

sEMG-based Gesture-Free Hand Intention Recognition: System, Dataset, Toolbox, and Benchmark Results

Hongxin Li, Jingsheng Tang*, Xuechao Xu, Wei Dai, Yaru Liu, Junhao Xiao, Huimin Lu, IEEE Member, and Zongtan Zhou

Abstract—In sensitive scenarios, such as meetings, negotiations, and team sports, messages must be conveyed without detection by non-collaborators. Previous methods, such as encrypting messages, eye contact, and micro-gestures, had problems with either inaccurate information transmission or leakage of interaction intentions. To this end, a novel gesture-free hand intention recognition scheme was proposed, that adopted surface electromyography (sEMG) and isometric contraction theory to recognize different hand intentions without any gesture. Specifically, this work includes four parts: (1) the experimental system, consisting of the upper computer software, self-conducted myoelectric watch, and sports platform, is built to get sEMG signals and simulate multiple usage scenarios; (2) the paradigm is designed to standard prompt and collect the gesture-free sEMG datasets. Eight-channel signals of ten subjects were recorded twice per subject at about 5-10 days intervals; (3) the toolbox integrates preprocessing methods (data segmentation, filter, normalization, etc.), commonly used sEMG signal decoding methods, and various plotting functions, to facilitate the research of the dataset; (4) the benchmark results of widely used methods are provided. The results involve single-day, cross-day, and cross-subject experiments of 6-class and 12-class gesture-free hand intention when subjects with different sports motions. To help future research, all data, hardware, software, and methods are open-sourced on the following website: [click here](#).

Index Terms—Gesture-free hand intention recognition, surface electromyographic (sEMG), open-source dataset, benchmark results, human-computer interaction.

I. INTRODUCTION

IN the scenario where collaborators and non-collaborators coexist, how to achieve message transmission between collaborators without being noticed by non-collaborators is of great significance for team cooperation in conference negotiations or team sports. Previous research has focused on the concealment of interactive information, ensuring that it is difficult for non-collaborators to decode. For example, Katriel et al. [1] utilize end-to-end encryption and decryption technologies to prevent messages from being stolen and exploited. Another typical approach is to develop a specialized code language for collaborators, such as eye contact [2] and gestures [3], [4], that

are difficult to understand for non-collaborators. Although the above methods prevent non-collaborators decode the meaning of the message, the process of sending the message is exposed, which may easily arouse the perception and vigilance of non-collaborators. In recent years, extensive research on human-computer interaction has satisfied the concealment, but there are shortcomings in other aspects. Typically, brain-computer interaction directly decodes human intentions for transmitting messages without body movements [5], but has problems such as dependence on stimulus paradigms, low accuracy, and poor wearing comfort. Gaze-based interaction only produces slight head displacement, but it is limited by the Midis touch problem [6] in high-eye load application scenarios. Therefore, it is necessary to design a novel interaction method, that can convey messages accurately without obvious body movements or explicit signals.

The gesture recognition system based on surface electromyography (sEMG) signals has been widely studied and achieved robust performance [4]. As a weak electrobiological signal, it reflects information related to surface muscle and bone activity in the human body. It is a motor unit action potential train sequence generated by the recruited motor units during muscle excitation, which is filtered by muscles, subcutaneous fat, and skin tissue, and superimposed on the electrode on the skin surface to form a signal. The sEMG signals are the neural electrical signals accompanying muscle contractions, which can truly reflect the intention of the human body to make movements, and the relationship between intention and movement is not sufficient or necessary. A typical case is the isometric contraction of muscles [7], where the muscles produce excitation and output force, but their length remains unchanged. The muscle isometric contraction refers to the contraction of muscles that maintain a constant length while experiencing changes in tension. In this contracted state, muscle tension can increase to its maximum. However, due to the absence of displacement, physically speaking, muscles do not perform any external work, yet still require a significant amount of energy consumption. The active contraction of muscle components generates significant tension, which causes the elastic components of the muscle to elongate and counteract. Therefore, regardless of whether the limbs produce movements or not, sEMG signals can always accurately determine whether a person has the intention to control limb movements. This is the theoretical basis for selecting sEMG signals to recognize gesture-free hand intentions in this work.

This work was supported in part by the Science and Technology Innovation 2030 “Brain Science and Brain-Like Research” Key Project of China under Grant 2022ZD0208504; in part by the National Natural Science Foundation of China under Grant U22A2059, Grant U1913202, and Grant 62203460; and in part by the Major Project of the Natural Science Foundation of Hunan, China, under Grant 2021JC0004.

All authors are with the College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China.

* Corresponding author (e-mail: mrtang@nudt.edu.cn).

Based on the isometric contraction theory, the experimental system is built to acquire sEMG data without any gesture. The system first designs a myoelectric wristband, that can acquire eight-channel sEMG signals and three-axis acceleration signals with a 500Hz sampling rate. The supporting host computer software has two pages, one is to help users control, store, and visualize the data, and another is used to control the experimental paradigm and prompt subjects to apply the correct hand force. Considering users may use this interaction technology in various sports scenarios, we have specially customized a sports platform that supports simulating usage at different sports speeds.

To standardize the recording process of sEMG, the corresponding experimental paradigm has been designed. The experimental procedure consists of 12 blocks \times 12 trials of data recording. The trial number corresponds one-to-one with the hand intention. Block denotes the current sports state of the subject, which includes three rounds of switching between four different sports states. Each trial records the sEMG signals for ten seconds with the prompts of voice, text, image, and progress bar. Ten subjects take part in the above procedure twice, with an interval of 5-10 days. The data format has 15 channels, which denotes the real-time sEMG data, IMU data, and information on the block, speed, time stamp, and trigger, respectively.

To promote future research, we have developed a toolbox for this dataset that supports the function for data preprocessing, sEMG classification, and visualization. The data preprocessing section converts the collected sEMG dataset into data samples for model training through steps such as data slicing, filtering, baseline correction, and dataset preparation. We have summarized and reproduced the representively sEMG signal decoding methods in recent years, and will use them as the baseline to test the classification performance of the gesture-free hand intention recognition task. Finally, the toolbox provides most of the visualization analyse methods, such as t-distributed stochastic neighbor embedding (t-SNE), confusion matrix, spectrogram, and etc.

Based on the dataset recorded by the experimental system and paradigm, and the analysis tools and methods supported by the toolbox, the mainstream sEMG classification methods have been tested and the benchmark results for the gesture-free hand intention recognition task have been provided. The experiments involve three types: single-day, cross-day, and cross-subject. To adapt to the different application scenarios, the length of the time window is set to 250ms, 500ms, and 750ms, respectively. The classification categories are set to 6 and 12. The former only includes samples of single-finger force and resting state, while the latter expands samples of multiple-finger combination force. In addition, we further obtained the details of the classification results by the confusion matrix tool and analyzed the impact of different sports modes on the task. Finally, we conducted online experiments to demonstrate the effectiveness of the proposed scheme.

