AdaptiveEdge: Adaptive Model Update for Motor-Intent Decoding With Knowledge Distillation and Efficient EMG Sensor System

Mustapha Deji Dere, Giwon Ku, Ji-Hun Jo, Saehyung Cheong, Sarfraz Ali, and Boreom Lee, *Member, IEEE*

***Abstract*— Recent advancements in electromyogram (EMG)-based gesture decoding have enabled the development of active rehabilitation devices and enhanced human-machine interaction capabilities. While production-grade EMG sensors offer improved signal-to-noise ratios, their technical complexity necessitate innovative solutions to address inherent limitations. Additionally, EMG-based motor-intent decoders are prone to performance degradation due to factors such as fatigue, electrode shifts, and varying acquisition conditions. To address these challenges, we propose a low-cost EMG sensor grid alongside an advanced decoding strategy named AdaptiveEdge. This adaptive model update strategy integrates offline training with real-time on-device parameter updates, facilitating seamless adaptation to diverse EMG disturbance scenarios. Our comprehensive experiments demonstrated significant accuracy improvements: AdaptiveEdge yielded 10.18% higher accuracy (88.66%) when both offline and on-device update were utilized compared to 78.48% without offline training. Furthermore, AdaptiveEdge not only enhances decoding accuracy but also optimizes memory usage and energy consumption, making it particularly suitable for on-device applications such as neuroprosthetics. These advancements collectively pave the way for more effective and practical EMG-based devices, thereby improving human-machine interaction capabilities. The code associated with this study can be accessed here: https://github.com/deremustapha/AdpativeEdge**

Received 21 July 2025; revised 30 September 2025; accepted 12 October 2025. Date of publication 16 October 2025; date of current version 24 October 2025. This work was supported by the National Research Foundation of Korea (NRF) grant funded by Korean Government (MSIT), under Grant RS-2023-00277220. *(Corresponding author: Boreom Lee.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by The Gwangju Institute of Science and Technology Investigation Review Board's under Application No. 20240215-HR-75-03-02.

Mustapha Deji Dere, Ji-Hun Jo, Saehyung Cheong, and Sarfraz Ali are with the Department of Biomedical Science and Engineering, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea.

Giwon Ku is with the AI Convergence Home Appliance Research Center, Korea Electronics Technology Institute (KETI), Gwangju 61011, Republic of Korea, and also with the Department of Biomedical Science and Engineering, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea.

Boreom Lee is with the Department of Biomedical Science and Engineering and the Department of AI Convergence, Gwangju Institute of Science and Technology, Gwangju 61005, South Korea (e-mail: leebr@gist.ac.kr).

Digital Object Identifier 10.1109/TNSRE.2025.3622132

***Index Terms*— Adaptive, edge, electromyography (EMG), hand gesture, human-machine interface (HMI).**

I. Introduction

ELECTROMYOGRAM (EMG) signals, which can be measured intramuscularly or from the skin's surface as surface EMG (sEMG) signals, provide valuable insights into motor information encoded by the central nervous system (CNS). The global EMG is a summation of motor unit action potentials that offer a glimpse into neural code transmission through the peripheral nervous system (PNS). Decoding sensorimotor neural codes from sEMG is crucial for developing next-generation human-computer interfaces (HCI), as demonstrated by recent advancements such as virtual keyboard typing [1], [2], bionic limb control [3], [4], neuro-rehabilitation [5], [6], [7], and biometrics authentication [8]. To develop a realistic active device that leverages CNS and PNS connectivity for control or machine interaction, researchers identified five key characteristics [9]: 1) a simple interface, 2) minimal calibration requirements, 3) high classification accuracy, 4) population sampling, and 5) long-term stability. However, low spatial selectivity and high volume conduction issues in sEMG-based approach could lead to failures in points 2), 3), and 5).

sEMG-based applications are frequently impeded by distribution shifts arising from individual anatomical differences, electrode-shifts, and fatigue [10], [11]. These factors substantially contribute to model collapse across trials, sessions, and long-term wear, thereby restricting their applications to research laboratories [12]. EMG decomposition [13], [14], domain-specific model architecture designs [15], [16], [17], [18], [19], [20], data engineering [21], [22], [23], and on-device continual learning [24], [25], [26] have been proposed to mitigate model collapse due to EMG disturbances. The emergence of on-device continual learning with EMG decoders represents a promising approach for mitigating data distribution shifts caused by EMG disturbances [2], [12], [24], [25]. This methodology stands out due to its ability to facilitate lifelong learning, a feature that is absent in other competing methods. Furthermore, the concept of on-device training as a solution to distribution shift remains a critical research topic within deep learning literature [2], [27].

The use of deep neural networks (DNNs) has shown high performance in decoding upper-limb motor-intent from raw

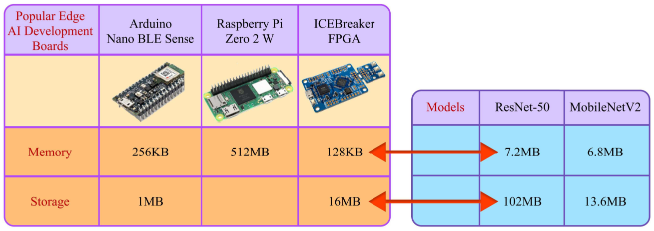

Fig. 1. A comparison of peak memory and storage requirements for deep learning models and resources on popular edge devices.

EMG data [28], [29]. Conventional decoding algorithms which includes convolution neural networks (CNNs) [28], [30], [31], [32], recurrent-neural networks (RNNs) [33], [34], transformers [2], [35], vision transformers [36], [37], [38], [39], and hybrid architectures [1], [40], have been proposed for upper-limb motor-intent decoding based on sEMG. However, a significant drawback of DNN for real-time gesture edge inference is the memory and computational resource constraints they impose. Specifically, the memory and computation requirements for popular DNN far exceed those available on popular extreme edge devices (Fig. 1), fundamentally limiting their deployment on resource-constrained devices. This limitation necessitates careful consideration of architecture design [41] and on-device training methods to ensure they meet computational demands without compromising performance or introducing issues such as catastrophic forgetting or model collapse. While full-layer fine-tuning avoids model collapse, it comes at the cost of significant memory and computational overhead due to the need to store intermediate activations. On the other hand, last-layer fine-tuning, while more computationally efficient, often leads to model collapse. To address these challenges, several approaches have been proposed, including sparsely updating specific layers or kernels as a trade-off between computational efficiency and accuracy [27], [42], [43], [44].

In our study, we present three key contributions to the field of EMG measurement and gesture decoding, addressing critical limitations in current systems.

1) We introduce FlexAdapt, a sensor system comprising a flexible, cost-effective EMG sensor grid fabricated using readily accessible materials with a total development cost below 300 USD. This design directly mitigates the constraints associated with bulky and expensive commercial sensors, facilitating broader accessibility and practical implementation.
2) We have developed a comprehensive dataset specifically designed to evaluate and benchmark EMG decoding algorithms under realistic conditions. This dataset incorporates critical sources of EMG disturbance – including variations due to electrode shifts, fatigue, donning/doffing procedures, and adaptation to new participants – providing a standardized resource for comparative analysis.
3) We propose AdaptiveEdge, a novel model update strategy, designed to mitigate the degradation typically observed in EMG signals arising from these disturbances. AdaptiveEdge's dynamic update process during real-world usage scenarios represents a significant advance towards robust and reliable human-machine interaction through EMG control.

The remainder of this paper is organized as follows. Section II details the wearable EMG system fabrication, data measurement, characterization, and processing techniques. Section III presents the experimental findings. Section IV discusses the results alongside potential future research directions. Finally, Section V concludes with a summary of the key contributions of this study.

II. Methods

A. Flexible EMG Fabrication

In this study, we designed and fabricated a flexible electrode grid we named FlexAdapt, evaluating three key criteria: 1) low cost ($<$ 300 USD), 2) competitive signal-to-noise ratio (SNR), and 3) a form factor under 100 g. The FlexAdapt overall dimensions measuring (115 mm $\times$ 60 mm), was designed by creating a schematic illustration (Fig. 2(a)) using Inkscape software. The design features a circular electrode with a diameter of 20 mm and an inter-electrode distance of 25 mm, partially adhering to recommendations from the SENIAM report [46]. The illustration was printed on clear A4 paper using Samsung printer (SL-C1453FW) and transferred to a transparent film via a black permanent marker.

