in both the class room traditional lecture and VR lecture. Figure 1 shows a participant in a traditional lecture. The same participant during a VR lecture session is shown in Figure 2.

Materials

Course selection

A Physiology and Anatomy course was chosen in this research because it is one of the best fields for education in the virtual reality. Biology department in University of Bahrain is teaching this course to students. They have used the conventional method that is used in most of the global universities, in which the professor explains the topic in a lecture hall and uses light projector to display slides with some explanatory graphs.

Participants

Data were collected from seven different (volunteers) students. These data were recorded by electrodes using 10-20 system. 14 electrodes were used to obtain raw data. For each student, three different signals were recorded on different real classes and three different signals were recorded on VR classes.



Figure 1: A participant during a Real-Lecture session.



Figure 2: A participant during a VR-Lecture session.

EEG recording

Raw EEG data were captured from participants three times for each student (male) in real lecture and virtual lecture. Five minutes for each time. The data were recorded when students were listening to the real lecture like a normal situation. For virtual lecture, students set in computer engineer labs and wear VR glasses to watch the lecture. The captured Raw EEG data were acquired by EMOTIV EPOC+ device with fourteen sensors, which are AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1 and O2. EMOTIV EPOC+ at 128 Hz sampling rate.

Method

In this study, we are proposing a new methodology in comparison field based on EEG signals that gives more accurate results than the traditional methods used. We analyze EEG signals to extract the engagement and emotion of a participant for comparison.

Signal preprocessing

EEG raw data need preprocessing before applying any classification. In this research, FFT algorithm is used to convert EEG raw data from time domain to frequency domain, which allows dividing the data in different frequency bands as required by further steps. EEG signals have common noise in all sensors over 500 times greater than the actual value.

Median filter was used to reject common noise between sensors. It calculates the median value of all fourteen sensors and subtracts it from each sensor value. Unsigned integer is used for data transmission, which makes large DC offset. First order high pass filter is used to remove the DC offset from the data. The cut-off frequency used for the first high pass filter is 0.16 Hz. The DC offset can also be eliminated by subtracting the background level from the signal sum through using the IIR filter. Slew rate limiting is a technique that is used to remove spikes in EEG signals, the rising and falling slew rate parameter is 50 μ V and -50 μ V, respectively.

Feature extraction and selection

According to the experiments done by the scientists, it was discovered that there are different electrical frequency bands that affect in different activities. Berger has found that there is a relation between those different frequency bands and the corresponding different activities. Berger noticed in his experiments that there is frequency band between 8 Hz to 12 Hz with clearly oscillating pattern in specific activities and he called it alpha waves. It was the first discovered frequency band in the brain waves field. The scientist noticed that there are five ranges of the frequency in the brain wave, they could distinguish between them, and the signals changed clearly in these different ranges depending on the activities as they noticed in their experiments. They called these different frequency bands delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) from lowest frequency range to highest. Berger introduced alpha and beta waves in 1929. Walter and Dovey introduced theta wave that is below the alpha range in 1944 and Walter found the delta rhythm below the theta wave in 1936 [7-9]. The highest range amongst brain waves is the gamma wave, as introduced by Jasper and Andrews in 1938. The frequency that ranges from 0 Hz to 4 Hz is called delta waves. Delta waves are the slowest waves and they occur in sleep and wake state as the scientists notice. The frequency that ranges from 4 Hz to 8 Hz is called theta waves. In many experiments, it was discovered that theta has oscillated and increased in response to higher workload and it distinguishes a wake and sleep state. Emotional stress such as frustration and disappointment affect in theta waves clearly.

The frequency that ranges from 8 Hz to 12 Hz is called alpha waves. It is changed clearly when the person is in wakeful relaxation state. Alpha waves increase when the person is in relaxation state having his eyes closed. The change occurs clearly in the specific regions in the brain. The frequency that ranges from 12 Hz to 30 Hz is called beta waves. The beta waves sometimes are divided into two ranges to be more specific. It is fast and small, focus concentration does affect in the beta waves. Frontal and central of the brain the best area that where beta waves are clearly functional. The frequency that ranges above 30 Hz is called gamma waves. Beta and gamma waves are related to attention, perception, and cognition. It is clearly functional during attention and memory relevant tasks. Gamma waves oscillations increment occurs during object recognition activity, which is part of the selective attention [10,11].

