

# Ultrasound-based Recognition of Finger Gestures using Spiking Neural Networks equipped with Spike-Timing-Dependent Plasticity

Antonios Lykourinas <sup>\*</sup>, Chinmay Pendse <sup>‡</sup>, Francky Catthoor <sup>†</sup>, Veronique Rochus <sup>‡</sup>,  
Xavier Rottenberg <sup>‡</sup> and Athanassios Skodras <sup>\*</sup>

<sup>\*</sup>Department of ECE, Faculty of Engineering, University of Patras, Patras, Greece,

<sup>†</sup>Department of ECE, National Technological University of Athens, Athens, Greece

<sup>‡</sup>IMEC, Leuven, Belgium

e-mail: <sup>\*</sup> alykourinas@ac.upatras.gr

**Abstract**—In recent years, researchers constantly attempt to derive Ultrasound-based (US-based) hand gesture recognition (HGR) solutions suitable for edge applications. This process involves improving several design aspects of US-based HGR systems such as the transducers, the wearable US acquisition systems and the algorithms employed in terms of energy consumption, computational complexity and robustness. The subject of this paper is the latter. In this paper, we present a spiking framework for US-based HGR. The proposed approach leverages a single-layer Spiking Neural Network (SNN) equipped with Spike-Timing-Dependent Plasticity (STDP) as a feature descriptor for Rate-based (RB) coded A-line US signals coupled with a lightweight linear support vector machine (SVM) classifier. According to our findings, our proposed approach achieves performance comparable to that of the state-of-the-art on the ultrasound-based adaptive prosthetic control (Ultra-Pro) dataset. Furthermore, we demonstrate that our feature descriptor exhibits inter-session generalization properties, i.e. re-training is not required between within-day sessions. The energy efficiency, robustness and low computational complexity benefits of our approach as well as the possibility of online on-chip learning offered by STDP-enabled neuromorphic hardware add significance to our results.

## I. INTRODUCTION

US has gained popularity as an alternative modality for prosthesis control (PC) and HGR due to its desirable characteristics. It is a non-invasive technique that enables the monitoring of musculoskeletal structures within the human body with high spatio-temporal resolution and can be employed in different modalities, namely, A-mode, B-mode, and M-mode US [1]. Recent works targeting PC and HGR applications, have demonstrated the superiority of US-based Human Machine Interfaces (HMIs) over their popular surface-Electromyography (sEMG) counterparts [2], [3], [4]. Moreover, US complements the electrophysiological nature of sEMG signals, offering performance improvements when fused, paving the way for hybrid HMIs [3], [5], [6]. Advances in integrated pMUT sensors makes that even more promising [7].

In recent years, research focus has shifted towards simpler A-mode US, as it has been demonstrated that a small subset of scanlines of a linear array (4 vs 128 *transducer elements*) can be used for HGR without hindering the HMI's performance [8]. The benefits arising from using a subset of scanlines, such as reduced computational complexity and power consumption, as well as advancements in the miniaturization of

the instrumentation have led to the proposal of novel US acquisition systems with the ultimate goal of deriving truly wearable US-based HGR and PC solutions. Notable examples include: WMAUS [9], WULPUS [10], a portable 24-channel US device operating under a *multiple-receiver approach* [4] and a distributed US configuration targeting 2-Degrees of Freedom (DoF) proportional control applications [11].