To fabricate FlexAdapt (Fig. 2(b)), a silk mesh screen printing frame was coated in a dark space which had indirect UV light using photosensitive emulsion (Azo-fix No.1), which was dried for five minutes. The illustration printed on the transparent film was placed on the coated frame and exposed to UV light (395-400 nm, 50 W 220 V) for five minutes while positioned 500 mm away from the lamp. After exposure, a conductive silver (Ag) paste (ELCOAT P-100) was applied to the PET substrate (Goodfellow's ES30-FM-000200) to form the electrode. Gold electrodes were glued onto the Ag to prevent movement and connected to OpenBCI Cyton analog-to-digital converter (ADC) board (Fig. 2(b)). The two 2 $\times$ 4 flexible EMG electrode arrays (8 electrodes each), configured for horizontal bipolar recording, yielding a total of 8 channels, (Fig. 2(c)) were interconnected with the Cyton board (OpenBCI), which operates at a sampling frequency of 250 Hz to ensure accurate data acquisition. The Cyton board was powered by a single lithium-ion battery with a voltage of 3.7 V. This lightweight design allows for easy attachment to any part of the upper limb using a medical adhesive tape. The total weight of the EMG sensor system was measured at 52 g with an overall cost of $<$ 300 USD. The SNR varied across channels and participants, achieving a maximum value of 59.70 dB from participant 1 during session 1 on channel 1. The SNR for FlexAdapt was comparable to that of similar flexible EMG sensors [24], [45] but slightly lower than that of MyoArmband [28] (Table I).

The characterization of the EMG properties was carried out on one participant while conducting a full power grasp and relaxing the muscles intermittently. The amplitude burst

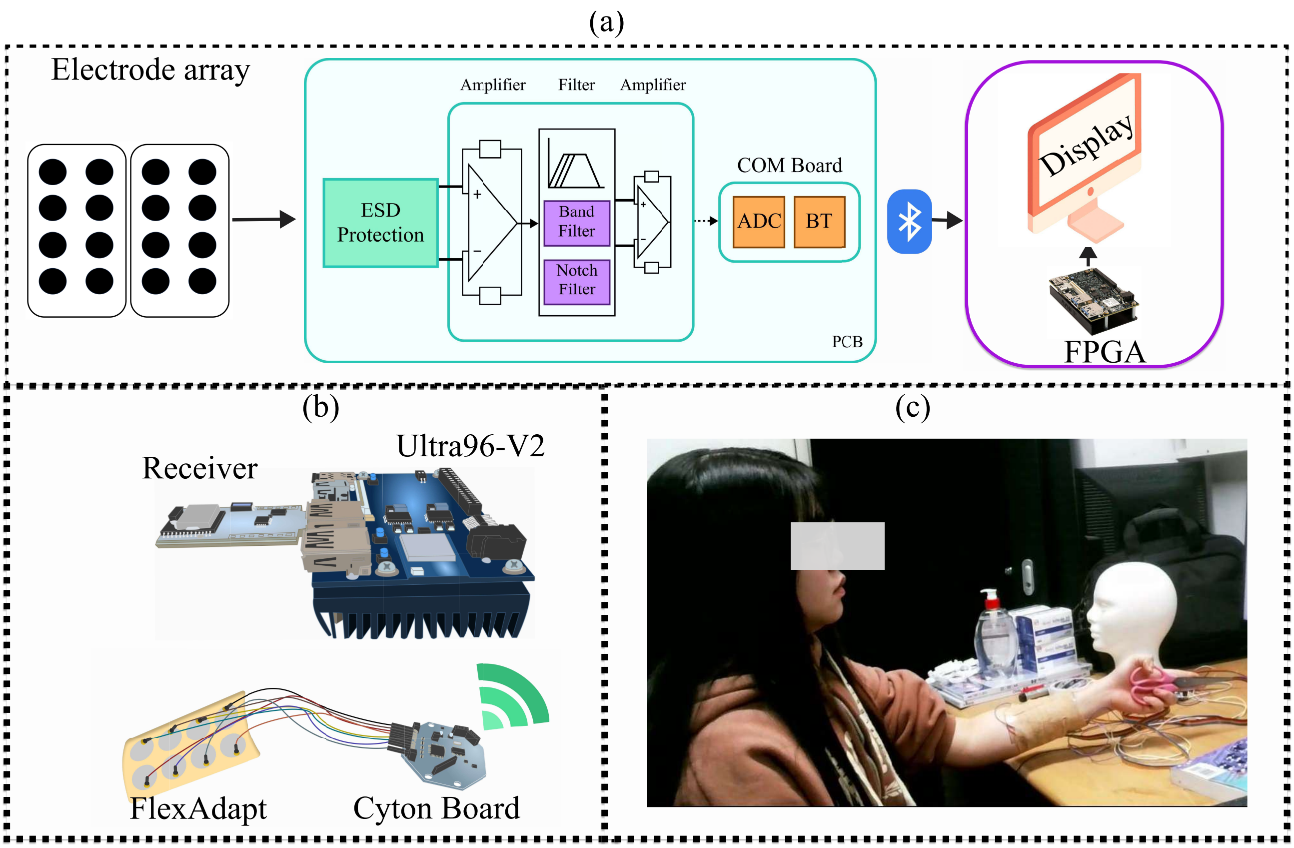

Fig. 2. FlexAdapt design. (a). System flow chart. (b). Illustration of data transmission to the deployment hardware. (c). Demonstration on the participant's executing a gesture.

TABLE I
COMPARISON OF SNR ESTIMATIONS IN DIFFERENT STUDIES

Studies	SNR (dB)	Estimation
[45]	24.07	This was reported in the study. Result from one channel, not the average across all channels.
[24]	0.23	Data from Participant 1 during Trial 1 in Experiment 1, middle finger extension (signal) with rest serving as background noise. The result is an average of all channels.
[28]	28.89	Data from Participant 1 during Trial 1 in Experiment 1, large diameter grasp (signal) with rest serving as background noise. The result is an average of all channels.
FlexAdapt	14.43	The average SNR (calculated across all channels) in Participant 3 under an electrode shift scenario was evaluated.

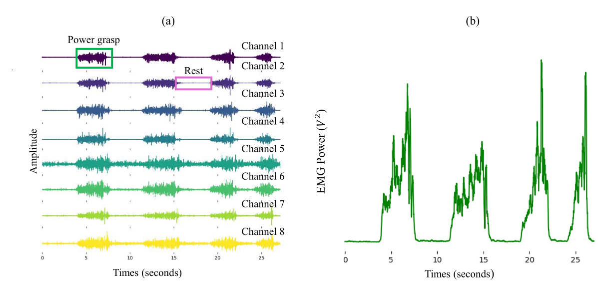

Fig. 3. Overview of raw surface electromyography (sEMG) data acquired from a single participant. (a). sEMG while the participant carried out power grasp and rest gesture intermittently. (b). Activation of the power grasp evaluated by averaging across all eight channels.

corresponding to instances of contraction and a flat signal representing relaxation of the muscle in the sEMG trace (Fig. 3(a)). The sensitivity of the FlexAdapt allowed for the difference across the gesture during dynamic gesture execution (Fig. 3(b)).

B. Overview of the System

1) Background: According to authors in [1], personalizing models through participant-specific fine-tuning significantly improves performance. They demonstrated that pre-trained models (trained on data from 25 participants, totaling 1900 minutes), and thereafter fine-tuned with just 20 minutes of data achieved comparable performance to generalized models trained on 14,000 minutes of data. However, the study raises questions about how to fine-tune models at the edge without sacrificing accuracy in situations where large-scale data collection is not feasible. To address the limitation of on-device model update, TinyTL was proposed to update bias with a residual bottleneck [44], while Tiny-Train [42] demonstrated that meta-learning effectively reduces knowledge forgetting during on-device updates. The concept of enhancing on-device adaptive model updates through meta-learning was also reinforced in [2]. Inspired by these studies, we introduce AdaptiveEdge, a framework designed for edge on-device training, as illustrated in Fig. 4.

