# Whole-Field Action Sensing via Wearable Single-Channel EMG Sensors and Resource-Efficient Motion Network

**Xuanming Jiang[1,5*], Dingyu Nie[2,5*], Baoyi An[2,5*], Yuzhe Zheng[2], Yichuan Mao[2], Jialie Shen[3], Xueming Qian[4,6], Zhiwen Jin[2], Wei Lan[2†], Guoshuai Zhao[1,6†]**

[1]School of Software Engineering, Xi'an Jiaotong University, Xi'an, China
[2]School of Physical Science and Technology, Lanzhou University, Lanzhou, China
[3]School of Science and Technology, City St George's, University of London, London, the United Kingdom
[4]School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China
[5]Xi'an Jiyun Technology Co., Ltd., Xi'an, China
[6]Shaanxi Yulan Jiuzhou Intelligent Optoelectronic Technology Co., Ltd., Xi'an, China
jiangxm24@stu.xjtu.edu.cn, niedy2025@lzu.edu.cn, anby2024@lzu.edu.cn, lanw@lzu.edu.cn, guoshuai.zhao@xjtu.edu.cn

## Abstract

The proliferation of collaborative training and multi-person sports has underscored the necessity for concurrent whole-field action sensing. However, Electromyography (EMG) recognition, which plays a pivotal role in Wearable Human Activity Recognition (WHAR) for analyzing muscle activity and decoding action intent, still faces challenges in achieving a balance between performance, cost, and efficiency in multi-person scenarios. Unlike current channel-expansion solutions, we propose a wireless wearable **S**ingle-Dimensional **S**parse **EMG** (**2SEMG**) **Sensor** for efficient personal sampling. These action-unaffected sensors leverage the proposed lightweight **O**ne-Dimensional **Mo**tion **Net**work (**OMONet**) to facilitate concurrent action sensing. Experiments demonstrate that OMONet achieves leading performance and efficiency in action signal recognition, and two real-world badminton matches further confirm the performance, robustness, and real-time efficiency of the whole-field action sensing network constructed via 2SEMG Sensors and OMONet.

## Introduction

With the rise of sports scenarios requiring multi-player interaction and team coordination, there is an unquestionable need for a comprehensive analysis of athletes' actions throughout the field (Carling et al. 2008; Zeng, Shi, and Zhou 2022). Current Human Activity Recognition (HAR) typically utilizes vision, WiFi, and other signals to capture and analyze multi-person action data (Zeng, Shi, and Zhou 2022; Abuhoureyah, Swee, and Chiew 2024). However, they inevitably face challenges such as visual occlusion, signal interference, and inadequate real-time efficiency (Huang et al. 2021). Therefore, Wearable Human Activity Recognition (WHAR) devices, which monitor physiological and kinematic signals in real-time, hold promise for gradually replacing fixed devices, enabling Whole-Field Human Activity Recognition (WFHAR) in collaborative team sports.

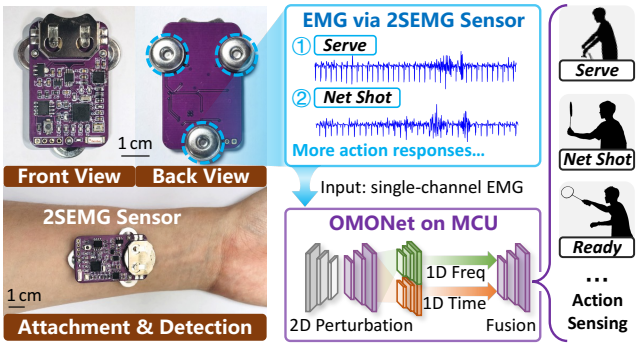


Figure 1: Overview of muscle activity recognition via wireless 2SEMG Sensor and on-chip OMONet.

Among various physiological signals, Electromyography (EMG) plays a key role by directly reflecting muscle activity and action intent (Xiong et al. 2024b). Methods using EMG for HAR have garnered widespread attention (Moin et al. 2021; Yang et al. 2025). However, current EMG approaches developed for single-person settings fail to meet the demands regarding synchronous acquisition of whole-field action data due to their lack of capability for parallel EMG collection and recognition (Xiong et al. 2024a).

To address EMG-based WFHAR's acquisition and transmission delays, a widely used strategy is to synchronize signal acquisition and inference within an embedded Microcontroller Unit (MCU) (Lu et al. 2024). However, the limited computational and memory resources of the MCUs significantly hinder the deployment of reliable models (Liu et al. 2024). These concurrent and irreconcilable issues in both hardware and algorithms severely constrain the large-scale deployment and sustained operation of EMG-based action sensing systems (Cárdenas-Valdez et al. 2023).

Therefore, we introduce a wireless, low-power and wearable **S**ingle-Dimensional **S**parse **EMG** (**2SEMG**) **Sensor**, and a lightweight **O**ne-Dimensional **Mo**tion **Net**work (**OMONet**). The main contributions of this paper include:

*These authors contributed equally.
†Corresponding author.

- **Wearable 2SEMG Sensor.** We develop a **7.2-g**, action-imperceptible sensor that enables high-fidelity EMG collection under 1 kHz native waveform. It operates in a low-power mode (**12 mW** during work) and supports wireless whole-field (∼**100 m**) EMG sampling terminals.
- **Efficient OMONet for Action Recognition.** We propose the time-freq-decoupled learning OMONet with **0.08M Params** and **0.17GFLOPs**, delivering leading performance and efficiency on open-source datasets and 2SEMG Sensor-derived **1-BEMG benchmark**.
- **Real-World Multi-Player Evaluation.** We evaluate the 2SEMG Sensors-OMONet integration in **two real badminton matches**, confirming the sensors' high-fidelity EMG capture in intense multi-player sports and validating OMONet for real-time concurrent action recognition.

## Related Work

### Sensor Design in WHAR

Recently, single-channel EMG-based systems for WHAR have gained attention due to their low power consumption and ease of deployment (Tavakoli, Benussi, and Lourenco 2017; Wu, Ruan, and Lee 2021; Wei et al. 2022). Although effective in detecting localized or isolated actions, such systems often fail to provide sufficient dimensional information when recognizing complex, full-body coordinated activities (Baskaran and Adams 2023). To address this limitation, current studies have explored high-density, multi-channel EMG arrays and hybrid systems that integrate inertial measurement units with EMG sensors. These approaches aim to improve recognition performance by capturing richer spatial relationships between muscle groups (Zhang, Sun, and Zou 2022; Chamberland et al. 2023; Lee et al. 2023). However, multi-channel systems introduce challenges such as increased hardware complexity and cost, difficulties in sensor node synchronization, limited real-time efficiency, and a particularly frustrating user experience (Ng et al. 2024).

### Deep Learning in Physiological Signal Recognition

Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) are prevalent in 1D physiological signal processing (Ma et al. 2021; Yang et al. 2024). Typically, a hybrid CNN-LSTM model leverages CNN to extract spatial patterns, followed by LSTM to model temporal dependencies, thus improving performance (Bao et al. 2020). However, LSTM or Transformer-based models commonly exhibit excessive complexity and high memory usage, which complicates their applications on resource-limited platforms (Montazerin et al. 2023). Recent methods such as HAR-Mamba use a dual-path structure based on selective state-space modeling, reducing computational and memory footprints while maintaining long-range dependency modeling (Dao and Gu 2024; Li et al. 2025). Other studies convert EMG into Short-Time Fourier Transform (STFT) spectrum and apply CNNs to retain local-global information (Yuan et al. 2025). Building on these insights, we propose Motion Attention that separates time-frequency extraction and synchronization while enhancing the long-/short-term representation ability of CNNs for efficient EMG recognition.

