Open Access Dataset, Toolbox and Benchmark Processing Results of High-Density Surface Electromyogram Recordings

Xinyu Jiang†, Xiangyu Liu†, Jiahao Fan†, Xinming Ye, Chenyun Dai*, *Member, IEEE*, Edward A Clancy, *Senior Member, IEEE*, Metin Akay, *Fellow, IEEE*, and Wei Chen*, *Senior Member, IEEE*

***Abstract*—We provide an open access dataset of *High densitY Surface Electromyogram* (HD-sEMG) *Recordings* (named "*Hyser*"), a toolbox for neural interface research, and benchmark results for pattern recognition and EMG-force applications. Data from 20 subjects were acquired twice per subject on different days following the same experimental paradigm. We acquired 256-channel HD-sEMG from forearm muscles during dexterous finger manipulations. This Hyser dataset contains five sub-datasets as: (1) pattern recognition (PR) dataset acquired during 34 commonly used hand gestures, (2) maximal voluntary muscle contraction (MVC) dataset while subjects contracted each individual finger, (3) one-degree of freedom (DoF) dataset acquired during force-varying contraction of each individual finger, (4) N-DoF dataset acquired during prescribed contractions of combinations of multiple fingers, and (5) random task dataset acquired during random contraction of combinations of fingers without any prescribed force trajectory. Dataset 1 can be used for gesture recognition studies. Datasets 2–5 also recorded individual finger forces, thus can be used for studies on proportional control of neuroprostheses. Our toolbox can be used to: (1) analyze each of the five datasets using standard benchmark methods and (2) decompose HD-sEMG signals into motor unit action potentials via independent component analysis. We expect our dataset, toolbox and benchmark analyses can provide a unique platform to promote a wide range of neural interface research and collaboration among neural rehabilitation engineers.**

***Index Terms*—HD-sEMG, neural interface, hand gesture recognition, prosthetic control.**

I. Introduction

SURFACE electromyogram (sEMG)-based neural interface techniques [1] have attracted increasing attention in recent years. Neural interfaces help amputees regain function via neuroprostheses which can be intuitively controlled by sEMG signals from residual muscles in the stump. In more general applications, intact users can use neural interfaces to enhance function via sEMG-based intuitive control of exoskeletons [2], or to manipulate mobile devices via gesture recognition [3].

In recent years, with the advancement of flexible sensor techniques, 2-dimensional (2-D) high-density sEMG (HD-sEMG) electrode arrays have substantially increased the spatial resolution and number of recording sites, compared with traditional sEMG electrodes. Specifically, HD-sEMG provides a high-resolution spatial activation image of the muscle group covered by the array. Furthermore, HD-sEMG allows decomposition of the global multi-channel HD-sEMG at the macroscopic level into motor unit (MU) spike trains at the microscopic level [4], using independent component analysis (ICA) [5], [6]. This breakthrough analysis method to decode information of MUs from HD-sEMG has been applied in diverse fields such as neuromuscular physiology [7], clinical neurophysiology [8], neuromuscular biometrics [9] and neural interface [2].

In the past decade, open-access sEMG datasets have made research on neural interfaces easier by saving a huge amount of time for researchers to acquire experimental data. Furthermore, researchers can easily compare the performance of their proposed methods using the same benchmark dataset. So far, there have been five open-access sEMG datasets recorded from the forearm: (1) Ninapro [10], [11], (2) CapgMyo [12], (3) CSL-HDEMG [13], (4) SEEDS [14] and (5) HIT-SIMCO [15]. A comparison of these datasets and our dataset is presented in Table I. The Ninapro dataset used traditional sEMG electrodes to acquire sEMG from both intact and amputated subjects. CapgMyo, CSL-HDEMG and SEEDS datasets used HD-sEMG electrode arrays (up to 192 channels) to acquire sEMG from intact subjects. The first four of these datasets acquired sEMG during prescribed movements or gestures. Prescribed movements or gestures limit the degrees of freedom (DoFs) of neuroprostheses because, in practical scenarios, users need to switch between arbitrary movements (gestures) or any combinations of DoFs. The lack of arbitrary switching between DoF combinations limits that ability of these datasets to simulate realistic applications. The HIT-SimCo Dataset acquired 8-channel sEMG under random combinations of 3 DoFs of the wrist, filling in the current gap to a certain extent. However, data acquired from only one day cannot support studies on cross-day prosthetic control. Furthermore, dexterous control of neuroprostheses also requires precise prediction of