This work first proposes the scheme of gesture-free hand intention recognition based on the sEMG signal and the isometric contraction theory. The detailed contribution and significance of this work lie in:

- 1) Designed a sEMG-based gesture-free hand intention recognition experimental system, which enhances the concealment of information transmission while ensuring accurate and efficient information transmission;
- 2) Based on this system, we designed the paradigm and recorded the gesture-free sEMG dataset containing ten subjects. Considering real-world application scenarios, the data is collected when the subjects are standing, slow, fast, and jogging;
- 3) For the convenience of data analysis, we have developed a data analysis toolbox, which includes functions such as dataset management, data processing, and visualization analysis;
- 4) To promote further research, benchmark results of representative sEMG signal processing algorithms were provided on this dataset.

II. EXPERIMENTAL SYSTEM

As illustrated in Fig. 1, the experimental system consists of three parts: 1) the self-conducted myoelectric wristband, which is designed for recording the real-time sEMG signals. The sub-figure on the right shows the details of external structure and PCB circuit design; 2) the matched host computer software, consists of two graphical user interface (GUI) pages. The first page is responsible for managing data, including the functions of data receiving, storage, and visualization, while the second page is responsible for managing paradigms, including the functions of paradigm control and task prompts; 3) Umay U3H sports platform, considering the different application scenario, we adopt a treadmill to simulate the various motion patterns possible for the user. The detailed introduction of the first two parts is as follows:

A. Self-conducted myoelectric wristband

Before designing the myoelectric wristband, we considered the problem of the sEMG source, which part of the body is used to collect the signals. Due to the well-developed muscle group in the forearm, previous studies have often recorded sEMG signals in this area, but this may not be effective for gestures involving fine finger movements. Botros et al. [8], [9] show that classifiers trained and tested using wrist EMG signals are significantly higher than the former in both single-finger and multi-finger gestures. This is because controlling the muscle group extension range of human fingers covers the wrist, which means that electromyographic signals representing human finger movement intention can be effectively collected in the wrist, demonstrating the theoretical basis for realizing gesture-free hand intention recognition. In addition, using wrist EMG signals can be more easily integrated into wristband devices, making consumers wear more comfortable.

From part A of Fig. 1, the myoelectric wristband consists of two parts: the host and the strap. The host is composed of a PCB circuit board and a corresponding outer shell, while the strap is made of soft silicone. The wires between each electrode and the host are hidden in an inner silicone composite molding technology cavity. The spindle adopts a hollow design, and the cables in the silicone band are connected to the

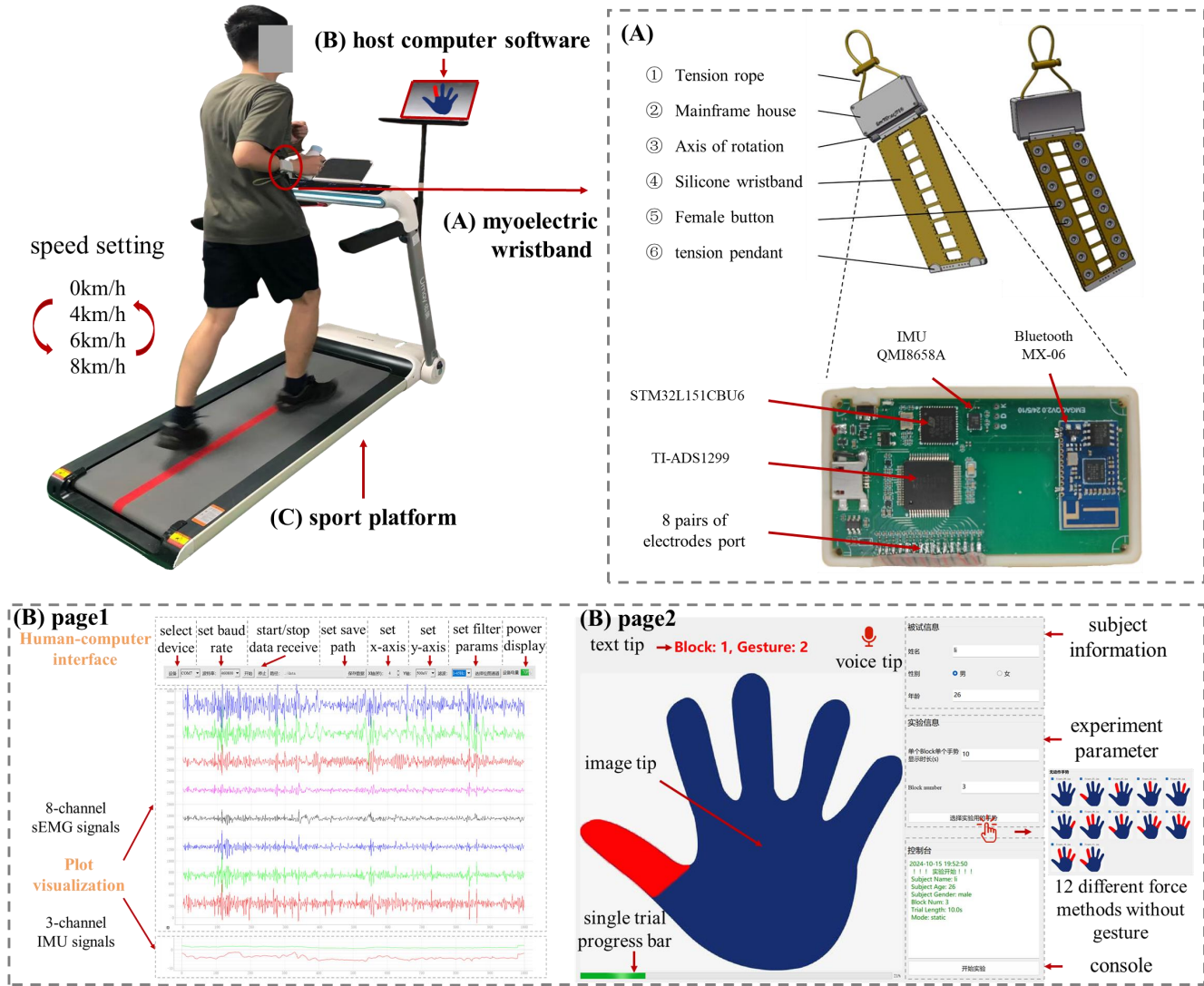


Fig. 1. Experimental system for sEMG-based gesture-free hand intention recognition. It comprises the self-conducted myoelectric wristband, the matched host computer software, and the Umay U3H sports platform. The myoelectric wristband is responsible for recording real-time sEMG signals; The host computer software has two pages, one is designed for data parameter setting and visualization, and another is for paradigm parameter settings and task prompts; The sports platform simulates different application scenarios.