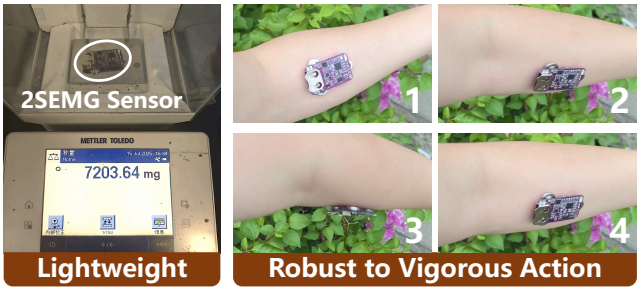


Figure 2: Sensor with negligible action interference.

## Methodology

The proposed wireless 2SEMG Sensor and lightweight OMONet constitute the hardware and software terminals of the EMG-based whole-field action sensing network.

### Single-Dimensional Sparse EMG Sensor

The 2SEMG Sensor uses a 45×28 mm Printed Circuit Board (PCB) as its substrate, integrating three 12 mm-diameter (Ø) electrode ports for silver chloride (AgCl) gel patches that adhere the lightweight sensor to the user's skin via replaceable skin-friendly gel, as shown in Figure 2. The sensor is powered by a 3V lithium coin battery (20 mm Ø, 49 mAh).

The EMG signal, denoted as $s(t) \in \mathbb{R}$, exhibits microvolt-scale amplitudes and a 10-500 Hz frequency spectrum, with amplitude modulation correlating to muscle activation levels. For EMG signal sampling, the 2SEMG Sensor utilizes a bipolar lead configuration consisting of two active electrodes with an interposed dedicated EMG AgCl electrode, as illustrated in the left panel of Figure 1.

The raw EMG with coupling noise is expressed as $v(t)$:

$$v(t) = (s * h_{\text{hp}})(t) + n(t), \quad n(t) \sim \mathcal{N}(0,1) \tag{1}$$

where $h_{\text{hp}}$ denotes the impulse response of a high-pass filter.

Subsequently, a high input impedance differential amplifier implemented by combining an instrumentation amplifier with an operational amplifier is used to amplify the differential EMG signal while suppressing common-mode interference. Next, an analog band-pass filter is designed to attenuate motion artifacts and high-frequency noise. These processes of obtaining $y_{\text{bpf}}(t)$ can be expressed as follows:

$$y_{\text{bpf}}(t) = [G\left(v^{+} - v^{-}\right) * h_{\text{bp}}](t) \tag{2}$$

where $G$ denotes the amplification factor, $h_{\text{bp}}$ denotes the band-pass filter response, and $v^{+}(t)$, $v^{-}(t)$ represent the signals captured by the active electrodes (positive/negative).

The conditioned analog signal is then digitized using a 12-bit Analog-to-Digital Converter (ADC) at a sampling rate of 1 kHz, satisfying the Nyquist sampling theorem $y_{\text{bpf}}[n] = y_{\text{bpf}}(nT_{\text{s}}), n{\in}\mathbb{Z}$ (Shannon 1949), where $T_{\text{s}}$ denotes the sampling period. Thus, the quantized $y_{\text{q}}[n]$ is calculated by:

$$y_{\text{q}}[n] = \text{Round}\left(\frac{y_{\text{bpf}}[n]}{\Delta}\right) \cdot \Delta, \quad \Delta = \frac{V_{\text{FS}}}{2^{b}} \tag{3}$$

where $V_{\text{FS}}$ is the full-scale voltage, $b$ is the ADC resolution in bits. Finally, the ADC output is transmitted via Bluetooth Low Energy (BLE) to an OMONet-running MCU.

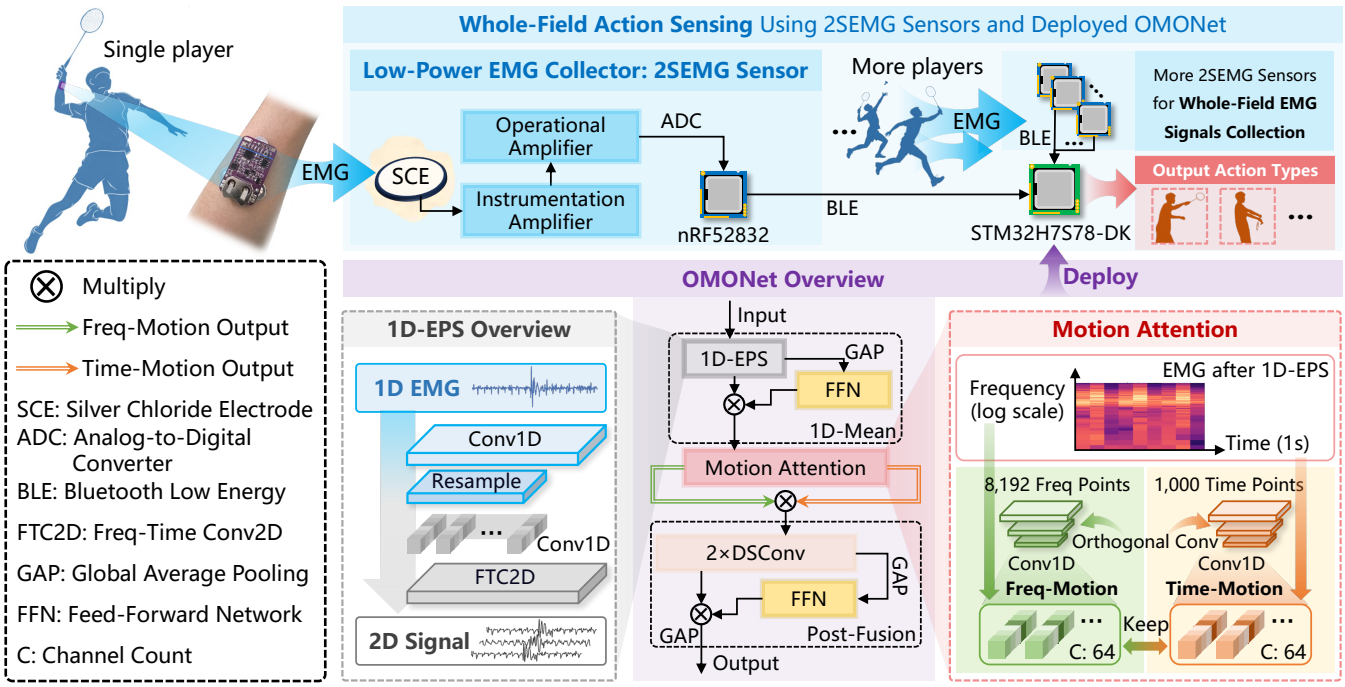


Figure 3: Overview of the whole-field action sensing network based on 2SEMG Sensors and OMONet.

| Metrics | Values | Metrics | Values |
|---|---|---|---|
| Power Consumption | 12 mW | Battery Life | >10 h |
| Resolution | 12-bit | Bandwidth | 500 Hz |
| Dynamic Range | 80 dB | Weight $_{\text{Battery-free}}$ | 7.2 g |
| Sampling Rate | 1 kHz | Wireless Range | ∼100 m |
| Sampling Delay | <1 ms | Wireless Delay | 20-50 ms |
| SNR | 27.4 dB | CMRR | 100 dB |

Table 1: 2SEMG Sensor performance during work.