The AdaptiveEdge framework (Fig. 4) proposed in this study is composed of two distinct phases: 1) offline model

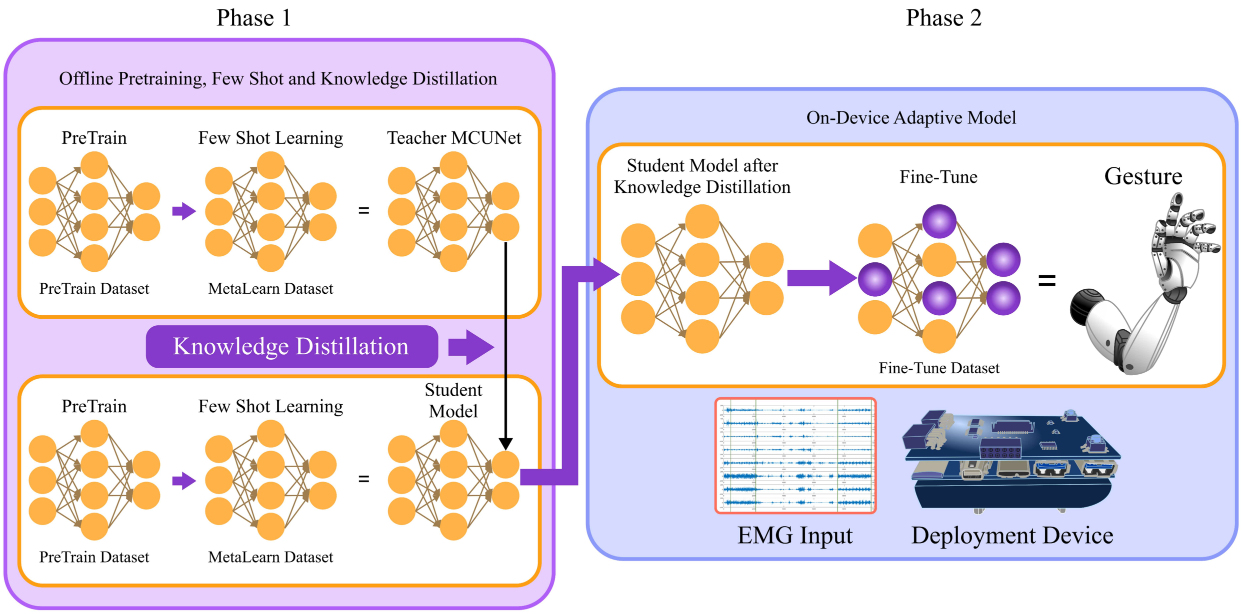

Fig. 4. Overview of the proposed AdaptiveEdge sparse model update.

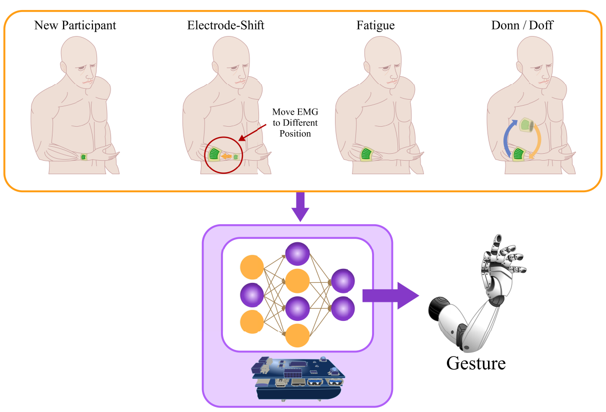

Fig. 5. Description of various EMG disturbances warranting adaptive model update.

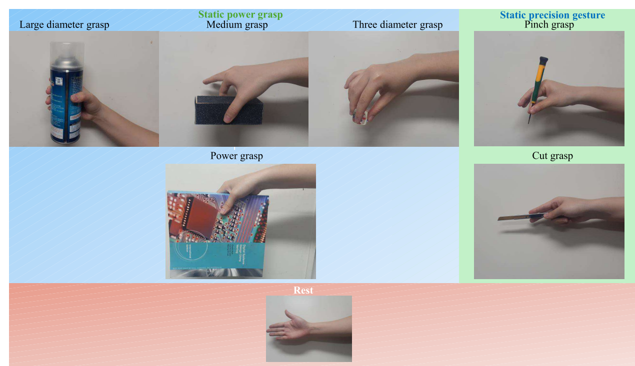

Fig. 6. Gestures acquired with FlexAdapt.

pre-training and 2) on-device model fine-tuning. The offline pre-training phase includes three sequential stages: training the backbone model to establish a foundational understanding of EMG data, implementing few-shot meta-learning to enable rapid adaptation from limited labeled examples, and employing knowledge distillation (KD) to effectively familiarize the backbone model with FlexAdapt data.

Following the first phase, on-device model fine-tuning occurs post-KD (student model is deployed to the edge). Adaptive updating process ensures that the model's weights are adjusted to minimize any degradation in performance during on-device inference, thereby maintaining robust accuracy and reliability. AdaptiveEdge was compared against three conventional approaches: FullTrain, which involves fine-tuning all model layers; LastLayer, which updates only the final layer; and TinyTL-B [44], which focuses on updating biases and last layer. In this study we refer to TinyTL-B as TinyTL.

C. Datasets and Preprocessing

Two open-access datasets ([28] and [30]) along with FlexAdapt data were employed in the offline pre-training (Phase 1). The [28] and [30] datasets each included seven discrete gestures acquired from 40 and 33 participants, respectively. While there are notable gesture differences between [28] and [30], the gestures recorded in this study (Fig. 6) closely align with those reported in [28]. The collection process for [28] and [30] involved an 8-channel, 200 Hz MyoArmband placed along the forearm to capture the necessary signals.

Data acquisition for FlexAdapt was facilitated by custom software outlined in [28]. Participants were recruited from a diverse group of 7 healthy individuals: 4 female participants aged 25.25 $\pm$ 1.08 years and 3 male participants aged 31.66 $\pm$ 3.39 years, with no prior upper limb injuries or neurological conditions reported. The study was conducted at the Gwangju Institute of Science and Technology, with approval from the Investigation Review Board (20240215-HR-75-03-02), following informed consent from all participants. The experimental protocol included four sessions designed to simulate EMG disturbances (Fig. 5), with each session requiring participants to replicate static gestures displayed on a screen (Fig. 6). Participants were provided with instructions before beginning the task. Three participants engaged in electrode-shift and fatigue experiments, while one participant participated exclusively in a donning/doffing experiment. All participants were

TABLE II
EDGE MODELS USED IN THIS STUDY

Model	Parameters	Multiply-Accumulate
MCUNet [50]	577,247	3,950,304
ProxylessNas [51]	2,937,959	7,552,416
MobileNetV2 [52]	2,232,295	5,572,576

considered novel to this study. Each gesture was replicated for nine trials. To simulate electrode shift, FlexAdapt was positioned at random distances (in millimeters) away from its baseline data acquisition location. Fatigue simulation involved having participants perform multiple movements before re-engaging with the experimental setup. Finally, FlexAdapt was removed and placed in a slightly different position to simulate the donning and doffing processes.

Power line interference was filtered using a notch filter with 60-Hz, followed by a high-pass filter at 10-Hz to eliminate artifacts such as tremors from the raw EMG data. Generally, raw neural signals are advantageous for decoding motor intent, but deep learning decoders sometimes benefit from domain-specific feature extraction [29]. The raw sEMG was processed with a sliding window technique that included overlapping segments to maintain temporal resolution [47], [48]. We selected a segmentation window of 200 ms for processing the open-access datasets ([28] and [30]), which operate at a sampling frequency of 200 Hz. This selection aligns with the window size reported in [28], resulting in a segment containing 40 datapoints (200 ms $\times$ 200 Hz). Given our device's higher sampling rate of 250 Hz, we adjusted the segmentation window to 160 ms with an 80% overlap ([24] used an equivalent overlap). This configuration yielded 40 datapoints per segment (160 ms $\times$ 250 Hz). The choice of a 160 ms window was made to ensure alignment between the segmented datapoints in the open-access dataset and those from FlexAdapt. Our implementation achieved a mean inference latency of 0.256 ± 0.02 ms, which falls within the optimal electromechanical delay requirements as suggested in [49].