PCB main board of the main body casing through the hollow spindle. Eight pairs of 4.0mm medical standard female buckles are arranged longitudinally on the inside of the wristband, which is compatible to be used with electrode pads of 3.9mm standard interface. Considering the requirement for low power consumption, the STM32L151CUB6 chip is selected to control the data acquisition and forwarding process. In terms of analog-to-digital converters (ADC), we have chosen ADS1299 from Texas Instruments (TI) company to collect weak surface electromyography signals, which is an integrated acquisition chip specifically designed for physiological electrical signals. Its features include an eight-channel 24 bit ADC analog-to-digital sampling module with a programmable gain programmable gain amplifier (PGA) added to the front-end, as well as a right leg drive circuit for denoising. In addition, a three-axis accelerometer supported by QMI8658A is built to measure x, y, and z three-dimensional acceleration signals. Based on the size of the comprehensive data transmission,

speed requirements, and power consumption, the system selects MX-06 and uses Bluetooth SPP protocol to transmit the sEMG data to the host computer software. Overall, the self-conducted myoelectric wristband supported to collect and transmit the eight-channel wrist sEMG signals for monitoring the tension/relaxation state of the hand muscles or the intention of the movement, and 2) three-axis acceleration signals for monitoring the movement state of the wearer.

B. Design of the host computer software

From part B of Fig. 1, the host computer consists of two pages that switch with a button: (1) data manager; and (2) paradigm manager, which uses a button to switch. These graphical user interfaces are developed to receive the sEMG data, visualize the multi-channel signals, and record necessary experimental information. The following is a detailed introduction to the functions of each page:

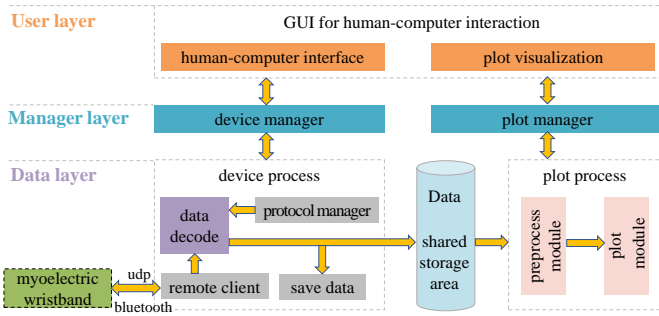


Fig. 2. Software architecture of the page1: data manager, which consists of three layers, i.e. the user layer for human-computer interface; the manager layer controls device and plot operations; and the data layer for data save, decode, preprocess, and plot.

1) *Page1: data manager:* The host computer has three real-time functions, i.e. receiving the byte stream data from the myoelectric wristband; decoding the sEMG signals based on the data protocol; and visualizing the multi-channel signals. Fig. 2 shows the hierarchical software architecture of this page, which consists of user, manager, and data layers. The GUI of the user layer is shown in Fig. 1(B1), implemented based on the PyQt library. The top column is operable and used to set parameters for various functions. The following two columns cannot be operated, they are used to display the waveforms of sEMG signals and acceleration signals, respectively. All functions are divided into device functions and plot functions, running on two different threads and sharing data through memory-sharing technology. To facilitate the management of functions, we have set up two managers to implement the classification management and calling of functions.

2) *Page2: paradigm manager:* This page is designed for paradigm parameter settings and task prompts. To facilitate information management and assist subjects in accurately completing experimental tasks, we have developed a graphical interface that prompts subjects to switch and maintain the corresponding gesture-free hand intention state from four aspects: images, text, sound, and progress bar. The GUI is shown in Fig. 1(B2), with the experimental prompt area on the left and the experimental settings area on the right.

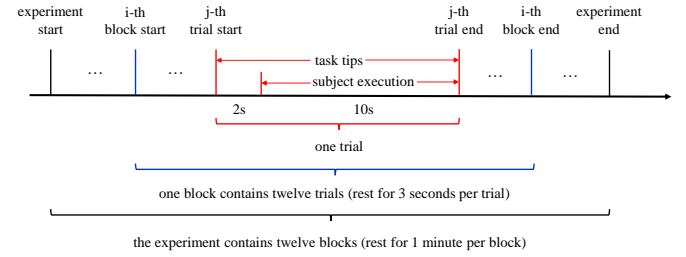
III. PARADIGM AND DATASET

Based on the above experimental system, an experimental paradigm was designed to construct the sEMG-based gesture-free hand intention recognition dataset. The paradigm collected eight-channel sEMG data and three-channel accelerometer signals with a 500Hz sampling rate from ten subjects.

A. Subjects

Before each experiment procedure, the subject needs to fill in personal information in the upper right corner of Fig. 1(B2). According to the statistics, ten subjects (6 males and 4 females, aged between 22-28 years old) participated in our experiments. All subjects had no muscle or nervous system injuries. The experiment strictly adheres to the provisions of the Helsinki Declaration, providing a detailed introduction to

the experimental process and signing an informed consent form before the experiment begins. It should be noted that each subject is required to complete two experiments, with an interval of 5-10 days, to construct a cross-day dataset.



No.	1	2	3	4	5	6	7	8	9	10	11	12
Trial												
Block	0 km/h	4 km/h	6 km/h	8 km/h	0 km/h	4 km/h	6 km/h	8 km/h	0 km/h	4 km/h	6 km/h	8 km/h

Fig. 3. The sEMG-based gesture-free hand intention recognition paradigm. Each experiment includes 12 blocks and each block includes 12 trials. The motion speed corresponds to each block and the hand force mode corresponds to each trial are different. The specific correspondence is shown in the table.

