Preprint Version. Manuscript submitted to and accepted by 2018 International Conference on Cyberworlds.

EEG-based Cadets Training and Performance Assessment System in Maritime Virtual Simulator

Yisi Liu, Zirui Lan, Olga Sourina
Fraunhofer Singapore
Singapore
{liuys, lanz0001, eosourina}@ntu.edu.sg

Hui Ping Liew, Gopala Krishnan
Maritime Institute @ Singapore Polytechnic
Singapore
{liew_hui_ping, gopala_krishnan}@sp.edu.sg

Dimitrios Konovessis
Singapore Institute of Technology
Singapore
Dimitrios.Konovessis@SingaporeTech.edu.sg

Hock Eng Ang
School of Mechanical and Aerospace Engineering,
Nanyang Technological University
Singapore
mheang@ntu.edu.sg

***Abstract*—Deep investment in the maritime industries has led to many cutting edge technological advances in shipping navigation and operational safety to ensure safe and efficient logistical transportations. However, even with the best technology equipped onboard, maritime accidents are still occurring with at least three quarters of them attributed to human errors. Due to the rising need to address the human factors in shipping operations, various human factors studies are conducted in maritime domain. In this paper, an Electroencephalogram (EEG)-based cadets training and performance assessment system is proposed and implemented that could be used in the maritime virtual simulator. The system includes an EEG processing and analyses part and an evaluation part. It could recognize the brain states such as mental workload, emotions, and stress from raw EEG signal recorded during the exercises in the simulator and then give an indicative recommendation on "pass", "retrain", or "fail" of the cadet based on the EEG recognition results and input of the level of the task difficulty performed.**

Keywords-EEG; human factors; maritime simulator; assessment; maritime training

I. INTRODUCTION

With at least 75% of maritime accidents attributed to human errors, many researches and studies have been done based on statistical analysis using previously published maritime accidents reports to ascertain the deep underlying causes of human errors. However, accidents are not due to one single fault but a combination of complex systemic failures that eventually lead to catastrophes [1], and it is difficult to understand or anticipate the ship officers' psychophysiological states during their voyages under different operational conditions and scenarios.

To address these challenges, various human factors studies are conducted in maritime domain. The current methods usually utilize a full mission simulator with real-world scenarios. With the help of the simulator, the human factor study can be done by post-exercise questionnaires, or by analyses of the bio-signals recorded when the cadets are performing the maritime tasks. Compared to the questionnaires, bio-signals have the advantage such as continuous monitoring of the cadets' psychophysiological states with high temporal resolution during the maritime tasks performance. Thus, bio-signals analyses technology can be used as an additional assessment tool. In [2-4], Electroencephalogram (EEG)-based brain states recognition algorithms were proposed to identify human mental workload levels, emotions, and stress levels. In this paper, we propose and implement the EEG-based cadets training and performance assessment system. The system consists of the EEG data processing and analyses part and cadet performance assessment part. With the help of this system, the instructors could monitor and evaluate the cadet psychophysiological states during the task performance and get an indicative recommendation on "pass", "retrain", or "fail" of the cadet based on the EEG recognition results and input level of the task difficulty.

The paper is organized as follows. In Section II, the related work including maritime simulators use, current trainee assessment in maritime academy, psychological tests in maritime domain, and EEG-based brain states evaluation algorithms are reviewed. In Section III, the designed and implemented EEG-based training and performance assessment system is introduced. In Section IV, the case study which utilizes the assessment system is presented. Section V concludes the paper.

II. RELATED WORK

A. *Maritime Simulators*

According to the current research [5], wrong judgement, failure to follow the rules and improper lookout are the causes of 80% accidents in the maritime domain. The underlying reasons leading to such situation could be due to stress, fatigue and/or lack of awareness of the crew.

The study of human factors in maritime domain conventionally relies on the accident reports. However, near misses and a lot of accidents often are not reported [6], which may lead to incomprehensive study. Due to this problem, it is important to have other methods, which allow collecting all information about the crew during the navigation of the ship. One of the methods is use of a full mission virtual maritime simulator.