D. Architectures

The choice of decoder in EMG-based motor-intent decoding has been extensively explored over the past decade, with various DNN-based decoder architectures being proposed. Considering the specialized requirements for edge deployment, we evaluated three distinct edge models (Table II). The input and output dimensions were carefully adjusted to match our study requirements. Specifically setting the input shape to a 3-dimensional vector with dimensions (1, 8, 40). Correspondingly, the decoder was configured to produce an output of size 7, corresponding to the seven motor gestures under investigation in this study.

E. Offline Pre-Training

Offline backbone models (Phase 1: Stage 1 and Stage 2) were trained using open-access datasets that feature same sampling frequencies. For pre-training the backbone models (Phase 1: Stage 1), we used the MyoArmband dataset [30]. We utilized training and evaluation data from all 40 participants (12 females and 28 males) for pre-training the models f_θ.

For few-shot meta-learning (Phase 1: Stage 2), we adopted the prototypical network approach [53]. Given a pre-trained network f_θ, for each episode, given a support set ($S = (x_i, y_i)$) with gesture classes (g), we computed the class prototype as the mean embedding of its support examples: $P_g \leftarrow \frac{1}{N_g} \sum_{(x_i,g_i)\in S_g} f_\theta(x_i)$, where P_g represents the centroid cluster for gesture g, N_g is the number of support samples (x_i) per gesture, S_g denotes the support set for each episode, and i an instance in the support. For a query set, class probabilities were obtained by computing a softmax over negative squared Euclidean distances to the prototypes. The episode loss was evaluated as the negative log-likelihood over the query set. Open-access EMG data [28] from the first 10 subjects were used. A 7-way 5 support-and-query setup with 2000 episodes and epochs ranging from 10 to 15 was employed for few-shot meta-learning.

Using KD [54], we hypothesize that student models trained on data drawn from the deployment distribution (FlexAdapt) will require minimal on-device fine-tuning (Phase 1: Stage 3). We trained the student model with the objective in (1), where y denotes the ground-truth label; p_s is the student's predictive distribution (from logits); and $p_t^{(T)}$, $p_s^{(T)}$ are the teacher and student softened distributions obtained with temperature $T > 0$. $\mathcal{L}_{\text{CE}}$ is the standard cross-entropy with hard labels, and $\mathcal{L}_{\text{KLD}}$ is the Kullback–Leibler divergence between the softened (soft-target) distributions. The factor T^2 compensates for the gradient scaling, while the mixing coefficient $\alpha \in [0, 1]$ balances the contributions of the soft-target and ground-truth label.

$$\mathcal{L}_{\text{KD}} = (1-\alpha)\,\mathcal{L}_{\text{CE}}(y, p_s) + \alpha\, T^2\, \mathcal{L}_{\text{KLD}}\left(p_t^{(T)} \,\|\, p_s^{(T)}\right) \tag{1}$$

F. Adaptive Model Update

We adapted the pre-trained KD model f_θ with parameters $\theta \in \mathbb{R}^d$ under the constraint that only $k \ll d$ coordinates may be updated. We used the aggregated batch gradient (2) for each mini-batch $\mathcal{B} = \{(x_i, y_i)\}_{i=1}^{B}$: where $t = 1, \ldots, T$ indexes the T probe mini-batches used to estimate sensitivity. We then estimate a per-coordinate gradient-variance proxy across mini-batches (3). V_j measures the stability of the batch-aggregated gradient for coordinate θ_j under input variability across mini-batches. Larger V_j indicates that θ_j is more responsive to changes in the data distribution and is thus a better candidate for adaptation. We rank all coordinates by V_j and update only the top-k. We treat k (update-ratio $r = k/d$) as the hyperparameter controlling the accuracy–compute trade-off; in our experiments we set $r = 0.60$.

$$\bar{g}_j^{(t)} = \frac{\partial}{\partial \theta_j} \frac{1}{B} \sum_{i=1}^{B} \mathcal{L}(f_\theta(x_i), y_i)\,, \tag{2}$$

$$V_j = \frac{1}{T} \sum_{t=1}^{T} \left(\bar{g}_j^{(t)}\right)^2 - \left(\frac{1}{T} \sum_{t=1}^{T} \bar{g}_j^{(t)}\right)^2. \tag{3}$$

G. Evaluation

We utilized a train-test split for time series (TSTS) [29] for offline and on-device model training. Classification accuracy

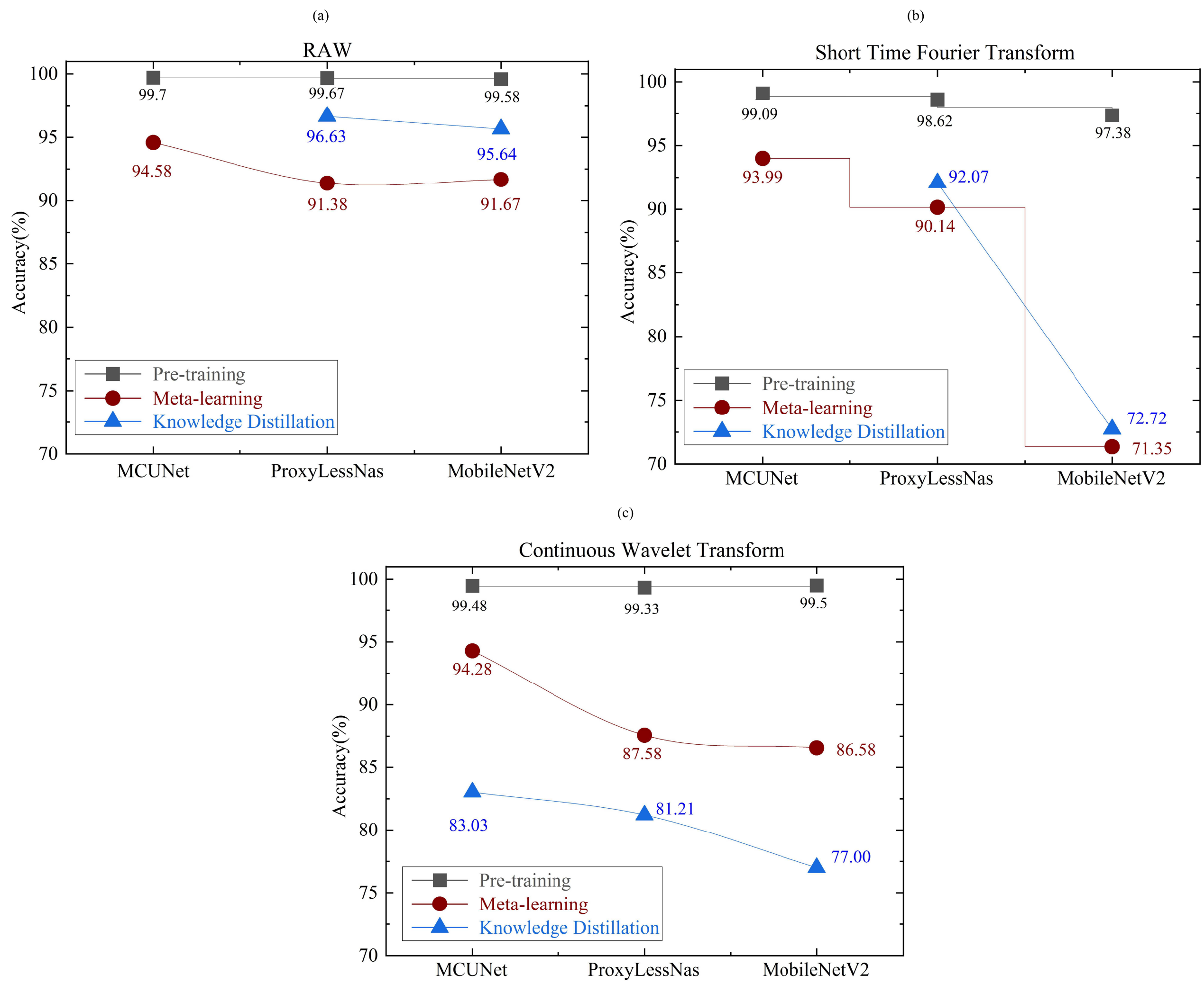

Fig. 7. Offline accuracy using different EMG data inputs. (a) Raw input without feature extraction. (b) Short time fourier transform. (c) Continuous wavelet transform.