As indicated in Table 1, the 2SEMG Sensor, weighing merely 7.2 g, enables sampling and processing with negligible impact on the wearer's actions. Not only does it achieve ultra-low 12 mW sampling power consumption, along with a >10 hours battery life and ∼100 m wireless transmission range sufficient for most sports, but its superior Common-Mode Rejection Ratio (CMRR) and high Signal-to-Noise Ratio (SNR) also ensure high-fidelity EMG acquisition.

## One-Dimensional Motion Network

OMONet processes the STFT spectrogram as input and leverages 1D-2D-mixed convolutions to capture variations in simulate-enhanced physiological signal activation across adjacent time-frequency frames. It employs Motion Attention to focus on frequency and temporal evolution via separated Freq-Motion and Time-Motion, and finally integrates both to recognize EMG patterns across urgency spectrums.

**Preliminary: STFT Mapping.** Surface EMG signals represent non-stationary, multi-frequency superpositions of motor unit action potentials. Although simple time-domain modeling cannot distinguish muscle contraction patterns, time-frequency representations provide an intuitive visualization of the energy distribution of physiological signals.

Since OMONet directly processes raw EMG signals, the input 1D timing signal $s_{1D}(t)$ undergoes Direct-Current (DC) offset removal and normalization as follows:

$$s_n(t) = \frac{s_{1D}(t) - \mu}{\max\{|s_{1D}(t) - \mu|\}}, \quad \mu = \frac{1}{L}\sum_{t=1}^{L} s_{1D}(t) \tag{4}$$

where $L$ denotes the length of $s_{1D}(t)$, $(s_{1D}(t) - \mu)$ serves as the DC-removed $s_{1D}(t)$, and $s_n(t)$ is the normalization result. Next, a 512-point Hann window with a 256-point hop size is applied for STFT computation as $\mathbf{STFT}(s_n)$.

Suppose that the extensive feature map $\mathcal{X}^{T\times F\times C}$ from 1D EMG has dimensions $T \times F \times C$, where $T$ denotes the number of time frames, $F$ indicates the number of frequency bins, and $C$ represents the number of channels. The mapping from $\mathbf{STFT}(s_n)$ to $\mathcal{X}$ can be expressed as follows:

$$\mathcal{X}[t, f, c] = \frac{\mathbf{STFT}(s_n)[\tau, \nu] - \min\{\mathbf{STFT}(s_n)\}}{\max\{\mathbf{STFT}(s_n)\} - \min\{\mathbf{STFT}(s_n)\}} \tag{5}$$

where $t, f, c$ represent the index of $\mathcal{X}^{T\times F\times C}$; $\tau$ and $\nu$ denote the time index and the frequency index, respectively.

**1D-EMG Perturbation Synthesis.** For the raw sequence $s_{1D}$, truncate it or pad it to a fixed length 2048 for further learnable enhancement of non-stationary EMG patterns.

To simulate potential physiological noise, define $\omega_p$ and $\beta_p$ as the perturbation weights and bias, respectively, thus the perturbated $s_{1D}$ based on 1D Convolution (Conv1D) within a specific time-frequency bin can be defined as follows:

$$\mathcal{S}_{1D} = \text{MaxPooling}(\text{ReLU}(\omega_p * s_{1D} + \beta_p)) \tag{6}$$

where $*$ denotes convolution. By analogy with Equation (4), the fitted mean $\mu_{\text{p}}$ and the variance $\sigma_{\text{p}}$ are defined as:

$$\mu_{\text{p}} = w_{\mu} \cdot \mathcal{S}_{\text{1D}} + \beta_{\mu}, \quad \sigma_{\text{p}} = \omega_{\sigma} \cdot \mathcal{S}_{\text{1D}} + \beta_{\sigma} \tag{7}$$

where $w_{\{\mu,\sigma\}}$ and $\beta_{\{\mu,\sigma\}}$ denote learnable weights and biases, respectively.

Subsequently, resampling is conducted via introducing subtle perturbations correlated with the time-frequency domain within a low-dimensional space as follows:

$$\mathcal{S}_{\text{z}} = \mu_{\text{p}} + \exp\left(0.5 \log(\sigma_{\text{p}})\right) \cdot \epsilon, \quad \epsilon \sim \mathcal{N}(0,1) \tag{8}$$

Despite increasing the diversity of time-frequency patterns, the basic $\mathcal{S}_{\text{1D}}$ patterns are maintained within $\mathcal{S}_{\text{z}}$.

Thereafter, the motion pattern-enhanced $\mathcal{S}_{\text{re1D}}$ is designed to be calculated by progressively reconstructing the time-domain signal through Conv1D and upsampling on $\mathcal{S}_{\text{z}}$. Therefore, STFT-based feature maps $\mathcal{X}[t,f,c]$ can be obtained based on $\mathcal{S}_{\text{re1D}}$ according to Equation (5).

To obtain 2D time-frequency representations, a dedicated 2D Convolution (Conv2D) $\mathcal{C}_k^a(\cdot)$ is designed as follows:

$$\begin{gathered}\mathcal{C}_k^a(\cdot) = \mathcal{X} \mapsto \left[\sum_{c=1}^{C} \sum_{u=-A_t}^{A_t} \sum_{v=-A_f}^{A_f} \omega_k^{(c)}[u,v] \cdot \mathcal{X}_{\text{idx}}\right]_{k=1}^{K} \\ \mathcal{X}_{\text{idx}} = \mathcal{X}[t-u, f-v, c],\ A_{\{t,f\}} = \lfloor (a_{\{t,f\}} - 1)/2 \rfloor\end{gathered} \tag{9}$$

where $a$ denotes the kernel size; $K$, $k$ denote the kernel number and the index of the output channel, respectively; $\omega_k^{(c)}$ denotes the weight of the $k$-th filter and $c$-th channel.

Next, the Frequency-Time Conv2D (FTC2D) uses two Conv2D ($\mathcal{C}_{32}^{3,3}$) to capture local energy blocks, and nonlinearly maps them to the expanded $\mathcal{X}_{\text{E}}^{T\times F\times K}$ that represents the motion energy distribution of time-freq combinations.

To reduce redundant frequency information and computational complexity, the max pooling and the ReLU activation are used to compress the frequency sequence while preserving the fewer temporal information. As shown in Figure 3, the output of the 1D-EPS resembles a spectrogram.

Finally, the average energy $\mathcal{S}_k$ of the spectrogram within the channel $k$ can be obtained based on channel-wise mean.

**1D-Mean.** The 1D-Mean independently compresses the temporal and frequency dimensions of the perturbed 2D signals to further combine features, equivalent to calculating the time-averaged power for each freq-band to reflect the activity level of muscle units in overlapping freq-bands.