Detection of useful information directly from the raw EEG data is difficult. The scientists have used preprocessing for raw EEG data to make it easy to extract features from the raw data. Extracting features means that the scientist takes the noticeable characteristics in the raw data. Features extraction helps researchers to describe the data and recognize some information from EEG signals such as emotion. The feature extraction is very useful in classification to classify numbers of classes as explained in the next section. All the fourteen sensors were used to extract the features. In each two seconds, 256 values were taken because the sampling rate is 128 Hz. The data is then converted to frequency domain and divided to the five frequency bands, delta, theta, alpha, beta and gamma. For each sensor, mean and standard deviation is applied to each frequency band in each two second. These features are used in classification, as explained in the next section.

Signal classification

In order to classify the signals, KNN was used as the classification method. Its job is to calculate the distance between the predicted data and training example data. In MATLAB, all types of SVM (Support Vector Machine) and KNN classification algorithms were tried in order to choose the best classification method based on the accuracy. Fine KNN was chosen because it has provided the best accuracy, which is approximately equal to 67.7%. The distance metric option used in this study was Euclidean and the Number of Neighbors is equal to 1, as the default parameters set by MATLAB. In this study, the classification is divided into two parts, classification with valence (high valence and low valence) and classification with arousal (high arousal and low arousal). Accuracy measuring of the classification is provided automatically by MATLAB after performing the train command on the example data.

Engagement score

Engagement is described as alertness and attention based on specific task [8]. The characterization of engagement is expressed by increased physiological arousal and beta waves along with attenuated alpha waves [8].

Engagement indices: Three engagement indices were used: $\beta/(\alpha+\theta)$, β/α , and $1/\alpha$. These indices were used by Pope et al. and were collected from the EEG literature on attention and vigilance [12-14]. According to Pope et al. study, EEG was recorded from most four effective electrode locations: Cz, Pz, P3, and P4. Fast Fourier Transform was used to extract bin powers from these four locations. The combined alpha, theta, and beta power were used to derive each index. Based on experiments, it is found that the first engagement index, $\beta/(\alpha+\theta)$, resulted better performance based on hypotheses[15].

Engagement score calculation: Eight channels F3, F4, FC5, FC6,

P7, P8, O1, and O2 were used from the EMOTIV neuroheadset. These channels were selected due to their closeness to those used by Pope et al. The program will start reading the data that comes from those channels and fill them into two-second window. This window of signals first goes through noise and artifacts removal step and then goes through different band-pass filters to obtain β , α , and θ , the combined power in the ranges of 13-30 Hz, 5-13 Hz and 4-8 Hz frequency.

The instantaneous user engagement is represented by the Engagement Index (EI) equation (1). This formula gives the best engagement as explained before. The window is continuously shifted and it calculates new instantaneous Engagement Index based on the EMOTIV's sampling rate (128Hz). To measure the lower bounds and upper bounds of a participant, lowest value and highest value are taken from over all the instantaneous engagement indices. These two values are used to scale the Engagement Score on 0 and 1 scale. The Engagement Score was an average of 32 instantaneous Engagement indices over a 2 second window. New engagement value is calculated every 250 ms as the EMOTIV neuroheadset has a sampling frequency 128 Hz.

In order to scale the Engagement Score on 0 and 1 scale, the following linear formula was used:

$$ES = \frac{EI - \min(EI)}{\max(EI) - \min(EI)}$$

ES=Engagement Score between 0 and 1.

EI=averaged Engagement Index over a 2 second window.

min(EI)=minimum Engagement Index.

max(EI)=maximum Engagement Index.

The Engagement Score was calculated for each participant. Each participant had an Engagement Score between 0 and 1. This value was multiplied by 100 to represent the Engagement Score in percentage for later comparison.

Types of statistics comparison

Two types of statistics comparison were used in this research, dependent paired samples t-test and independent paired samples t-test. The details instructions will be described in the following sections.

Discussion

Positive emotion statistics

Table 1 shows the results of positive emotions for each student during real classes and VR classes, the first Column shows the student Id, second column shows the 3 different positive emotions that were collected among 3 different real classes, third column shows the positive emotions that were collected among 3 different VR classes. The data are shown according to the percentage value of the positive emotions. This data have been analyzed and is eligible to be involved in further statistics.

Student ID	Real classes			VR classes		
	Emotion 1	Emotion 2	Emotion 3	Emotion 1	Emotion 2	Emotion 3
Student 1	52.77%	21.67%	45.18%	17.69%	20.00%	30.09%
Student 2	83.07%	22.05%	29.70%	23.70%	26.27%	33.33%
Student 3	26.15%	17.24%	31.53%	11.11%	6.29%	24.80%
Student 4	32.62%	13.19%	15.62%	11.95%	18.03%	18.64%
Student 5	44.44%	17.36%	11.18	31.66%	22.75	8.46%
Student 6	9.09%	35.83%	19.69%	10.41%	10.56%	2.98%
Student 7	42.30%	67.14%	-	15.73%	26.31%	30.61%

Table 1: Percentage of positive emotion for each student in real class and in VRclass.