The virtual maritime simulators are suitable for human factors study as the environment could be fully controlled. For example, a weather condition and traffic density can be changed. Additionally, the simulation could be designed to any time length according to the purposes of the exercise. Lastly, cadets/trainees could go through the same simulation settings and comparison could be made between subjects [6].

B. *Cadets/Trainees Assessment in Maritime Academy*

The methods to assess maritime cadets/trainees in maritime academies currently include written exam, achievement reports and interviews provided by the lecturers and ship training staff [7], the evaluation of observed actions of the cadets/trainees in the simulators [8], etc.

The written exam covers the knowledge that is taught in the classroom, and it is considered as an objective measurement, while the achievement report and interviews with lecturers and ship training staff are subjective. For example, it is shown in [7] that the ship staffs are prone to giving positive decision on some outstanding students who draw the attention during their stay on board. The ship staff tries to avoid negative decision since they do not work closely with the cadets/trainees. As a result, the subjective assessment could have the problems of unreliability, invalidity, and unfairness.

Another way to assess the cadets/trainees is to observe their actions in the simulator. The indicators of performance in the simulator could be correct action, correct sequence of steps, reaction time, subsystem knowledge, exchange of voice messages among operators, knowledge regarding the impact of abnormal situation, or speed of fault diagnosis depending on the difficulty level of the tasks in the simulator [8]. It is also mentioned in [8] that capability to handle stress is important to be assessed as the cadets/trainees should be able to manage any abnormal situation. However, the current assessment has only limited focus on psychophysiological states of the cadets/trainees.

C. *Psychological Tests in Maritime*

In this section, psychological tests used in maritime are reviewed. However, these tests are mainly used for screening and selecting during cadets/officers recruitment or promotion.

SeaPert [9] is one of these tests which includes behavioral, cognitive and personality analysis of the cadets and officers. It can provide information about personality, leadership and stress adaptability of the testees by a series of tests on leadership style, interpersonal style, mechanical reasoning, and abstract reasoning. This assessment is used in recruiting new cadets, training, and promotion of officers.

Ability Profiling Program [10] is another example of the psychological tests used in maritime industry. The test consists of seven types of tests to evaluate how the subject observes, processes and acts on information. It is claimed that the test results are able to predict a subject’s ability and performance.

Both above-mentioned tests are used in selection or screening during promotion or recruitment. Thus, the current trainee assessment still lacks a reliable method to evaluate the psychophysiological states of the trainee while he/she is performing the maritime tasks.

III. SYSTEM DESIGN

With the development of bio-signal sensors, it is possible to integrate bio-signals such as EEG, heart rate, etc. into the assessment of maritime trainees. In this paper, EEG-based brain states recognition algorithms are used to identify mental workload levels, emotions, and stress levels of trainees. The implemented EEG-based brain state recognition tools provide the maritime instructor with continuous high temporal resolution monitoring of the cadet psychophysiological states. It is also an objective brain states assessment as it can reflect the true and inner feelings of the trainees while they are performing maritime tasks. Current main limitation of the EEG-based technology use is that the trainee has to wear the EEG device mounted on his/her head while he/she is performing maritime tasks.

Studies have shown that higher mental demands led to the increase of human fatigue and lower concentration, which resulted in poor judgements [11-14]. While low levels of stress can aid in boosting performance, for example, in sports, high levels of stress can cause an inverse effect such as degrading performance due to the affected cognitive functions [15]. As mental workload reflects the amount of efforts needed to perform a particular task [16], it could be used as an indicator of the cadet’s acquired skills in the simulator. The more skilled the trainee is, the lower mental workload he/she has during the task performance. Thus, in our system, we evaluate the brain states such as mental workload and stress in relation to the level of the task difficulty to give additional information to the instructor/examiner on “pass”, “retrain”, or “fail” of the cadet based on the EEG processing and analyses results and the input level of the task difficulty.