Given a Feed-Forward Network (FFN) (Vaswani et al. 2017), which assigns physiological importance weights to different frequency patterns based on combined features as:

$$\text{FFN}(\mathcal{S}_k) = \max\{0, \mathcal{S}_k \cdot \omega_{\text{F1}} + \beta_{\text{F1}}\} \cdot \omega_{\text{F2}} + \beta_{\text{F2}} \tag{10}$$

where $\omega_{\{\text{F1,F2}\}}$ and $\beta_{\{\text{F1,F2}\}}$ represent learnable weights and biases, respectively.

This process comprises two stages: (1) the **reduction stage** fuses data using a fully connection layer to capture collaborative relationships throughout the full spectral range; and (2) the **elevation stage** maps the modulated full-spectrum information back to the original freq-bands.

Finally, the integrated 2D time-frequency coupling features $\mathcal{X}_{\text{Mean}}$ is designed to be obtained by multiplying the $\mathcal{X}_{\text{E}}$ with the activated $\text{FFN}(\mathcal{S}_k)$. Therefore, the 1D-Mean module is designed to amplify the highly correlated motion bands in the physiological signals while attenuating the subtle variations arising from individual differences.

**Motion Attention.** To decouple time-frequency information for enriching feature patterns, two separated motion perception paths are designed to process time-frequency information via the operation defined in Equation (9), which can be expressed in time ($\mathcal{X}_{\text{t}}$) and frequency ($\mathcal{X}_{\text{f}}$) forms as:

$$\mathcal{X}_{\text{t}} = \mathcal{C}_{64}^{1,3}(\mathcal{X}_{\text{Mean}}), \quad \mathcal{X}_{\text{f}} = \mathcal{C}_{64}^{3,1}(\mathcal{X}_{\text{Mean}}) \tag{11}$$

The temporal convolution employs a $1 \times 3$ kernel to emphasize interframe dynamics, capturing the continuity of the time-oriented movement. In contrast, convolution in the frequency domain applies a $3 \times 1$ kernel to focus on activation traits of distinct muscle fiber types. After undergoing a series of max pooling and ReLU activation along the frequency and time domains to reduce model complexity, the two decoupled motion paths are recombined into a feature map ($\mathcal{X}_{\text{re}}$) with enhanced time-frequency interactions:

$$\mathcal{X}_{\text{re}}[t,f,k] = \mathcal{X}_{\text{t}}[t,f,k] \cdot \mathcal{X}_{\text{f}}[t,f,k] \tag{12}$$

The fused $\mathcal{X}_{\text{re}}$, where each channel represents a distinct freq-band, enables convolutional sliding within these bands to capture localized time-frequency variation patterns.

**Post-Fusion.** Given that rapid muscle contractions typically induce abrupt spikes in the high-frequency range, two Depthwise Separable Convolutions (DSConv) (Chollet 2017) with 128 filters (doubling the output channels of a single motion perception path in Equation (11)) are devised to amplify sharp changes and localized high-energy fluctuations within individual freq-bands during cross-channel mixing. These processes can be expressed as reconstructing the 2D time-frequency synchronization $\mathcal{X}_{\text{2D}}$ as:

$$\mathcal{X}_{\text{2D}} = \sum\nolimits_{\Delta t} \sum\nolimits_{\Delta f} \omega^{(\text{2D})} \cdot \mathcal{X}_{\text{re}}[t - \Delta t, f - \Delta f] \tag{13}$$

where $\omega^{(\text{2D})}$ denotes a single-channel convolution kernel. This process combines the local responses of single freq-bands to learn multi-frequency cooperative patterns.

Subsequently, another 2D coupling stage is designed as:

$$\mathcal{X}_{\text{fusion}} = \text{GAP}(\mathcal{X}_{\text{2D}} \otimes \text{FFN}(\text{GAP}(\mathcal{X}_{\text{2D}}))) \tag{14}$$

where GAP denotes Global Average Pooling (GAP), while $\otimes$ represents multiply. Through global statistical analysis and adaptive rescaling, this top-down fusion emphasizes the most physiologically pertinent time-frequency features.

# Experiment

## Experimental Setup

**Comparison Consideration:** To validate the performance and efficiency of OMONet, we compared it with leading EMG-specific models, such as **DDG** (Ye et al. 2023), **DCNN** (Rehman et al. 2024), and **DACNN** (Wattanasiri et al. 2025), as well as advanced temporal signal processing models, including **MILLET** (Early et al. 2024), **TS-GAC** (Wang et al. 2024), and **ML-TCN** (Le et al. 2025). All models were exported as ONNX files to ensure a fair on-board comparison.

| Metrics | | DDG | DCNN | MILLET | TS-GAC | DACNN | ML-TCN | OMONet |
|---|---|---|---|---|---|---|---|---|
| **Params** (M) ↓ | | 0.123 | 3.790 | 0.267 | 0.159 | **0.034** | 34.485 | 0.078 |
| **FLOPs** (G) ↓ | | 0.586 | 0.536 | 1.604 | 0.245 | **0.048** | 14.554 | 0.168 |
| **MACCs** (G) ↓ | | 0.314 | 0.269 | 0.866 | 0.123 | **0.024** | 7.280 | 0.085 |
| **Flash Usage** (KB) ↓ | | 543.2 | 3686.4 | | 648.5 | **149.2** | | 326.6 |
| **RAM Usage** (KB) ↓ | | **335.0** | 558.8 | Overflow | 577.0 | 503.8 | Overflow | 555.5 |
| **Inference Latency** (s) ↓ | | 0.28 ±0.07 | 0.59 ±0.03 | | 4.29 ±0.13 | **0.25** ±0.07 | | 0.33 ±0.02 |
| 1-BEMG | **Acc.** (%) ↑ | 93.75 ±4.37 | 87.50 ±0.51 | 92.27 ±0.94 | 99.41 ±0.49 | 92.06 ±0.90 | 98.36 ±1.64 | **99.47** ±0.41 |
| | **Recall** (%) ↑ | 91.95 ±4.50 | 87.51 ±0.48 | 92.30 ±0.97 | 99.38 ±0.51 | 92.09 ±0.94 | 98.34 ±1.66 | **99.51** ±0.29 |
| | **F1-Score** (%) ↑ | 92.17 ±4.19 | 87.44 ±0.47 | 92.36 ±0.89 | 99.42 ±0.48 | 92.15 ±0.86 | 98.35 ±1.65 | **99.47** ±0.37 |
| Myo | **Acc.** (%) ↑ | 69.81 ±2.77 | 78.04 ±0.82 | 79.10 ±0.88 | 77.05 ±1.64 | 67.54 ±1.28 | **85.98** ±1.19 | 83.33 ±0.97 |
| | **Recall** (%) ↑ | 69.78 ±2.74 | 78.07 ±0.81 | 79.22 ±0.82 | 76.75 ±1.75 | 67.48 ±1.21 | **85.84** ±1.26 | 83.36 ±0.85 |
| | **F1-Score** (%) ↑ | 70.10 ±2.61 | 78.15 ±0.76 | 79.23 ±0.81 | 76.55 ±1.59 | 67.43 ±1.17 | **86.05** ±1.15 | 83.33 ±0.95 |
| Yaseen18 | **Acc.** (%) ↑ | 93.69 ±2.07 | 92.50 ±1.03 | 93.17 ±1.32 | 91.25 ±1.25 | 91.00 ±1.54 | 94.50 ±1.37 | **95.50** ±1.13 |
| | **Recall** (%) ↑ | 93.63 ±2.02 | 92.51 ±1.08 | 93.15 ±1.24 | 91.04 ±1.43 | 91.12 ±1.49 | 94.37 ±1.54 | **95.50** ±1.04 |
| | **F1-Score** (%) ↑ | 93.58 ±1.97 | 92.47 ±0.99 | 93.16 ±1.29 | 91.16 ±1.27 | 90.95 ±1.47 | 94.38 ±1.54 | **95.39** ±1.08 |