(CA) (4) was used to evaluate 20% of the test data. We applied statistical analysis to evaluate the impact of offline pre-training (Phase 1) before on-device model update. All statistical analyses were performed using R.

$$CA = \frac{\text{Number of correct predictions}}{\text{Total number of true events}} \times 100 \tag{4}$$

III. Results

A. Selection of Teacher and Student Models

Selecting an appropriate teacher model is critical for achieving strong deployment performance. As shown in Fig. 7, MCUNet yielded the highest CA across diverse datasets during both pre-training (99.70%) and meta-learning (94.58%) without domain-specific feature extraction (Fig. 7(a)). We further conducted a comparative analysis using alternative extracted input features as proposed in [29]. The results (Fig. 7(b) and Fig. 7(c)) were comparable to those obtained with raw inputs, with MCUNet consistently outperforming MobileNetV2 and ProxylessNAS in both pre-training and meta-learning. Based on these observations, we selected MobileNetV2 and ProxylessNAS as the student models due to their comparatively lower performance relative to MCUNet. Participant 1's Flex-Adapt data collected during the electrode-shift experiment was used to perform offline KD training and evaluation, not on-device fine-tuning. Following KD, ProxylessNAS and MobileNetV2 achieved accuracy of 96.63% and 95.64%, respectively, using raw EMG input.

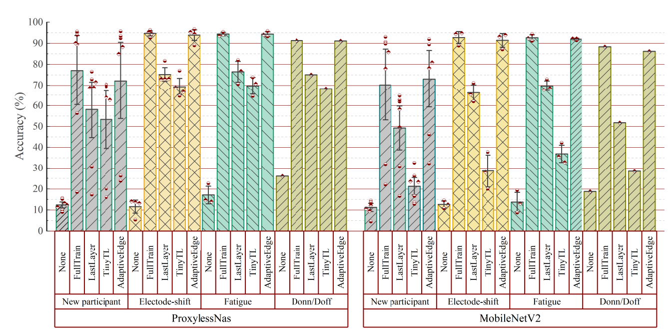

Fig. 8. Different fine-tuning methods with only pre-training the backbone.

B. On-Device Adaptive Model Update

The results in Fig. 8 and Table III highlights task-specific fine-tuning substantially outperforming the frozen, "None"

TABLE III
THE ACCURACY RESULTS OF DIFFERENT FINE-TUNING METHODS WITH ONLY PRE-TRAINING

Model	Method	New Participant	Electrode Shift	Fatigue	Donn Doff	Mean
ProxylessNAS	None	12.38±2.29	11.56±3.57	17.08±3.93	26.05	16.77±5.75
	FullTrain	77.05±27.12	94.66±1.33	94.37±0.76	91.26	89.33±7.21
	LastLayer	58.09±21.97	75.18±3.74	76.46±4.66	75.06	71.20±7.58
	TinyTL	53.28±22.48	69.27±4.70	69.70±3.88	68.39	65.16±6.87
	AdaptiveEdge	72.09±29.95	93.85±2.92	94.34±1.08	91.10	87.85±9.17
MobileNetV2	None	11.17±3.15	12.62±1.52	13.72±4.38	18.75	14.07±2.85
	FullTrain	70.17±27.81	92.70±2.70	92.62±1.33	88.32	85.96±9.28
	LastLayer	49.23±17.17	66.15±3.78	69.72±2.27	51.71	59.20±8.87
	TinyTL	21.21±6.72	28.58±7.22	36.89±3.93	28.44	28.77±5.53
	AdaptiveEdge	72.97±22.21	91.37±3.06	91.93±0.40	86.18	85.61±7.63

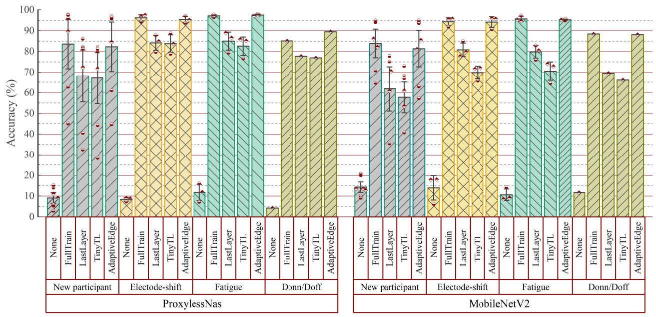

Fig. 9. Different fine-tuning methods after pre-training and applying meta-learning to the backbones.

TABLE IV
THE ACCURACY RESULTS OF FINE-TUNING METHODS WITH PRE-TRAINING AND META-LEARNING

Model	Method	New Participant	Electrode Shift	Fatigue	Donn Doff	Mean
ProxylessNAS	None	9.06±4.29	8.38±0.77	11.71±3.66	4.21	8.34±2.69
	FullTrain	83.53±19.46	96.19±1.36	97.12±0.52	85.07	90.48±6.20
	LastLayer	68.13±20.46	84.15±3.38	84.95±4.07	77.89	78.73±6.73
	TinyTL	67.07±20.19	83.75±4.04	82.51±4.15	76.89	77.56±6.58
	AdaptiveEdge	82.22±19.49	95.34±1.39	97.49±0.44	89.52	91.14±5.91
MobileNetV2	None	14.26±4.16	13.94±5.54	10.60±2.62	11.68	12.62±1.53
	FullTrain	83.86±11.19	94.33±1.77	95.68±1.11	88.40	90.57±4.74
	LastLayer	61.84±17.52	80.79±3.03	79.77±2.88	69.50	72.97±7.79
	TinyTL	57.71±12.12	69.78±2.89	70.41±4.26	66.00	65.97±5.05
	AdaptiveEdge	81.32±14.29	94.09±2.45	95.24±0.32	88.09	89.69±5.53

baseline across all four EMG disturbance scenarios (new participants, electrode-shifts, fatigue, and donning/doffing). For both models (Table III), FullTrain achieves the highest mean CA (89.33±7.21 for ProxylessNAS and 85.96±9.28 for MobileNetV2) at the cost of retraining every parameter. AdaptiveEdge resulted in mean CA of (87.85±9.17 for ProxylessNAS and 85.61±7.63 for MobileNetV2). Lightweight approaches such as LastLayer and TinyTL recovered much of the accuracy loss in contrast to None; however, they did not surpass a mean of 67.68±7.84 for ProxylessNAS and 43.99±16.91 for MobileNetV2 across all EMG disturbance scenarios. Additionally, the result in Fig. 8 also indicate ProxylessNas outperforms MobileNetV2.

The integration of few-shot meta-learning after pre-training (Fig. 9 and Table IV) substantially enhanced the adaptation capabilities of both model architectures (ProxylessNAS and MobileNetV2), with notable improvements across all fine-tuning methods. The FullTrain approach achieves the highest mean accuracy (Table IV) for all scenarios (90.48±6.20 for ProxylessNAS and 90.57±4.74 for MobileNetV2), while AdaptiveEdge follows closely behind (91.14±5.91 for ProxylessNAS and 89.69±5.53 for MobileNetV2) with minimal

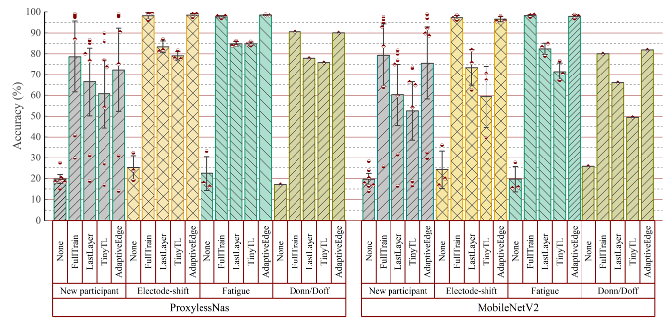

Fig. 10. Impact of different on-device fine-tuning methods on backbones after pre-training, meta-learning and knowledge distillation.