A. *Overview of the EEG-based Training and Performance Assessment System*

In this paper, an EEG-based cadets training and performance assessment system is designed and implemented for maritime instructors/examiners use to provide additional information based on the EEG for evaluation of the trainees’ performance in the simulator. The system consists of two parts – a processing program and an evaluation program.

The flow chart of the system is presented in Fig. 1. Currently, the EEG-based brain state evaluation algorithms are subject-dependent. A calibration session is needed before the classifier can be applied to identify the brain states during the maritime exercise. The features extracted from the calibration data are used to train the classifier. The trained model is then used to identify the brain states. The recognized brain states such as workload, emotion, and stress are the input to the evaluation program, based on which a final indicative recommendation on "pass", "fail", or "re-train" of the trainee is given to instructors/examiners. The details of the EEG processing part and evaluation part of the system are given in the following sections.

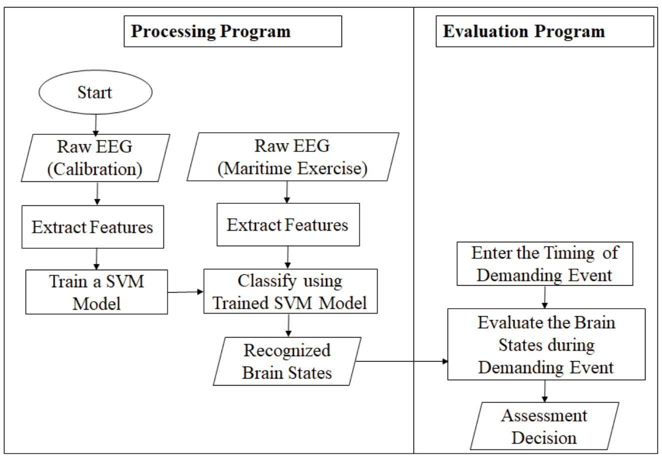

Figure 1. Flow chart of the proposed assessment system.

B. *EEG-based Brain State Recognition Algorithms*

In [3], an EEG-based emotion recognition algorithm was proposed and validated. Statistical features and fractal dimension features were extracted from raw EEG data. Support Vector Machine (SVM) was used as the classifier. Up to eight emotions can be recognized with an accuracy of 69.53%. In this paper, the targeted emotions to be recognized are negative, neutral, and positive emotional states. The best algorithm accuracy for 3 emotions is 72.22%.

Mental workload is defined as the mental efforts needed to complete the task [16]. In [2], statistical features and fractal dimension features were used while different classifiers such as SVM and k-Nearest Neighbors were compared. The best accuracy obtained is 80.09% for 4 workload levels achieved using SVM.

Since both EEG-based emotion and workload recognition algorithms are subject-dependent, a calibration is needed to train the classifier for each individual. Sound clips from IADS database [17] are selected and used in the implemented system. The sounds are played to the subject to evoke the corresponding emotions. The EEG data are recorded when the subjects are exposed to the stimuli and are used as training data. Stroop color test is used to evoke four levels of workload during workload algorithm calibration session. The recorded EEG data are used to train the SVM classifier for the workload recognition.

In both algorithms, the EEG data are filtered with a 2-42 Hz bandpass filter and processed using 4-second sliding window with 1 second overlapping to extract the corresponding features. A SVM classifier is then trained with the features extracted from the calibration data for 3 emotions and for 4 workload levels. After that, the trained SVM models are used to classify the subjects' emotions and workload levels during the maritime exercise.

Based on the recognized emotions and levels of workload from EEG, [4] proposed an algorithm to identify stress levels by combining emotional and workload states. As the result, with 3 levels of emotional states and 4 levels of mental workload states, 8 levels of stress can be recognized.

The described EEG-based workload, emotions and stress recognition algorithms are used in the EEG processing part of the implemented system.

C. *EEG Processing*

The processing part of the system has to process the raw EEG data and recognize different brain states such as workload, emotions, and stress of the maritime trainees from their EEG applying the state-of-art algorithms described in the previous section. The program is implemented in C++ on Visual Studio 2017. The GUI is shown in Fig. 2.