Quantitative complexity analysis is grounded in three critical metrics, including (1) total number of parameters (**Params**) (Jiang et al. 2025), (2) floating point operations (**FLOPs**), and (3) multiply-accumulate operations (**MACCs**) (Pau and Aymone 2024).

Table 2: Performance and efficiency comparison between OMONet and leading methods on physiological signal recognition.

| Badminton Actions | DDG | DCNN | MILLET | TS-GAC | DACNN | ML-TCN | OMONet | Average Acc. (%) |
|---|---|---|---|---|---|---|---|---|
| *Ready* | 100.00 | 99.53 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 99.93 |
| *Clear Shot* | 99.58 | 89.35 | 94.94 | 100.00 | 92.42 | 100.00 | 100.00 | 96.61 |
| *Forehand Serve* | 99.17 | 98.15 | 99.70 | 100.00 | 98.86 | 99.37 | 100.00 | 99.32 |
| *Backhand Serve* | 96.25 | 90.28 | 98.51 | 96.73 | 82.57 | 92.17 | 100.00 | 93.79 |
| *Intercept* | 79.58 | 64.81 | 73.55 | 98.81 | 79.27 | 96.97 | 98.61 | 84.51 |
| *Lift* | 88.75 | 90.74 | 93.75 | 100.00 | 99.62 | 99.75 | 98.47 | 95.87 |
| *Net Shot* | 88.33 | 97.68 | 100.00 | 99.70 | 93.56 | 98.61 | 100.00 | 96.84 |
| *Smash* | 98.33 | 69.44 | 77.68 | 100.00 | 90.15 | 100.00 | 98.70 | 90.61 |
| **F1-Score** (%) ↑ | 92.17 | 87.44 | 92.36 | 99.42 | 92.15 | 98.35 | 99.47 | - |

Table 3: Acc. (%) comparison of OMONet and leading methods across badminton actions in the proposed 1-BEMG benchmark.

**Datasets:** OMONet was evaluated for EMG-based action recognition on the proposed **1**D-**B**adminton **EMG** and the open-source Myo (Michael Lohr 2019) datasets, and its generalization to physiological signals was further tested through the open-source Yaseen18 (Yaseen, Son, and Kwon 2018) dataset. For overview: (1) the **1-BEMG** comprises 960 EMG clips from six experienced badminton players (3 male, 3 female), recorded via 2SEMG Sensors on their dominant arms during eight representative badminton actions (provided in Figure 4), each repeated 120 times; (2) the **Myo** contains 1,890 surface EMG clips from 13 subjects performing rock-paper-scissors gestures; (3) the **Yaseen18** provides 1,000 heart sound clips in 5 classes: normal and four pathological cardiac conditions. All datasets were randomly divided into 6: 2: 2 for training, validation, and testing.

**Training Details:** OMONet was optimized using a batch size of 32, the Adam optimizer (Kingma and Ba 2014), and the Cross-Entropy Loss. The initial learning rate was set to $10^{-3}$, with a minimum threshold of $10^{-12}$. The learning rate was reduced by 10% if the loss decreased by $<10^{-4}$ for 3 consecutive epochs, and the training was terminated if the decrease in loss was $<10^{-4}$ lasted for 10 consecutive epochs.

Each model was independently trained 20 times. Experiments were conducted using Python 3.9 (Ubuntu 20.04), TensorFlow 2.13.0 (Keras 2.13.1), and CUDA 11.8. Training and testing were performed on an NVIDIA GeForce RTX 3060 Ti GPU, while inference latency and memory usage records were executed on an STM32H7S78-DK MCU.

## Main Results

The comparison evaluates leading temporal signal processing models and those oriented toward EMG or similar physiological signals. Table 2 reveals that OMONet achieves an accuracy gap of $<3\%$ compared to ML-TCN on the EMG datasets, while using only $\sim$0.2% of its parameters and $\sim$1% of its FLOPs and MACCs, verifying its exceptional efficiency advantage over models of high complexity while maintaining comparable performance. Moreover, OMONet exhibits uniformly superior performance compared to ML-TCN on the non-EMG Yaseen18 dataset, validating its generalization advantage. Compared to lightweight models such as STM32-deployable DACNN, OMONet exhibits higher F1-score across all datasets, thus confirming its significant performance superiority over high-efficiency models.

| Ablation Targets | Model Complexity | Deployment Efficiency | Datasets | Acc. (%) | F1-Score (%) |
|---|---|---|---|---|---|
| w/o Motion Attention | Params (M): 0.036$_{(-0.042)}$<br>FLOPs (G): 0.096$_{(-0.072)}$<br>MACCs (G): 0.048$_{(-0.037)}$ | Flash Usage (KB): 235.4$_{(-91.2)}$<br>RAM Usage (KB): 429.4$_{(-126.1)}$<br>Inference Latency (s): 0.25$_{(-0.08)}$ | 1-BEMG<br>Myo<br>Yaseen18 | 90.65$_{(-8.82)}$*<br>73.80$_{(-9.53)}$*<br>85.50$_{(-10.00)}$* | 89.91$_{(-9.56)}$*<br>73.62$_{(-9.71)}$*<br>85.49$_{(-9.90)}$* |
| w/o Freq-Motion | Params (M): 0.055$_{(-0.023)}$<br>FLOPs (G): 0.115$_{(-0.053)}$<br>MACCs (G): 0.058$_{(-0.027)}$ | Flash Usage (KB): 244.0$_{(-82.6)}$<br>RAM Usage (KB): 436.3$_{(-119.2)}$<br>Inference Latency (s): 0.25$_{(-0.08)}$ | 1-BEMG<br>Myo<br>Yaseen18 | 97.07$_{(-2.40)}$*<br>76.98$_{(-6.35)}$*<br>87.00$_{(-8.50)}$* | 97.18$_{(-2.29)}$*<br>77.02$_{(-6.31)}$*<br>87.04$_{(-8.35)}$* |
| w/o Time-Motion | Params (M): 0.060$_{(-0.018)}$<br>FLOPs (G): 0.115$_{(-0.053)}$<br>MACCs (G): 0.058$_{(-0.027)}$ | Flash Usage (KB): 252.6$_{(-74.0)}$<br>RAM Usage (KB): 437.7$_{(-117.8)}$<br>Inference Latency (s): 0.27$_{(-0.06)}$ | 1-BEMG<br>Myo<br>Yaseen18 | 95.24$_{(-4.23)}$*<br>77.77$_{(-5.56)}$*<br>80.16$_{(-15.34)}$* | 95.17$_{(-4.30)}$*<br>76.32$_{(-7.01)}$*<br>80.07$_{(-15.32)}$* |
| w/o 1D-Mean | Params (M): 0.056$_{(-0.022)}$<br>FLOPs (G): 0.225$_{(+0.057)}$<br>MACCs (G): 0.116$_{(+0.031)}$ | Flash Usage (KB): 238.5$_{(-88.1)}$<br>RAM Usage (KB): 1269.8$_{(+714.3)}$<br>Inference Latency(s): 10.45$_{(+10.12)}$ | 1-BEMG<br>Myo<br>Yaseen18 | 93.75$_{(-5.72)}$*<br>74.07$_{(-9.26)}$*<br>66.66$_{(-28.84)}$* | 93.28$_{(-6.19)}$*<br>74.01$_{(-9.32)}$*<br>66.66$_{(-28.73)}$* |
| w/o Post-Fusion | Params (M): 0.047$_{(-0.031)}$<br>FLOPs (G): 0.159$_{(-0.009)}$<br>MACCs (G): 0.081$_{(-0.004)}$ | Flash Usage (KB): 209.9$_{(-116.7)}$<br>RAM Usage (KB): 552.3$_{(-3.2)}$<br>Inference Latency (s): 0.32$_{(-0.01)}$ | 1-BEMG<br>Myo<br>Yaseen18 | 90.33$_{(-9.14)}$*<br>75.13$_{(-8.20)}$*<br>85.01$_{(-10.49)}$* | 90.21$_{(-9.26)}$*<br>74.98$_{(-8.35)}$*<br>84.92$_{(-10.47)}$* |