TABLE V
THE ACCURACY RESULTS OF FINE-TUNING METHODS WITH PRE-TRAINING, META-LEARNING AND KNOWLEDGE DISTILLATION

Model	Method	New Participant	Electrode Shift	Fatigue	Donn Doff	Mean
ProxylessNAS	None	19.68±3.47	25.13±5.52	22.41±7.71	17.05	21.07±3.01
	FullTrain	78.62±27.84	98.18±1.52	97.65±0.55	90.58	91.26±7.88
	LastLayer	66.41±26.61	83.36±2.65	84.85±1.04	77.92	78.14±7.24
	TinyTL	60.69±26.80	79.08±1.88	84.80±0.82	75.86	75.11±8.91
	AdaptiveEdge	72.26±32.66	98.50±0.77	98.54±0.25	90.05	89.16±8.96
MobileNetV2	None	19.56±4.56	24.24±8.53	19.60±5.56	25.73	22.28±2.75
	FullTrain	79.30±24.52	97.26±0.97	98.21±0.47	80.10	88.72±9.03
	LastLayer	60.22±24.15	73.42±8.12	82.39±65.96	65.96	70.50±8.30
	TinyTL	52.46±22.83	59.27±13.88	71.40±11.29	49.39	58.13±8.45
	AdaptiveEdge	75.51±28.32	96.54±1.17	97.86±0.78	81.90	87.96±9.53

difference between FullTrain and AdaptiveEdge. Offline training of the model using meta-learning reduces both the performance gap between LastLayer and TinyTL and the dispersion caused by various EMG disturbances, thereby demonstrating enhanced efficiency in knowledge transfer within parameter-efficient fine-tuning strategies. These results align with prior studies [2], [42] indicating that meta-learning frameworks can effectively address the substantial variability in data by learning to quickly adapt to new calibration data.

Incorporating pre-training, meta-learning, and knowledge distillation (Fig. 10 and Table V) achieves slightly enhanced performance across the evaluation test set, particularly in addressing domain shifts such as electrode-shift and fatigue (98.52±0.02 for ProxylessNAS and 97.21±0.66 for MobileNetV2). The overall performance of AdaptiveEdge across all EMG disturbances (89.84±10.71 for ProxylessNAS and 87.96±9.53 for MobileNetV2) not only maintains strong performance comparable to FullTrain but also demonstrates superior robustness relative to TinyTL and LastLayer update approaches. These consistent improvements across electrode-shift and fatigue experiments further highlight the resilience of the combined approach in addressing common sources of variability. Additionally, Fig. 8, Fig. 9, and Fig. 10 suggest that participant experience has a minimal effect on model performance, with the result from the first session (new participant) being notably lower than those from other experimental sessions. Also, the outlier (dots in Fig. 10) during session one can be attributed to noisy data from participants 5 (male) and 6 (female) in their first experimental session. When we excluded their data in session one, FullTrain achieved a mean accuracy of 95.60±3.01 for ProxylessNAS and 92.18±7.23 for MobileNetV2, while AdaptiveEdge yielded 94.81±3.78 for ProxylessNAS and 92.40±6.28 for MobileNetV2. Fig. 11

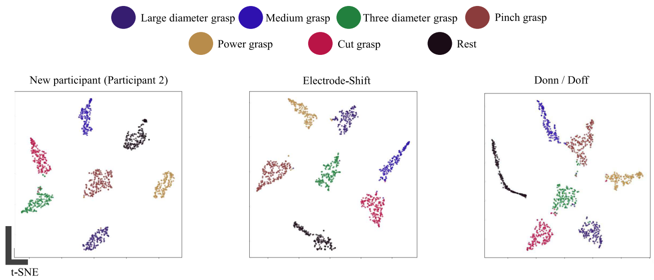

Fig. 11. ProxylessNAS t-SNE embedding activation for Participant two.

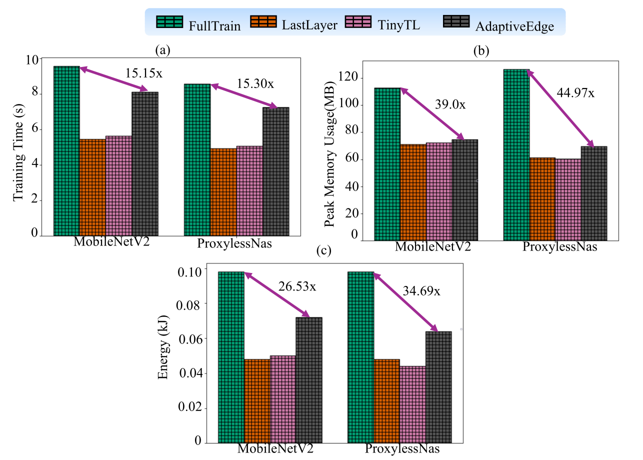

Fig. 12. On-device performance comparison of different fine-tuned methods with a batch size of 4 and r = 0.6. (a) Average epoch training time in seconds (s). (b) Peak latency in megabytes (MB). (c) Average epoch energy consumption in kilo joules (kJ).

demonstrates the behavior of the embedding activation in the ProxylessNas last layer under varying disturbance scenarios after an AdaptiveEdge model update.

Beyond CA, we evaluated critical deployment metrics including training time, peak memory usage, and energy consumption across all fine-tuning methods. Our experiments were conducted using a batch size of 4 on the FPGA. As shown in Fig. 12, AdaptiveEdge demonstrated superior performance across all deployment metrics compared to FullTrain. When compared to TinyTL and LastLayer, AdaptiveEdge exhibited higher memory usage (3)–9 MB) while consuming slightly greater energy consumption than the average of TinyTL and LastLayer across both models. These results underscore AdaptiveEdge balance between accuracy and resource utilization during deployment.

C. Ablation Experiment

We conducted ablation studies to compare fine-tuning methods with and without an offline training phase. To evaluate performance differences, we conducted paired-sample t-tests and Wilcoxon signed-rank tests, and measured effect sizes using Cohen's d_z. Both parametric (t-test) and nonparametric (Wilcoxon signed-rank test) approaches were employed to ensure statistical robustness. Our results rejected the null hypothesis in both cases (paired t-test, $p < 0.0001$; Wilcoxon signed-rank test, $p = 0.001$), providing strong evidence that offline training significantly improves model accuracy. The

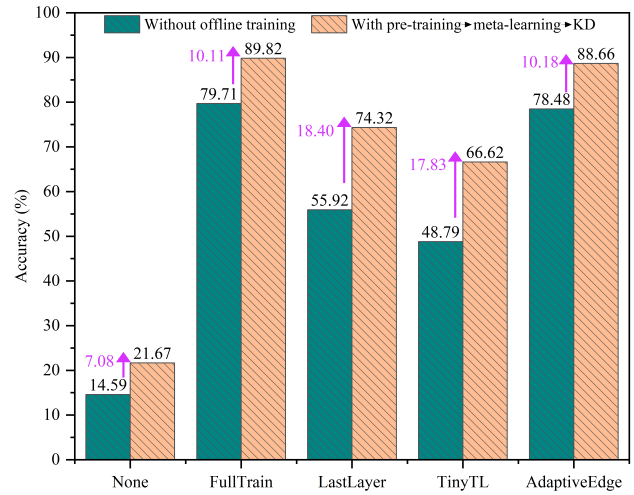

Fig. 13. Mean accuracy of ProxylessNas and MobileNetV2 across different on-device fine-tuning methods.

observed effect sizes, as measured by Cohen's d_z (d_z = 2.53), indicate a very large, standardized effect. As shown in Fig. 13, this combined offline training consistently resulted in substantial performance improvements across all fine-tuning methods. The highest improvement was observed in the TinyTL and LastLayer approaches, where accuracy increased by 17.83% and 18.40%, respectively. This highlights the effectiveness of the proposed method. The FullTrain yielded highest CA (89.82%), followed by AdaptiveEdge (88.66%).

