# Artificial Neural Network (ANN) for Performance Assessment in Virtual Reality (VR) Simulators: From Surgical to Maritime Training

**Tusher, Hasan Mahbub; Nazir, Salman; Mallam, Steven; Munim, Ziaul Haque**
Department of Maritime operations - University of South-Eastern Norway

This is an Accepted Manuscript of an article published by IEEE in *2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* on December 26, 2022, available online: https://doi.org/10.1109/IEEM55944.2022.9989816

Tusher, H. M., Nazir, S., Mallam, S. & Munim, Z. H. (2022, December 7-10). Artificial Neural Network (ANN) for Performance Assessment in Virtual Reality (VR) Simulators: From Surgical to Maritime Training [Paper presentation]. 2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Kuala Lumpur. https://doi.org/10.1109/IEEM55944.2022.9989816

# Artificial Neural Network (ANN) for Performance Assessment in Virtual Reality (VR) Simulators: From Surgical to Maritime Training

H.M. Tusher[1], S. Nazir[1*], S. Mallam[1], Z.H. Munim[1]
[1]Department of Maritime Operations, University of South-Eastern Norway, Horten, Norway
(salman.nazir@usn.no)

*Abstract* – **Simulator training is an integral part of seafarer education and training. Maritime Virtual Reality (VR) simulators have added a new dimension to the range of available state-of-the-art training tools in recent years. The lack of appropriate pedagogical intervention including inadequate performance assessment frameworks for the trainees are few of the limitations of maritime VR simulators. In this study, a performance assessment framework utilizing Artificial Neural Network (ANN) in VR training from the healthcare domain is adapted through literature review. This framework could be operationalized in maritime training for aiding the performance assessment of seafarers and in turn increasing the pedagogical efficiency of maritime VR simulators. The implication of such adaption is also discussed considering the human factors and the technical dimensions of maritime training.**

*Keywords* – **performance assessment, seafarer, deep learning, machine learning, artificial neural network**

## I. INTRODUCTION

The continuously evolving interplay between humans and machines currently accounts for most of the dynamic challenges in complex socio-technical systems [1]. Being a safety-critical industry, maritime workplaces are complex where technical and social artefacts interact with each other in a dynamic environment. In such environments, incidents and accidents occur due to broadly two categories of failures: (i) active failures occurring at the "sharp end" of the system involving operators and field workers, and (ii) whereas latent failures occur at the higher order in the hierarchy of system [2]. Training is often cited as one of the safeguards against these failures in the maritime domain [3], [4]. Therefore, developing a competent workforce through enhanced training as a means of reducing risks in maritime workplaces is emphasized.

The context of developing maritime competence, as mandated by international maritime regulations, has evolved simultaneously with the advancement of technologies in workplaces and changes in organizational and operational demands. Consequently, the integration of state-of-the-art training tools such as Virtual Reality (VR), Augmented Reality (AR) and cloud-based simulators for enhanced training output have been envisaged in the maritime domain [5].

### A. VR simulators in maritime training

Besides being a cost-effective and space-saving technology, VR simulators can provide realistic experiences to the trainees which support different learning theories such as experiential learning, situated learning and constructivist learning etc.[5], [6]. VR training is regarded as useful for training cognitive skills, spatial memory skills, and procedural as well as psychomotor skills in different safety-critical domains [7]. It also provides realistic training while reducing learning time for seafarers compared to the other types of traditional simulators [8], [9]. In addition, emergencies such as fire, flooding, steering gear failure etc can be simulated in a safe environment in VR which may not be possible to demonstrate in a real ship. Therefore, the cost of training, as well as the risk to the trainees reduce considerably in VR simulator training. VR simulation has also been used evidently in procedural knowledge and technical skill training operations in the maritime domain [10].

Differing maritime stakeholders have employed 3D VR solutions in differing ways to assess their technological suitability for maritime operations and training [11]. Most maritime safety training utilizing VR simulators operates in a selected few areas, such as command bridge operations, machine room operations, crane and fire safety operations [11].

