



Bio-physiological-signals-based VR cybersickness detection

Chenxin Qu¹ · Xiaoping Che¹ · Siqi Ma¹ · Shuqin Zhu¹

Received: 29 November 2021 / Accepted: 27 March 2022 / Published online: 23 May 2022
© China Computer Federation (CCF) 2022

Abstract

With the gradual maturity of virtual reality (VR) technology in recent years, VR industry is in a trend of rapid growth, providing new possibilities for content design. Although VR technology has been able to provide users with excellent immersive experience, side effects that affect the user experience still exist, especially the cybersickness. It would cause extreme physical discomfort to the users and the discontinuation of use. Many researchers have tried to find the inducement of cybersickness and to detect and limit the occurrence of this symptom, but most of the current detection and analysis methods rely on subjective questionnaires to collect users' posterior states, such as dizziness, nausea, cold sweats, disorientation, eyestrain and so on. There is no mature real-time cybersickness detection system for VR developers to evaluate the susceptibility of their products to cybersickness so far, which has hindered the adoption of VR to some extent. The purpose of this study is to implement the real-time monitoring of cybersickness using physiological sensors to measure data and quantify the influence factors of cybersickness through deep learning model. Besides, we have developed a VR experimental platform and passive navigation task to induce user cybersickness. During the experiment, to train the LSTM Attention neural network model, we collected the user's real-time physiological signals, including skin electrical activity (EDA) and electrocardiogram (ECG), as well as the position and bone rotation data of the users' virtual avatar. The model can detect the level of users' cybersickness in real-time during VR experience. And the model has been verified by the fivefold cross-validation that the average accuracy of 96.85% was achieved for classification of cybersickness level, showing great performance compared with other relevant studies. The results show the feasibility of accurate classification of cybersickness using the model we proposed. Also the model can provide reference for VR researchers and developers to improve the user experience.

Keywords Virtual reality · User experience · VR interaction · Cybersickness · Bio-signals · LSTM-attention

1 Introduction

In recent years, with the gradual maturity of virtual reality (VR) technology, VR research and development is booming, and the scale of VR market is expanding. Since many downstream application markets have not been fully opened, it is expected that VR industry will still remain in a period

of rapid growth in the next few years. According to Goldman Sachs, VR/AR market revenue will reach \$80 billion by 2025. At the same time, lower costs and increased availability of VR hardware further promote its popularity. At present, we have made great progress in VR hardware technology. In June 2019, the wireless all-in-one Oculus Quest was released. Since then, HTC, Huawei, Pico and other manufacturers have launched VR All-in-one Headset. Before that, the penetration rate of consumer VR in the game market has been very low, with less than 1% of Steam players owning VR devices until the end of 2019. However, according to the latest data released on 5 May 2020, by the end of April 2020, this percentage had jumped to 1.91%. Within four months, the number of VR players on Steam has more than doubled, which marks the first time that VR to C has entered the positive cycle of "hardware content bicycle evolution" in the game and entertainment market.

✉ Xiaoping Che
xpche@bjtu.edu.cn
Chenxin Qu
21112201@bjtu.edu.cn
Siqi Ma
20122012@bjtu.edu.cn
Shuqin Zhu
18301151@bjtu.edu.cn

¹ School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China

However, while providing users with immersive experience, VR also brings many unpleasant side effects, among which the most obvious one is cybersickness. At present, there are many studies on cybersickness, which prove that cybersickness will continue to exist in the process of users' VR experience, making users feel uncomfortable, causing users to have resistance and aversion to VR, which further affects their willingness to continue using it. It is also showed that cybersickness even caused a large number of users to abandon the use of VR. In addition, researchers focus mostly on finding causes and user responses to cybersickness but only a handful of them are working on how to prevent or predict the occurrence of cybersickness. Most current detection and analysis methods rely on subjective questionnaires to collect a user's posterior status, such as dizziness, nausea, cold sweat, etc (Green 2016; McCauley and Sharkey 1992; Harms et al. 2015). In the meantime, there is no mature real-time cybersickness detection system for VR developers to evaluate the susceptibility of their products to cybersickness. This brings up the following three research questions (RQ) of this paper:

- RQ1: Shortcomings and trends of existing detection methods for cybersickness
- RQ2: How to quantify cybersickness detection in real-time
- RQ3: How to improve the accuracy of cybersickness detection method

Aiming at the above three RQs, this paper summarizes a large amount of research literature and analyzes the existing cybersickness detection methods to solve RQ1. In this way, physiological signals are used to detect the level of motion sickness when the user is immersed in a virtual environment without real-time feedback from the user, and solve RQ2 and RQ3. On this premise, a real-time quantified VR cybersickness detection framework is proposed. Real-time physiological signals and subjective feedback from users are used to train the LSTM-ATTENTION model for cybersickness detection, which is integrated into the VR platform. In this way, without real-time feedback from users, physiological signals are used to detect the cybersickness level when users are immersed in the virtual environment, so as to solve RQ2 and RQ3. The experiment results prove that the model proposed in this paper can detect the severity of users' cybersickness in real-time, and the accuracy is higher than that of other research works. VR developers can adopt different technical measures according to the severity of user cybersickness to reduce the discomfort, thereby improving the user's VR experience.

2 Related work

Aiming at to solve RQ1, this paper analyzes the existing research literature on cybersickness detection to summarize the symptoms, theories and influencing factors of cybersickness. We briefly describe the advantages and disadvantages of different detection methods in the following.

2.1 Definition of cybersickness

Cybersickness, one of the side effects of VR, is an important factor affecting users' VR experience. The term "cybersickness" is used to describe motion-sickness-like experiences in VR, the main symptoms of which are closely similar to transportation motion sickness such as: dizziness, nausea, cold sweating, disorientation and eye strain (Green 2016; McCauley and Sharkey 1992; nalivaiko et al. 2015). Many researchers have tried to figure out cause of it, but currently, there is no exact theory to explain clearly why and how it occur. Scientists have discussions around three main possible theories: the sensory conflict theory, poison/intoxication theory and the postural instability theory.

2.1.1 Sensory conflict theory

The oldest and most accepted theory is Sensory Conflict Theory, proposed by Reason and Brand (1975). This theory is based on the premise that discrepancies between the human senses which provide information about the body's orientation and motion causes a perceptual conflict which the body does not know how to handle. Those sensory conflicts arise when the sensory information is not the stimulus that the user expects based on his/her real world experience. The main disadvantages of this theory are the impossibility to predict the occurrence of cybersickness in every situation to explain why two participants do not have the same symptoms from the same stimulation condition as highlighted in Iavoli (2000).

2.1.2 Poison/intoxication theory

This theory tries to provide the explanations for the occurrence of cybersickness from an evolutionary standpoint. According to Treisman (1977), when the human body experiences abnormal coordination of the visual, vestibular, and other sensory inputs, the nervous system mistakes these circumstances with poison digestion and responds by emptying the stomach. This corporal response of the nervous system can cause symptoms like nausea, vomiting and discomfort.

2.1.3 Postural instability theory

This theory refutes the sensory conflict theory. Riccio and Stoffregen (2010) made an argument that “one of the primary behavioral goals in humans is to maintain postural stability in the environment. In this case, postural stability is defined as the state in which uncontrolled movements of the perception and action systems are minimized”. Postural stability depends from the perception of the surrounding environment and the prediction of action consequences (for example, imagine somebody walking on concrete or on ice). Every time a walker encounters changes in the environment, he/she has to quickly adapt his/her general body behavior to maintain postural stability (for example, you go on concrete and suddenly step on the ice) Iavola (2000). When such a situation occurs, the organism reacts as an emergency situation to prevent from falling, which increases organism stress condition.

In addition to the above theories of motion sickness, many studies have attempted to classify the causes of cybersickness. According to Jin et al. (2018), Davis et al. (2014), Rebenitsch (2015), the factors can be categorized into hardware, software, and individual factors. Based on an extensive literature review and prioritization of factors in the literature (Rebenitsch 2015), the most common factors associated with cybersickness are listed in Table 1. In our work, the factors we take into consideration are motion in the scene (StanneyKay 1998) and user's profile (Kim et al. 2019).