Values in () represent changes relative to full OMONet, with * **marking statistically significant differences (Wilcoxon p-value$<$0.05)**.

Table 4: Ablation on the attention mechanism with two sub motion paths and the other innovative modules in OMONet.

Furthermore, the deployment results demonstrate that OMONet efficiently utilizes memory with a Flash cost $\sim$60% of its Random Access Memory (RAM) usage, and outperforms the non-deployable MILLET by achieving $\sim$5% higher average accuracy. Compared to other efficient models that can also be deployed on STM32, OMONet exhibits a maximum accuracy advantage of $\sim$16%, verifying its feasibility in resource-constrained wearable devices.

As indicated in Table 3, OMONet stands out as the **only model** that maintains $>$98% accuracy in all eight action categories in 1-BEMG dataset, and achieves 100% recognition accuracy in *ready*, *clear shot*, *forehand serve*, *backhand serve*, and *net shot*, while other models only achieve $<$85% average accuracy on *intercept*. These results underscore OMONet's superior robustness and detection balance, validating its efficacy in learning complex EMG patterns.

Overall, many up-to-date advanced models still face bottlenecks due to the similarity of actions and model complexity. OMONet's superiority stems from its time-freq-decoupled learning architecture, which captures fundamental temporal features via a 1D-Mean module and couples them with a dual-stream attention network. This network deconstructs and automatically fuses critical time and frequency combinations, such as rhythm and explosive force, accurately distinguishing actions commonly confused by other models. The independent fast-slow decoupling and recoupling of time-frequency features make this design particularly suitable for detecting badminton actions with varying intensities and amplitudes but overlapping information.

## Ablation Study

**Effect of Motion Attention.** As shown in Table 4, Motion Attention accounts for $\sim$50% of OMONet's parameters and $\sim$40% of its FLOPs and MACCs, yet it contributes to $<$25% of inference latency, validating the efficiency of the dual-stream architecture for parallel time-frequency decoupling. Removing Motion Attention significantly degrades OMONet's performance due to the loss of discrete time-frequency features (reducing the 2D-oriented Post-Fusion module to conventional convolution layers) and results in $\sim$10% accuracy and F1-score drops across all datasets. This highlights the critical role of Motion Attention in enhancing OMONet's generalization across various physiological signal recognition tasks under different action patterns.

**Effect of Freq-Motion and Time-Motion.** Table 4 indicates consistent RAM in models with a single motion perception path *vs.* Motion Attention, validating the efficiency of the time-frequency dual-stream architecture. It also shows that removing a motion perception path that accounts for $\sim$20% of memory usage, significantly degrades model performance, with the removal of Time-Motion causing an up-to-2$\times$ drop in accuracy and F1-score compared to removing Freq-Motion. These results suggest that relying solely on a single motion perception path reduces feature informativeness by neglecting the other dimension, underscoring the necessity of independent time-frequency learning and subsequent recoupling for 1D signals. Moreover, the results show that temporal continuity outweighs frequency-domain energy distribution in influencing action recognition for physiological signals, as evidenced by OMONet's greater performance drop when excluding the Time-Motion.

**Effect of 1D-Mean.** As indicated in Table 4, removing the 1D-Mean module, which accounts for $\sim$30% of OMONet's parameters, results in a more pronounced performance degradation compared to eliminating a single motion perception path. Despite a $<$30% reduction in Flash footprint, inference latency and RAM usage increased by $\sim$3000% and $\sim$130%, respectively. These observations indicate that 1D-Mean not only accelerates inference by compressing temporal and frequency intervals to compress the dual-stream input size, but also proves that signals processed by 1D-Mean introduce more distinctive time-frequency features for further learning by decoupled Motion Attention.

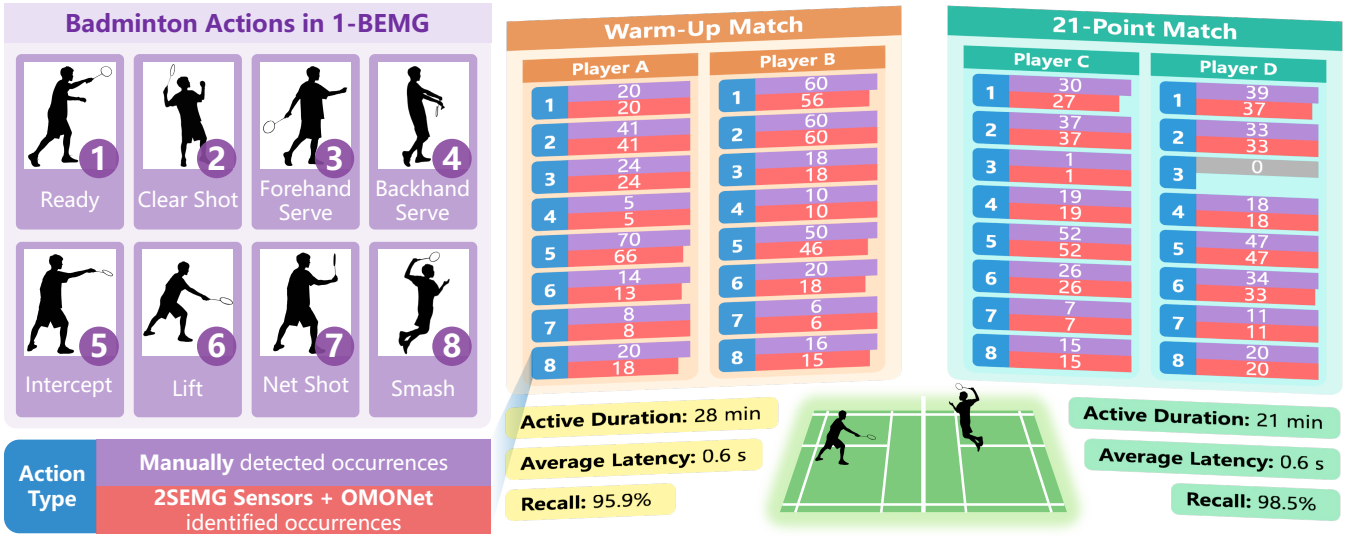


Figure 4: Eight badminton action categories in the proposed 1-BEMG dataset and validation of the 2SEMG Sensor and OMONet-based whole-field action sensing network through real-world warm-up and 21-point badminton matches.