Despite the tremendous success of the aforementioned US acquisition systems in several applications, including simultaneous finger gesture recognition and wrist angle estimation [12], a *semi-supervised framework* for 2-DoF virtual prosthesis operation [13] (WMAUS), and regressing 3-DoFs hand kinematics in an online setting [4] (24-channel device), they can not be considered truly wearable. Moreover, the aforementioned solutions leverage either power-hungry algorithms [12] or handle the large dimensionality of US data using dimensionality reduction algorithms [4] hindering the integration of the applications directly on the edge. In an effort to surpass the aforementioned issues, *Vostrikov et al.* [14] integrated a 1D convolutional auto-encoder (CAE) into the microcontroller (MCU) of WULPUS [10], a novel wearable US acquisition device, for feature extraction from raw US signals achieving a compression rate of  $\times 20$ . The feature extractor was coupled with an XG Boost (XGB) classifier, successfully recognizing 4 different hand gestures with an accuracy of 96%. Their feature extractor had only 26 trainable parameters and demonstrated inter-subject generalization properties requiring only the XGB classifier to be re-trained on 1-D CAE-encoded features for each new subject without compromising performance.

Following the work of [14], we believe that a successful integration of US HGR applications on the edge researchers should focus on the: *i)* design of wearable US systems that meet the requirements of *compact, lightweight, re-configurable, ultra-low power* with the ability of *wireless data transmission* [15], *ii)* design of efficient and non-computationally intensive unsupervised US-data feature extraction algorithms, characterized by the property of performing subject-independent operations and *iii)* coupling the system with a *lightweight* classifier for predicting/regressing gestures/hand kinematics. Our main motivation for the work presented in this paper is that we believe that both design aspects *i)* and *ii)* could be improved by drawing inspiration from neuromorphic computing (NC).

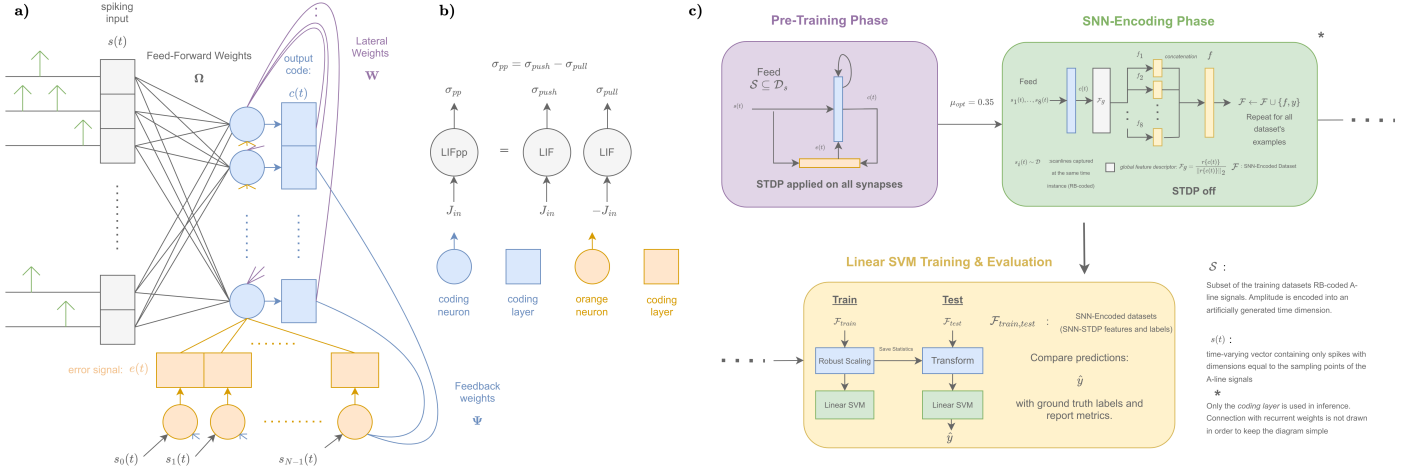


Fig. 1. The SNN-STDP topology that acts as spiking substrate for solving the joint DLBP problem through a) its specific wiring and b) use of LIF push-pull neurons that enable us to approximate LASSO regression. c) The spiking framework for US-based HGR, where the SNN-STDP topology is incorporated as a feature descriptor for RB-coded A-line signals