B. Procedure

Before the paradigm, some preparatory work needs to be completed, such as checking the experimental equipment, setting the host computer software parameters correctly, cleaning the subjects' hands, and wearing the self-conducted myoelectric wristband. Like a smartwatch, the myoelectric wristband is placed directly above the back of the wrist. The experiment consists of twelve blocks, and each block contains twelve trials. In this work, we designed twelve types of gesture-free hand intention states, each corresponding to a trial. Considering the application requirements of decoding intention without gesture in sports scenes, the paradigm sets motion speeds of 0km/h, 4km/h, 6km/h, and 8km/h to simulate four motion states: stationary, slow walking, fast walking, and jogging. The experimental system would select corresponding image tips and motion platform speed settings based on the block and trial number, as shown in the table below Fig. 3. The fingers marked in red indicate that this gesture requires the finger to maintain a vigorous state, while the fingers marked in blue remain relaxed. During the experiment, the subjects were required to hold a cylindrical object (simulate holding things in hand) with their hands and maintain a tight grip on the object with their fingers during the force application process to ensure that sEMG signals were collected gesture-free.

C. Dataset format

The sEMG-based gesture-free hand intention recognition dataset includes two experiments from ten subjects in two days. In each experiment, the system records a source file of shape (T, C) with a .dat suffix. The parameter T denotes the sampling points during the experiments, which is the product of recording time and sampling rate. At each sampling point, the experimental system would record 15-channel signals. As

shown in Fig. 4, channels 1-8 record the eight-channel sEMG signals, and channels 9-11 record the axis-x, axis-y, and axis-z of the IMU data. Channel 12 is the sampling timestamp of the data for these 15 channels. Channel 13-15 denote the parameters of the paradigm, i.e. trigger, block, and speed information.

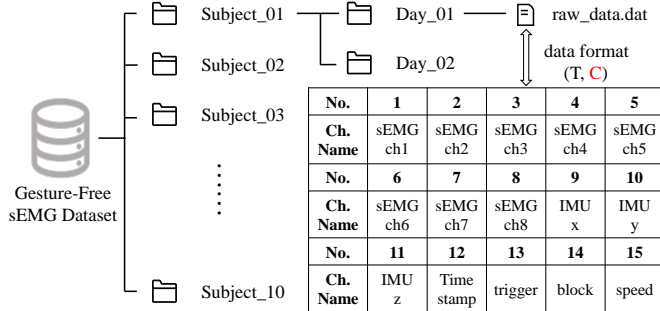


Fig. 4. Dataset format. The dataset includes data from two experiments involving ten subjects. Each experiment records a source file of shape (T, C) , where T and C denote the number of sampling points and channels respectively. There are a total of 15 channels, and the meaning of each channel is shown in the table.

IV. DATA ANALYSE TOOLBOX

A. Preprocess Tool

To help future researchers in processing and analyzing data, we have developed multiple data preprocessing methods, which are described in detail as follows:

- 1) **Data segmentation:** Obtain all sample data that meet the conditions from the raw data based on the label of the gesture, as well as the set time window size and window movement step size. The function head code can be expressed as `def data_segmentation(raw_data, gesture_id, window_size, step)`;
- 2) **Data filter:** According to the cutoff frequency, various common filters such as low-pass, high pass, band-pass, and notch are provided. It should be noted that all filters are zero phases to ensure the integrity of phase information. The pseudo-code of the function head is `def data_{filtername}_filter(raw_data, [cut_off_params])`;
- 3) **Baseline correction:** sEMG data is non-stationary and exhibits significant time drift. To solve this problem, baseline correction methods are often used to ensure distribution consistency between signals collected at different times. The calculation method is to subtract the signal mean in a calm state at the beginning of the trial from the current time data. The pseudo-code of the function head is `def data_baseline_correction(raw_data, baseline_len)`;
- 4) **Train / test dataset preparation:** This function manage all recorded sEMG data. Based on the type of experiment, such as single-day, cross-day, or cross-subject experiments, it would produce the training and testing datasets, which are required for the decoding algorithm. The pseudo-code of the function head is `def data_divide_{experiment-type}(dataset, [subject_id, day_id, and etc.])`.

B. Decoding Method

By researching existing literature, we have implemented the following ten widely used sEMG signal decoding methods. The traditional methods are LDA [10]), NaiveBayes [11], k-nearest neighbor (KNN) [12], SVM [13], and Random Forests [14] and the deep learning methods are LSTM [15], 1DCNN [16], 2DCNN [17], CNN-LSTM [18]), and L-EMGNet [19], respectively. For traditional classification methods, the time-frequency characteristics of sEMG signals are extracted as classification features, including signal mean, root mean square (RMS) value, average frequency, and spectral density. It should be noted that all methods can run end-to-end to meet practical application needs, and the code is open-sourced on the project website.

C. Visualization Tool

Data visualization can help us assess signal quality, discover objective patterns, and further optimize research. For this purpose, we have developed various function scripts and implemented commonly used visualization methods, including (1) 1D / 2D waveform plotting, for observing and comparing sEMG temporal dynamics; (2) The spectrum diagram used to analyze the main frequency components of sEMG signals and design filters accordingly; (3) confusion matrix is used to analyze the details of algorithm classification results, to screen out the most discriminative force application method; (4) t-SNE distribution, used to observe the distribution of sEMG features, and etc.

V. EXPERIMENTS AND BENCHMARK RESULTS

A. Experiments setting

To evaluate the performance of the gesture-free hand intention recognition system, we designed single-day, cross-day, and cross-subject experiments to test the existing sEMG classification algorithms. In the single-day experiment, the first eight blocks of data are used for training, while the rest are used for testing. On the one hand, it can avoid the class imbalance problem, and on the other hand, it ensures that the test data is collected after the training data to simulate actual application scenarios [20]. The cross-day experiment uses the data from the first experiment as the training set to test the performance of the model in the second experiment. As for cross-subject experiments, Leave-One-Out cross-validation [21] has been adopted to test the performance of the gesture-free hand intention recognition system. Considering that the requirements for interactive instructions vary in different application scenarios, the experiment is divided into 6 and 12 classes of finger force modes. Among them, the 6 classes of force modes only include resting state and single-finger force, while the 12 classes of force modes add 6 classes of combined force modes on the basis of the former. Through the literature review, the window sizes of sEMG signals were set to 250ms, 500ms, and 750ms, respectively, with a moving step size of 250ms to ensure real-time interaction response [22]. All classification methods in our toolbox have been tested. Consistent with most studies, we only use accuracy to measure the performance of model classification.

TABLE I

BENCHMARK RESULTS OF SINGLE-DAY, CROSS-DAY, AND CROSS-SUBJECT EXPERIMENTS WITH DIFFERENT WINDOW SIZES (MEAN \pm STANDARD DEVIATION). THE BASIC AND ADVANCED GESTURE-FREE HAND INTENTION RECOGNITION TASKS INCLUDE 6 AND 12 POWER MODES RESPECTIVELY. THE LENGTHS OF THE TIME WINDOWS ARE 250MS, 500MS, AND 750MS, RESPECTIVELY.