2.2 Subjective and objective evaluation methods of VR user experience

Existing VR user experience evaluation methods include subjective evaluation method, objective evaluation method and a combination of the two methods.

The subjective evaluation method is mainly the user's self-evaluation. The direct experience of various aspects of VR products can be evaluated through oral evaluation or questionnaire survey (Harms et al. 2015). Questionnaire can effectively evaluate users' subjective experience of products of different sizes (Schrepp et al. 2017).

Paper (Chertoff et al. 2010) designed a survey tool Virtual Experience Test (VET), which uses the question scoring mode of 1 (very disagree) to 5 (very agree) to measure the overall virtual environment experience, but the experiment was only conducted in one gaming environment. Paper (Kim et al. 2018) modified the existing cybersickness simulator questionnaires and developed a VR cybersickness questionnaire as a measurement index in VR environment. Paper (Trindade et al. 2018) evaluates usability and user experience in a 3D VR environment model of a beach. It evaluated participants' interactions in the virtual world through their behavior, and evaluated their presence and responses to cybersickness simulators through questionnaire.

The objective evaluation method is mainly to monitor the user's brain wave, heart rate, eye movement and other physiological signals (Healey and Picard 2005). Heartbeat activity can reflect the changes of users' emotion, so heart rate monitoring is generally considered as an effective method to measure users' emotional state (Yao et al. 2014). They developed a new method to test cybersickness in VR game users: electrocardiogram (ECG) signals and brain functional connectivity (FC). The results showed that FC has significant difference for two VR games with different usability, and so does the gamma band, which proved that ECG is a good tool to analyze the cybersickness.

In addition, a large number of researches have focused on the combination of subjective and objective evaluation. Since most of the user experience in immersive virtual environment (IVE) can be measured by questionnaires, paper (Tcha-Tokey et al. 2016) focuses on the questionnaire method. However, researchers still believe that the best way to measure users' experience in IVE is to collect the results of appropriate subjective and objective methods and compare them. As an example, in paper (Yu et al. 2018), participants were asked to answer Igroup Presence Questionnaire (IPQ) after experiencing VR scenarios, so that objective visual performance data can be compared with subjective performance evaluation tools, and the combination of questionnaire surveys and heart rate can evaluate user experience more comprehensively and convincingly.

Table 1 Related factors of cybersickness

Hardware	Software	Individual
Head movement	Motion in a scene	Video game experience
	Scene texture	VR game experience
	Color in a scene	Susceptibility to motion sickness
Field of view	Sitting vs. Standing	Duration of VR exposure
	Independent visual background	Postural instability
Resolution	The degree of control	History of headache or migraine
	Scene content	Age
	Change of color	Gender

Therefore, this paper chooses to use a combination of subjective and objective evaluation to study the relationship among user characteristics, cybersickness and users' experience in VR. At the same time, given that the existing research does not use objective methods to verify the final conclusion, this paper uses ECG data to objectively verify the user experience.

2.3 Qualitative and quantitative research methods of VR cybersickness

Cybersickness estimation attracts significant research interest both in industrial and academic sectors, and has obtained some successful results. Researchers classify the measures for cybersickness into two categories: qualitative and quantitative (Harm 1990).

2.3.1 Qualitative method

Qualitative test scores are used to gather psychological descriptions or reports of signs and symptoms from experimenters and test participants. There are different existing questionnaires including Pensacola Motion Sickness Questionnaire (MSQ), SSQ, Motion sickness assessment questionnaire (MSAQ), Virtual reality sickness questionnaire (VRSQ) (Kim et al. 2018), Pensacola Diagnostic Criteria (PDC) based on widely used criteria of Pensacola Diagnostic Report Scale (PDRS) (Green 2016), 11-point Misery scale (MISC) or Joyfulness scale (JOSC) (Ng et al. 2020), and other methods of sickness index calculation. Gavvani et al. (2018) claim that cybersickness symptoms are not different from motion sickness ones. Thus, SSQ, which excludes sopite symptoms of MSQ, may not correctly include relevant symptoms of cybersickness at all.

Alternatively, another method was proposed by Keshavarz and Hecht (2011) called Fast Motion Sickness Scale (FMS) which consists of requesting the participants to verbally give their evaluation of sickness level on a scale ranging from 0 to 20, with zero representing no discomfort and twenty representing frank sickness. These kinds of questionnaires can prove useful as they are fast and efficient and can be performed online in the virtual environments as well. According to Sevinc and Berkman (2020), Cybersickness questionnaire (CSQ) and VRSQ have better psychometric qualities for assessing cybersickness in Head-mounted display (HMD)-based VR applications and provide a well-rounded approach to measure its symptoms and calculate its subjective aspect.

2.3.2 Quantitative method

In parallel, quantitative assessments through physiological body signals arising from simulator sickness provide researchers with opportunities to have precise direct

comparisons between and within-participants in an experiment. Such measurements will not be clouded by relatively past experiences and will provide us invaluable data about how the body reacts to different stimuli in a virtual environment. Moreover, a lot of research works successfully demonstrate the success of such approach (Wang et al. 2019; Martin et al. 2018; Nam et al. 2018; Hristova et al. 2009), where they were able to measure different physiological signals using various methods such as ECG, EEG, EDA, GSR, electrogastrogram (EGG), electromyography (EMG), photoplethysmogram (PPG), electrooculogram (EOG), to name a few. These sensors can also be used to measure various physical activities as well as different responses based on the virtual stimuli.

Methods that use bio-signals for automatic measurement of stress and objective data collection have achieved practical results (Cho et al. 2017; Bakker et al. 2011). Heart rate and blood pressure measurements are commonly used for analysis (Rebenitsch et al. 2016), but some research groups have managed to expand their horizons by integrating other physiological signals such as brain activity, skin reaction, etc. Kiryu et al. (2007) used ECG, blood pressure and respiration in their study to figure out the trigger factors and accumulation factors in cybersickness. The power frequencies of the physiological signals are used to estimate the sensation intervals and the onset of cybersickness. Dennison et al. (2016) examined ECG, EGG, EOG, PPG, breathing rate and GSR to study the feasibility of using physiological signals to predict cybersickness. In contrast to the eye and body movement data that changes according to personal intention, bio-signal information such as blood pressure and heart rate can provide a more objective and quantitative feedback from users.

Most current methods for assessing cybersickness use the various questionnaires mentioned above, which are collected before and after each task or immersive experience, so users can only report their physical state before and after the experience. A more objective measurement requires statistics on its continuous impact on the user's reaction time and physical discomfort. Compared with traditional questionnaire method, physiological sensors can measure users' physiological state more objectively and accurately during the entire experience. Moreover, answering the questionnaire during the task execution will disturb the user's immersive experience and prolong the total time. Furthermore, questionnaires answered can not accurately reflect the level of discomfort and cybersickness of users during the experience. On the contrary, physiological sensors can not only help researchers more accurately determine the physiological variables related to cybersickness, but also can predict when cybersickness will occur. Based on previous research works (nalivaiko et al. 2015; Kim et al. 2019; Jeong et al. 2019; Islam et al. 2020), in order to solve RQ2, this paper

will use skin electrical activity (EDA), electrocardiogram (ECG) and user's interactive feedback in VR environment to measure the level of user's cybersickness, so as to achieve the purpose of real-time monitoring.

2.4 Machine learning for physiological signals

Recently, machine learning has been adopted in various fields such as emotion recognition and pattern analysis by analyzing complex physiological data (Jeong et al. 2019). Naive Bayes (NB), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) classifiers used for emotion recognition were tested in Hinkle et al. (2019), among which the highest overall leave-one-out accuracy of 80% was achieved using a SVM and five features extracted from bio-signal: meanHR, magPPV, slopeGSR, mECGHR, HRV. In Cho et al. (2017), three physiological signals (PPG, EDA, and skin temperature (SKT)) were measured in a stressful VR environment. The average classification accuracy for stress levels was over 95% using a kernel-based extreme-learning machine (K-ELM) and the integrated feature including HRV, skin conductance (SC) and SKT.