NC attempts to imitate computational principles of living organisms in order to improve the energy efficiency of computing systems. In recent years, NC has found enormous success in applications, including simultaneous localization and mapping (SLAM) and people detection using drones [16], where both computational and power resources are limited. Neuromorphic applications leverage neuromorphic sensors that mimic biological sensors which transcribe signal changes to *asynchronous ON/OFF events* instead of performing periodic sampling at pre-defined time intervals. Their operation offers several advantages compared to conventional *frame-based* sensors. For example, Dynamic Vision Sensor (DVS) cameras, can achieve higher temporal resolution, Dynamic Range (DR), Signal-to-Noise Ratio (SNR), lower power consumption and hardware-based lossless compression [17].

For interpreting the event streams generated by neuromorphic sensors, a class of algorithms called neuromorphic algorithms have been developed. These often involve hand-crafted online feature descriptors for event-based streams [18] and SNNs. SNNs operate directly on spike trains (binary activations) instead of floating point values processed by artificial neural networks (ANNs), rendering them both hardware friendly and energy efficient [19]. Furthermore, advancements have been made in SNNs trained using the biologically plausible STDP rule which updates the synaptic weights of each neuron *locally* based on the relative timing between their *pre-synaptic* and *post-synaptic* spikes. It is believed that studying SNNs equipped with STDP will enable online on-chip learning and a generation of ultra-low-power and -area neuromorphic systems [20].

In this work, we propose a framework, equipped with a SNN-STDP feature descriptor [16] coupled with a lightweight linear SVM classifier for US-based HGR classification. According to our findings, the SNN-STDP feature descriptor demonstrates inter-session generalization capabilities without the need for re-training for within-day sessions and satisfies both design aspects *ii*) and *iii*), paving the way for US-based HGR on the edge. Our main contributions can be summarized as follows:

- 1) We introduce a SNN-STDP network described under the theory of joint dictionary learning and basis pursuit (DLBP) as a feature descriptor for RB coded A-line signals, accompanied with guidelines for optimizing its performance. The proposed feature descriptor when coupled with a *lightweight* linear SVM classifier can achieve comparable performance with the state-of-the-art.
- 2) We show that the proposed framework for US-based HGR demonstrates inter-session generalization properties, i.e. a pre-trained SNN-STDP network can be effectively used as a feature descriptor for RB coded A-line signals in subsequent within-day sessions of the same subject without hindering the HMI's performance.

The rest of the paper is organized as follows. In Section II, we introduce the SNN-STDP topology and a STDP kernel design methodology. In Section III, we describe our experimental setup, which includes a dataset description and our spiking framework for US-based HGR. In Section IV, we present our results in US-based HGR recognition experiments. Finally, in Section V, we conclude our work.

## II. NOVEL METHODS

### A. SNN-STDP topology

In this section, we will describe the SNN-STDP topology used as a feature descriptor for US-based HGR. The SNN-STDP topology used in this work introduces the use of push-pull Leaky Integrate & Fire (LIFpp) neuron pairs and a specific wiring configuration on its *neural layers*, enabling it to act as a spiking substrate for solving the joint DLBP problem via LASSO coding [16] (See Fig. 1),

$$C, \Phi = \arg \min_{C, \Phi} \sum_{k=1}^K \frac{1}{2} \|\Phi \bar{c}_k - \bar{s}_k\|_2^2 + \lambda_1 \|c_k\|_1 \quad (1)$$

in which our goal is to jointly *i*) learn a dictionary  $\Phi$  of size  $N \times M$  and *ii*) infer output codes  $\bar{c}_k$  of size  $M$  from a dataset

of  $N$ -dimensional vectors  $\bar{s}_k, \forall k$  (where  $C = [\bar{c}_1, \bar{c}_2, \dots, \bar{c}_k]$  contains all output vectors  $\bar{c}_k$ ) and  $\lambda_1$  is a sparsity-defining hyperparameter. This problem is typically solved by alternating between performing a *proximal distant step to the  $l_1$  norm* to infer output codes  $c_k$  and a *stochastic gradient step* (SGD) for learning of  $\Phi$ .

