IEEE VR 2023 Workshop: Datasets for developing intelligent XR applications (DATA4XR)

Yuyang Wang, Jean-Rémy Chardonnet, Lik-Hang Lee, Pan Hui

HAL Id: hal-04592274
https://hal.science/hal-04592274

Submitted on 29 May 2024

IEEE VR 2023 Workshops

Workshop: Datasets for developing intelligent XR applications (DATA4XR)

ORGANIZERS:

Yuyang Wang
Hong Kong University of Science and
Technology (Guangzhou), China

Jean-Rémy Chardonnet
Arts et Métiers Institute of Technology, France

Lik-Hang Lee
The Hong Kong Polytechnic University, Hong
Kong SAR

Pan Hui
Hong Kong University of Science and
Technology (Guangzhou), China

ABSTRACT:

The 2nd workshop on Datasets for Developing Intelligent XR Applications (DATA4XR) aims to address the challenges of public datasets and reproducibility in Extended Reality, also known as XR (Augmented Reality, Virtual Reality, and Mixed Reality) research. The workshop brings together experts to discuss the availability, privacy concerns, and ethics related to open-sourcing the datasets used in XR research for algorithm training and user behavior analysis. By examining the ethical, moral, and privacy concerns and categorizing open-source datasets for XR, the workshop expects to advance XR research by incorporating artificial intelligence and defining a research agenda for XR in open science. In addition, the workshop focuses on involving the Human-Computer Interaction (HCI) community in exploring its role in the future of XR research. We invite submissions of technical papers, position papers, and research papers on topics related to XR and open science.

MESSAGE FROM THE DATA4XR WORKSHOP ORGANIZERS:

ABOUT THE WORKSHOP:

Data is a crucial resource for AI researchers in developing models. Access to standard datasets makes it easier for researchers to develop cutting-edge AI algorithms. Some countries ask researchers to publish their data alongside their papers in open science repositories. For example, the European Union's Horizon 2020 program demands that data produced during funded research projects be available for reuse. Also, the National Institutes of Health (NIH) in the United States should deposit the data from funded research projects into the PubMed Central repository. These policies expect to promote transparency, facilitate scientific collaboration, and increase research impact by allowing others to replicate published findings.

Meanwhile, XR technologies, including AR, VR, and MR, are maturing and becoming widely adopted. However, when XR researchers seek to implement advanced AI algorithms in developing interactive applications, the need for more publicly available datasets poses a challenge. Furthermore, while XR researchers often open-source their work, public datasets on user behavior are rare, hindering the replication of XR research based on user behavior. In the HCI community, there is an ongoing debate on ensuring the replicability and reproducibility of research results. However, many user studies involve specific populations and locations, leading to results that may need to be replicable in different cultural contexts. Additionally, authors may be hesitant to share additional materials, such as proprietary software, hindering replicative research and open science.

This workshop will focus on open-source datasets used in XR research for algorithm training and user behavior analysis. However, data availability and privacy concerns may pose challenges to replicative research. Therefore, the workshop aims to continue discussions from the IEEE VR 2022, including technical discussions on cutting-edge datasets for XR intelligent interaction applications and open-source datasets for replicating XR user experience techniques.

OBJECTIVES:

Artificial Intelligence (AI) techniques have been a promising area of research in predicting cybersickness, by utilizing large datasets to uncover implicit patterns and correlations. For instance, AI algorithms can be utilized to evaluate physiological signals, such as heart rate and skin conductance, to link to the existence and intensity of cybersickness. Kim et al. (2019) [1] employed a deep learning approach to examine data from the Simulator Sickness Questionnaire (SSQ) and identify the relationship between user behavior and cybersickness. The findings indicated that the deep learning algorithm could predict cybersickness with an accuracy rate of over 85%. Islam et al. (2020) [2] proposed a method for predicting the severity of cybersickness experienced by users in virtual reality environments using deep neural networks. The proposed method involves physiological signals from the user, including EEG, ECG, and GSR, and using these signals as input to an AI model, achieving an accuracy of 97.44% for measuring current cybersickness severity and 87.38% for estimating future cybersickness severity. Similar research can be found in [3] in which a LSTM model is employed to measure cybersickness. However, replication of these studies is restricted to the dataset, and peer researchers have to employ similar ideas for other studies. On

the other hand, inconsistent data collection can make the comparison between baselines and newly proposed methods more complicated and infeasible. As such, it is important to pursue a standardized way of sharing datasets for AI-driven XR studies, which is supported by the openness of the data and its interoperability among various studies.

The 2nd workshop on Datasets for Developing Intelligent XR Applications (DATA4XR), hosted by the IEEE VR 2023, provides a valuable platform for domain researchers to access resources and establish collaborations across labs. This workshop brings experts from various fields to address the challenges of public datasets and reproducibility in XR. The focus will be on involving the HCI community and exploring its role in the future of XR research. The workshop will help categorize open-source datasets for XR and define new challenges through discussions and reflections. In addition, examining ethical, moral, and privacy concerns will help develop a research agenda for XR in open science.

Numerous publications have been on datasets such as MINIST, ImageNet, and CIFAR-10. However, to fully unleash the potential of XR research, the community requires standardized datasets to create data-driven models using machine and deep learning. The workshop calls for researchers to submit the following types of papers: 1) technical papers that introduce their accessible and usable datasets for data-driven model development, 2) position papers outlining ethics and early concepts for creating standardized, sharable datasets, 3) research papers integrating data-driven models into XR research, and 4) short survey with insights and critiques that can trigger discussion among the researchers in the VR community.

REFERENCE:

[1] J. Kim, W. Kim, H. Oh, S. Lee, and S. Lee, "A Deep Cybersickness Predictor Based on Brain Signal Analysis for Virtual Reality Contents," in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), Oct. 2019, pp. 10579–10588. doi: 10.1109/ICCV.2019.01068.

[2] R. Islam et al., "Automatic Detection and Prediction of Cybersickness Severity using Deep Neural Networks from user's Physiological Signals," Proc. - 2020 IEEE Int. Symp. Mix. Augment. Reality, ISMAR 2020, pp. 400–411, 2020, doi: 10.1109/ISMAR50242.2020.00066.

[3] Y. Wang, J.-R. Chardonnet, and F. Merienne, "VR Sickness Prediction for Navigation in Immersive Virtual Environments using a Deep Long Short Term Memory Model," in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Mar. 2019, pp. 1874–1881. doi: 10.1109/VR.2019.8798213.