Table 1 shows the data paired, because the same individuals were involved in both trials (Real classes and VR) and the number of individuals is less than 10. For these reasons, the Paired Samples T Test will be more appropriate than other types of tests. Paired Samples T Test is used to know if there is any difference between two different paired tests. In this case, we will check if there is any difference between VR classes and normal classes.

One of the important conditions for the use of Paired Samples T Test is the normal distribution. The average of positive emotion percentage of the student will be used in this test. Table 2 below shows the average of each student.

Normal distribution was calculated based on average of the real and VR classes. The Confidence Interval for mean was 95% so $\alpha=0.05$. Table 3 shows useful statistics of the data. It shows that the mean of real classes is 32.98% and the mean of VR classes is 19.11%. The standard deviation of the real classes is 13.47 and VR is 6.76%.

Shapiro-Wilk Test found that the data is normally distributed. Based on this, the condition of Paired Samples T Test is met. Table 4 shows the result of Shapiro-Wilk test. The value of this test for real

Student ID	Average of real classes	Average of VR classes
Student 1	39.87%	22.59%
Student 2	44.94%	27.77%
Student 3	24.97%	14.07%
Student 4	20.48%	16.20%
Student 5	24.33%	20.96%
Student 6	21.54%	7.98%
Student 7	54.72%	24.22%

Table 2: Average percentage of positive emotion for each student in real class and in VR class.

	Statistic	Std. Error
Class		
Mean	32.9786	5.09259
95% Confidence Interval for Mean: Lower Bound	20.5175	
Upper Bound	45.4397	
5% Trimmed Mean	32.4651	
Median	24.9700	
Variance	181.541	
Std. Deviation	13.47372	
Minimum	20.48	
Maximum	54.72	
Range	34.24	
Interquartile Range	23.40	
Skewness	.735	.794
Kurtosis	-1.189-	1.587
VR		
Mean	19.1129	2.55598
95% Confidence Interval for Mean: Lower Bound	12.8586	
Upper Bound	25.3671	
5% Trimmed Mean	19.2504	
Median	20.9600	
Variance	45.731	
Std. Deviation	6.76248	
Minimum	7.98	
Maximum	27.77	
Range	19.79	
Interquartile Range	10.15	
Skewness	-.530-	.794
Kurtosis	-.371-	1.587

Table 3: Useful statistics of the data.

classes was 0.154 and for VR classes was 0.894. α is Much less than 0.154, so the condition is met.

Paired Samples Correlations test shows that there is a strong correlation between this data. The relationship between this data is 0.781, which is strongly noticeable. Table 5 shows the correlations result of the data.

In order to use Paired Samples T test, we have established suitable null and alternative hypotheses:

Null Hypothesis H0: Class=VR

Alternate Hypothesis HA: Class>VR.

Paired Samples T test was applied, it shows the mean of (real classes – VR classes) as 13.87, the error of mean as 3.49 and the standard deviation as 9.22. Table 6 below shows this data from SPSS software.

According to the above table, the t is 3.978 and the significance level is 0.007. It is smaller than α . Because of that, the Null Hypothesis will be rejected. So, the Alternate Hypothesis will be accepted. The mean of the real classes is much larger than VR classes. Therefore, the real classes will be more effective in terms of positive emotion and it always gives much better result.

Engagement score statistics

Table 7 demonstrates the results of engagement for each student during real classes and VR classes, the primary column demonstrates the student Id, second column demonstrates the 3 different engagements that were collected among 3 different real classes, third column

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Class	0.295	7	0.065	0.861	7	0.154
VR	0.179	7	0.200	0.969	7	0.894

Table 4: Result of Shapiro-Wilk test.

	N	Correlation	Sig.
Pair 1 Class and VR	7	.781	.038

Table 5: Correlations result of the data.

	Paired Differences				
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the Difference	
				Lower	Upper
Pair 1 Class and VR	13.86571	9.22144	3.48538	5.3373	22.39413
Paired Samples Test					
	t	df	sig. (2-tailed)		
Pair 1 Class and VR	3.978	6	0.007		

Table 6: Data from SPSS software.