Since cybersickness can affect the physical state of the user, it is suitable to use physiological data for the analysis. Although electroencephalogram (EEG) contains many noise (Kim et al. 2019; Jeong et al. 2019), some researches still try to use neural network to evaluate cyber sickness based on EEG signals. Pane et al. (2018) proposed the identification of cybersickness severity level using the features extracted from electroencephalogram (EEG) signals. Using a rules-based algorithm, i.e., CN2 Rules Induction, the classification yields the best accuracy of 88.9%. It is outperforming other tested classifiers' accuracies such as decision tree (72.2%) and SVM (83.3%). In Li et al. (2019), participants' subjective evaluations (mild, moderate, and severe feelings of motion sickness) were recorded, as well as EEG, center of pressure(COP), and head and waist motion trajectories. The voting classifier they proposed utilized four types of base classifiers: KNN, Logistic Regression (LR), Random Forest (RF) and Multi Layer Perceptron Neural Network (MLPNN). The averaged accuracy of the classifier were 91.1%. Previously cybersickness analysis with physiological data has been presented with machine learning algorithms that produce the accuracy up to 88.9% (Cho et al. 2017). Therefore, it is worth investigating whether the accuracy can be improved when measuring cybersickness by analyzing physiological data with deep learning algorithms, such as Deep Neural Network (DNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), which are the most commonly used deep learning algorithms for physiological data. Kim et al. (2017) have used the CNN deep learning algorithm for cybersickness studies. They applied the CNN algorithm to VR videos, because the abrupt motion within a video is an important factor to indicate cybersickness.

Jin et al. (2018) extracted features from heterogeneous data sources including the VR visual input, the head movement, and the individual characteristics. They trained three different networks including the CNN, the Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN), and the Support Vector Regression (SVR). The results indicated that the best performance of predicting cybersickness was obtained by the LSTM-RNN. However, neither of the works applied the deep learning algorithms to physiological data. Islam et al. (2020) used heart rate variability (HRV) and skin electrical response (GSR). However, they did not get users' feedback on cyber sickness during the experiment, which means that their research results are not convincing enough.

Islam et al. (2020) considered four features (i.e., min, max, runningavg, and percentage of change from resting condition) from each of four physiological signals (HR, BR, HRV, and GSR). They proposed a labeled physiological data that can be used in a Long-short-term-memory (LSTM) regression analysis to predict user cybersickness. The mean absolute error (MAE) of the test data of the regression was 8.7%. In Kim et al. (2019), they developed an EEG driven VR cybersickness level prediction model. In the first stage, the EEG data is transformed into multi-channel spectrogram which accounts for the correlation of spectral and temporal coefficient. Then, a CNN is applied to encode the cognitive representation of the EEG spectrogram. In the second stage, they train a cybersickness prediction model on the VR video sequence by designing a RNN. Here, the encoded cognitive representation is transferred to the model to train the visual and cognitive features for cybersickness prediction. The proposed framework can achieve 90.48% accuracy. Jeong et al. (2019) applied and compared DNN and CNN deep learning algorithms for objective cybersickness measurement from EEG data. The experiments showed that there is no significant difference in terms of the accuracy between CNN with DNN (both around 98%), but DNN structure is better regarding the computational cost while it still has room for improvement in accuracy. Therefore, in order to solve RQ3, we chose LSTM as the basic network component with an added Attention layer. To compare different results and choose a better structure for the whole network, three sets of structures are tested on the physiological data and avatar's motion data separately, which are only using LSTM layer, putting the Attention layer in front of or after the LSTM layer. The three different combinations of LSTM and attention are shown in Fig. 1.

3 Real-time VR cybersickness detection

We proposed a combination of subjective and objective cybersickness detection method using the user's subjective evaluation data, objective physiological signals, and the

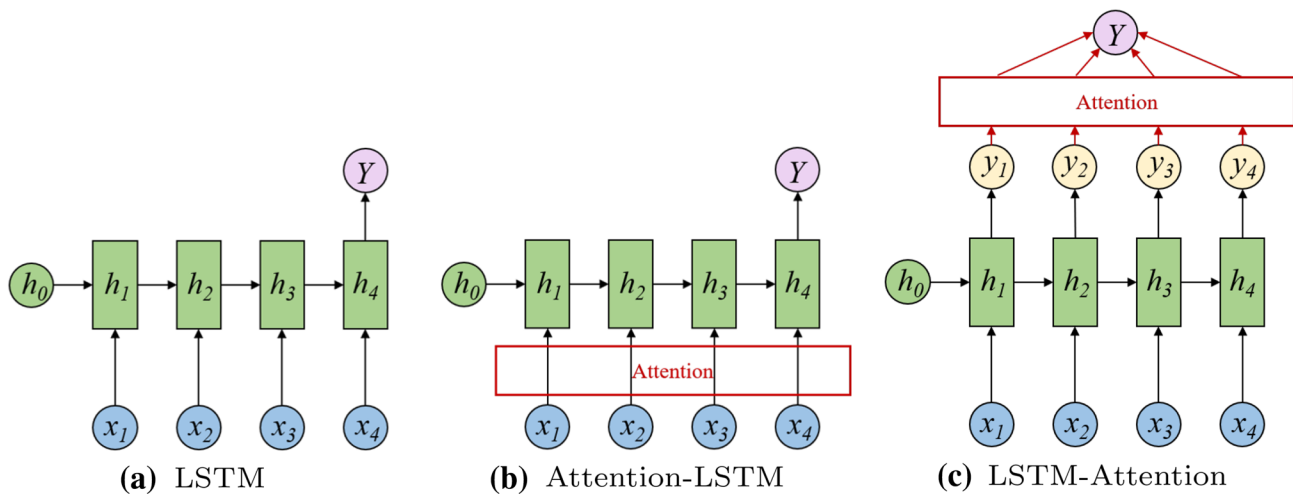


Fig. 1 Three combinations of LSTM and attention

user's virtual avatar's relative data to train the neural network model to detect cybersickness in real-time and solve RQ2. We have developed a VR platform that can provide users with cybersickness experience and integrated it with bio-signal acquisition system (BIOPAC), also we added a feedback and logging modules to the platform to record the real-time feedback from users. Next, during the experiment, physiological data (EDA and ECG), as well as features extracted from ECG (Heart Rate and RR interval), avatar's body motion data and user's feedback are collected. After that, these raw data are preprocessed and fed to the neural network model for training. Finally, the LSTM -Attention model which can detect the level of cybersickness is obtained to solve RQ3. And the model will be integrated into the real-time cybersickness prediction system described below.

The architecture of our proposed real-time cybersickness prediction system is shown in Fig. 2. When we talk about

real-time, it means the system is able to give feedback with a quite low latency. The system consists in two steps:

- The first step: The users put the HMD and biosensors on and get immersed in the virtual environment, which allow us to collect their physiological signals during the immersive experience. After the data acquisition, the LSTM Attention model will be trained to learn the features of the input signals and learn at which cases the users will likely get cybersickness. As our model is built with raw data (without many preprocessing) getting from the system directly, the model can be deployed directly in other immersive systems for the cybersickness prediction.
- The second step: The pre-trained model will be deployed in an immersive system. Biosensors can keep tracking data of the user and send the information to the LSTM-Attention model which will predict online the cybersickness