### B. Performance assessment in maritime VR training

Available performance assessment measures in VR simulators often lack evidence-based data to support their validity and reliability [12]. Empirical investigations related to VR simulator training in different safety-critical domains mostly use NASA-TLX in measuring usability, presence and other subjective self-reported performance measures through questionnaires [7]. Other types of performance measures also include different basic measurements such as time to complete a task, retention time, response time, number of errors, eye-tracking etc. as quantitative measures [7], [11]. Bayesian networks have been used to reduce the subjectivity of performance assessment during pilotage operation in the maritime full-mission bridge simulator training [13]. However, performance assessment in other diverse maritime scenarios especially during VR training is left unexplored.

*C. Performance assessment using Artificial Neural Network (ANN)*

Artificial Intelligence (AI), especially Artificial Neural Network (ANN) has been identified as a state-of-the-art solution for its use in objective human performance assessment in differing contexts [14], [15]. In education, AI-powered tools could behave like humanoid robots replacing instructors or aiding them in different learning and assessment scenarios [16]. Four (04) major categories of AI applications have been identified in [17]: (1) profiling and prediction, (2) assessment and evaluation, (3) adaptive systems and personalization, and (4) intelligent tutoring systems. A range of AI assessment implementation include automated essay scoring system, computerized adaptive tests, providing adaptive education and training material to the students, computerized formative, or summative assessments [18]. Luckin et al. (2016) refers AI integration in students' assessments as "renaissance in assessment" since AI can provide just-in-time assessment to the students bypassing the stop-and-test type of assessment frameworks. The application of AI techniques especially Artificial Neural Network (ANN) in VR has been prevalent in different domains, such as teleconferencing, human-computer interaction, behaviour recognition, education, agriculture, transport and tourist domains to name a few [20]. However, utilizing ANN in VR for performance assessment is comparatively new in literature [21] which could bring novel possibilities if investigated in maritime VR training contexts. The application of such framework may widely differ along with the context of training. Therefore, investigating only one type of application, e.g., surgical VR training in healthcare domain where ANN algorithms are used only for performance assessment could provide deeper insight into this one-dimensional utility of those algorithms.

*D. Goal of this study*

This study aims to investigate instances in literature where ANN performance assessment models are being used for surgical VR training. The results could reveal novel utility of these ANN algorithms in maritime VR training. In addition, other human factors and technological implications of utilizing ANN algorithm for maritime performance assessment is discussed in brief.

## II. LITERATURE REVIEW

*A. Methods*

A scoping literature review is performed to locate ANN utilization examples for performance assessment in VR simulators during surgical training. The following search string was applied in two (02) scientific databases *Scopus* and *Web of Science* during June 2022 without setting any time range:

*("virtual reality" AND "simulat*" AND "artificial neural network" AND "surgical" AND "assessment")*

*B. Results*

The results of the literature review reveal differing ways ANN is utilized for surgical performance assessment in VR training simulators [22]–[27]. The goal of such ANN-VR applications was to either categorize the level of expertise of surgeons or to propose efficient techniques for individual performance assessment.

In one instance, surgeons selected several surgical performance matrices in a specific surgical task and developed an ANN algorithm that was trained with the data from a simulated VR scenario involving real surgeons with varying level of expertise [22]. The trained algorithm showed 80% accuracy in categorizing performance in the novice to the expert range.

On similar instance, scientists made efforts to quantify psychomotor skills performed during surgical procedures [27]. Collected data consisted of 250 simulated surgical tasks carried out in VR by 50 surgeons with expertise ranging from 1 to 25 years. The analysis proposed that the ANN algorithm could predict surgical performance on four different expert levels with almost 90% accuracy.

Both instances above depict a common framework of ANN utilization for surgical assessment during simulated training in VR. The overall process of ANN application in VR training scenario could be summarized in three stages: metric generation from raw data, training the ANN algorithm and validation [22].