†Xinyu Jiang, Xiangyu Liu and Jiahao Fan contributed equally to this work.
*Corresponding authors: Chenyun Dai (chenyundai@fudan.edu.cn), Wei Chen (w_chen@fudan.edu.cn)
Xinyu Jiang, Jiahao Fan, Chenyun Dai and Wei Chen are with the Center for Intelligent Medical Electronics, School of Information Science and Technology, Fudan University, Shanghai 200433, China.
Xiangyu Liu is with School of Art Design and Media, East China University of Science and Technology, Shanghai 200237, China.
Xinming Ye is with School of Sports Science and Engineering, East China University of Science and Technology, Shanghai 200237, China.
Edward A. Clancy is with Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA.
Metin Akay is with Department of Biomedical Engineering, University of Houston, Houston, TX, 77204, USA.
This work is supported by Shanghai Municipal Science and Technology Major Project (Grant No. 2017SHZDZX01) and Shanghai Pujiang Program (Grant No. 19PJ1401100).
Manuscript received XX XX, XXXX; revised XX XX, XXXX.

TABLE I: Comparison between previous sEMG datasets and ours.

	Ninapro	CapgMyo	CSL-HDEMG	SEEDS	HIT-SimCo	Hyser
No. Subjects	10–40	10–18	5	25	8	20
Amputees?	Yes	No	No	No	No	No
No. Gestures /Movements	Up to 53	8, 12, 2	27	13	11 (Involving 3 DoFs)	34
No. Electrodes	Up to 16	128 (HD)	192 (HD)	126 (HD) + 8	8	256 (HD)
Sampling Rate (Hz)	100–2000	1000	2048	2048	1926	2048
Kinematics or Force?	Kinematics	No	No	Kinematics	Wrist Force	Finger Force
Cross-Day?	No	Yes	Yes	No	No	Yes
Random Switch Between Combinations of DoFs?	No	No	No	No	Yes	Yes
Research Paradigms Using the Dataset	Gesture Recognition + Proportional Control of a Target Gesture	Gesture Recognition	Gesture Recognition	Gesture Recognition + Proportional Control of a Target Gesture	Gesture Recognition + Proportional Control of 3-DoF Wrist Force	Gesture Recognition + Proportional Control of Finger Force (Both 1 DoF and N DoF Control)

the force corresponding to each individual finger, which has not been provided in previous datasets. Additionally, sEMG data under dynamic and isometric contractions are not balanced in previous datasets, with data under dynamic contractions more common.

In this work, we provide open access *High densitY Surface Electromyogram Recordings* (*Hyser*) to fill in the gaps in existing datasets. We acquired 256-channel HD-sEMG from forearm muscles contributing to dexterous finger manipulations. Our Hyser dataset contains five sub-datasets to fulfill the demands of different applications. Dataset 1, named the pattern recognition (PR) dataset, was acquired under 34 hand gestures in common daily use. Both dynamic hand movements and gesture maintenance tasks were involved. Dataset 2, named the maximal voluntary muscle contraction (MVC) dataset, can be used to evaluate MVC of muscles corresponding to each finger. Dataset 3, named the one-DoF dataset, was acquired during isometric contraction (ranging from 30% MVC flexion to 30% MVC extension) of muscles corresponding to each finger. Dataset 4, named the N-DoF dataset, was acquired during isometric contraction (ranging from 30% MVC flexion to 30% MVC extension) of muscles corresponding to several prescribed combinations of multiple fingers. Dataset 5, named the random task dataset, was acquired during isometric contraction of muscles corresponding to any random combination of fingers without any prescribed force trajectory. Subjects switched between any combinations of DoFs at any time. Dataset 1 can be used to develop gesture recognition-based prosthetic control. Datasets 2–5, with both HD-sEMG and synchronized force labels, can be used for proportional estimation of the force corresponding to each single finger, further contributing to controlling dexterous prosthetic hands. Twenty subjects participated in our experiment with the data of each subject acquired twice on different days following the same experimental paradigm. The inter-day data can be used to simulate cross-day factors (e.g., the cross-day variation of sEMG characteristics and the shift of electrode arrays) in practical applications. Application results are reported from all datasets using standard (benchmark) analysis methods or emerging deep learning-based methods. Besides, we provide a toolbox for HD-sEMG analysis, which performs: (1) analysis of the pattern recognition dataset using linear discriminant analysis (LDA)-based and deep learning-based hand gesture classification, (2) analysis of datasets 2–4, i.e., EMG-force regression, (3) decomposition of HD-sEMG signals into MU spike trains using ICA. All analyses in our toolbox were implemented via Matlab. We expect our dataset, toolbox and benchmark analyses can provide a unique platform to promote a wide range of neural interface research and collaboration among neural rehabilitation engineers in the future.