Fig. 2 Real-time cybersickness detection system

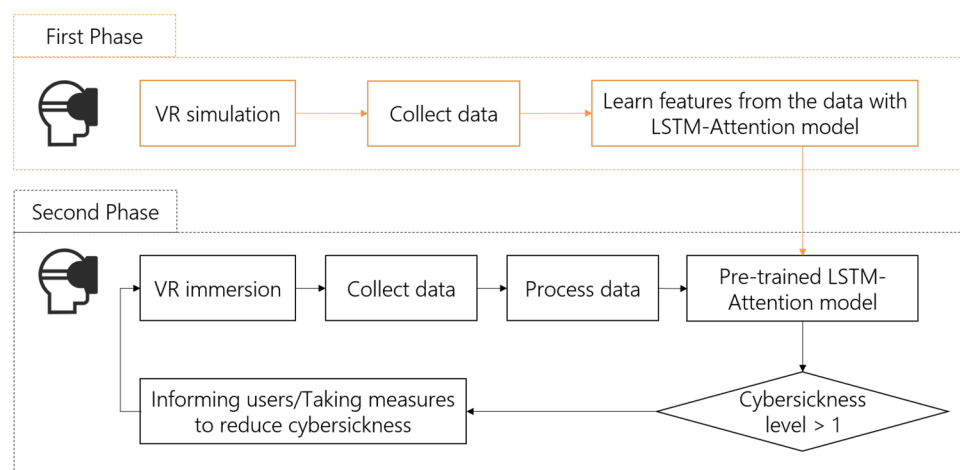
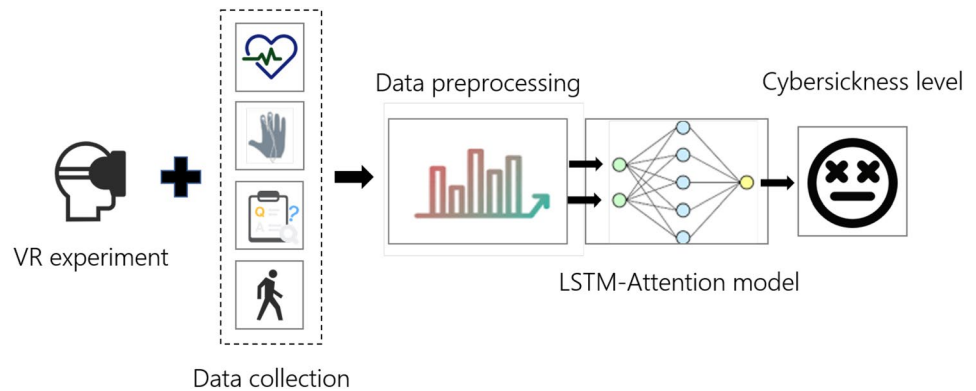


Fig. 3 Framework for the experiment procedure



ness level according to the inputs. The system can then alert the user of cybersickness occurs or take possible measures to reduce it.

3.1 Cybersickness detection platform

In order to solve RQ2 and RQ3, this paper designs a VR simulation experiment platform and intends to perform the experiment on various configurations of the cybersickness factors as shown in Table 1 and different experiment conditions which may have effect on bio-signals (sitting/standing, active/passive motion, moving/staying still). Due to limitation of time, so far we only ran a preliminary experiment on one of the conditions which required the participants stand still in the real world while doing the passive navigation tasks in the virtual environment. Next, we recruited participants for experiment. During the experience, we collected data related to cybersickness generated by participants, and trained the neural network with these data to establish a network model that could detect cybersickness in real-time according to user's physiological signals and body movement data. The framework for our experiment procedure is shown in Fig. 3.

In the main virtual environment of the system, participant can see their virtual avatar and its reflection in the mirror, showed in Fig. 4. This provides the user a visual feedback on what he/she is doing and allows him/her to see and interact with the virtual environment from a first-person viewpoint which may reduces the cognitive load of the participant. In order to animate the rigged avatar, we need to achieve a full body tracking while the user is in the virtual environment. Inverse Kinematics (IK) provided by Unity 3D is used. There are totally six HTC Vive trackers, two for feet (attached to the user's ankle), one for torso (attached on the back of their waist); two controllers for hands and one head-mounted display (HMD) for head tracking. The device setup for the user is illustrated in Fig. 5. The six trackers are used to track the user's outer skeleton in real-time and on top of



Fig. 4 Virtual avatar seen the mirror



Fig. 5 Setup of trackers on a user

that, IK is applied on all the inner joints to get an exact body motion of the user.

We induce cybersickness by adding a moving scene setting based on previous literature reviews (Davis et al. 2014). In the experiment, we require users to perform passive navigation tasks to obtain the physiological data related to cybersickness. Besides, the virtual environment is simplified to limit influence on the bio-signals caused by other kinds of side effects, such as cognitive load and stress. Once the passive navigation task begins, the bio-signal recording will start and will set a flag on the data each time when the user clicks a specific button on the controller to indicate their level of cybersickness. After each passive navigation task, the avatar's position and bone rotation will be output and recorded, as well as user's interactive feedback on cybersickness level.

3.2 User data and questionnaire design

The study was conducted on 9 healthy volunteers (average age of 28 from 22 to 35 years old) of both genders (2 females and 7 males). Before the experiment, participants filled in an anonymous demographic questionnaire which includes several simple questions. Studies have shown that the degree of cybersickness varies according to individual differences (e.g. previous experience, susceptibility, gender, age, etc.) (Kim et al. 2019) Therefore, we consider participant's own characteristics. In the personal characteristic questionnaire, we collected the participants' personal characteristic and specific background knowledge related to the experiment, including age, gender, dominant hand (left or right), 3D video game experience, previous VR experience and VR background knowledge. Among them, 3D video game experience is divided into four levels: daily, twice a week, once or twice a month, and almost never. VR background is also divided into four levels: none, understanding, intermediate and expert. The values of all items in the questionnaire and their numerical values are shown in Table 2.

During the experiment, participants were equipped with HTC Vive VR devices and BIOPAC biosensors to perform three passive navigation tasks at different speeds. The experiment consisted of three phases: Welcome phase, which helped participants prepare for the experiment; Calibration phase, to achieve comprehensive tracking of participants' body movements and help participants to get familiar with the virtual platform environment; Evaluation phase, which includes passive navigation tasks (each task lasts 2–3 minutes) and experimental data collecting.

Table 2 Background questionnaire item value

Questionnaire	Item value	Numerical value
Age	22–35	
Gender	Male	1
	Female	0
Left or right handed	Left	1
	Right	0
3D video game experience	Almost never	0
	More than once a month	1
	One or two times a week	2
	Everyday	3
Previous VR experience	Yes	1
	No	0
VR knowledge	None	0
	Basic	1
	Intermediate	2
	Expert	3

3.2.1 Welcome phase

Considering the Covid-19 pandemic, in order to avoid possible risks of infection, all the participants were asked to clean their hands with alcohol-based sanitizer before beginning the experiment and wear face masks during the whole experiment to avoid direct skin contact with the HMD. Then they were required to read the instructions, sign the consent forms and fill in background questionnaires asking for some general information

3.2.2 Calibration phase

Once the trackers were set up by the experimenter, the participant could begin the calibration phase in Fig. 6. In the initial scene as shown in Fig. 7, there is a natural background with trees, pile of stones and grasses. The nature scene could help the participant to relax before the next steps (Gerber et al. 2017) The participant could see a virtual avatar and its mirrored image in the virtual environment. The participants had to manually walk towards the virtual avatar and align themselves to base position of the avatar in Fig. 8. Once the participants felt that the avatar's height corresponds to theirs, they could click the trigger button while standing in a 'T' pose. This was required because when the arms and legs were extended, the trackers could provide the approximate limb lengths of the participant. After finishing the calibration, the participant would be able to move freely in the virtual space, and could perform different body movements and observe the same movements via a properly configured virtual avatar shown in the mirror, shown in Fig. 9.



Fig. 6 Calibration



Fig. 7 VR scene



Fig. 8 User operation

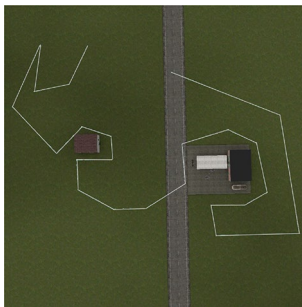


Fig. 9 Moving route

3.2.3 Evaluation phase

Before the start of the evaluation phase, the participants were required to take off the HMD and take a short break until they were ready to begin, we then put the biosensors on their body to record the baseline of the participants' bio-signal. After that, participants can start the experimental evaluation phase independently. They are prohibited from speaking during the experiment, to avoid affecting physiological signals. However, participants can ask any questions before and after each task, and they can also ask to stop the experiment when necessary.

Afterwards, the participants would put on the HMD again and the evaluation phases started. The evaluation virtual scene is the same as in the calibration phase but several passive navigation tasks with different speed conditions were included. First the participants would stand still, point at "Move" button, and click the trigger button on the controller to start the movement in the virtual scene. During the passive navigation tasks, the participants could click on the upper part of the controller's touchpad if they felt sick (if they felt worse, they could click the button several times), and click on the lower part of the touchpad if they felt better (button shown in Figure 10). However, if the participants felt that they could not continue, they could ask to stop the experiment immediately. For each task, after the participant clicked on the "Move" button, the bio-signal recording would start.

