

# Continual and federated learning of neuromorphic gesture interfaces for industrial welding robots

M. Neumeier<sup>(1)</sup>, B. Schindler<sup>(2)</sup>, J. Lecomte<sup>(1)</sup>, A. Kloss<sup>(3)</sup>, A. von Arnim<sup>(1)</sup> and F. Röhrbein<sup>(2)</sup>

<sup>(1)</sup> Fortiss GmbH, <sup>(2)</sup> Technische Universität Chemnitz, <sup>(3)</sup> NEURA Robotics GmbH

Human–robot interaction (HRI) is a key enabler of collaborative robotics, spanning presence detection, action recognition, and intuitive user interfaces. Among these, gesture recognition offers significant advantages over touch and voice interfaces in industrial environments characterized by e.g. glove usage and high acoustic noise. However, practical deployment requires low-latency inference, on-device processing for privacy, energy efficiency for mobile platforms, and adaptability to novel gestures and domain shifts.

Conventional approaches based on RGB cameras and deep neural networks face challenges in meeting these constraints [1], while alternative sensing modalities such as radar suffer from limited spatial resolution [2]. Furthermore, existing systems often underutilize temporal dynamics and static background suppression. Deployment is further complicated by MLOps requirements, including continuous monitoring and retraining, which are difficult to realize on resource-constrained embedded systems [3]. Federated learning partially addresses these issues but remains challenging in such settings [4].

This work proposes a neuromorphic gesture interface for robotic systems, leveraging event-based cameras, spiking neural networks (SNNs), and neuromorphic hardware. The approach exploits sparse, low-latency processing with inherent background suppression and high temporal resolution [5]. To enable adaptability, we extend the system with a continual learning head based on continual learning prototypes (CLP) [6], supporting supervised and semi-supervised class-incremental learning, as well as novel federated learning and knowledge consolidation mechanisms.

We evaluate the proposed system on the DVSGesture dataset [7] and a newly recorded dataset under supervised and semi-supervised settings. Deployed on Intel Loihi 2, the system achieves recognition or adaptation of 1 s gesture samples within <5 ms at an energy cost of 4 mJ. Long-term and federated deployment are assessed in a simulated multi-day, multi-user scenario, considering accuracy, precision, recall, update frequency, and memory footprint. Finally, the system is integrated with a robotic arm for gesture-based control in a smart welding use case, demonstrating its practical applicability and potential for industrial adoption.

[1] Z. Hu et al., ICISCAE, 2019.

[2] M. Yellamma, ICTBIG, 2025.

[3] S. Leroux et al., IEEE IPDPS, 2022.

[4] K. Kopparapu et al., PerCom Workshops, 2022.

[5] G. Gallego et al., IEEE TPAMI, 2019.

[6] E. Hajizada et al., IROS, 2024.

[7] A. Amir et al., CVPR, 2017.