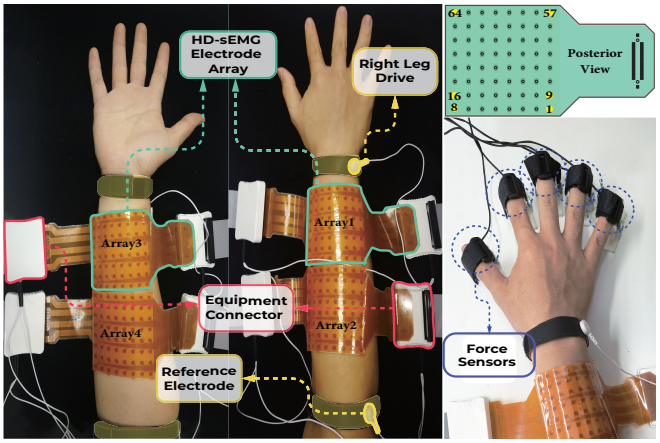

Fig. 1: Electrode Placement and Experiment Setup. We show photos of the same arm.

II. Open Access Dataset: Apparatus and Data Collection Methods

A. Subjects

Twenty intact subjects participated in this study. The detailed information of recruited subjects is presented in Table A1 (in Appendix). All subjects were informed about the research purpose and experimental procedure. All subjects signed written informed consent before the experiment. This study was reviewed and approved by the ethics committee of Fudan University (approval number: BE2035).

B. Data Acquisition

To reduce skin-electrode impedance, subjects' right forearm was carefully cleaned with abrasive gel and then wiped using

an alcohol pad. Four electrode arrays were mounted about the forearm. Each electrode array consists of 64 gelled elliptical electrodes (5-mm major axis, 2.8-mm minor axis) with a 10-mm inter-electrode distance (center-to-center), arranged in an 8×8 electrode layout. Four such arrays were used, two placed on each of the extensor and flexor muscles (Fig. 1). The 256 channels were arranged by successively concatenating the 64 channels of array 1, 2, 3 and 4 labeled in Fig. 1. The arrangement of the 64 channels of each array was presented in the top right corner of Fig. 1. Note that the HD-sEMG electrode arrays are difficult to be placed exactly on the extensor or flexor muscles. Accordingly, signals generated from other muscle groups may also be captured. However, all these captured signals together form the unique patterns of the extracted feature set. The radial and ulnar aspects of the forearm, the humeroulnar joint and the head of the ulna together formed the boundaries of the area to place electrode arrays. On each forearm side, two 8×8 electrode arrays construct a 16×8 electrode array. We aligned the center of the area to place electrode arrays and the center of the 16×8 electrode array, with the long axis of the constructed 16×8 electrode array along the long axis of subjects' forearm. A right leg drive electrode was placed on the head of the ulna. A reference electrode was placed on the olecranon. During data acquisition, subjects sat in a comfortable chair, following the instructions shown on a computer screen in front of them to perform the required experimental tasks. A total of five datasets, comprised of the (1) PR dataset, (2) MVC dataset, (3) one DoF dataset, (4) N DoF dataset and (5) random task dataset, were acquired in our experiment. The 256-channel HD-sEMG signals were acquired using the Quattrocento system (OT Bioelettronica, Torino, Italy), with a gain of 150, an ADC resolution of 16 bits (second-order high-pass cut-off at 10 Hz; low-pass cut-off at 500 Hz having a transition bandwidth of ∼25 Hz and a stop band attenuation $>$ 100 dB), and a sampling rate of 2048 Hz. For the MVC dataset, one DoF dataset, N DoF dataset and random task dataset, the ground truth force trajectories of the five fingers were acquired using five separate sensor-amplifier pairs (sensor: SAS, Huatran, Shenzhen, China; amplifier: HSGA, Huatran, Shenzhen, China) with a sampling rate of 100 Hz. The detailed parameters of used equipment and sensors are available in the "equipment_info.pdf" file in our dataset. The acquired HD-sEMG and force data were synchronized by transmitting a synchronization trigger signal to both HD-sEMG and force acquisition systems at the onset of each task. For each subject, acquisition of all five datasets was conducted on two different days, with an interval varying from 3 to 25 days (8.50 ± 6.72 days on average), following the same experiment paradigm.