For the workload recognition, there is a button for the user to select the data recorded during calibration by clicking on the "Open Workload File" button (framed in orange in Fig. 2). Multiple files can be selected in the pop-up window. After choosing the training data, a message dialog box appears with the following message, "Proceed to feature extraction for training/test file". When the feature extraction is completed, another message dialog box appears with the following message, "Training/Test data feature extraction done". For emotion recognition, similar operations are applied. The user needs to select the training data (framed in blue in Fig. 2) and wait until the feature extraction is done. Then, the user can proceed to train the Support Vector Machine by clicking on the "Train SVM Model" button. A message dialog box appears with the following message "Train SVM done" when it is completed.

After the SVM model is trained and saved, the next step is feature extraction from the EEG data recorded during the maritime exercise and classification. The EEG data files recorded during the maritime exercises can be chosen by clicking "Open Test Files" button as framed in green in Fig. 2. Once the feature extraction is completed, by clicking the "Classify Test Data" button, the system starts to classify the brain states such as emotion and workload based on the trained SVM model. The level of stress is also recognized by combining the emotional and workload states as described in previous Section. A message dialog box appears showing "Classifying test data done" when the processing is completed. To save the classified result data, the "Save Classification Result" button should be clicked.

D. Evaluation

The evaluation program is used to assess the maritime trainees based on the brain state recognition results from EEG and give a final recommendation whether the trainees pass/fail the assessment or need retraining. The program is implemented using HTML and JavaScript.

In this system, first, the user (instructor/examiner) needs to input the basic information about the trainee such as his/her name and age. The user also has to input the timing of the demanding events and the corresponding difficulty levels each event as shown in Fig. 3. The next step is retrieving the files of the processing program results and predefined thresholds (Fig. 4). Once all information has been input into the system, the background algorithm starts the processing as follows: EEG-based brain states recognition results corresponding to the demanding events are obtained and an average mental workload is calculated over every 2-minute interval. Then, the average workload is compared with the corresponding thresholds, and if it falls within the range defined by the two thresholds, it is considered as the "pass" result; otherwise it is the "fail" result. To make a final decision, the percentage of "pass" intervals and "fail" intervals is counted in each event and summed up across all events. If the percentage is not less than 62.5%, the decision is "pass", if it is equal or more than 50% but less than 62.5%, the decision is "retrain", and otherwise the decision is "fail". The thresholds and percentile of recommendation are decided based on the database of EEG data collected from maritime trainees and can be updated by the maritime instructors/examiners. The pseudocode of the evaluation algorithm is given as Algorithm 1. The system screenshot of the final evaluation result shown in Fig. 5 includes the text of the final decision (pass/fail/retrain), pie charts which visually illustrates the proportion of fail and pass intervals, the subject's information such as name and age, and the details of each event. In the example shown in Fig. 5, there are four events: leaving port and joining TSS, leaving TSS mid-point, navigating in the TSS during poor visibility, and crossing the TSS in a precautionary area. The subject managed to pass the assessment, as an overall pass rate is 88.89%. The user can also select to view the cadet's workload/stress change during the particular event as shown in Fig. 6 with the time resolution of 1 second. For evaluation of the stress levels, the procedure is the same as the workload procedure except stress data are retrieved instead of the workload data.

Thus, the instructor/examiner gets indicative recommendations from the system based on objective assessment of the trainee workload and stress.

Figure 2. Screenshot of the processing program.