The SNN-STDP topology iterates over the time-domain (using a simulation timestep  $dt$ ) of the spike encoded data for solving both *i*) and *ii*) and consists of two distinct *neural layers*, a *coding layer* and an *error layer* (See Fig. 1 a). Each layer is formed by an ensemble of LIFpp neuron pairs, which are composed of two LIF neurons that share the same *synaptic weights* but receive the input current  $\bar{J}_{in}$  with opposite polarity ( $\bar{J}_{in}$  and  $-\bar{J}_{in}$  for push and pull neuron, respectively). The LIFpp neurons' *synaptic weights* are learned locally via *STDP*. The use of LIFpp neurons enables the network to approximate LASSO regression (See Fig. 1 b). The *coding layer* possesses two sets of weights, namely *feed-forward weights*  $\Omega$  and *recurrent weights*  $W$ , whereas the *error layer* possess a set of feedback weights  $\Psi$  (See Fig. 1 a). Wiring the input  $\bar{s}$  to the neurons in the *coding layer* as follows:

$$\bar{J}_{in}^c = \mathcal{PSC}\{\eta_1 \Omega \bar{s} - W \bar{c}\} \quad (2)$$

where  $\mathcal{PSC}\{\cdot\}$  denotes the *post-synaptic current* kernel used to estimate the local spiking rate and  $\eta_1$  denotes the coding rate, solves the *LASSO* part *ii*) of eq. 1. Wiring the input to the neurons in the *coding layer* as:

$$\bar{J}_{in}^e = \mathcal{PSC}\{\Psi \bar{c} - \bar{s}\} \quad (3)$$

and wiring back their output  $\bar{e}$  to the *coding layer* through weights  $\Omega$  solves the *dictionary learning* part *i*) of eq. 1 when STDP is applied to all synapses. Under the assumption that spikes are Poisson distributed, STDP with homeostatic regularization, performed by configuring appropriately in each layer the topology's *internal signals*  $\langle s(t), c(t), f(t), e(t) \rangle$  as *pre-synaptic* and *post-synaptic* spike trains, is used as a spiking substitute for the SGD step *i*) performed for solving the dictionary learning problem, allowing weight matrices  $\Omega$ ,  $W$  and  $\Psi$  to converge to  $\Phi^T$ ,  $\eta_1 \Phi^T \Phi - I_M$  and  $\Phi$ , respectively during training [21].

### B. STDP kernel Design Methodology

According to [16], the ability to learn crisp atoms is directly linked with STDP kernel's ability to *match* the important frequency band of the spike distribution of the dataset. Since this information cannot be known *a priori*, we developed a method for designing different STDP kernels, based on the double exponential rule, that exhibit the frequency response of a *bandpass filter* (BPF).

The location of the poles and the zeros of an STDP kernel based on the double exponential rule are given by the following equations [16]:

$$p_+ = \tau_+^{-1} \quad (4)$$

$$p_- = \tau_-^{-1} \quad (5)$$

$$z = \frac{\tau_-^{-1} \alpha \tau_+^{-1}}{1 + \alpha} \quad (6)$$

where  $\tau_{-,+}$  are the depression and potentiation time constants and  $\alpha = \frac{A_-}{A_+}$  with  $A_-$  and  $A_+$  being the depression and

potentiation weight, respectively. As can be seen in eq. 4 and 5, the locations for the poles are determined solely by the potentiation and depression time constants, thus passband regions are also defined by their values. Furthermore, in order for the STDP kernel to exhibit a BPF frequency response, the zero of the kernel should be placed at  $s = 0$  ( $s$  represents  $s$ -plane at Laplace domain). By setting the kernel's potentiation weight equal to  $A_+ = 1$ , we can solve for the depression weight  $A_-$  by setting eq. 6 to zero, which yields:

$$A_- = A_+ \frac{\tau_+}{\tau_-} = \frac{\tau_+}{\tau_-} \quad (7)$$

The proposed STDP kernel design technique allowed us to explore kernels based on the double exponential rule with different passband regions.