**Effect of Post-Fusion.** The impact of removing Post-Fusion on accuracy and F1-score closely resembles that of eliminating Motion Attention, as indicated in Table 4. Moreover, the Post-Fusion induces ∼0 extra RAM usage and inference time, verifying its lightweight design that does not impose additional computation burden on the OMONet with Motion Attention. These synchronized results also highlight the indispensable link between the time-frequency decoupling-recoupling of Motion Attention and Post-Fusion, where Post-Fusion stage reconstructs 2D time-frequency synchronization from two motion perception paths to explore time-frequency cooperative patterns.

## Real-World Evaluation

**Evaluation Settings.** It includes two badminton matches: (1) a one-on-one warm-up match that incorporates free movement to evaluate actions seldom observed in formal contests, and (2) a one-on-one 21-point formal competition.

All athletes wore a 2SEMG Sensor on their dominant arms for continuous acquisition of EMG signals, with data transmitted to an STM32H7S78-DK MCU hub for real-time concurrent action recognition. Figure 4 illustrates the eight types of badminton action that OMONet needs to detect.

**Case Study.** As illustrated in Figure 4, both badminton matches demonstrate athlete-specific action distributions, which align with the frequency patterns of actions observed in other badminton competitions. This alignment verifies the integrity of the EMG signals captured via 2SEMG Sensors.

The 2SEMG Sensors-OMONet network did not attain 100% recognition accuracy for detecting *ready*, *intercept*, *lift*, and *smash* actions, consistent with the results in Table 3, except for the *ready* position. Given the pronounced inherent ambiguity and individual habit variations in the *ready* position compared to other actions, the stable missing rate of the *ready* position in both matches is understandable. For *smash*, while achieving ∼92% recall in the warm-up match, it reaches 100% recall in the formal match. This indicates that the completeness rate of the action influences the recognition results, as many actions during warm-ups are often not fully executed. Therefore, the *intercept* and *lift* actions, which exhibit ∼7% missing rates in the warm-up match but exhibit near perfect recognition in the formal match (recall>99%), resulted not only from their inherent similarity (as the left panel of Figure 4 reveals the recognition challenges caused by the action similarity), but also from the lower rate of the action completeness in warm-ups.

Overall, the real-time, high-fidelity EMG acquisition capability of the 2SEMG Sensor enables OMONet to achieve >98% recall in formal match scenarios with <1 s end-to-end latency, thus validating the practical viability of the proposed solution for real-world whole-field action sensing.

## Conclusion

This study aims to scale action sensing from individual to more ubiquitous group-level applications. Therefore, we developed a wearable and low-power 2SEMG Sensor to capture athletes' EMG signals, and introduced a lightweight time-freq-decoupled learning OMONet that enables reliable concurrent action signal recognition with minimal computational resource consumption. Among them, the 7.2-g wireless 2SEMG Sensor has negligible impact on wearers, and the OMONet demonstrates leading performance and efficiency on various physiological signal datasets. To validate the feasibility of the proposed 2SEMG Sensors-OMONet network for whole-field action sensing, the widely participated badminton characterized by vigorous, dynamic actions among multiple players was selected as a representative case study. Finally, the action sensing results from two badminton matches confirm the real-time capability and reliability of the proposed solution, while revealing the feasibility of extending this study to other multi-person scenarios.

## Acknowledgements

We sincerely appreciate the recognition of this work by the Program Chairs, Area Chairs, Senior Program Committee, and Program Committee. This project was jointly supported by the National Natural Science Foundation of China (Grant Nos. 62372364, 62272380, 12247101, 62374077, and 62404088), the Technical Innovation Guidance Plan of Shaanxi Province, China (Grant No. 2024QCY-KXJ-199), the Key Research and Development Projects in Gansu Province, China (Grant No. 25YFFA030), the Joint Research Fund of Gansu Province, China (Grant No. 24JRRA818), the Key Project of Natural Science Foundation of Gansu Province, China (Grant No. 24JRRA395), the Natural Science Foundation of Gansu Province, China (Grant Nos. 25JRRA659 and 22JR5RA389), the '111 Center' under Grant No. B20063, and the Fundamental Research Funds for the Central Universities (Grant Nos. lzujbky-2024-jdzx06 and lzujbky-2024-it48).

## References

Abuhoureyah, F.; Swee, S. K.; and Chiew, W. Y. 2024. Multi-user human activity recognition through adaptive location-independent WiFi signal characteristics. *IEEE Access*, 12: 112008–112024.

Bao, T.; Zaidi, S. A. R.; Xie, S.; Yang, P.; and Zhang, Z.-Q. 2020. A CNN-LSTM hybrid model for wrist kinematics estimation using surface electromyography. *IEEE Transactions on Instrumentation and Measurement*, 70: 1–9.

Baskaran, P.; and Adams, J. A. 2023. Multi-dimensional task recognition for human-robot teaming: Literature review. *Frontiers in Robotics and AI*, 10: 1123374.

Cárdenas-Valdez, J. R.; Corral-Domínguez, Á. H.; de Jesús García-Ortega, M.; Calvillo-Téllez, A.; Hurtado-Sánchez, C.; Inzunza-González, E.; et al. 2023. EMG signal transmission system under RF schemes. *Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*, 11: 277–282.

Carling, C.; Bloomfield, J.; Nelsen, L.; and Reilly, T. 2008. The role of motion analysis in elite soccer: Contemporary performance measurement techniques and work rate data. *Sports Medicine*, 38(10): 839–862.

Chamberland, F.; Buteau, É.; Tam, S.; Fortier, P.; Boukadoum, M.; Campeau-Lecours, A.; and Gosselin, B. 2023. EMaGer: A wearable full-circumference HD-EMG sensor and data augmentation method for robust hand gesture recognition. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1–5. IEEE.

Chollet, F. 2017. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1251–1258. IEEE.

Dao, T.; and Gu, A. 2024. Transformers are SSMs: Generalized models and efficient algorithms through structured state space duality. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, 10041–10071. PMLR.

Early, J.; Cheung, G. K.; Cutajar, K.; Xie, H.; Kandola, J.; and Twomey, N. 2024. Inherently interpretable time series classification via multiple instance learning. In *The Twelfth International Conference on Learning Representations*. OpenReview.