Algorithms 1 Evaluation algorithm

```
Input: eventInfo {The time points and the difficulty level of the demanding
  events}
n ← size(eventInfo, 1)
for all i such that 1≤ i ≤ n do
    {Retrieve the time points of a demanding event}
    timePoints ← eventInfo_i1
    {Retrieve the difficulty level of the demanding event}
    difficultyLvl ← eventInfo_i2
    {Retrieve a pair thresholds for the corresponding difficulty level of the
     demanding event}
    thresholds ← getThresholds(difficultyLvl)
    {Retrieve EEG-based recognition results of the demanding event}
    brainState ← getEEG(timePoints)
    {Calculate average brain state over 2-minute interval}
    meanBrainState ← averageEEG (brainState)
    m ← size(meanBrainState, 1)
    for all j such that 1≤ j ≤ m do
        if thresholds_1≤meanBrainState_j≤ thresholds_2
            passCounter←passCounter+1
        else
            failCounter←failCounter+1
        end if
    end for
end for
{Calculate the percentage of pass events}
finalConclusion ← passCounter/n x 100
{Make final decsion about pass, fail, or retrain according to the pass
 percentage}
if finalConclusion ≥ 62.5
    decision ← pass
else 50≤ finalConclusion < 62.5
    decision ← retrain
elseif
    decision ← fail
end if
```

IV. Case Study on Workload Assessment

In this section, the implemented EEG-based assessment system is applied to analyze performance of 18 maritime trainees based on their workload recognition from EEG and give indicative decision on their performance.

A. Data Collection

Eighteen trainees were recruited. They all had similar maritime background and training. Each trainee went through 4 bridge simulation exercises in the simulator. The exercises took about 20-30 minutes each and weather, and traffic condition varied in different exercises to achieve different levels of difficulty. For example, the visibility during navigation can be low, and extreme weather condition like heavy rain can be added.

Before each exercise, the cadets signed the consent form and were briefed by the instructor. Information such as ship type, location starting points and endpoint was given. The trainee then had time to plan the route on the map. Once they completed the planning, calibration of EEG-based emotion and workload recognition algorithms was done in the simulator, followed by the maritime exercise. EEG data were recorded during each exercise as shown in Fig. 7.

The EEG data were recorded by Emotiv [18] device with 14 electrodes. The recorded data were used as the input for the assessment system. Videos were also taken to study the trainees' behaviors when they were performing the maritime exercises.

B. *Results*

The EEG data were processed by the processing program and then fed into the evaluation program to get the final decision on the performance.

Among all exercises, the cadets/trainees had six collisions as follows: trainee 2, trainee 12, and trainee 17 - in exercise 2; trainee 3, and trainee 10 - in exercise 3; trainee 16 – in exercise 4. The final indicative decision by the EEG-based evaluation system for each trainee is given in Table I. Here, the exercises with collision are highlighted in bold. In summary, two trainees who had collisions got "fail" recommendation by the system, two trainees got "retrain" recommendation, and two trainees got "pass" result by the EEG-based assessment system. When their mental workload graphs were analyzed it was shown that all trainees demonstrated high workload during the collision and after it.

The video recordings of all trainees were further studied in order to find the correlation between their behaviors and EEG-based assessment results.

First, videos of trainee 2 and 3 who experienced ship collision but got "pass" recommendation by the system were analysed. From the videos and from the workload graphs, it can be seen that both trainee 2 and 3 were confident before collision during the task performance. Their workload started to increase only right before the collision occurred and stayed high during and after the collision. More analyses are needed on their psychophysiological states. Other factors such as emotions and vigilance could be considered besides the mental workload and stress to understand the cause of their underperformance.

The video of trainee 12 and 17 who experienced collision but got "retrain" recommendation were analyzed. From the videos and from the workload graphs it can be seen that trainee 12 was too relaxed until the collision happened for a longer period whereas trainee 17 experienced higher workload throughout the entire exercise, and especially after collision.

Finally, trainees who experienced collision and obtained a failure indicative recommendation based on the EEG were investigated. From the videos it can be seen that both trainee 10 and 16 were very nervous and stressed throughout the simulation exercise. Their mental workload graphs showed very high levels of workload during the entire exercise which means that they did not acquire yet enough skills.

For the trainees who did not have collision and got "pass" recommendation from the system, the observations from the videos showed that they were attentive and confident during the tasks performance. It corresponds to the results from the EEG-based assessment that showed their optimal psychophysiological state as well. For the trainees who ended up with "retrain" recommendation based on the EEG results, the videos showed that they were doing well at the start of the exercise but along the way they became more and more nervous which may be caused by the increase in skills demands and finally higher workload values. For those trainees who got "fail" indicative recommendation, the video showed that they were too nervous during the entire exercise.