### C. Spiking Framework for US-based HGR

In this work, we propose important extensions to the base framework published in [21]. In particular, we equip it with a novel US-oriented SNN-STDP descriptor for RB-coded A-line signals. Our novel feature descriptor produces contextually rich and sparse feature representations for RB-coded A-line signals, allowing us to efficiently tackle the HGR problem with a lightweight linear SVM classifier. The framework operates on multiple phases (See Fig. 1 c) which are extensively described below:

- 1) **Pre-Training Phase:** A subset  $\mathcal{S}$  from the RB-coded A-line signals is selected for training the SNN-STDP topology. Then, they are fed sequentially to the network while STDP is applied to all synapses [21].
- 2) **SNN Dataset Encoding Phase:** The SNN-STDP system is coupled with a global feature descriptor [21] for encoding the RB-coding signals into feature vectors. Features  $f_i$  ( $i = 1, \dots, N_{transducers}$ ) of A-line signals captured at the same time instance are concatenated to form the final feature vector  $\mathbf{f}$ .
- 3) **Linear SVM Training & Evaluation Phase:** A linear SVM classifier is trained on *scaled* SNN-encoded training dataset  $F_{train}$ . The trained linear SVM classifier is then used to make predictions  $\hat{y}$  on new samples drawn from either  $F_{val}$  or  $F_{test}$ .

## III. EXPERIMENTAL SETUP

### A. Ultra-Pro Dataset

For our experiments, we used the publicly available Ultra-Pro dataset [12], which features US RF, sEMG and inertial measurement unit (IMU) sensor data from four subjects (all transradial amputees). Each subject performed 6 finger gestures, namely, Rest (RS), Power Grip (PG), Fine Pinch (FP), Index Point (IP), Tripod Grip (TG) and Key Grip (KG) with concurrent wrist rotation of an approximate frequency of  $0.5Hz$ , in three within-day sessions. In each session, the subjects performed the gestures sequentially, each gesture was held for 50 seconds while the subjects performed concurrent wrist rotation. The US RF signals were captured using a customized wearable armband featuring 8 evenly spaced transducers, operating at a frequency of  $5MHz$ , driven by a customized US device [9]. Data streams from all sensors were synchronized using specialized software.

TABLE II. COMPARISON OF FINGER GESTURE RECOGNITION PERFORMANCE IN TERMS OF ACCURACY BETWEEN THE PROPOSED SNN-STDP & LINEAR SVM SYSTEM AND THE STATE-OF-THE-ART CNN ARCHITECTURE. BOLD ENTRIES INDICATE THE BEST PERFORMING MODEL FOR EACH CASE

Subject-Session	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3	4-1	4-2	4-3
Ours	<b>73.80%</b>	<b>79.42%</b>	70.47%	75.47%	74.24%	73.96%	<b>64.73%</b>	<b>58.87%</b>	<b>60.06%</b>	84.64%	79.11%	80.89%
CNN [22]	72.90%	79.09%	<b>73.97%</b>	<b>84.42%</b>	<b>74.50%</b>	<b>81.25%</b>	61.27%	52.80%	55.77%	<b>85.13%</b>	<b>83.70%</b>	<b>84.30%</b>

TABLE III. PERFORMANCE COMPARISON OF A PRE-TRAINED FEATURE DESCRIPTOR (SESSION 1) WITH A FEATURE DESCRIPTOR TRAINED FROM SCRATCH ON SUBSEQUENT SESSIONS OF EACH SUBJECT OF THE ULTRA-PRO DATASET. BOLD ENTRIES INDICATE THE BEST PERFORMING CASE