Huang, J.; Liu, B.; Miao, C.; Lu, Y.; Zheng, Q.; Wu, Y.; Liu, J.; Su, L.; and Chen, C. W. 2021. PhaseAnti: An anti-interference WiFi-based activity recognition system using interference-independent phase component. *IEEE Transactions on Mobile Computing*, 22(5): 2938–2954.

Jiang, X.; An, B.; Zhao, G.; and Qian, X. 2025. M3Net: Efficient time-frequency integration network with mirror attention for audio classification on edge. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(17): 17644–17652.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. In *The Second International Conference on Learning Representations*. OpenReview.

Le, K.-N. T.; Byun, G.; Raza, S. M.; Le, D.-T.; and Choo, H. 2025. Respiratory anomaly and disease detection using multi-level temporal convolutional networks. *IEEE Journal of Biomedical and Health Informatics*, 29(7): 4834–4846.

Lee, H.; Lee, S.; Kim, J.; Jung, H.; Yoon, K. J.; Gandla, S.; Park, H.; and Kim, S. 2023. Stretchable array electromyography sensor with graph neural network for static and dynamic gestures recognition system. *npj Flexible Electronics*, 7(1): 20.

Li, S.; Zhu, T.; Duan, F.; Chen, L.; Ning, H.; Nugent, C.; and Wan, Y. 2025. HARMamba: Efficient and Lightweight Wearable Sensor Human Activity Recognition Based on Bidirectional Mamba. *IEEE Internet of Things Journal*, 12(3): 2373–2384.

Liu, D.; Tian, X.; Bai, J.; Wang, S.; Dai, S.; Wang, Y.; Wang, Z.; and Zhang, S. 2024. A wearable in-sensor computing platform based on stretchable organic electrochemical transistors. *Nature Electronics*, 7(12): 1176–1185.

Lu, C.; Xu, X.; Liu, Y.; Li, D.; Wang, Y.; Xian, W.; Chen, C.; Wei, B.; and Tian, J. 2024. An embedded electromyogram signal acquisition device. *Sensors*, 24(13): 4106.

Ma, C.; Lin, C.; Samuel, O. W.; Guo, W.; Zhang, H.; Greenwald, S.; Xu, L.; and Li, G. 2021. A bi-directional LSTM network for estimating continuous upper limb movement from surface electromyography. *IEEE Robotics and Automation Letters*, 6(4): 7217–7224.

Michael Lohr, B. 2019. Myo Dataset. https://github.com/michidk/myo-dataset.

Moin, A.; Zhou, A.; Rahimi, A.; Menon, A.; Benatti, S.; Alexandrov, G.; Tamakloe, S.; Ting, J.; Yamamoto, N.; Khan, Y.; et al. 2021. A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition. *Nature Electronics*, 4(1): 54–63.

Montazerin, M.; Rahimian, E.; Naderkhani, F.; Atashzar, S. F.; Yanushkevich, S.; and Mohammadi, A. 2023. Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals. *Scientific Reports*, 13(1): 11000.

Ng, C. L.; Reaz, M. B. I.; Crespo, M. L.; Cicuttin, A.; Shapiai, M. I. B.; Ali, S. H. B. M.; and Chowdhury, M. E. H. 2024. A versatile and wireless multichannel capacitive EMG measurement system for digital healthcare. *IEEE Internet of Things Journal*, 11(11): 20120–20137.

Pau, D. P.; and Aymone, F. M. 2024. Mathematical formulation of learning and its computational complexity for transformers' layers. *Eng*, 5(1): 34–50.

Rehman, A.; Moussa, M.; Saleh, H.; Werghi, N.; Khraibi, A.; and Khandoker, A. 2024. Chin EMG sscalogram-based deep CNN for OSA screening. In *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1–4. IEEE.

Shannon, C. E. 1949. Communication in the presence of noise. *Proceedings of the IRE*, 37(1): 10–21.

Tavakoli, M.; Benussi, C.; and Lourenco, J. L. 2017. Single channel surface EMG control of advanced prosthetic hands: A simple, low cost and efficient approach. *Expert Systems with Applications*, 79: 322–332.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L. u.; and Polosukhin, I. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Wang, Y.; Xu, Y.; Yang, J.; Wu, M.; Li, X.; Xie, L.; and Chen, Z. 2024. Graph-aware contrasting for multivariate time-series classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(14): 15725–15734.

Wattanasiri, P.; Wilson, S.; Huo, W.; and Vaidyanathan, R. 2025. Gesture recognition through mechanomyogram signals: An adaptive framework for arm posture variability. *IEEE Journal of Biomedical and Health Informatics*, 29(4): 2453–2462.

Wei, C.; Wang, H.; Hu, F.; Zhou, B.; Feng, N.; Lu, Y.; Tang, H.; and Jia, X. 2022. Single-channel surface electromyography signal classification with variational mode decomposition and entropy feature for lower limb movements recognition. *Biomedical Signal Processing and Control*, 74: 103487.

Wu, Y.-D.; Ruan, S.-J.; and Lee, Y.-H. 2021. An ultra-low power surface EMG sensor for wearable biometric and medical applications. *Biosensors*, 11(11): 411.

Xiong, B.; Chen, W.; Li, H.; Niu, Y.; Zeng, N.; Gan, Z.; and Xu, Y. 2024a. Patchemg: Few-shot emg signal generation with diffusion models for data augmentation to improve classification performance. *IEEE Transactions on Instrumentation and Measurement*, 73: 1–14.

Xiong, D.; Zhang, D.; Chu, Y.; Zhao, Y.; and Zhao, X. 2024b. Intuitive human-robot-environment interaction with EMG signals: A review. *IEEE/CAA Journal of Automatica Sinica*, 11(5): 1075–1091.

Yang, J.; Shibata, K.; Weber, D.; and Erickson, Z. 2025. High-density electromyography for effective gesture-based control of physically assistive mobile manipulators. *npj Robotics*, 3(1): 2.

Yang, J.; Soh, M.; Lieu, V.; Weber, D. J.; and Erickson, Z. 2024. EMGBench: Benchmarking out-of-distribution generalization and adaptation for electromyography. In *Advances in Neural Information Processing Systems*, volume 37. Curran Associates, Inc.

Yaseen; Son, G.-Y.; and Kwon, S. 2018. Classification of heart sound signal using multiple features. *Applied Sciences*, 8(12): 2344.

Ye, Y.; He, Y.; Pan, T.; Dong, Q.; Yuan, J.; and Zhou, W. 2023. Cross-subject EMG hand gesture recognition based on dynamic domain generalization. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1–4. IEEE.

Yuan, Y.; Dong, A.; Xu, W.; Han, Y.; Yu, J.; and Zhou, Y. 2025. TransGER: Transformer-based CNN-BiGRU architecture for sEMG gesture recognition in time-frequency domain. In *International Conference on Wireless Artificial Intelligent Computing Systems and Applications*, 297–306. Springer.

Zeng, X.; Shi, Y.; and Zhou, A. 2022. Multi-har: Human activity recognition in multi-person scenes based on mmwave sensing. In *2022 IEEE 8th International Conference on Computer and Communications (ICCC)*, 1789–1793. IEEE.

Zhang, T.; Sun, H.; and Zou, Y. 2022. An electromyography signals-based human-robot collaboration system for human motion intention recognition and realization. *Robotics and Computer-Integrated Manufacturing*, 77: 102359.