TABLE I. ASSESSMENT RESULTS OF 18 SUBJECTS

Trainee ID	**Exercise**			
	1	***2***	***3***	***4***
1	F	F	P	R
2	P	**P**	P	F
3	P	P	**P**	P
4	P	R	P	P
5	R	R	F	F
6	R	F	P	P
7	R	R	F	F
8	F	F	F	F
9	F	F	F	F
10	F	F	**F**	P
11	F	R	P	P
12	R	**R**	F	F
13	R	R	F	P
14	F	P	F	F
15	P	P	F	R
16	F	R	F	**F**
17	F	**R**	P	P
18	P	P	R	F

V. CONCLUSION

In this paper, we presented a novel EEG-based assessment system that could help maritime training instructors/examiners to assess trainees. The system provides information about psychophysiological states of cadets during the maritime task performance in virtual simulator and gives an indicative recommendation on the cadet/trainee performance assessment such as "pass", "retrain" or "fail" based on objective EEG-based measurements. Different brain states such as mental workload, emotions, and stress can be monitored with 1-second temporal resolution.

A case study which utilizes the implemented system to assess 18 maritime trainees is also presented. The analyses of the assessment results using both videos and EEG-based workload recognition data showed that further study is needed to improve accuracy of the underlying algorithms and the overall system. The results also showed that the implemented system can be used as an additional tool to help training instructors in spotting the areas where the trainee still has some lack of skills. In the next step, the system will be further tested with more cadets before its first trial use for cadets training in the maritime simulator.

Currently, we are also working on the improving the accuracy of subject-independent workload and stress recognition algorithms and its integration into a new version of the system that would allow the instructors/examiners using the system without time-consuming calibration.

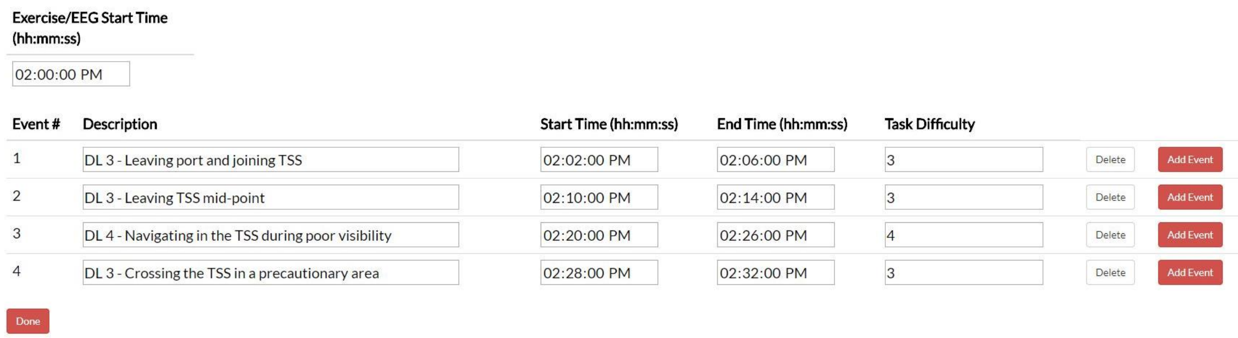

Figure 3. Screenshot of the evaluation system.