First, surgical procedures were performed in a simulated environment, then data collection and pre-processing phase of the ANN-VR assessment framework starts. Following the raw data acquisition from VR training task, metric generation commences to identify important performance metrics that are required for efficient task performance. Afterwards, appropriate metrics are selected for a specific task context which in turn narrows down the group of metrics. The final selected metrics are then fed as inputs to the neural network training algorithm (see Fig. 1).

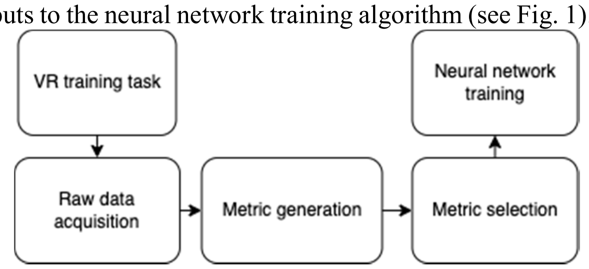


Fig. 1. Framework of implementing ANN in VR surgical training task (adopted from Mirchi, Bissonnette, Ledwos, et al., 2020)

For training the ANN algorithm, data from each performance metric are fed into the input layer ($x_1$, $x_2$, $x_3$…. $x_n$) of the neural network (see ***Error! Reference source not found.***), which are connected to several neurons in the hidden layer ($y_1$, $y_2$… $y_m$). The connections between the neurons would be supported by different weights ($w_{1,1}$, $w_{1,2}$…. $w_{1,m}$) which constitutes how much the neurons of the hidden layer (y) will be sensitive to the inputs (x). Each hidden layered neuron (y) is connected to three possible outputs represented by $z_1$, $z_2$, $z_3$ for novice, semi-expert and expert performers respectively (see ***Error! Reference source not found.***).

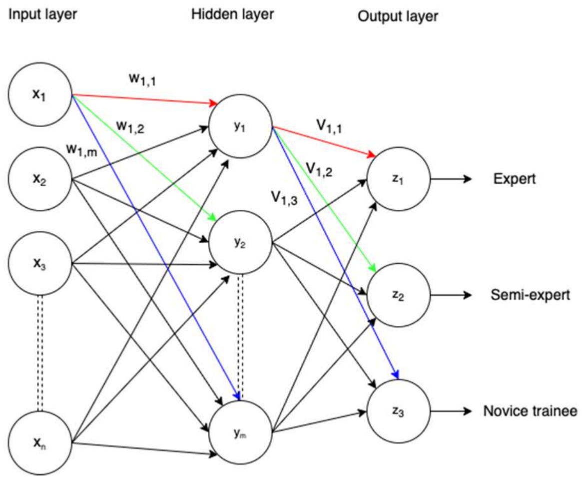


Fig. 2. Framework for multi-layered ANN in surgical performance assessment (adopted from Mirchi, Bissonnette, Ledwos, et al., 2020)

## III. IMPLICATION OF ANN-VR ASSESSMENT FRAMEWORK IN MARITIME TRAINING

### *A. Reducing subjectivity in performance assessment*

ANN-VR performance assessment framework identified from surgical training instances could be utilized to reduce the subjectivity in performance assessment in maritime VR training. Differing sets of performance matrices for different training scenarios (e.g., navigation, fire-fighting, enclosed space entry etc.) need to be developed first in order to build tailored assessment frameworks for maritime VR training. Afterwards, raw data collection from VR training scenarios corresponding to those developed metrics could be used for training the ANN algorithm.

It would be possible to provide more accurate adaptive training to the trainees according to their level of expertise as identified by the ANN-VR framework. In addition, the trainee portfolios generated during simulator training sessions could be compared against their real-world performance to extract crucial behavioural information related to individual trainees. The trained ANN algorithm for a specific maritime training task in VR could be utilized for future assessments of unlabelled raw performance data from similar tasks. The development process of such an algorithm could follow previous studies derived from the healthcare domain [22], [24], [27].