IV. Discussion

Recent advancements in rehabilitation technology have focused on improving devices for individuals with upper limb impairments. A key challenge in translating laboratory innovations into real-world applications is addressing model degradation limitations while simultaneously designing sensors that are both aesthetically pleasing and affordably priced. To address these issues, we developed a flexible EMG sensor grid called FlexAdapt. This sensor system was designed using easily accessible materials with a total development cost under 300 USD, making it more accessible to individuals who may lack specialized expertise or resources. The fabrication process not only enhances the usability of EMG technology but also empowers users to customize their sensors independently, reducing bureaucratic barriers. Although FlexAdapt exhibited varying SNR, its average performance was comparable to other flexible and commercial EMG devices. The primary advantage lies in its cost-effectiveness and form-size, which allows for easier adaptation to various arm sizes.

Participant-specific model training and inference strategies are considered the gold standard in HCI studies. Due to practical limitations in collecting a large, generalizable dataset across multiple participants [1], adaptive approaches becomes essential when EMG disturbance scenarios arise. AdaptiveEdge framework serves as an alternative strategy that updates only a subset of model weights based on gradient variance. Utilizing the ProxylessNas model, AdaptiveEdge yielded an accuracy of 72.26%, 98.50%, 98.54%, and 90.05% for new participants, electrode-shift scenarios, fatigue conditions, and

TABLE VI
COMPARISON WITH SIMILAR WEARABLE GESTURE RECOGNITION SYSTEMS

	Moin et al. (2021) [24]	Lee et al. (2023) [45]	Chamberland et al. (2023) [21]	Gu et al. (2023) [3]	Kaifosh et al. (2024) [1]	Liu et al. (2024) [26]	This work
Wearable form factor	✓	✓	✓	✓	✓	✓	✓
Number of gestures	21	18	6	4	9	5	7
Channels	64	8	64	4	48	8	8
Classifier	HD-Algorithm	GNN	CNN	LDA	LSTM+CNN	CNN	CNN
Adaptive update	✓	✗	✗	✗	✓	✓	✓

EMG disturbance due to donning and doffing respectively. Compared to TinyTL [44] and LastLayer update approaches, AdaptiveEdge demonstrated superior accuracy but fell short of achieving FullTrain fine-tuning performance. AdaptiveEdge provides a practical balance between model adaptability in real-world HCI applications.

A comparison of FlexAdapt with state-of-the-art (SOTA) sEMG sensors are presented in Table VI. Our EMG system has a wearable form factor with a compact design of 8 channels. Moin et al. [24] employed a hyper-dimensional computing (HD) algorithm approach with hand-crafted features. While this design highlighted high model adaptation, its transferability to DNN approach are limited. In contrast, ours leverages a adaptive model that can be used for any type of DNN. Lee et al. [45] proposed a graph neural network (GNN) for sEMG signal processing. However, their focus on algorithmic performance came at the cost of real-time adaptability, which is essential in many bio-inspired applications. Conversely, our system prioritizes both hardware simplicity and algorithmic performance, ensuring that it can be effectively deployed in resource-constrained environments. Furthermore, Kaifosh and Reardon [1] used full layer fine-tuning on a large dataset. While this approach demonstrates high accuracy and generalization, it requires significant resources to execute. Our system's unique combination of wearability and adaptive capabilities sets it apart from SOTA sEMG sensors. By prioritizing both hardware simplicity and algorithmic performance, we propose a solution that is both data-effective and edge-deployable to address EMG disturbance scenarios in real-time.

In this study, several limitations were identified that could impact the interpretation of the results. First, the small sample size coupled with restrictive demographic profiles may limit the generalizability of our findings. To address this, we plan to increase both the sample size and representativeness in future studies to enhance the robustness and applicability of our design. The FlexAdapt system currently does not meet aesthetic product standards, which is a notable limitation. Additionally, this study failed to systematically investigate critical mechanical characteristics such as flexibility, which may hinder direct comparisons with existing SOTA EMG systems [55]. While we utilized a commercial ADC (cyton) for digital conversion and Bluetooth technology during the development phase, the pricing was not considered a priority. To ensure long-term viability and cost-effectiveness, future iterations of FlexAdapt will incorporate alternative, more affordable ADC and Bluetooth technologies. Moreover, further investigation is warranted to examine the performance of AdaptiveEdge under low-bit quantization conditions, which will contribute to refining the design and functionality of our final product.

V. CONCLUSION

In this study, we have designed a low-cost EMG sensor grid which is easy to fabricate. The constructed sensor not only meets but also challenges off-the-shelf non-medical grade EMG devices in terms of SNR. We address issues of model degradation from factors such as new participants, electrode-shifts, fatigue, and changes in the acquisition device's state during donning or doffing. To solve these issues, we propose a pipeline of offline training and on-device model update techniques. Our results demonstrate that the AdaptiveEdge model outperforms competing fine-tuning methods (LastLayer and TinyTL) in terms of accuracy across various degradation scenarios. Furthermore, while the AdaptiveEdge model maintains high accuracy akin to full-layer fine-tuning, it offers significant advantages in terms of memory efficiency and energy consumption relative to FullTrain, making it particularly suitable for applications requiring on-device parameter updates, such as EMG-based neuro-prosthetic devices.

REFERENCES

[1] P. Kaifosh and T. R. Reardon, "A generic non-invasive neuromotor interface for human-computer interaction," *Nature*, vol. 645, pp. 1–10, Sep. 2025.

[2] K. K. Kim et al., "A substrate-less nanomesh receptor with meta-learning for rapid hand task recognition," *Nature Electron.*, pp. 64–75, Dec. 2022.

[3] G. Gu et al., "A soft neuroprosthetic hand providing simultaneous myoelectric control and tactile feedback," *Nature Biomed. Eng.*, vol. 7, no. 4, pp. 589–598, Aug. 2021.

[4] D. Farina et al., "Toward higher-performance bionic limbs for wider clinical use," *Nature Biomed. Eng.*, vol. 7, no. 4, pp. 473–485, May 2021.

[5] H. Lorach et al., "Walking naturally after spinal cord injury using a brain–spine interface," *Nature*, vol. 618, no. 7963, pp. 126–133, 2023.

[6] H. Song et al., "Continuous neural control of a bionic limb restores biomimetic gait after amputation," *Nature Med.*, vol. 30, no. 7, pp. 2010–2019, Jul. 2024.

[7] S. J. Lee, K.-M. Park, K.-T. Kim, E.-Y. Seo, D. T. Gemechu, and O. Kim, "Developing a multi-directional lower-limb training system toward aging in place rehabilitation: A preliminary feasibility study on healthy individuals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 975–983, 2025.

[8] A. Pradhan, J. He, H. Lee, and N. Jiang, "Multi-day analysis of wrist electromyogram-based biometrics for authentication and personal identification," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 5, no. 4, pp. 553–565, Oct. 2023.

[9] A. Holobar and D. Farina, "Noninvasive neural interfacing with wearable muscle sensors: Combining convolutive blind source separation methods and deep learning techniques for neural decoding," *IEEE Signal Process. Mag.*, vol. 38, no. 4, pp. 103–118, Jul. 2021.

[10] H. Xu and A. Xiong, "Advances and disturbances in sEMG-based intentions and movements recognition: A review," *IEEE Sensors J.*, vol. 21, no. 12, pp. 13019–13028, Jun. 2021.

[11] B. Zhu, D. Zhang, Y. Chu, Y. Gu, and X. Zhao, "SeNic: An open source dataset for sEMG-based gesture recognition in non-ideal conditions," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1252–1260, 2022.

[12] K. R. Pyun et al., "Machine-learned wearable sensors for real-time hand-motion recognition: Toward practical applications," *Nat. Sci. Rev.*, vol. 11, no. 2, p. 298, Jan. 2024.

[13] D. Farina et al., "Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation," *Nature Biomed. Eng.*, vol. 1, no. 2, p. 0025, Feb. 2017.

[14] C. Chen, Y. Yu, X. Sheng, D. Farina, and X. Zhu, "Simultaneous and proportional control of wrist and hand movements by decoding motor unit discharges in real time," *J. Neural Eng.*, vol. 18, no. 5, Oct. 2021, Art. no. 056010.