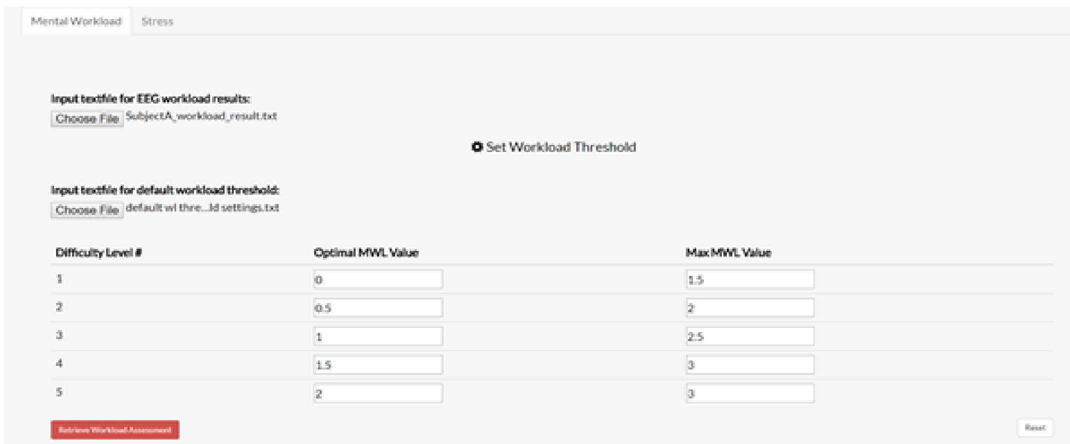

Figure 4. Screenshot of the evaluation system for retrieving workload results and threshold.

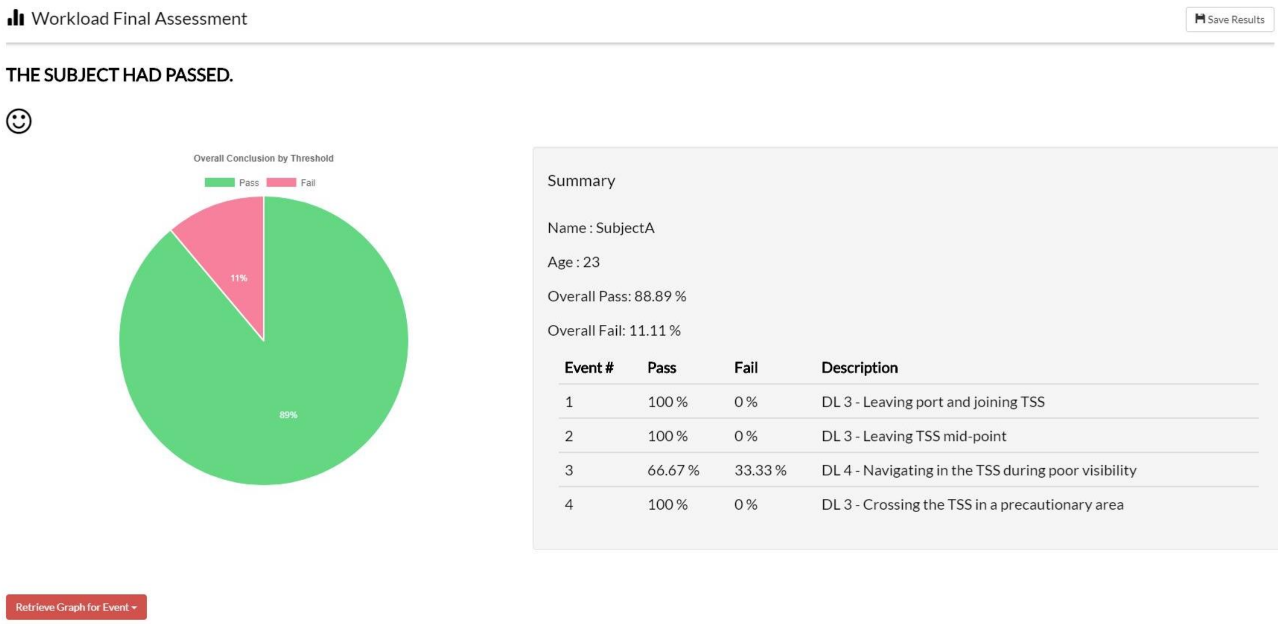

Figure 5. Screenshot of final recommendation by the evaluation system.