At the end of the evaluation phase, the participants were asked to fill out a subjective questionnaire. According to Sevinc and Berkman (2020), CSQ and VRSQ have better psychometric qualities for assessing HMD-based VR applications and provide a well-rounded approach to measure cybersickness symptoms and calculate the subjective aspect of cybersickness. CSQ uses a scoring method based on item weights while VRSQ employs a simpler scoring method. Besides, VRSQ proposed by Kim et al. (2018) extracts the types of motion sickness items from SSQ related to the VR headset which means it is more aligned with our experimental design. Hence, we decided to use VRSQ to collect feedback from the participants after the experiment and designed Table 3.

The VRSQ items are shown in Table 4. The most prominent feature of VRSQ is that it is comprised of two components, namely the oculomotor and disorientation components. Oculomotor component includes General discomfort, Fatigue, Eyestrain, Difficulty focusing. Disorientation

Fig. 10 Controller button configuration used in the experiment

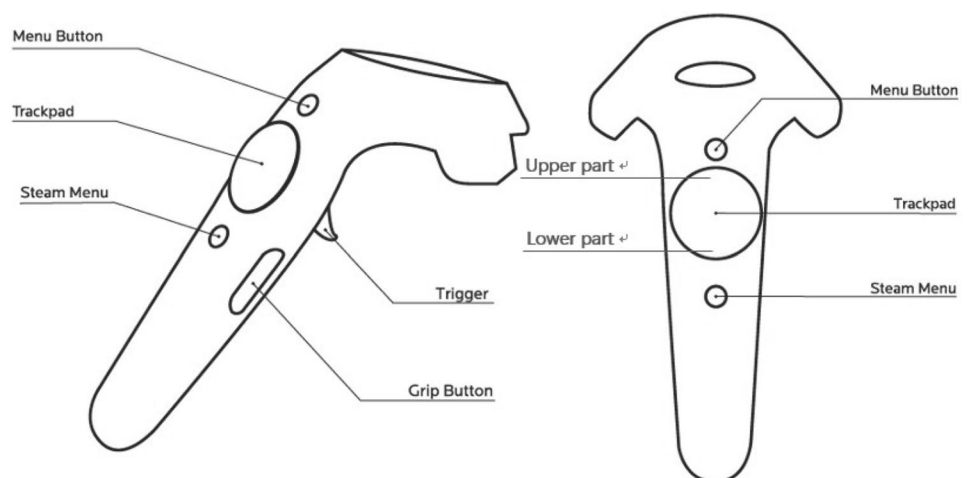


Table 3 VR sickness questionnaire(VRSQ)

VRSQ symptom	Oculomotor	Disorientation
General discomfort	*	
Fatigue	*	
Eyestrain	*	
Difficulty focusing	*	
Headache		*
Fullness of head		*
Blurred vision		*
Dizzy(eyes closed)		*
Vertigo		*
Total	[1]	[2]

*Represents for the VRSQ symptom will occur in oculomotor or disorientation

Table 4 Computation score of VRSQ

VRSQ components	Computation
Oculomotor	$([1]/12)*100$
Disorientation	$([2]/15)*100$
Total	$(\text{Oculomotor} + \text{Disorientation score})/2$

component includes Headache, Fullness of head, Blurred vision, Dizzy (eyes closed) and Vertigo. Also we show the formula used to calculate the score of eye movement and

orientation disorder and the final total score of VRSQ in Table 3. The final total score of the questionnaire is the arithmetic mean of the scores of the two parts.

The data in the cybersickness level log file includes timestamp and the sickness level value from 0 to 2: 0 means no sickness at all, 1 means slight sickness and 2 means severe sickness. In the file, the first line is the timestamp for the start of each navigation task. The cybersickness level values were used to label all the data, including the avatar's position and rotation as well as the physiological data. Therefore we needed to match its sequence length according to the data we used for different network layers and the timestep we set for each layer. It means we needed to expand or reduce the length of the sequence without changing its statistical properties. For example, for the avatar's position and rotation data (assuming the sequence length of the data is L), we use LSTM as the first layer and set timestep to 257. The length of the cybersickness level sequence should be round $L/257$.

3.3 Sensor based user data collection

As is shown in Fig. 11, physiological data including EDA and ECG raw data, RR interval and Heart rate was recorded by the BIOPAC system.

The capture of signal was not fully started from the beginning. Hence, we had to remove those lines as well as abnormal values in all types of signal. The unit of EDA is microsiemens (μS), and normally the value ranges from 5 to 50. Normally the heart rate range is 60–100 bpm (corresponding

**Fig. 11** Data acquisition interface of Acqknowledge software

R–R interval is 1–0.6s). To normalize the data, we removed the lines with heart rate over 150 or equal to 0.

4 LSTM-attention based VR cybersickness detection

In order to solve RQ3, this paper will use neural network model to detect the severity of user cybersickness from the multi-source data combining subjective and objective. In most current studies, subjective questionnaires (such as CSQ and VRSQ) are used to analyze the severity of cybersickness experienced by users after their VR experience, which means that feedback and treatment cannot be provided at the first time when users feel uncomfortable. On the other hand, due to the limitation of labeled data, there are few researches to automate the process of cybersickness detection through deep learning methods (Islam et al. 2020). Therefore, in this paper, we tried to establish a neural network model based on recurrent neural network LSTM, and added ATTENTION layer to optimize feature extraction effect. This section mainly introduces the processing process of input data and the design and optimization of LSTM-ATTENTION model structure.

4.1 Data pre-processing

After establishing the data set related to user cybersickness, we need to further preprocess the data before inputting the data into the neural network model we designed. According to Jeong et al. (2018), the learning accuracy is higher when the Z-score standardization is applied to the feature data than the min-max normalization. Also after standardization/normalization, the speed of gradient descent to find the optimal solution is accelerated, and the accuracy may be improved. Hence, to prepare the input data for my network, Z-score is applied to the feature data (avatar's position and rotation & physiological data). The StandardScaler class in the sklearn package is used to do the standardization. The advantage of using this class is that the standardization is done for each feature dimension, but not for all samples, and the parameters, e.g. mean, variance, in the training set can be saved and directly used to convert the test set data. The Z-score standardization is computed as:

$$x = (\text{data} - \text{mean}) / \text{standard deviation} \quad (4-1)$$

We used the cybersickness level (CL) value to label the physiological data and avatar's motion data (referred to as data below). We hypothesised that, if a participant rated $CL_t = n$, at time t on a scale from [0–2], then for all the data within $[t-1, t]$, the corresponding sickness level will be n . The data was then formulated with their corresponding CL_t rating. An example data point at time t is shown below:

$$[EDA, ECG, RR, HR] \rightarrow CL_t$$

$$[pos_x, pos_y, pos_z, rot_x, rot_y, rot_z] \rightarrow CL_t$$

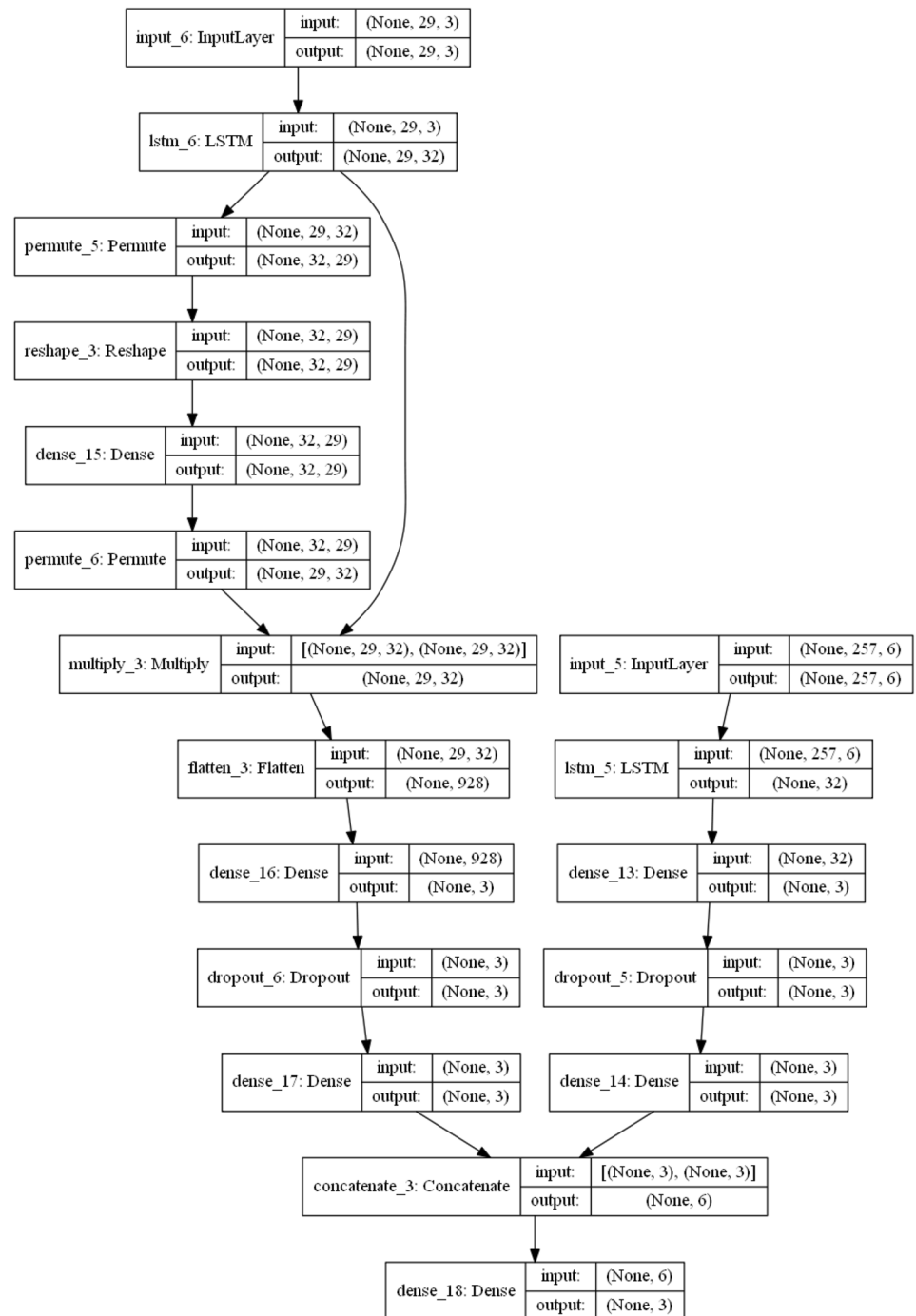
Since the label data (cybersickness level value) are classified values of 0, 1 and 2, one-hot encoding is applied. One-hot encoding, also known as one bit effective coding, whose principle is to use N-bit status registers to encode n different status values. Each state has its own independent register bits. At any time, only one of the register bits is effective. One-hot encoding can convert classification variables into data formats that are easy to use by deep learning algorithms.

The main reason for using it is that the output layer of my multiclassification network uses softmax function, which gives a probability distribution output. Therefore, the input label is also required to appear in the form of a probability distribution. One-hot encoding can be used to convert discrete labels into binary vectors. For example, the cybersickness levels are 0 (none), 1 (slight), 2 (severe), then after converting to one-hot format, the labels are 100, 010, 001.

4.2 LSTM-attention based model

Our LSTM-Attention network model is composed of two parts, the first dealing with two different inputs (physiological data and avatar's motion data) using two models, and the second combining the two models' outputs to a single output to give the classification result. We used the Keras functional API to create the non-linear topology multiple input model. One of the model structures we tested is shown in Fig. 12 in detail.

In the first part of the whole model, the first input (avatar's motion data) size is $6 \times m$, where 6 is the number of the dimension of avatar's motion data, and m is the number of time steps. The second input (physiological data) size is $4 \times n$, where 4 is the number of the feature extracted from the physiological data, and n is the number of time steps. There are 32 neurons in the LSTM unit. An Attention layer consisted of several Keras layers is added in front of or after the LSTM layer. The output is two Dense layers both with 3 neurons. For all the dense layers except the output one, relu activation function is used, and for the output dense layer, softmax is used. After the first Dense layer, a dropout layer with rate of 50% is added to overcome overfitting during training. It works by ignoring half of the feature detectors (leaving half of the hidden layer node value as 0) in each training batch. This method can reduce the interaction between feature detectors (hidden nodes) and make the model more general. Since we performed a multi-classification task, in

Fig. 12 One of the proposed LSTM-attention structures

the second part of the whole model, the softmax activation function is used for the final output Dense layer to discriminate features onto the cybersickness level. The algorithm of model is shown in Algorithm 1.

Algorithm 1 LSTM-Attention-based motion sickness detection algorithm

Require: D_1 (virtual avatar location bone & rotation), D_2 (physiological data)

Ensure: y

(Detection results)

- 1: $D \leftarrow D_1, D_2$;
(Combine the two initial datasets into vector D by RNN)
- 2: **for** $i = 1$ **to** k **do**
- 3: $h_1, h_2 \dots h_t \leftarrow LSTM(D)$;
(Enter machine readable vector D into LSTM and obtain the output of the corresponding hidden layer.)
- 4: Calculate the attention probability distribution values $\alpha_1, \alpha_2 \dots \alpha_t$;
- 5: Attention normalization;
- 6: $v \leftarrow f(\alpha, h)$;
(h are multiplied by the weight and added.)
- 7: $d \leftarrow Dropout(v, level = 0.5)$;
- 8: **end for**
- 9: $(trainSet, testSet) \leftarrow Split(dataSet)$;
(The dataset is splitted into two as 80% training set and 20% test set.)
- 10: $(trainSet, validationSet) \leftarrow Split(trainSet)$;
(20% of the train set is reserved for the validation set.)
- 11: Train (trainSet);
- 12: $(accuracyRate, loss) \leftarrow Test(testSet)$;
- 13: $y \leftarrow Softmax(d)$;

We trained the LSTM-Attention using the Adam (a stochastic optimization method) optimizer with a batch size of 100 examples and 50 epochs before testing the model. According to former experience, softmax function is widely used with the categorical cross entropy loss to test multi-classification model. However, to determine which estimator function was better for my model, categorical cross entropy and binary cross entropy loss were both tested. To ensure the performance validity of my model, the overall experimental evaluations were performed with the validation set using 5-fold cross-validation method. The model was trained based on the data of 9 participants. We mixed and split the entire dataset into two parts, 80% of the data used as training set and the remaining data as testing set, then split the training set randomly into 5 folds, and fit the model using 4 folds while validated the model using the 1 remaining fold. We noted down the accuracy and loss, repeated this process until every fold served as the validation set. Then we took the average of the recorded accuracy as the accuracy of the model on the training set. Finally we used testing set to test the performance of the model.

5 Experiment results

We will present and discuss the performance of our network model as well as the questionnaires analysis results in this section.

5.1 Model performance

Table 5 shows the results using physiological data only to make the prediction of cybersickness level. We applied different combinations of LSTM and Attention on the model for comparison. From the results, we can see that

Table 5 Accuracy of different models on physiological data

	LSTM	LSTM-Attention	Attention-LSTM
Train Accuracy	80.20%	83.6%	85.07%
Test Accuracy	90.59%	92.48%	91.48%

The best result of test accuracy 92.48%, which is shown in bold

Table 6 Accuracy of different models on avatar's motion data

	LSTM	LSTM-Attention	Attention-LSTM
Train Accuracy	93.15%	94.54%	94.70%
Test Accuracy	93.55%	92.52%	90.25%

The best result of test accuracy 93.55%, which is shown in bold

Table 7 Accuracy of different combined models on whole dataset

Motion data & Physiological data	Train accuracy	Test accuracy
LSTM&LSTM	91.88%	94.85%
LSTM&LSTM-Attention	90.04%	96.58%
LSTM&Attention-LSTM	89.63%	94.91%
LSTM-Attention&LSTM	92.54%	90.68%
LSTM-Attention&LSTM-Attention	92.08%	95.26%
LSTM-Attention&Attention-LSTM	78.07%	92.62%
LSTM-Attention&LSTM-Attention	90.40%	94.78%
Attention-LSTM&LSTM-Attention	92.40%	94.66%
Attention-LSTM&Attention-LSTM	92.13%	91.97%

The best result of test accuracy 96.58%, which is shown in bold

Table 8 Pearson correlation coefficient between VRSQ score and real-time feedback

	VRSQ score	Oculomotor	Disorientation
Cybersickness Level	0.80	0.84	0.71

for physiological data, using LSTM-Attention obtained the best result of test accuracy 92.48%, which is shown in bold.