[15] P. Shi, X. Zhang, W. Li, and H. Yu, "Improving the robustness and adaptability of sEMG-based pattern recognition using deep domain adaptation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 11, pp. 5450–5460, Nov. 2022.

[16] Y. Liu et al., "A novel unsupervised dynamic feature domain adaptation strategy for cross-individual myoelectric gesture recognition," *J. Neural Eng.*, vol. 20, no. 6, Dec. 2023, Art. no. 066044.

[17] L. Wu, A. Liu, X. Zhang, X. Chen, and X. Chen, "Electrode shift robust CNN for high-density myoelectric pattern recognition control," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[18] H. Deng et al., "Online unsupervised adaptation of latent representation for myoelectric control during user-decoder co-adaptation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 1026–1037, 2025.

[19] T. Su, X. Tan, X. Jiang, X. Liu, B. Hu, and C. Dai, "A dynamic balanced single-source domain generalization model for cross-posture myoelectric control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 255–265, 2025.

[20] H. Lee, M. Jiang, J. Yang, Z. Yang, and Q. Zhao, "Decoding gestures in electromyography: Spatiotemporal graph neural networks for generalizable and interpretable classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 33, pp. 404–419, 2025.

[21] F. Chamberland et al., "Novel wearable HD-EMG sensor with shift-robust gesture recognition using deep learning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 17, no. 5, pp. 968–984, Oct. 2023.

[22] X. Lin, X. Zhang, X. Zhang, X. Chen, and X. Chen, "DSDAN: Dual-step domain adaptation network based on bidirectional knowledge distillation for cross-user myoelectric pattern recognition," *IEEE Sensors J.*, vol. 23, no. 21, pp. 26765–26775, Nov. 2023.

[23] M. D. Dere and B. Lee, "A novel approach to surface EMG-based gesture classification using a vision transformer integrated with convolutive blind source separation," *IEEE J. Biomed. Health Informat.*, vol. 28, no. 1, pp. 181–192, Jan. 2024.

[24] A. Moin et al., "A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition," *Nature Electron.*, vol. 4, no. 1, pp. 54–63, Dec. 2020.

[25] S. Duan et al., "Highly sensitive and mechanically stable MXene textile sensors for adaptive smart data glove embedded with near-sensor edge intelligence," *Adv. Fiber Mater.*, vol. 6, no. 5, pp. 1–13, Oct. 2024.

[26] Y. Liu, R. Wang, Y. Li, and Y. Wang, "A novel dual-model adaptive continuous learning strategy for wrist-sEMG real-time gesture recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 4186–4196, 2024.

[27] F. Paissan et al., "Structured sparse back-propagation for lightweight on-device continual learning on microcontroller units," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Benin, Jun. 2024, pp. 2172–2181.

[28] M. D. Dere, J.-H. Jo, and B. Lee, "Event-driven edge deep learning decoder for real-time gesture classification and neuro-inspired rehabilitation device control," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023.

[29] J. Yang, M. Soh, V. Lieu, D. Weber, and Z. Erickson, "Emgbench: Benchmarking out-of-distribution generalization and adaptation for electromyography," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 37, 2024, pp. 50313–50342.

[30] U. Côté-Allard et al., "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019.

[31] S. Tam, M. Boukadoum, A. Campeau-Lecours, and B. Gosselin, "A fully embedded adaptive real-time hand gesture classifier leveraging HD-sEMG and deep learning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 2, pp. 232–243, Apr. 2020.

[32] A. T. Nguyen et al., "A portable, self-contained neuroprosthetic hand with deep learning-based finger control," *J. Neural Eng.*, vol. 18, no. 5, Oct. 2021, Art. no. 056051.

[33] M. Simão, P. Neto, and O. Gibaru, "EMG-based online classification of gestures with recurrent neural networks," *Pattern Recognit. Lett.*, vol. 128, pp. 45–51, Dec. 2019.

[34] T. M. Bittibssi, A. H. Zekry, M. A. Genedy, and S. A. Maged, "SEMG pattern recognition based on recurrent neural network," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 103048.

[35] M. Montazerin, E. Rahimian, F. Naderkhani, S. F. Atashzar, S. Yanushkevich, and A. Mohammadi, "Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals," *Sci. Rep.*, vol. 13, no. 1, p. 11000, Jul. 2023.

[36] Q. Hu, G. A. Azar, A. Fletcher, S. Rangan, and S. F. Atashzar, "ViT-MDHGR: cross-day reliability and agility in dynamic hand gesture prediction via HD-sEMG signal decoding," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 3, pp. 419–430, Apr. 2024.

[37] R. V. Godoy et al., "Electromyography-based, robust hand motion classification employing temporal multi-channel vision transformers," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10200–10207, Oct. 2022.

[38] R. V. Godoy, A. Dwivedi, and M. Liarokapis, "Electromyography based decoding of dexterous, in-hand manipulation motions with temporal multichannel vision transformers," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2207–2216, 2022.

[39] S. Shen, X. Wang, F. Mao, L. Sun, and M. Gu, "Movements classification through sEMG with convolutional vision transformer and stacking ensemble learning," *IEEE Sensors J.*, vol. 22, no. 13, pp. 13318–13325, Jul. 2022.

[40] N. K. Karnam, S. R. Dubey, A. C. Turlapaty, and B. Gokaraju, "EMGHandNet: A hybrid CNN and bi-LSTM architecture for hand activity classification using surface EMG signals," *Biocybernetics Biomed. Eng.*, vol. 42, no. 1, pp. 325–340, Jan. 2022.

[41] P. Ren et al., "A comprehensive survey of neural architecture search: Challenges and solutions," *ACM Comput. Surveys*, vol. 54, no. 4, pp. 1–34, May 2022.

[42] Y. D. Kwon, R. Li, S. Venieris, J. Chauhan, N. D. Lane, and C. Mascolo, "TinyTrain: Resource-aware task-adaptive sparse training of DNNs at the data-scarce edge," in *Proc. Int. Conf. Mach. Learn.*, 2024, pp. 25812–25843.

[43] J. Lin, L. Zhu, W.-M. Chen, W. Wang, C. Gan, and S. Han, "On-device training under 256KB memory," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 22941–22954.

[44] H. Cai, C. Gan, L. Zhu, and S. Han, "TinyTL: Reduce activations, not trainable parameters for efficient on-device learning," 2020, *arXiv:2007.11622*.

[45] H. Lee et al., "Stretchable array electromyography sensor with graph neural network for static and dynamic gestures recognition system," *npj Flexible Electron.*, vol. 7, no. 1, p. 20, Apr. 2023.

[46] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for SEMG sensors and sensor placement procedures," *J. Electromyogr. Kinesiol.*, vol. 10, no. 5, pp. 361–374, Oct. 2000.

[47] R. N. Khushaba and K. Nazarpour, "Decoding HD-EMG signals for myoelectric control–how small can the analysis window size be?," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 8569–8574, Oct. 2021.

[48] A. Tigrini et al., "Decoding transient sEMG data for intent motion recognition in transhumeral amputees," *Biomed. Signal Process. Control*, vol. 85, Aug. 2023, Art. no. 104936.

[49] K. Englehart, B. Hudgin, and P. A. Parker, "A wavelet-based continuous classification scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 3, pp. 302–311, Mar. 2001.

[50] J. Lin, W.-M. Chen, Y. Lin, J. Cohn, C. Gan, and S. Han, "MCUNet: Tiny deep learning on IoT devices," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 11711–11722.

[51] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

[52] H. Cai, L. Zhu, and S. Han, "ProxylessNAS: Direct neural architecture search on target task and hardware," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–13.

[53] J. Wang and Y. Zhai, "Prototypical Siamese networks for few-shot learning," in *Proc. IEEE 10th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jul. 2020, pp. 178–181.

[54] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, *arXiv:1503.02531*.

[55] G. Rolandino, C. Zangrandi, T. Vieira, G. Luigi Cerone, B. Andrews, and J. J. FitzGerald, "HDE-array: Development and validation of a new dry electrode array design to acquire HD-sEMG for hand position estimation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 4004–4013, 2024.