Subject-Session	1-2	1-3	2-2	2-3	3-2	3-3	4-2	4-3
Pre-trained	<b>79.89%</b>	70.47%	74.00%	<b>74.87%</b>	58.56%	59.42%	<b>80.00%</b>	<b>81.27%</b>
From scratch	79.42%	70.47%	<b>74.24%</b>	73.96%	<b>58.87%</b>	<b>60.05%</b>	79.11%	80.89%

## B. Preprocessing & Spike Encoding

For preprocessing the US RF signals, we followed the standard procedure of *bandpass filtering*, *time-gain compensation*, *envelope detection* and *log compression* using parameters recommended in [23]. Following [12], the first and the last 20 samples of the A-line signals were discarded. For converting the static A-line signals of the dataset to the event domain, we used RB coding [19], i.e. amplitude information is encoded in an artificially generated time dimension through spike trains by means of *spiking activity*. During training and inference, the RB-coded signals were iterated along time dimension (50 timesteps are used in this work) using a simulation step  $dt$  of  $5ms$ .

## C. Tuning the SNN-STDP Topology

1) *Training & Neural Parameters*: The SNN-STDP topology’s *neural* and *training parameters* were optimized by monitoring the inner loss  $\|r\{e(t)\}\|_2$ , a direct measure of the re-projection quality of the model [16]. Furthermore, the optimal neuron threshold  $\mu_{opt}$  was discovered via grid search (values ranging from 0.001 to 0.7). The network’s neural and training parameters used in this work are depicted in Table I.

TABLE I. NEURAL AND TRAINING PARAMETERS FOR THE SNN-STDP TOPOLOGY

$n_c n_d$	$\mu$	$\tau_s$	$M$	$ S $	$\mu_{opt}$
1 0.001	0.01V	10ms	4096	384 scanlines	0.35

2) *STDP Parameters*: We used the STDP kernel design methodology presented in Section II for constructing STDP kernels with different passband regions. In order to determine the appropriate time constants for the kernel, we used both *random search* [24] and *tree-structured parzen estimators* [25]. We discovered that, for our application, the optimal configuration for the parameters of the STDP double exponential rule was  $A_+ = 1$ ,  $A_- = 1.778$ ,  $\tau_+ = 3.2ms$ ,  $\tau_- = 1.8ms$ .

## IV. RESULTS

### A. Classification Results

In this section, we will assess the performance of the spiking framework across different subjects-sessions of the Ultra-Pro dataset. To allow for a fair comparison, we followed the train-test-validation split proposed in the original paper [12]. Each experiment was repeated 10 times using different *random seeds* and performance was assessed by means of classification accuracy (averaged across all 10 experiments). The performance comparison with the state-of-the-art [22] is depicted in Table II.

The proposed framework provides on-par results compared with the state-of-the-art and it even performs better in 5/12 cases. Furthermore, there are several advantages of using the proposed spiking framework over other approaches found in literature [12], [22]. These include: *i)* the fact that SNN-STDP feature extractor is a *single-layer* network, *ii)* that its synapse weights are updated *locally* using the biologically plausible STDP learning rule instead of the computationally intensive backpropagation of error algorithm and *iii)* that it operates on spike trains (binary activations) instead of floating point values. Thus, our spiking framework satisfies both design aspects *ii)* and *iii)* rendering it appropriate for US-based HGR applications on the edge.

### B. Inter-session Classification Results

In these experiments, we used the pre-trained SNN-STDP topology on Session 1 for each subject’s subsequent sessions. The results are reported in Table III.

We observed that the SNN-STDP feature descriptor demonstrated inter-session generalization properties, i.e. a pre-trained SNN-STDP feature descriptor could be directly used to encode RB-coded A-line signals on subsequent subjects’ sessions. This shows a clear promise to reduce the additional training/online learning effort to move between sessions. Our results demonstrate the ability of the SNN-STDP feature descriptor to capture both rich and invariant RB-coded A-line signal representations, an often significant challenge to overcome in biomedical applications [26], [27], while using a minimal training examples. In future work, we want to explore the potential of leaving that even fully unsupervised.