Student ID	Real classes			VR classes		
Student 1	61.60%	71.50%	58.82%	74.028%	56.27%	55.54%
Student 2	49.13%	65.36%	66.21%	69.83%	52.53%	60.31%
Student 3	66.48%	60.18%	59.17%	69.35%	57.96%	61.17%
Student 4	68.19%	52.31%	71.64%	63.58%	58.29%	68.36%
Student 5	63.25%	62.15%	67.40%	69.91%	73.24%	62.07%
Student 6	58.92%	64.74%	78.04%	60.13%	68.36%	60.02%
Student 7	55.52%	45.98%	-	64.01%	58.62%	44.71%

Table 7: Percentage of engagement score for each student in real class and in VR class.

demonstrates the engagements that were collected among 3 different VR classes. Note: The data is shown according to the percentage value of the engagement. This data has been analyzed and is eligible to be involved in further statistics.

Table 7 demonstrates this data paired, with reference to the fact that the same students were included in both trials (real classes and VR) and the number of students is under 10. Consequently, the Paired Samples T-Test will be more suitable than other sort of tests. For this situation, we will check if there is any difference between VR classes and real Classes. One of the essential conditions for the utilization of Paired Samples T Test is the normal distribution. The average of the student will be utilized as a part of this test. Table 8 below demonstrates the average of each student.

Normal distribution was figured using average of the real and VR classes. The confidence Interval for mean was 95% so the $\alpha=0.05$. Table 9 shows useful statistics of the data. It demonstrates the mean of real classes as 61.78% and the mean of VR classes as 62.30%, the Std. Deviation of the real classes as 5.32% and VR as 3.74%.

Student ID	Average of real classes	Average of VR classes
Student 1	63.97%	61.95%
Student 2	60.23%	60.98%
Student 3	61.94%	62.83%
Student 4	64.05%	63.41%
Student 5	64.27%	68.41%
Student 6	67.23%	62.84%
Student 7	50.75%	55.78%

Table 8: Average percentage of engagement score for each student in real class and in VR class.

Class	Statistic	Std. Error
Mean	61.7771	2.01152
95% Confidence Interval for Mean: Lower Bound	56.8551	
Upper Bound	66.6691	
5% Trimmed Mean	62.0868	
Median	63.9700	
Variance	28.323	
Std. Deviation	5.32198	
Minimum	50.75	
Maximum	67.23	
Range	16.48	
Interquartile Range	4.04	
Skewness	-1.772	.794
Kurtosis	3.770	1.587
VR		
Mean	62.3014	1.41217
95% Confidence Interval for Mean: Lower Bound	58.8460	
Upper Bound	65.7569	
5% Trimmed Mean	62.3244	
Median	62.8300	
Variance	13.960	
Std. Deviation	3.73626	
Minimum	55.78	
Maximum	68.41	
Range	12.63	
Interquartile Range	2.52	
Skewness	-.226-	.794
Kurtosis	2.279	1.587

Table 9: Useful statistics of the data.

Shapiro-Wilk Test found that the data is normally distributed. Based on that, the condition of Paired Samples T Test is met. Table 10 demonstrates the result of Shapiro-Wilk test. The value of this test for real classes was 0.064 and for VR classes was 0.439. α is less than 0.064, so the condition is met.

Paired Samples Correlations test results indicate that there is a strong correlation among this data. The correlation between this data is 0.789, which is strongly noticeable. Table 11 demonstrates the correlations result of the data.

It must determine a Null hypothesis and Alternative hypothesis for a paired sample t-test.

Null Hypothesis H0: Class=VR

Alternate Hypothesis HA: VR>Class

We assume in Alternate Hypothesis VR>Class because the mean of VR was larger than Class. Paired Samples T test applied, it demonstrates the mean of (real classes – VR classes) as -0.52, the error of mean as 1.2 and the standard deviation as 3.3. Table 12 demonstrates this information from SPSS software.

As indicated by above table, the t is -0.42 and the significance level is 0.689. It is much bigger than α . According to this result, there is not enough evidence to reject the null hypothesis.

As a result of that, the Null Hypothesis will be accepted. In the result, involving more samples might lead to find a difference between engagements.

If we consider the samples as independents, the independent sample T-Test can be used. In this case, the final result will be the same for both engagement and positive emotion. In positive emotion, the t is equal to 2.4 and the significance is equal to 0.018, which is less than α . Based on the above, the final result will match the result of Paired T-Test. For engagement, the t is -0.218 and the significance will be 0.485, which is bigger than α , so the null hypothesis is accepted. Hence, the final result will remain the same regardless the approach used.