Method (Year)	Type	6-classes			12-classes		
		250ms	500ms	750ms	250ms	500ms	750ms
LDA (2019)	SD	0.8618 \pm 0.0807	0.8877 \pm 0.0769	0.8925 \pm 0.0834	0.6227 \pm 0.0921	0.6690 \pm 0.0956	0.6869 \pm 0.0953
	CD	0.6352 \pm 0.1776	0.6572 \pm 0.1815	0.6590 \pm 0.1818	0.3885 \pm 0.1354	0.4066 \pm 0.1442	0.4113 \pm 0.1459
	CS	0.5546 \pm 0.1608	0.5700 \pm 0.1604	0.5755 \pm 0.1623	0.2978 \pm 0.0936	0.3100 \pm 0.0984	0.3070 \pm 0.1000
NaiveBayes (2021)	SD	0.7438 \pm 0.1105	0.7657 \pm 0.1160	0.7778 \pm 0.1178	0.4668 \pm 0.1029	0.5055 \pm 0.1099	0.5250 \pm 0.1114
	CD	0.5675 \pm 0.1769	0.5712 \pm 0.1827	0.5709 \pm 0.1838	0.3441 \pm 0.1224	0.3564 \pm 0.1319	0.3618 \pm 0.1408
	CS	0.4878 \pm 0.1413	0.5122 \pm 0.1498	0.5090 \pm 0.1302	0.2566 \pm 0.0978	0.2649 \pm 0.1022	0.2674 \pm 0.1025
KNN (2024)	SD	0.7459 \pm 0.1502	0.7554 \pm 0.1542	0.7557 \pm 0.1555	0.4838 \pm 0.1281	0.5094 \pm 0.1401	0.5102 \pm 0.1433
	CD	0.5072 \pm 0.1608	0.5040 \pm 0.1716	0.5005 \pm 0.1721	0.2858 \pm 0.1059	0.2908 \pm 0.1228	0.2881 \pm 0.1234
	CS	0.4624 \pm 0.1158	0.4674 \pm 0.1184	0.4612 \pm 0.1130	0.2231 \pm 0.0535	0.2264 \pm 0.0541	0.2251 \pm 0.0551
SVM (2012)	SD	0.7952 \pm 0.1215	0.8105 \pm 0.1252	0.8114 \pm 0.1270	0.5520 \pm 0.1291	0.5845 \pm 0.1392	0.5899 \pm 0.1426
	CD	0.5511 \pm 0.1668	0.5495 \pm 0.1693	0.5476 \pm 0.1762	0.3354 \pm 0.1302	0.3359 \pm 0.1352	0.3359 \pm 0.1376
	CS	0.5161 \pm 0.1181	0.5199 \pm 0.1304	0.5304 \pm 0.1254	0.2499 \pm 0.0751	0.2562 \pm 0.0798	0.2575 \pm 0.0787
Random Forests (2015)	SD	0.8872 \pm 0.0633	0.9018 \pm 0.0627	0.9050 \pm 0.0666	0.6577 \pm 0.0916	0.6941 \pm 0.0959	0.7129 \pm 0.0931
	CD	0.6685 \pm 0.2080	0.6814 \pm 0.2174	0.6861 \pm 0.2169	0.4361 \pm 0.1639	0.4519 \pm 0.1725	0.4697 \pm 0.1741
	CS	0.5808 \pm 0.1290	0.5933 \pm 0.1388	0.5988 \pm 0.1300	0.3184 \pm 0.0867	0.3275 \pm 0.0897	0.3253 \pm 0.0876
LSTM (2018)	SD	0.8677 \pm 0.0709	0.8907 \pm 0.0610	0.8824 \pm 0.0744	0.6639 \pm 0.1003	0.7216 \pm 0.0965	0.7304 \pm 0.1022
	CD	0.5963 \pm 0.1553	0.6030 \pm 0.1587	0.6069 \pm 0.1795	0.3862 \pm 0.1357	0.3936 \pm 0.1557	0.3884 \pm 0.1709
	CS	0.5734 \pm 0.1285	0.6066 \pm 0.1236	0.6192 \pm 0.1206	0.3044 \pm 0.0797	0.3287 \pm 0.0863	0.3485 \pm 0.1024
1DCNN (2017)	SD	0.7521 \pm 0.1168	0.7692 \pm 0.1194	0.7722 \pm 0.1335	0.4911 \pm 0.1163	0.5230 \pm 0.1328	0.5205 \pm 0.1280
	CD	0.4991 \pm 0.1226	0.5257 \pm 0.1493	0.5236 \pm 0.1456	0.3155 \pm 0.1104	0.3204 \pm 0.1341	0.3113 \pm 0.1189
	CS	0.4873 \pm 0.1234	0.5087 \pm 0.1250	0.4952 \pm 0.1151	0.2149 \pm 0.0483	0.2127 \pm 0.0497	0.2064 \pm 0.0515
2DCNN (2021)	SD	0.9047 \pm 0.0595	0.9273 \pm 0.0625	0.9242 \pm 0.0830	0.6971 \pm 0.1075	0.7439 \pm 0.0951	0.7542 \pm 0.1098
	CD	0.6409 \pm 0.1902	0.6248 \pm 0.1970	0.6209 \pm 0.2197	0.4002 \pm 0.1753	0.4214 \pm 0.1839	0.4082 \pm 0.1658
	CS	0.6044 \pm 0.1477	0.6216 \pm 0.1575	0.6138 \pm 0.1411	0.3222 \pm 0.1174	0.3327 \pm 0.1268	0.3517 \pm 0.1367
CNN-LSTM (2018)	SD	0.8420 \pm 0.1704	0.8507 \pm 0.1708	0.8515 \pm 0.1730	0.6355 \pm 0.1253	0.6771 \pm 0.1111	0.6604 \pm 0.1684
	CD	0.6051 \pm 0.1687	0.6057 \pm 0.1942	0.5810 \pm 0.1936	0.3620 \pm 0.1275	0.3576 \pm 0.1525	0.3538 \pm 0.1623
	CS	0.5830 \pm 0.1366	0.5895 \pm 0.1312	0.6011 \pm 0.1347	0.3141 \pm 0.0850	0.3292 \pm 0.0937	0.3523 \pm 0.1146
L-EMGNet (2024)	SD	0.9107 \pm 0.0534	0.9254 \pm 0.0550	0.9367 \pm 0.0483	0.7051 \pm 0.0783	0.7386 \pm 0.0786	0.7917 \pm 0.0740
	CD	0.6803 \pm 0.2099	0.7136 \pm 0.1938	0.6905 \pm 0.2065	0.4451 \pm 0.1688	0.4548 \pm 0.1855	0.4465 \pm 0.1803
	CS	0.5559 \pm 0.1875	0.5513 \pm 0.1748	0.5613 \pm 0.1670	0.3143 \pm 0.1151	0.3290 \pm 0.1039	0.3321 \pm 0.1218

* The type SD, CD, and CS denote single-day, cross-day, and cross-subject experiments, respectively;

* The results highlighted in bold red/green/blue indicate that the model achieved the first/second/third ranking on the respective metric.