In the stepwise approach of the ANN-VR framework, the generation of performance metrics could reveal valuable insights throughout the whole training process. For example, investigating the intricate relationships among individual performance metrics, their comparative importance in the entire task as well as differing approaches by trainees to accomplish those tasks could provide in-depth knowledge and understanding of VR training.

### *B. The Human Element Aspects*

Discussions related to bias, identity and displacement of human labor could be some of the emerging issues in this context. Negative attitudes towards robots [28] could pose a hindrance in adopting the ANN-VR training framework as AI agents in our traditional learning environments, whereas the potential benefits emerging from the unique VR affordances could open new avenues of research. On the other hand, the students may lack the feeling of enjoyment while learning or lose motivation due to the absence of social aspects of learning associated with the presence of a human instructor [29]. In addition, perpetuating human biases in AI (e.g., ANN-VR framework) could not be ignored, since human inputs are essential during the development of such models [30].

Therefore, creating a harmonized environment would be a challenge where both types of agents – humans and AI can coexist to ensure an optimum learning environment [29]. Multidisciplinary collaboration consisting of maritime professionals, AI-engineers, VR-developers, and human factors researchers could be utilized for such a harmonized environment as envisaged.

### *C. Technological Implications*

AI-powered tools could be utilized for ships' behaviour prediction, risk prediction, surveillance and enhanced decision support systems in the context of future digitalized shipping [31]–[34]. Therefore, integrating deep learning algorithms such as ANN for seafarer training is the logical next step. Though utilizing VR for seafarer training is not new, there is no direct precedent of integrating ANN techniques for performance assessment in these training. Thus, the scope of technological advancement in the maritime training context will have greater potential with cross-domain collaboration (e.g., with the healthcare domain). Moreover, the respective neural network utilized for capturing ships' behavior corresponding to the navigators' actions is generally used for enhancing the self-driving capabilities of autonomous ships [35], which could potentially be used to train the ANN algorithms integrated with VR for future seafarers' training.

## IV. CONCLUSIONS AND FUTURE DIRECTIONS

This study adopts a conceptual framework from the healthcare domain utilizing ANN in VR simulator training and explores its potential applicability in maritime simulator training contexts. This attempt, if successful would open new doors to merging two state-of-the-art technologies, VR and AI in the maritime training context. Although the proposed framework is envisaged to be utilized in a different technical skill training scenario, this integrated platform for performance assessment and categorization could eventually be used in all scenarios including behavioural or non-technical skills assessment in the maritime domain.

The emerging implications resulting from the integration of the ANN-VR framework in maritime training would imply significant transformation in the existing maritime simulator training practices. Challenges related to technology acceptance, ethical consideration as well as organizational readiness factors need to be evaluated simultaneously. Eventually, the social and organizational compatibility of such novel technological solutions (e.g., ANN-VR assessment framework) would largely determine their potential for success in the future.

ACKNOWLEDGMENT

The authors would like to thank the Centre of Excellence in Maritime Simulator Training and Assessment (COAST) in Norway, funded by the Directorate for Higher Education and Competence (HK-dir) for their cooperation and facilitation for this research work.

REFERENCES

[1] S. Dekker, *The field guide to understanding 'human error'*. CRC press, 2017.

[2] J. Reason, 'Understanding adverse events: human factors.', *Qual. Saf. Health Care*, vol. 4, no. 2, pp. 80–89, Jun. 1995, doi: 10.1136/qshc.4.2.80.

[3] T. T. Türkistanli and B. Kuleyin, 'Game-based learning for better decision-making: A collision prevention training for maritime transportation engineering students', *Comput. Appl. Eng. Educ.*, vol. 30, no. 3, pp. 917–933, 2022.

[4] C.-J. Chae, K. H. Kim, and S. Y. Kang, 'Limiting ship accidents by identifying their causes and determining barriers to application of preventive measures', *J. Mar. Sci. Eng.*, vol. 9, no. 3, p. 302, 2021.