Conclusion

In this research, three-different EEG datasets were used for

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Class	0.243	7	0.200	0.818	7	0.62
VR	0.240	7	0.200	0.916	7	0.439

Table 10: Result of Shapiro-Wilk test.

	N	Correlation	Sig.
Pair 1 Class and VR	7	.789	.035

Table 11: Correlations result of the data.

	Paired Differences				
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the Difference	
				Lower	Upper
Pair 1 Class and VR	-.52429	3.30419	1.24887	-3.58016	2.53158
Paired Samples Test					
	t	df	sig. (2-tailed)		
Pair 1 Class and VR	-.420	6	.689		

Table 12: Information from SPSS software.

emotion classification (SEED, DEAP, and a dataset provided by a researcher). The dataset that is provided by the researcher Nadzeri et al. presents the best accuracy for this research usage. In classification, several classification algorithms (all types of SVM and KNN) were tested, and the best classification algorithm was used amongst all. Fine KNN algorithm was the best classification algorithm based on its accuracy, which is equal to 67.7%. Regarding the calculation of engagement score, a formula was used to calculate the engagement score. This formula provides the most effective results for calculating the engagement, which is confirmed by a research supported by NASA. Dependent paired samples t-test and independent paired samples t-test are two types of statistics comparison were used in this research. Both of them led to the same result. Mean percentage in positive emotion for real classes was 32.98% and for VR class was 19.11%. However, the standard deviation for real classes was 13.47% and for VR class was 6.76%. The result shows that the positive emotion in real classes is better than the VR classes. In statistics of engagement score, mean percentage was founded as 61.78% in real classes and 62.30% in VR class. Moreover, the standard deviation for real class was calculated as 5.32% and for VR as 3.74%. The result shows that both classes have almost same engagement from students.

Therefore, the study proves that the real lectures cannot be replaced with VR sessions but the VR sessions could help students for self-study to make some kind of topics clearer. Besides that, it could help students who are looking for study by affiliation. Also, it could help teachers to explain complex visual objects through VR easily. In future work, it would be better if more samples were participated on different courses instead of Physiology and Anatomy course.

References

1. Burdea GC, Coiffet P (2003) *Virtual reality technology*. (2nd edn.), John Wiley and Sons, NY.
2. Crosier JK (2000) *Experimental Comparison of Virtual Reality with Traditional Teaching Methods for Teaching Radioactivity*. *Education and Information Technologies* 5: 329-343.
3. Cantwell LB (2004) *A Comparison of Learning: Integration of A Virtual and Traditional Field Trip into an Introductory Environmental Geology Course*.
4. Moazami F, Bahrapour E, Azar MR, Moattari M (2014) *Comparing two methods of education (virtual versus traditional) on learning of Iranian dental students: a post-test only design study*. *BMC Medical Education* 14:45.
5. Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ (2006) *Can virtual reality improve anatomy education? A randomized controlled study of a computer-generated three-dimensional anatomical ear model*. *Medical Education* 40: 1057-1150.
6. Ng CL, Liu X, Chee SCJ, Ngo RYS (2015) *An Innovative 3-dimensional Model of the Epitympanum for Teaching of Middle Ear Anatomy*. *Otolaryngology-Head and Neck Surgery* 153: 832-837.
7. Codd AM, Choudhury B (2011) *Virtual Reality Anatomy: Is it Comparable with Traditional Methods in the Teaching of Human Forearm Musculoskeletal Anatomy?* *Anatomical Sciences Education* 4: 119-125.
8. Emotiv (2011) *Emotiv Homepage*.
9. Xu XL, Xu B, He B (2004) *An alternative subspace approach to EEG dipole source localization*. *Phys. Med. Biol* 49: 327-343.
10. Larsen E (2011) *Classification of EEG Signals in a Brain-Computer Interface System*.
11. Hettich D (2016) *Classification of Affective States in the Electroencephalogram*.
12. Streitberg B, Rohmel J, Herrmann W (1987) *COMSTAT Rule for Vigilance Classification Based on Spontaneous EEG Activity*.
13. Davidson R (1988) *EEG Measures of Cerebral Asymmetry: Conceptual and Methodological Issues*. *International Journal of Neuroscience* 39: 71-89.
14. Lubar J (1991) *Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders*. *Biofeedback and Self-Regulation* 16: 201-225.
15. Freeman F, Mikulka P, Prinzel L, Scerbo M (1999) *Evaluation of an adaptive automation system using three EEG indices with a visual tracking task*. *Biological Psychology* 50: 61-76.