B. Benchmark Results

Before analyzing the benchmark results, we used the Shapiro-Wilk test [23] to verify that the classification results on benchmark methods followed the normal distribution hypothesis. Therefore, we adopted mean \pm standard deviation to record the experimental results. Tab. I shows all classification results of single-day, cross-day, and cross-subjects with ten benchmark methods.

1) *Results of single-day experiments:* focus on the first row of each classification method, the model L-EMGNet achieves the best performance with 250ms and 750 ms window size, and the 2DCNN achieves the best in 500ms. Their performance advantages relative to the second and third rank is as follows: 0.6% and 2.35% in the 6-classes experiment with a 250ms window. The gaps in 500ms and 750ms are 0.19%, 2.55% and 1.25% and 3.17%. In the 12-classes experiment, the advantages are 0.8%, 0.53%, and 3.75% compared to the second rank and 4.12%, 2.23%, and 6.13% compared to the third rank. Generally, the sEMG-based gesture-free hand intention recognition task achieves excellent classification performance, especially in 6-class experiments.

2) *Results of cross-day experiments:* observe the second row, the model L-EMGNet achieves the best performance except for the 12-classes experiment with the 750ms window compared to Random Forests. The performance between the top-2 is relatively close, the gap is 1.18%, 3.22%, 0.44%, 0.9%, 0.29%, and 2.32%, respectively. The third rank of performance distributes scattered, including LDA, 2DCNN, and CNN-LSTM. The performance gap is 3.94%, 5.64%, 3.15%, 4.49%, 3.34%, and 5.86% compared to the best method.

3) *Results of cross-subject experiments:* the third row of Tab. I shows the cross-subject results. The results of L-EMGNet are no longer excellent, while the results of 2DCNN are better, achieving four first ranks and two second ranks. Specifically, the performance gap from the second rank to the best method is 2.13%, 1.5%, 0.54%, 0.38%, 0.35%, and 0.06%. Similar to the cross-day experiment, the third rank is distributed, appearing in four different methods. Compared to cross-day experiments, the performance gap is not significant, 2.36%, 2.83%, 1.81%, 0.79%, 0.37%, and 0.38%.

4) *Comparative analysis:* comprehensively compare the results of single-day, cross-day, and cross-subject experiments,

we have the following three observations:

- The results of cross-day and cross-subject experiments have significantly decreased. Compared to single-day experiment, the performance gap based on the best method is 23.04%, 21.37%, 24.62%, 26%, 28.91%, and 32.2% in cross-day experiments and 30.63%, 30.57%, 31.75%, 38.29%, 41.12% and 43.94% in cross-subject experiments. This may be due to displacement deviation caused by two wearing attempts, as well as physiological differences among different subjects and on different days.
- L-EMGNet achieved the best performance in the single-day and cross-day experiments but lagged behind 2DCNN in the cross-subject experiment. The reason may be that L-EMGNet is a lightweight model that performs well in experiments with small amounts of data. The cross-subject experiment involves a very large amount of data, which is more suitable for 2DCNN with a larger number of parameters.
- Increasing the window size is beneficial for improving recognition accuracy, but it also increases the system's latency. Using the best methods as statistical objects, the performance of the 750ms window is higher than that of the 250 window is 2.6%, 1.02%, and 1.48% in 6-classes experiments. The improvement is 8.66%, 2.46%, and 3.01% in 12-classes experiments. This is because a longer time window is beneficial for obtaining more stable sEMG data.

VI. DISCUSSION AND ANALYSES

To ensure fairness, all discussed algorithms use L-EMGNet, with data sourced from single-day experiments and sliced based on a 500ms window length.

A. Evaluation of sEMG signal quality

To evaluate the sEMG signal quality, we choose the following three metrics: (1) Signal-to-Noise Ratio (SNR), which is used to quantify the increase in sEMG signal amplitude during activation compared to background noise level recorded when the muscle is not contracting; (2) Signal-to-Motion Artifact Ratio (SMR): Motion artifacts are low-frequency noises that contaminate sEMG signals and are introduced by fine electrode movements and changes in contact at the electrode-skin interface; and (3) Power Spectrum Deformation (Ω): which is sensitive to symmetry and peaking in the power spectrum and to additive disturbances in the low- and high-frequency ranges. The calculation formula for all metrics is as follows:

$$\text{SNR} = 20\log(\text{rms}_{\text{Activation_filtered}}/\text{rms}_{\text{Resting_filtered}})$$

$$\text{SMR} = 10\log\left(\frac{\sum_{f=0}^{500} \text{PSD}_{\text{filtered}}}{\sum_{f=0}^{20} \text{PSD}_{\text{AboveLine_raw}}}\right) \quad (1)$$

$$\Omega = 10\log\left[\frac{(M_2/M_0)^{1/2}}{M_1/M_0}\right], M_i = \sum_{f=0}^{500} \text{PSD} \cdot f^i$$

Regarding force mode 1 as a resting state, we calculated the SNR, SMR, and PSD of five other single-finger force

modes, the results shown in Tab. II. The results in Botros et al. [8] research are also included in the table as a reference to determine the quality of our collected sEMG data. On the most critical SNR metric, the results are relatively close, but on SMR and Ω metric, the results are relatively poor, indicating that the data contains severe motion artifacts, possibly due to our data being collected in multiple motion modes. In addition, we found that the signal quality of the middle finger, ring finger, and little finger is better, which may be due to the fact that when holding a cylindrical object, its gravity will involuntarily drive the thumb and index finger to exert force.

TABLE II
SIGNAL QUALITY OF THE COLLECTED SEMG DATA, WITH THE 6-CLASSES FORCE MODES.