## V. CONCLUSION

In this paper, we present a spiking framework for US-based HGR that leverages a *single-layer* biologically plausible SNN-STDP feature descriptor and a lightweight SVM classifier. The proposed approach achieves performance comparable to that of the state-of-the-art (our spiking framework performs better in 5/12 cases) on the publicly available Ultra-Pro dataset. Furthermore, we have reported inter-session experiment results that highlight the robustness of the SNN-STDP feature descriptor to the variability of the US data, while trained on only a few A-line signals.

## ACKNOWLEDGMENT

This research was funded by IMEC (Belgium) under the contract ORD-372666-C3B9V.

## REFERENCES

- [1] X. Yang, C. Castellini, D. Farina, and H. Liu, "Ultrasound as a Neurorobotic Interface: A Review," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2024.
- [2] J.-Y. Guo, Y.-P. Zheng, Q.-H. Huang, X. Chen, J.-F. He, and H. Lai-Wa Chan, "Performances of One-Dimensional Sonomyography and Surface Electromyography in Tracking Guided Patterns of Wrist Extension," *Ultrasound in Medicine & Biology*, vol. 35, no. 6, pp. 894–902, 2009. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0301562908005826>
- [3] X. Yang, J. Yan, and H. Liu, "Comparative Analysis of Wearable A-Mode Ultrasound and sEMG for Muscle-Computer Interface," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 9, pp. 2434–2442, 2020.
- [4] B. G. Sgambato, M. H. Hasbani, D. Y. Barsakcioglu, J. Ibáñez, A. Jakob, M. Fournelle, M.-X. Tang, and D. Farina, "High Performance Wearable Ultrasound as a Human-Machine Interface for wrist and hand kinematic tracking," *IEEE Transactions on Biomedical Engineering*, pp. 1–10, 2023.
- [5] J. Zeng, Y. Zhou, Y. Yang, J. Wang, and H. Liu, "Feature Fusion of sEMG and Ultrasound Signals in Hand Gesture Recognition," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, pp. 3911–3916.
- [6] S. Wei, Y. Zhang, and H. Liu, "A Multimodal Multilevel Converged Attention Network for Hand Gesture Recognition with hybrid sEMG and A-Mode Ultrasound Sensing," *IEEE Transactions on Cybernetics*, vol. 53, no. 12, pp. 7723–7734, 2023.
- [7] M. Pandit, J. Aerts, D. Barbosa, M. Forlin, P. Deruytere, E. Georgitzikis, T. Kinoshita, V. Rochus, X. Rottenberg, and E. Hijzen, "Developing a phased pmut array patch for cardiac health monitoring," in *2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium (UFFC-JS)*, 2024, pp. 1–4.
- [8] N. Akhlaghi, A. Dhawan, A. A. Khan, B. Mukherjee, G. Diao, C. Truong, and S. Sikdar, "Sparsity Analysis of a Sonomyographic Muscle-Computer Interface," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 688–696, 2020.
- [9] X. Yang, Z. Chen, N. Hettiarachchi, J. Yan, and H. Liu, "A Wearable Ultrasound System for Sensing Muscular Morphological Deformations," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 6, pp. 3370–3379, 2021.
- [10] S. Frey, S. Vostrikov, L. Benini, and A. Cossettini, "WULPUS: a Wearable Ultra Low-Power Ultrasound probe for multi-day monitoring of carotid artery and muscle activity," in *2022 IEEE International Ultrasonics Symposium (IUS)*, 2022, pp. 1–4.
- [11] B. G. Sgambato, H. Hakami, X. Yang, D. Y. Barsakcioglu, A. Jakob, M. Fournelle, A. H. McGregor, M.-X. Tang, and D. Farina, "Towards Natural Multi-Dof Prosthetic Control with Distributed Ultrasound," in *2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium (UFFC-JS)*, 2024, pp. 1–6.