[5] T. Kim *et al.*, 'The continuum of simulator-based maritime training and education', *WMU J. Marit. Aff.*, vol. 20, no. 2, pp. 135–150, Jun. 2021, doi: 10.1007/s13437-021-00242-2.

[6] S.-K. Loke, 'How do virtual world experiences bring about learning? A critical review of theories', *Australas. J. Educ. Technol.*, p. 11, 2015.

[7] S. kumar Renganayagalu, S. C. Mallam, and S. Nazir, 'Effectiveness of VR Head Mounted Displays in Professional Training: A Systematic Review', *Technol. Knowl. Learn.*, Jan. 2021, doi: 10.1007/s10758-020-09489-9.

[8] H. Shen, J. Zhang, B. Yang, and B. Jia, 'Development of an educational virtual reality training system for marine engineers', *Comput. Appl. Eng. Educ.*, vol. 27, no. 3, pp. 580–602, May 2019, doi: 10.1002/cae.22099.

[9] S. C. Mallam, S. Nazir, and S. K. Renganayagalu, 'Rethinking Maritime Education, Training, and Operations in the Digital Era: Applications for Emerging Immersive Technologies', *J. Mar. Sci. Eng.*, vol. 7, no. 12, p. 428, Nov. 2019, doi: 10.3390/jmse7120428.

[10] J. Jung and Y. J. Ahn, 'Effects of interface on procedural skill transfer in virtual training: Lifeboat launching operation study: A comparative assessment interfaces in virtual training', *Comput. Animat. Virtual Worlds*, vol. 29, no. 3–4, p. e1812, May 2018, doi: 10.1002/cav.1812.

[11] E. Markopoulos and M. Luimula, 'Immersive Safe Oceans Technology: Developing Virtual Onboard Training Episodes for Maritime Safety', *Future Internet*, vol. 12, no. 5, p. 80, Apr. 2020, doi: 10.3390/fi12050080.

[12] M. M. Hammoud *et al.*, 'To the point: medical education review of the role of simulators in surgical training', *Am. J. Obstet. Gynecol.*, vol. 199, no. 4, pp. 338–343, 2008.

[13] J. Ernstsen and S. Nazir, 'Performance assessment in full-scale simulators–A case of maritime pilotage operations', *Saf. Sci.*, vol. 129, p. 104775, 2020.

[14] Y. Zhao and C. Smidts, 'CMS-BN: A cognitive modeling and simulation environment for human performance assessment, part 2—Application', *Reliab. Eng. Syst. Saf.*, vol. 213, p. 107775, 2021.

[15] E. Oster and M. Knapp, 'Artificial Intelligence, Microservices, xAPI, Human Performance Assessment, and Adaptive Training: How Five Buzzwords are Applied to Real Problems'.

[16] L. Chen, P. Chen, and Z. Lin, 'Artificial intelligence in education: A review', *Ieee Access*, vol. 8, pp. 75264–75278, 2020.

[17] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, 'Systematic review of research on artificial intelligence applications in higher education – where are the educators?', *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, p. 39, Dec. 2019, doi: 10.1186/s41239-019-0171-0.

[18] J. Gardner, M. O'Leary, and L. Yuan, 'Artificial intelligence in educational assessment: "Breakthrough? Or buncombe and ballyhoo?"', *J. Comput. Assist. Learn.*, vol. 37, no. 5, pp. 1207–1216, Oct. 2021, doi: 10.1111/jcal.12577.

[19] R. Luckin, W. Holmes, M. Griffiths, and L. B. Forcier, 'Intelligence unleashed: An argument for AI in education', 2016.

[20] T. Ribeiro de Oliveira *et al.*, 'Systematic Review of Virtual Reality Solutions Employing Artificial Intelligence Methods', in *Symposium on Virtual and Augmented Reality*, Virtual Event Brazil, Oct. 2021, pp. 42–55. doi: 10.1145/3488162.3488209.