Table 6 shows the results using avatar's motion data only to make the prediction of cybersickness level. Different combinations of LSTM and Attention were also applied on the model for comparison. From the results, we acknowledge that for motion data, only using LSTM obtained the best result of test accuracy 93.55%, which is shown in bold.

Table 7 shows the final results of the model with different combinations in the first part. First, we compared the results in first three rows, where LSTM was applied on the motion data, and three different combinations of LSTM and Attention were applied on the physiological data. We find that for physiological data, using LSTM Attention obtained the best result of test accuracy 96.58%, which is shown in bold. This result consisted with the result we got earlier with single physiological data input, which

supports that for physiological data, using LSTM-Attention can provide a good accuracy.

From Table 7 we can also find that using Attention before or after LSTM gave different results for both data. And among all the results, the combination of LSTM for motion data and LSTM-Attention for physiological data provided the highest accuracy of 96.58%.

5.1.1 Questionnaire results

Pearson correlation coefficient, also known as Pearson product-moment correlation coefficient (PPMCC), is used to measure the linear correlation between two variables X and Y, and its value is between -1 and 1. The correlation coefficient can be calculated by the following formula:

$$R_{xy} = \frac{\sum_{k=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{k=1}^n (x_i - \bar{x})^2 \sum_{k=1}^n (y_i - \bar{y})^2}} \quad (4-2)$$

Table 8 shows the Pearson correlation coefficients between the VRSQ score and cybersickness level (CL), decomposing the VRSQ score into three categories (total, oculomotor and disorientation). We found the VRSQ total score was significantly correlated with the average CL reported by each participant during the experiment (the Pearson correlation coefficient is 0.8). The average CL was also found correlated with oculomotor component and disorientation component total sub-score. The correlation between the average CL and oculomotor component sub-score is 0.84, and the correlation between the average CL and disorientation component sub-score is 0.71. The results indicate that oculomotor symptom is more related with cybersickness.

In addition, since all the correlation values range from 0.7 to 0.9, it implies that the use of CL to label the physiological data and avatar's motion data in our experiment is efficient. This way of obtaining labeled data not only provides the feasibility for the later establishment of a large-scale user cybersickness related database, but also provides the possibility for training a more accurate and reliable neural network model.

Table 9 Comparison of Our Proposed LSTM-Attention Network With Previous Approached

References	Task	Signals	Method	Accuracy
Pane et al. (2018)	Cybersickness	EEG	CN2	88.90%
Hinkle et al. (2019)	Emotion	ECG&GSR	SVM	80.00%
Kim et al. (2019)	Cybersickness	EEG	CNN+LSTM	90.48%
Sarkar et al. (2019)	Cybersickness	ECG	DMNN	89.4%
Jeong et al. (2019)	Cybersickness	EEG	DNN	94.02%
Ours	Cybersickness	ECG&EDA	LSTM-Attention	96.58%

The accuracy which achieves 96.58% are shown in bold italic value

5.2 Results analysis

Although there are some existing works on classification of cybersickness based on EEG data, we decided not to take into consideration EEG for our experiment due to the high price and difficulty in integration with our VR system. In the study, we used ECG and EDA, which can be measured using simpler and less intrusive wearable devices. Table 9 illustrates that our proposed model performs robustly when being compared to several other studies in the field, even with studies using EEG, which achieves 96.58% accuracy shown in bold italic value.

As we stated earlier, the information extracted by the Attention layer is closely related to its position in the network. The Attention layer directly followed the input layer allows us to gain an understanding of the importance of the input's feature space. The Attention layer put after LSTM can make the final decision of the model more focused on the effective features, and assign the main decision weights to the feature dimensions that really help on the final classification. The feature dimensions as an input to the Attention layer have already been abstracted by LSTM, accordingly the interpretability of the features is relatively poor.

Secondly, when comparing different results of the models using physiological data in Table 5 we find that applying Attention after LSTM give the best performance. For the physiological data, the Attention after LSTM can pay more attention to the potentially useful features extracted from raw bio-signal data at the LSTM layer, which maybe was the reason of the good result. However, regarding results using motion data in Table 6, Attention was not that useful because the intention of Attention is using weight parameters to filter out irrelevant features, while there is no difference in importance of the coordinates.

Meanwhile, we can find in Table 7 that the test accuracy is higher than train accuracy on almost all the models. We conjecture that because of the use of the dropout layer, it turns the neural network to a combination of a large set of weak classifiers. During training, dropout will randomly deactivate sets of classifiers, which may affect the training accuracy. However, when testing the model, dropout will be automatically ignored and all the weak classifiers will be allowed to work in the testing process, hence the test accuracy will be higher. Normally, dropout can guarantee a better test accuracy, sometimes even better than training accuracy.

Finally, categorical cross entropy and binary cross entropy (hereinafter referred to as CE and BCE) were both tested to find a better estimator function. Usually CE is used for multi-classification model, and BCE is suitable for two-class classification. However, BCE can also be used for multi-class single label classification problems. After our test, there is not much difference between the two and the performance of BCE is slightly better than CE.

6 Conclusion

The main contribution of this paper is to propose an approach combining subjective and objective measurements to estimate cybersickness using both subjective evaluation data, objective physiological signal and neural network model. Firstly, we developed a VR platform that can provide users with motion sickness experience. Then we integrated the platform with bio-signal acquisition system (BIOPAC) and added a feedback and logging modules to the platform to record the real-time feedback from the users. Next, during the experiment, physiological data (EDA and ECG), as well as features extracted from ECG (Heart Rate and RR interval), avatar's body motion data and users' feedback are collected. After that, these raw data are preprocessed and fed to the neural network model for training. Finally, the LSTM attention model which can detect the level of cybersickness is obtained, and can be integrated into the real-time cybersickness prediction system described below.

A fivefold cross-validation scheme was used to evaluate the performance validity of our model. Average accuracy of 96.58% was achieved for classification of level of cybersickness, showing great performance when being compared to other related studies (see Table 9). The results show the feasibility of accurate classification of cybersickness using our cybersickness prediction system. Although the training and testing of the network model are offline in this work, the model can be integrated into the VR platform for online real-time detection of motion sickness in the future.

Funding This work was supported by Fundamental Research Funds for the Central Universities (grant number- 2021JBM028).

References

- Bakker, J., Pechenizkiy, M., Sidorova, N.: What's your current stress level? Detection of stress patterns from gsr sensor data. *ICDM Workshops*, 573–580 (2011)
- Chertoff, B.D., Goldiez, B., LaViola, J.J.: Virtual experience test: A virtual environment evaluation questionnaire. *VR*, 103–110 (2010)
- Cho, D., Ham, J., Oh, J., Park, J., Kim, S., Lee, N.-K., Lee, B.: Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine. *SENSORS* **17**(10), 2435 (2017)
- Davis, S., Nesbitt, K., Nalivaiko, E.: A systematic review of cybersickness. *I3E*, 8–189 (2014)
- Dennison, S.M., Wisti, Z.A., D'Zmura, M.: Use of physiological signals to predict cybersickness. *Displays* **44**, 42–52 (2016)
- Gavani, M.A., Walker, R.F., Hodgson, M.D., Nalivaiko, E.: A comparative study of cybersickness during exposure to virtual reality and "classic" motion sickness: are they different? *J. Appl. Physiol.* **125**, 1670–1680 (2018)
- Gerber, S.M., Jeitiner, M.-M., Wyss, P., Chesham, A., Urwyler, P., Müri, R.M., Jakob, S.M., Nef, T.: Visuo-acoustic stimulation that