- [12] X. Yang, Y. Liu, Z. Yin, P. Wang, P. Deng, Z. Zhao, and H. Liu, "Simultaneous Prediction of Wrist and Hand Motions via Wearable Ultrasound Sensing for Natural Control of Hand Prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2517–2527, 2022.
- [13] X. Yang, J. Yan, Z. Yin, and H. Liu, "Sonomyographic prosthetic interaction: Online simultaneous and proportional control of wrist and hand motions using semisupervised learning," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 2, pp. 804–813, 2023.
- [14] S. Vostrikov, M. Anderegg, L. Benini, and A. Cossettini, "Unsupervised Feature Extraction From Raw Data for Gesture Recognition With Wearable Ultralow-Power Ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 7, pp. 831–841, 2024.
- [15] S. Vostrikov, J. Tille, L. Benini, and A. Cossettini, "TinyProbe: A Wearable 32-Channel Multimodal Wireless Ultrasound Probe," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 72, no. 1, pp. 64–76, 2025.
- [16] A. Safa, L. Keuninckx, G. Gielen, and F. Catthoor, *Neuromorphic Solutions for Sensor Fusion and Continual Learning Systems: Applications in Drone Navigation and Radar Sensing*, 1st ed. Springer Cham, Jul. 2024.
- [17] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 db Dynamic Range Frame-Free PWM Image Sensor with Lossless Pixel-Level Video Compression and Time-Domain CDS," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 259–275, 2011.
- [18] H. Ding, J. Jiang, and R. Yan, "A Time-Surface Enhancement Model for Event-based Spatiotemporal Feature Extraction," in *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024, pp. 1–7.
- [19] J. K. Eshraghian, M. Ward, E. O. Neftci, X. Wang, G. Lenz, G. Dwivedi, M. Bennamoun, D. S. Jeong, and W. D. Lu, "Training Spiking Neural Networks Using Lessons From Deep Learning," *Proceedings of the IEEE*, vol. 111, no. 9, pp. 1016–1054, 2023.
- [20] A. Vigneron and J. Martinet, "A critical survey of STDP in Spiking Neural Networks for Pattern Recognition," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–9.
- [21] A. Safa, I. Ocket, A. Bourdoux, H. Sahli, F. Catthoor, and G. G. Gielen, "Event Camera Data Classification Using Spiking Networks with Spike-Timing-Dependent Plasticity," in *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022, pp. 1–8.
- [22] A. Lykourinas, X. Rottenberg, F. Catthoor, and A. Skodras, "Unsupervised Domain Adaptation for Inter-Session Re-Calibration of Ultrasound-Based HMIs," *Sensors*, vol. 24, no. 15, 2024. [Online]. Available: <https://www.mdpi.com/1424-8220/24/15/5043>
- [23] X. Yang, Y. Zhou, and H. Liu, "Wearable Ultrasound-Based Decoding of Simultaneous Wrist/Hand Kinematics," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 9, pp. 8667–8675, 2021.
- [24] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. null, p. 281–305, Feb. 2012.
- [25] S. Watanabe, "Tree-Structured Parzen Estimator: Understanding Its Algorithm Components and Their Roles for Better Empirical Performance," 2023. [Online]. Available: <https://arxiv.org/abs/2304.11127>
- [26] U. Côté-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Laviolette, and B. Gosselin, "Deep Learning for Electromyographic Hand Gesture Signal Classification using Transfer Learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 760–771, 2019.
- [27] E. Rahimian, S. Zabihi, A. Asif, D. Farina, S. F. Atashzar, and A. Mohammadi, "FS-HGR: Few-Shot Learning for Hand Gesture Recognition via Electromyography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1004–1015, 2021.