Mode	SNR \uparrow	SMR \uparrow	Ω \downarrow
mode 2	12.3 (6.6) / 8.0 (3.6)	11.9 (6.2) / 1.5 (2.8)	1.04 (0.64) / 3.10 (0.52)
mode 3	11.5 (4.7) / 8.3 (2.4)	11.4 (5.2) / 1.6 (2.3)	1.03 (0.48) / 3.00 (0.41)
mode 4	7.8 (4.9) / 12.6 (3.7)	7.8 (5.2) / 5.0 (2.7)	1.34 (0.89) / 2.24 (0.54)
mode 5	11.4 (5.2) / 14.0 (3.4)	11.2 (4.8) / 6.4 (3.4)	0.97 (0.42) / 2.05 (0.60)
mode 6	11.0 (4.2) / 11.8 (3.9)	10.7 (4.3) / 4.2 (3.4)	0.98 (0.36) / 2.48 (0.65)

* The results on the left are from [8], and the results on the right are from this study;
* Higher values of SNR and SMR and lower values of Ω are better.

B. Details of the classification results

It is important to explore the classification details of the gesture-free hand intention recognition task, this is beneficial for screening out poor force modes, thereby improving the robustness of the system. The confusion matrices for 6 and 12 classes on a single-day experiment, as shown in Fig. 5. Observations have shown that the accuracy of patterns that include the use of index finger force is relatively low. For example, in the single-finger experiment of class 6, the index finger performed the worst. In the combined finger experiment, the results of the third, seventh, eleventh, and twelfth force modes were poor, while patterns that do not include index finger force were relatively good, except for the eighth force mode. This suggests that in future research, we can increase the use of the middle finger, ring finger, and little finger, and reduce the use of the index finger.

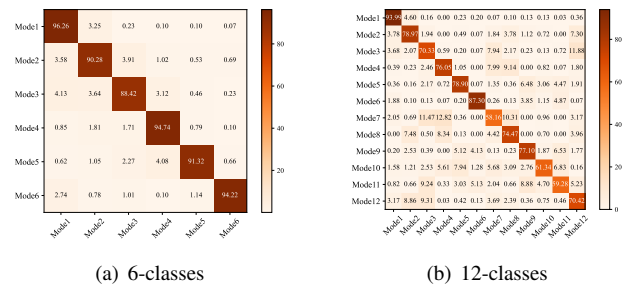


Fig. 5. Confusion matrix in single-day experiments with the 500ms window size using L-EMGNet. (a) 6-classes force modes; (b) 12-classes force modes.

C. The factor of different speeds

Considering multiple application scenarios, it is necessary to explore the interaction accuracy of the system in different

motion modes. To achieve this purpose, we recalculate the results according to its sports speeds, e.g. 0, 4, 6, 8 km/h. As illustrate in Tab. III, we found that there is no significant difference in classification performance among various motion modes. Two reasons are considered: on the one hand, deep learning has the ability to extract features from sEMG signals containing motion noise; On the other hand, motion noise is mainly distributed in low frequencies (below 20Hz) and is filtered out in the preprocessing stage.

TABLE III
PERFORMANCE WITH DIFFERENT SPORTS SPEED.

Speed	6-classes	12-classes
0 km/h	0.9279 ± 0.0521	0.7360 ± 0.0879
4 km/h	0.9251 ± 0.0583	0.7386 ± 0.0851
6 km/h	0.9238 ± 0.0597	0.7383 ± 0.0696
8 km/h	0.9248 ± 0.0567	0.7416 ± 0.0797
Mixed	0.9254 ± 0.0550	0.7386 ± 0.0786

D. The factor of different force intensities

Different finger force intensities may affect the recognition accuracy of the system, which is worth exploring for their relationship. As shown in Fig. 6, we have designed a force sensor to measure the force intensity of each finger. Considering that it is difficult to control force, the experiment only includes 6-classes of experiments using a single finger. Experiments were conducted on three subjects, S04, S06, and S08, which are the best, average, and worst subjects in benchmark experiments. The force intervals include as follows: [0, 100N], (100N, 200N], (200N, 300N], (300N, 400N], (400N, 500N], (500N, +∞). The accuracy of system recognition will be calculated based on the interval to which the force belongs. As shown in Fig. 7, the accuracy continues to improve and the trend of improvement gradually slows down with the finger force increases. This is because after exerting a certain degree of force, muscles have already contracted in place.

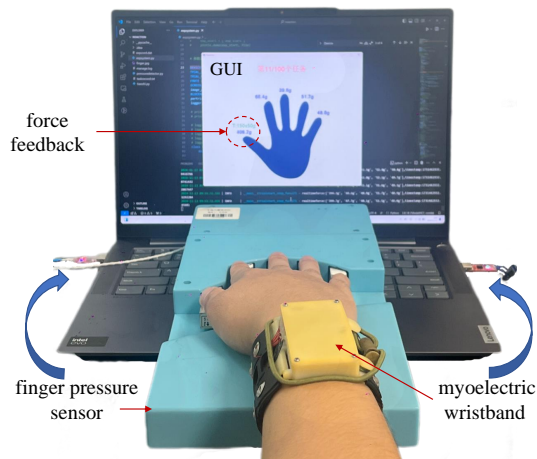


Fig. 6. Schematic diagram of the force pressure test system.

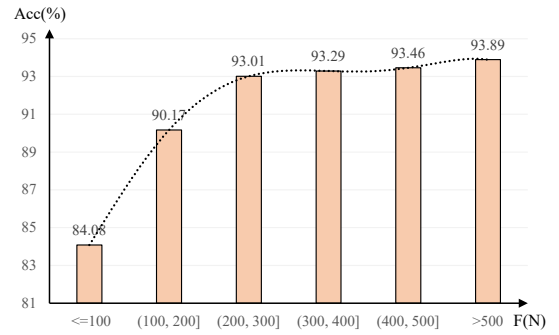


Fig. 7. Average results of each level force intensities. The curve is fitted by a 5th-order polynomial.

E. Online experiments

Accuracy and system response time are the two most important aspects of online experiments. This experiment is based on the same three subjects, and the testing system is shown in Fig. 8. Based on the keyboard experiment, the average reaction time of the three subjects was about 0.4s, which denotes the time from visual tips to the subject's reactions and should not be included in the system's reaction time. Thus, the system response time can be equated as $\Delta t = t_3 - t_1 \approx t_3 - t_0 - 0.4s$, because the t_1 is unable to accurately calculate. Each subject must complete 50 random force mode tests. Each test randomly prompts one force mode except for the resting state, and the first non-resting state prediction of the system is the predicted result and completion time, i.e. t_3 . The accuracies of the three subjects are 94%, 90%, and 88%, and system response times are 0.35s, 0.31s, and 0.3s, respectively. The excellent online results further validated the robustness in practical applications.