[21] G. Sharma, S. Chandra, S. Venkatraman, A. Mittal, and V. Singh, 'Artificial Neural Network in Virtual Reality : A Survey', *Int. J. Virtual Real.*, vol. 15, no. 2, pp. 44–52, Jan. 2016, doi: 10.20870/IJVR.2016.15.2.2873.

[22] S. Alkadri *et al.*, 'Utilizing a multilayer perceptron artificial neural network to assess a virtual reality surgical procedure', *Comput. Biol. Med.*, vol. 136, p. 104770, Sep. 2021, doi: 10.1016/j.compbiomed.2021.104770.

[23] V. Bissonnette *et al.*, 'Artificial Intelligence Distinguishes Surgical Training Levels in a Virtual Reality Spinal Task', *J. Bone Jt. Surg.*, vol. 101, no. 23, p. e127, Dec. 2019, doi: 10.2106/JBJS.18.01197.

[24] N. Mirchi *et al.*, ‘Artificial Neural Networks to Assess Virtual Reality Anterior Cervical Discectomy Performance’, *Oper. Neurosurg.*, vol. 19, no. 1, pp. 65–75, Jul. 2020, doi: 10.1093/ons/opz359.

[25] N. Mirchi, V. Bissonnette, R. Yilmaz, N. Ledwos, A. Winkler-Schwartz, and R. F. Del Maestro, ‘The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine’, *PLOS ONE*, vol. 15, no. 2, p. e0229596, Feb. 2020, doi: 10.1371/journal.pone.0229596.

[26] A. Winkler-Schwartz *et al.*, ‘Machine Learning Identification of Surgical and Operative Factors Associated With Surgical Expertise in Virtual Reality Simulation’, *JAMA Netw. Open*, vol. 2, no. 8, p. e198363, Aug. 2019, doi: 10.1001/jamanetworkopen.2019.8363.

[27] A. Winkler-Schwartz *et al.*, ‘Artificial Intelligence in Medical Education: Best Practices Using Machine Learning to Assess Surgical Expertise in Virtual Reality Simulation’, *J. Surg. Educ.*, vol. 76, no. 6, pp. 1681–1690, Nov. 2019, doi: 10.1016/j.jsurg.2019.05.015.

[28] T. Nomura, T. Suzuki, T. Kanda, and K. Kato, ‘Measurement of anxiety toward robots’, in *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, 2006, pp. 372–377.

[29] C. Edwards, A. Edwards, P. Spence, and X. Lin, ‘I, teacher: using artificial intelligence (AI) and social robots in communication and instruction’, *Commun. Educ.*, vol. 67, no. 4, pp. 473–480, 2018, doi: 10.1080/03634523.2018.1502459.

[30] G. S. Nelson, ‘Bias in artificial intelligence’, *N. C. Med. J.*, vol. 80, no. 4, pp. 220–222, 2019.

[31] S. C. Christos, T. Panagiotis, and G. Christos, ‘Combined multi-layered big data and responsible AI techniques for enhanced decision support in Shipping’, in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 669–673.

[32] J. Ma, W. Li, C. Jia, C. Zhang, and Y. Zhang, ‘Risk prediction for ship encounter situation awareness using long short-term memory based deep learning on intership behaviors’, *J. Adv. Transp.*, vol. 2020, 2020.

[33] B. Murray and L. P. Perera, ‘An AIS-based deep learning framework for regional ship behavior prediction’, *Reliab. Eng. Syst. Saf.*, vol. 215, p. 107819, 2021.

[34] D. Nguyen, R. Vadaine, G. Hajduch, R. Garello, and R. Fablet, ‘A multi-task deep learning architecture for maritime surveillance using AIS data streams’, in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 2018, pp. 331–340.

[35] L. P. Perera, ‘Deep learning toward autonomous ship navigation and possible COLREGs failures’, *J. Offshore Mech. Arct. Eng.*, vol. 142, no. 3, 2020.