

LivePose Portal: Democratize immersive arts by having participants train both themselves and deep-learning models for camera-based group interaction before entering the show !

Victor Rios
vrios@sat.qc.ca

Société des Arts Technologiques
Montréal, Québec, Canada

Emile Ouellet-Delorme
eodelorme@sat.qc.ca
Société des Arts Technologiques
Montréal, Québec, Canada

Valentine Auphan
vauphan@sat.qc.ca

Société des Arts Technologiques
Montréal, Québec, Canada

Emmanuel Durand
edurand@sat.qc.ca
Société des Arts Technologiques
Montréal, Québec, Canada

Bruno Colpron
bcolpron@sat.qc.ca

Société des Arts Technologiques
Montréal, Québec, Canada

Christian Frisson
cfrisson@sat.qc.ca

Société des Arts Technologiques
Montréal, Québec, Canada

The crux of LivePose Portal: a deployable box with depth camera, projector, edge computing device, and open-source software.



(a) Viewers can see their own detection mirrored through the portal.

(b) All audiences can train the model towards inclusive detection.

Figure 1: Scenography of the LivePose Portal.

ABSTRACT

The LivePose Portal is a project that aims to democratize immersive arts by allowing participants to train both themselves and deep-learning models for camera-based group interactivity in a portal before entering the show. The LivePose Portal is a deployable transportable portal designed to further help address artificial intelligence biases issue by providing a more diverse and inclusive

experience for all participants but also empowers participants to engage with immersive art on a more personal level.

CCS CONCEPTS

- Applied computing → Media arts;
- Computing methodologies → Activity recognition and understanding;
- Computer systems organization → Embedded systems.

KEYWORDS

multimedia arts, pose detection, edge computing, telepresence

ACM Reference Format:

Victor Rios, Valentine Auphan, Bruno Colpron, Emile Ouellet-Delorme, Emmanuel Durand, and Christian Frisson. 2023. LivePose Portal: Democratize immersive arts by having participants train both themselves and deep-learning models for camera-based group interaction before entering the show !. In *Proceedings of SIGGRAPH '23 Labs (SIGGRAPH '23)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/XXXXXXX.XXXXXXX>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGGRAPH '23, August 06–10, 2023, Los Angeles, CA

© 2023 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

We developed LivePose [2] to support group interaction in immersive arts, relying on RGB-D cameras (capturing color plus depth streams) and state-of-the-art deep learning techniques that we democratize by distributing pre-packaged software optimized for edge devices and by enabling their inter-operability with standard protocols including OpenSoundControl and WebSocket. We are aware that these deep learning techniques are developed mainly by visible majorities and thus work best exclusively with people that look like them, and that includes us, but that poses serious issues in equity, diversity, inclusion and accessibility. Figure 2 shows variable results in pose detection obtained with two different backends on scenes photographed during a hackathon that we organized on the theme of making immersive arts more accessible.

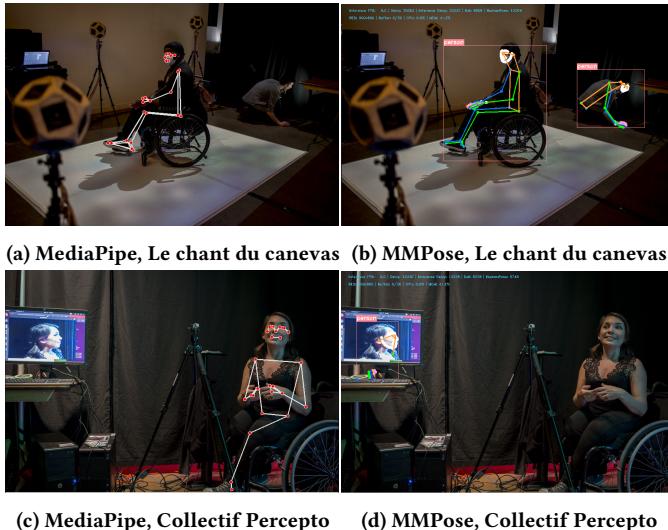


Figure 2: Four examples to show how two pose detection backends (Google MediaPipe and OpenMMLabs' MMPose) integrated in LivePose, applied on two different scenes, can generate variable results.

We have been re-training state-of-the-art deep learning models for pose and action recognition with our own recordings of people who are reproducing a set of actions designed for adding interactivity and bringing scenographies to immersive art installations. During that process, we noticed that small amounts of data can cause many biases to appear. Recognition can be weak on people who "diverge" from the majority and struggles to operate on people with visual features that "diverge" from these of people who are thus statistical "inliers". This observation strengthened our awareness of the issue of this lack of diversity.

We propose LivePose Portal, a deployable transportable portal that democratizes immersive arts even further by having participants train both themselves and the deep-learning models for camera-based group interactivity in a portal before entering the show !

Bias is everyone's responsibility. It hurts those discriminated against, of course, and it also hurts everyone by reducing people's

ability to be part of the technological revolution led by artificial intelligence. That way, we can reduce biases and make the algorithms fairer. The portal is designed to be deployed in various locations and offers a unique experience where participants can actively contribute to the creation of the immersive art they are about to encounter.

2 LIVEPOSE PORTAL

Our solution relies on the following hardware and software components:

- a physical portal prototype with UX designed to make the experience playful and educative;
- NVIDIA Jetson Xavier NX edge devices as cores of the physical portal for portable inferences and transportable deployment;
- LivePose ¹: our open-source tool for democratizing pose detection for multimedia arts and telepresence applications on open edge devices [2], and updated since then to support recording poses and actions to rosbag files (a file format designed for the Robot Operating System (ROS) ²;
- our recent fork OpenMMLab's Next Generation Video Understanding Toolbox and Benchmark MMAAction2 ³: a tool that we have recently learned to own, to train action detection with databases of video movements recorded with LivePose, with our recognition configuration currently relying on the Two-Stream Inflated 3D ConvNet (I3D) [1];
- mpa-focal-arm64-jetson ⁴: our distribution of LivePose and dependencies (including: Intel librealsense and OpenCV for camera support, Google MediaPipe and NVIDIA trt_pose and OpenMMLab mmpose and pytorch for pose detection) packaged and optimized for NVIDIA Jetson edge devices.

3 FUTURE WORK

Our future work includes:

- refining the user experience of people entering the portal: how do we onboard people to record their poses and actions,
- addressing challenges in agency and privacy: how do we properly inform portal participants that they can retract their recordings,
- extending the portal as introduction to virtual experiences: how do we adapt LivePose Portal to be an entry point to metaverse settings.

REFERENCES

- [1] Joao Carreira and Andrew Zisserman. 2017. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI, (July 2017), 4724–4733. ISBN: 978-1-5386-0457-1. doi: 10.1109/CVPR.2017.502.
- [2] Christian Frisson, Gabriel N. Downs, Marie-Ève Dumas, Farzaneh Askari, and Emmanuel Durand. 2022. LivePose: Democratizing Pose Detection for Multimedia Arts and Telepresence Applications on Open Edge Devices. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology (VRST '22)*. Association for Computing Machinery, New York, NY, USA, (Nov. 29, 2022), 1–2. ISBN: 978-1-4503-9889-3. doi: 10.1145/3562939.3565660.

¹<https://gitlab.com/sat-mtl/tools/livepose>

²<https://www.ros.org>

³<https://gitlab.com/sat-mtl/tools/forks/mmaction2/>

⁴<https://gitlab.com/sat-mtl/distribution/mpa-focal-arm64-jetson>