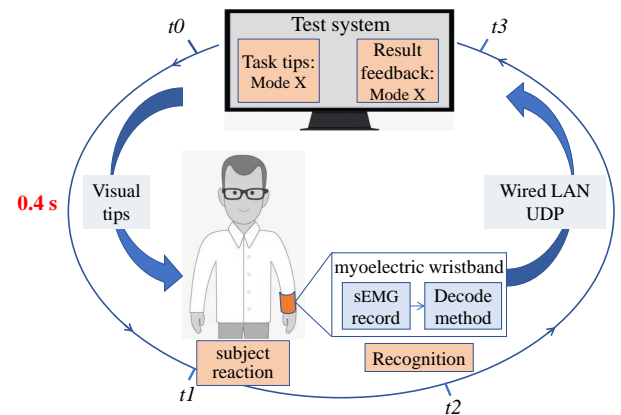


Fig. 8. Schematic diagram of the online experiment system.

VII. CONCLUSION

This work proposes the sEMG-based gesture-free hand intention recognition, which is a novel scheme for covert message transmission between collaborators in scenarios where collaborators and non-collaborators coexist. This scheme hides the action of sending messages, making it difficult for non-collaborators to detect, thereby protecting privacy and information security. To verify the feasibility, we designed the

experimental system and paradigm, collected the dataset, and developed the signal processing toolbox, ultimately obtaining benchmark results on various decoding methods. All data, software, and hardware designs and code had been open-sourced to promote the development of this research field. The benchmark results and discussion analysis indicate that the system performs well in a single-day experiment, but needs to be improved in cross-day and cross-subject experiments. Therefore, methods such as domain adaptation and testing adaptation that are suitable for solving distribution shift problems will be attempted in the future.

REFERENCES

- [1] K. Cohn-Gordon, C. Cremers, L. Garratt, J. Millican, and K. Milner, "On ends-to-ends encryption: Asynchronous group messaging with strong security guarantees," in *Proc. of the ACM SIGSAC Conf. Comput. and Comm. Security*, 2018, pp. 1802–1819.
- [2] L. S. Bohannon, A. M. Herbert, J. B. Pelz, and E. M. Rantanen, "Eye contact and video-mediated communication: A review," *Displays*, vol. 34, no. 2, pp. 177–185, Apr. 2013.
- [3] H. Chen, H. Shi, X. Liu, X. Li, and G. Zhao, "SMG: A micro-gesture dataset towards spontaneous body gestures for emotional stress state analysis," *Int. J. Comput. Vis.*, vol. 131, no. 6, pp. 1346–1366, Feb. 2023.
- [4] P.-G. Jung, G. Lim, S. Kim, and K. Kong, "A wearable gesture recognition device for detecting muscular activities based on air-pressure sensors," *IEEE Trans. Ind. Informat.*, vol. 11, no. 2, pp. 485–494, Feb. 2015.
- [5] S.-H. Lee, M. Lee, and S.-W. Lee, "Neural decoding of imagined speech and visual imagery as intuitive paradigms for BCI communication," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2647–2659, Nov. 2020.
- [6] R. J. Jacob, "What you look at is what you get: eye movement-based interaction techniques," in *Proc. of SIGCHI Conf. Hum. factor. Comput. Syst.*, 1990, pp. 11–18.
- [7] I. A. Brody, "Regulation of isometric contraction in skeletal muscle," *Exp. Neurol.*, vol. 50, no. 3, pp. 673–683, Mar. 1976.
- [8] F. S. Botros, A. Phinyomark, and E. J. Scheme, "Electromyography-based gesture recognition: Is it time to change focus from the forearm to the wrist?" *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 174–184, Dec. 2020.
- [9] S. Jiang et al., "Feasibility of wrist-worn, real-time hand, and surface gesture recognition via sEMG and IMU sensing," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3376–3385, Dec. 2017.
- [10] C. W. Antuvan and L. Masia, "An LDA-based approach for real-time simultaneous classification of movements using surface electromyography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 552–561, Feb. 2019.
- [11] Y. Narayan, "Comparative analysis of SVM and naive bayes classifier for the sEMG signal classification," *Mater. Today*, vol. 37, pp. 3241–3245, Oct. 2021.
- [12] P. Sahu, B. K. Singh, and N. Nirala, "Optimized k-nearest neighbors for classification of prosthetic hand movements using electromyography signal," *Eng. Appl. Artif. Intel.*, vol. 133, p. 108390, Jul. 2024.
- [13] A. Alkan and M. Günay, "Identification of EMG signals using discriminant analysis and SVM classifier," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 44–47, Jan. 2012.
- [14] E. Gokgoz and A. Subasi, "Comparison of decision tree algorithms for EMG signal classification using DWT," *Biomed. Signal Process. Control*, vol. 18, pp. 138–144, Apr. 2015.
- [15] A. Samadani, "Gated recurrent neural networks for EMG-based hand gesture classification. a comparative study," in *Proc. 40th IEEE Int. Conf. Med. and Bio. Soc. IEEE*, 2018, pp. 1–4.
- [16] Y. Cheng, G. Li, M. Yu, D. Jiang, J. Yun, Y. Liu, Y. Liu, and D. Chen, "Gesture recognition based on surface electromyography-feature image," *Concurr. Comput.*, vol. 33, no. 6, p. e6051, Oct. 2021.
- [17] Y. Du, W. Jin, W. Wei, Y. Hu, and W. Geng, "Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation," *Sensors*, vol. 17, no. 3, p. 458, Feb. 2017.
- [18] B. Xie, B. Li, and A. Harland, "Movement and gesture recognition using deep learning and wearable-sensor technology," in *Proc. of the Int. Conf. Artif. Intell. Pat Rec.*, 2018, pp. 26–31.
- [19] X. Xu et al., "Command and control method of grouping based on covert gesture interaction technology," *J. Command Control*, vol. 10, no. 3, pp. 331–338, Jun. 2024.
- [20] F. Li et al., "Decoupling representation learning for imbalanced electroencephalography classification in rapid serial visual presentation task," *J. Neural Eng.*, vol. 19, no. 3, p. 036011, May. 2022.
- [21] F. Di Nardo, C. Morbidoni, A. Cucchiarelli, and S. Fioretti, "Influence of EMG-signal processing and experimental set-up on prediction of gait events by neural network," *Biomed. Signal Process. Control*, vol. 63, p. 102232, Jan. 2021.
- [22] R. N. Khushaba and K. Nazarpour, "Decoding HD-EMG signals for myoelectric control-how small can the analysis window size be?" *IEEE Rob. Autom.*, vol. 6, no. 4, pp. 8569–8574, Sep. 2021.
- [23] Z. Hanusz, J. Tarasinska, and W. Zielinski, "Shapiro-wilk test with known mean," *Revstat Stat. J.*, vol. 14, no. 1, pp. 89–100, Feb. 2016.