- helps you to relax: a virtual reality setup for patients in the intensive care unit. *Sci. Rep.* **7**(1), 1–10 (2017)
- Green, P.: Motion sickness and concerns for self-driving vehicles: a literature review. *Tech. Rep. UMTRI-2016*, Ann Arbor, MI Univ. Michigan Transp. Res. Institute. (2016)
- Harm, D.L.: Physiology of motion sickness symptoms. In Crampton, G.H. (ed.) *Motion and space sickness*. CRC Press, Boca Raton (1990)
- Harms, J., Seitz, D., Wimmer, C., Kappel, K., Grechenig, T.: Low-cost gamification of online surveys: improving the user experience through achievement badges. *CHI PLAY* 109–113 (2015)
- Healey, J., Picard, R.: Detecting stress during real-world driving tasks. *IEEE Transact. Intell. Transport. Syst.* **6**, 156 (2005)
- Hinkle, L., Khoshhal, K., Metsis, V.: Physiological measurement for emotion recognition in virtual reality. In: 2019 2nd International Conference on Data Intelligence and Security (ICDIS), pp. 136–143 (2019). IEEE
- Hristova, E., Grinberg, M., Lalev, E.: Biosignal based emotion analysis of human-agent interactions. *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions*, 63–75 (2009)
- Islam, R., Lee, Y., Jaloli, M., Muhammad, I., Zhu, D., Quarles, J.: Automatic detection of cybersickness from physiological signal in a virtual roller coaster simulation. *VR Workshops*, 649–650 (2020)
- Jeong, D.K., Yoo, S., Jang, Y.: Vr sickness measurement with eeg using dnn algorithm. In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–2 (2018)
- Jeong, K.D., Yoo, S., Jang, Y.: Cybersickness analysis with eeg using deep learning algorithms. *VR*, 827–835 (2019)
- Jin, W., Fan, J., Gromala, D., Pasquier, P.: Automatic prediction of cybersickness for virtual reality games. *GEM*, 1–9 (2018)
- Keshavarz, B., Hecht, H.: Validating an efficient method to quantify motion sickness. *Human Factors* **53**, 415–426 (2011)
- Kim, H.G., Baddar, W.J., Lim, H.-t., Jeong, H., Ro, Y.M.: Measurement of exceptional motion in vr video contents for vr sickness assessment using deep convolutional autoencoder. In: *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–7 (2017)
- Kim, K.H., Park, J., Choi, Y., Choe, M.: Virtual reality sickness questionnaire (vrsq): motion sickness measurement index in a virtual reality environment. *Appl. Ergon.* **69**, 66–73 (2018)
- Kim, J., Kim, W., Oh, H., Lee, S., Lee, S.: A deep cybersickness predictor based on brain signal analysis for virtual reality contents. *ICCV*, 10579–10588 (2019)
- Kiryu, T., Uchiyama, E., Jimbo, M., Iijima, A.: Time-varying factors model with different time-scales for studying cybersickness. *HCI* **14**, 262–269 (2007)
- Laviola, J.J.: A discussion of cybersickness in virtual environments. *ACM SIGCHI Bull.* **32**, 47–56 (2000)
- Li, Y., Liu, A., Ding, L.: Machine learning assessment of visually induced motion sickness levels based on multiple biosignals. *Biomed Signal Process Control* **49**, 202–211 (2019)
- Martin, N., Mathieu, N., Pallamin, N., Ragot, M., Diverrez, J.-M.: Automatic recognition of virtual reality sickness based on physiological signals. In: *IBC* (2018)
- McCauley, E.M., Sharkey, J.T.: Cybersickness: perception of self-motion in virtual environments. *Teleoperators Virtual Environ.* **1**, 311–318 (1992)
- Nalivaiko, E., Davis, S., Blackmore, K., Vakulin, A., Nesbitt, K.: Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time. *Physiol. Behav.* **151**, 583–590 (2015)
- Nam, S.-H., Lee, Y.J., Kim, Y.J.: Biological-signal-based user-interface system for virtual-reality applications for healthcare. *J. Sens* 9054758–1905475810 (2018)
- Ng, K.T.A., Chan, K.Y.L., Lau, Y.K.H.: A study of cybersickness and sensory conflict theory using a motion-coupled virtual reality system. *Displays* **61**, 101922 (2020)
- Pane, E.S., Khoirunnisa, A.Z., Wibawa, A.D., Purnomo, M.H.: Identifying severity level of cybersickness from eeg signals using cn2 rule induction algorithm. In: 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), vol. 3, pp. 170–176 (2018). IEEE
- Reason, J.T., Brand, J.J.: *Motion sickness*. Academic press, Cambridge (1975)
- Rebenitsch, L.R.: *Cybersickness prioritization and modeling*. Michigan State University, Michigan (2015)
- Rebenitsch, L., Owen, B.C.: Review on cybersickness in applications and visual displays. *Virtual Reality* **20**, 101–125 (2016)
- Riccio, E.G., Stoffregen, A.T.: An ecological theory of motion sickness and postural instability. *Ecol. Psychol.* **3**, 195–240 (2010)
- Sarkar, P., Ross, K., Ruberto, A.J., Rodenburg, D., Hungler, P., Etemad, A.: Classification of cognitive load and expertise for adaptive simulation using deep multitask learning. In: 8th International Conference on Affective Computing and Intelligent Interaction, *ACII 2019*, Cambridge, United Kingdom, September 3–6, 2019, pp. 1–7. IEEE, ??? (2019)
- Schrepp, M., Cota, P.M., Gonçalves, R., Hinderks, A., Thomaschewski, J.: Adaption of user experience questionnaires for different user groups. *Universal Access in the Information Society*, 629–640 (2017)
- Sevinc, V., Berkman, I.M.: Psychometric evaluation of simulator sickness questionnaire and its variants as a measure of cybersickness in consumer virtual environments. *Applied Ergonomics* **82**, 102958 (2020)
- StanneyKay, M.: *HashPhillip: Locus of user-initiated control in virtual environments*. *Teleoperators & Virtual Environments*, Presence (1998)
- Tcha-Tokey, K., Loup-Escande, E., Christmann, O., Richir, S.: A questionnaire to measure the user experience in immersive virtual environments. *VRIC*, 19–1195 (2016)
- Treisman, M.: Motion sickness: an evolutionary hypothesis. *Science* **197**, 493–495 (1977)
- Trindade, Y., Rebelo, F., Noriega, P.: Tourism and virtual reality: user experience evaluation of a virtual environment prototype. In: Marcus, A., Wang, W. (eds.) *Design, user experience, and usability: users, contexts and case studies*, pp. 730–742. Springer, Cham (2018)
- Wang, Y., Chardonnet, J.-R., Mérienne, F.: Vr sickness prediction for navigation in immersive virtual environments using a deep long short term memory model. *VR*, 1874–1881 (2019)
- Yao, L., Liu, Y., Li, W., Zhou, L., Ge, Y., Chai, J., Sun, X.: Using physiological measures to evaluate user experience of mobile applications. *Lecture Notes in Computer Science*, 301–310 (2014)
- Yu, Q., Che, X., Ma, S., Pan, S., Yang, Y., Xing, W., Wang, X.: A hybrid user experience evaluation method for mobile games. *IEEE Access* **6**, 49067–49079 (2018)



Chenxin Qu was born in Xinzhou, Shanxi Province, China in 1998. She received her bachelor's degree at the School of Software, Beijing Jiaotong University, China in 2020. Now she is a Ph.D. student at the School of Software, Beijing Jiaotong University. Her research interests include Virtual Reality and Human-Computer Interaction.



Siqi Ma was born in Xingtai, HeBei Province, China in 1995. She received the bachelor's degree and master degree at School of Software, Beijing Jiaotong University, China in 2017 and 2020. From 2019 to 2020, she spent one year in France and obtained her second master degree at Telecom Sudparis on Computer Science for Communication Networks. Her research interests include data analysis, machine learning, testing and user experience.



Xiaoping Che was born in 1986. He received the BS degree in Network Engineering from Beijing University of Posts and Telecommunications, China, in 2009, and the MS degree in Computer and Communication Networks from Telecom SudParis, France, in 2011. From 2011 to 2014, he worked in French CNRS laboratory SAM-OVAR, and he got the Ph.D. degree from Institute Mines-Telecom/Telecom SudParis in 2014. His main research interests include VR, AR, User Experience and Crowd Sensing. Now, he serves as the vice dean of School of Engineering, Beijing Jiaotong University.



Shuqin Zhu was born in Beijing, China in 2000. Now he is a bachelor student at School of Software, Beijing Jiaotong University. His research interests include data analysis, machine learning and user experience.