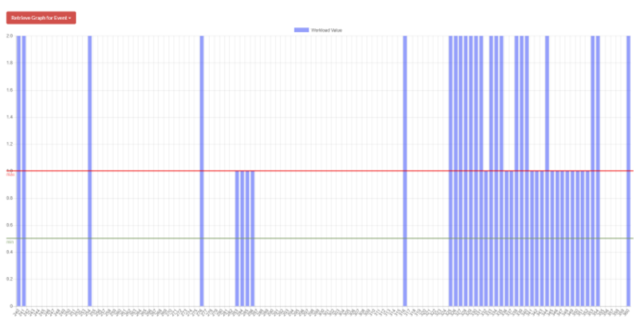

Figure 6. Screenshot of the evaluation system for retrieving detailed changes of workload overtime.

Figure 7. Maritime trainee wearing the EEG device is conducting maritime tasks in the simulator.

Acknowledgment

This research is supported by Singapore Maritime Institute and by the National Research Foundation, Prime Minister's Office, Singapore under its international Research Centres in Singapore Funding Initiative.

References

[1] A. M. Rothblum, "Human error and marine safety," in National Safety Council Congress and Expo, Orlando, FL, 2000, p. 7.

[2] W. L. Lim, O. Sourina, L. Wang, and Y. Liu, "EEG-based Mental Workload Recognition Related to Multitasking," in Proceeding of the Int Conf on Information, Communications and Signal Processing (ICICS), 2015, pp. 1-4.

[3] Y. Liu and O. Sourina, "Real-Time Subject-Dependent EEG-Based Emotion Recognition Algorithm," in Transactions on Computational Science XXIII, vol. 8490(Lecture Notes in Computer Science: Springer Berlin Heidelberg, 2014, pp. 199-223.

[4] W. L. Lim et al., "EEG-Based Mental Workload and Stress Monitoring of Crew Members in Maritime Virtual Simulator," in Transactions on Computational Science XXXII: Special Issue on Cybersecurity and Biometrics, M. L. Gavrilova, C. J. K. Tan, and A. Sourin, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, pp. 15-28.

[5] Y. Liu, X. Hou, O. Sourina, D. Konovessis, and G. Krishnan, "Human Factor Study for Maritime Simulator-based Assessment of Cadets," in Proceedings of the 35th International Conference on Ocean, Offshore and Arctic Engineering, Busan, South Korea, 2016.

[6] T. Koester, "Human factors and everyday routine in the maritime work domain," in Human Factors in Transportation, Communication, Health, and the Workplace. Human Factors and Ergonomics Society Europe Chapter Annual Meeting, 2001.

[7] A. P. D. E. Demirel, "A Study on the assessment of sea training as an integral part of maritime education and training," ICQH 2015, p. 47, 2015.

[8] S. Nazir, K. I. Øvergård, and Z. Yang, "Towards Effective Training for Process and Maritime Industries," Procedia Manufacturing, vol. 3, pp. 1519-1526, 2015/01/01/ 2015.

[9] SeaPert -- Maritime Test. Available: https://www.psychometrica.in/maritime-test.html

[10] Ability Profiling (APRO). Available: https://www.seagull.no/Recruitment-tools/APRO

[11] D. Gopher and E. Donchin, "Workload: An examination of the concept," 1986.

[12] S. Fan, X. Yan, J. Zhang, and J. Wang, "A review on human factors in maritime transportation using seafarers' physiological data," in Transportation Information and Safety (ICTIS), 2017 4th International Conference on, 2017, pp. 104-110: IEEE.

[13] B. Cain, "A review of the mental workload literature," Defence Research And Development Toronto (Canada)2007.

[14] Y. Y. Yurko, M. W. Scerbo, A. S. Prabhu, C. E. Acker, and D. Stefanidis, "Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool," Simulation in healthcare, vol. 5, no. 5, pp. 267-271, 2010.

[15] M. Mendl, "Performing under pressure: stress and cognitive function," Applied Animal Behaviour Science, vol. 65, no. 3, pp. 221-244, 1999.

[16] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: mental workload in ergonomics," Ergonomics, vol. 58, no. 1, pp. 1-17, 2015.

[17] Bradley M. M. and L. P.J., "The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual," University of Florida, Gainesville2007.

[18] Emotiv. http://www.emotiv.com.