C. Experimental Paradigm

1) PR Dataset: Subjects were instructed to perform the 34 hand gestures shown in Fig. 2. Each subject performed two repeated trials for each single gesture before they continued to the next one, following the sequence order shown in Fig. 2. In each trial, three dynamic tasks (1 s duration, from subjects' relaxing state to the required gesture) and one maintenance task (4 s duration, from subjects' relaxing state to the required gesture followed with maintenance at that gesture) were performed. Subjects were provided with a 2 s inter-task resting period and a 5 s inter-trial resting period to avoid the impact of muscle fatigue on sEMG properties. An audible "beep" queued each task. For each subject, HD-sEMG signals during 204 dynamic tasks (34 gestures × 2 trials × 3 tasks) and 68 maintenance tasks (34 gestures × 2 trials × 1 task) were acquired. If subjects missed a specific task or performed a wrong task, they were asked to inform the experiment assistant. The missed and wrong tasks were removed from the dataset. On average, 2.30 ± 2.71 dynamic tasks and 0.85 ± 1.05 maintenance tasks in each experiment were removed from the final dataset.

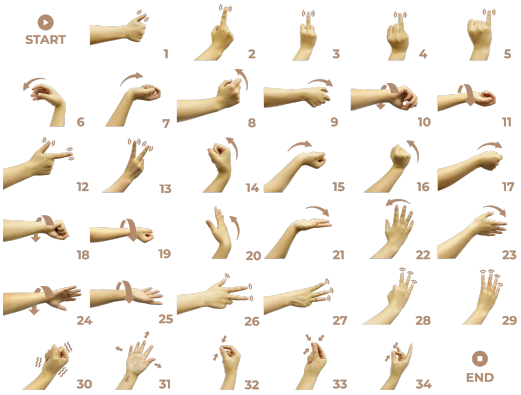

Fig. 2: All involved gestures: (1) thumb extension, (2) index finger extension, (3) middle finger extension, (4) ring finger extension, (5) little finger extension, (6) wrist flexion, (7) wrist extension, (8) wrist radial, (9) wrist ulnar, (10) wrist pronation, (11) wrist supination, (12) extension of thumb and index fingers, (13) extension of index and middle fingers, (14) wrist flexion combined with hand close, (15) wrist extension combined with hand close, (16) wrist radial combined with hand close, (17) wrist ulnar combined with hand close, (18) wrist pronation combined with hand close, (19) wrist supination combined with hand close, (20) wrist flexion combined with hand open, (21) wrist extension combined with hand open, (22) wrist radial combined with hand open, (23) wrist ulnar combined with hand open, (24) wrist pronation combined with hand open, (25) wrist supination combined with hand open, (26) extension of thumb, index and middle fingers, (27) extension of index, middle and ring fingers, (28) extension of middle, ring and little fingers, (29) extension of index, middle, ring and little fingers, (30) hand close, (31) hand open, (32) thumb and index fingers pinch, (33) thumb, index and middle fingers pinch, (34) thumb and middle fingers pinch.

2) MVC Dataset: Subjects were instructed to perform their MVC (isometric contractions) flexion and extension of each finger. Subjects' hands were not in the finger constraints for the PR data collection, but were in the constraints (with a subject's hand in a natural state) for all other data collection experiments, as shown in the right part of Fig. 1. The finger force sensors were secured to the table so that they could measure subjects' force in both flexion and extension directions. Note that the thumb is relatively complicated compared with the other four

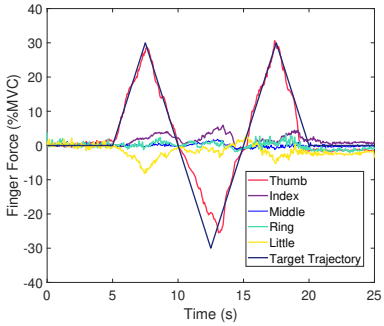

Fig. 3: Visual instruction for the one DoF dataset experiment. Force tracking using the thumb is shown. Positive %MVC refers to extension forces.

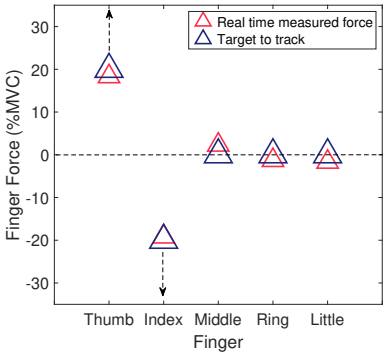

Fig. 4: Visual instruction for the experiment of N DoF dataset. Force tracking using the thumb and index finger is shown. Positive %MVC refers to extension forces. In the presented trial, thumb and index fingers were activated with opposing extension vs. flexion contraction efforts.

fingers due to its multi-DOF trapeziometacarpal (TM) and metacarpophalangeal (MCP) joints. We only considered flexion and extension of the thumb because these two contraction directions can support most daily activities, and to be consistent with the other four fingers. Subjects performed 2 successive trials to measure their flexion and extension MVC values of one target finger and then continued to the next finger. The MVC corresponding to the five fingers were measured following the order of thumb, index finger, middle finger, ring finger and little finger. In each trial, subjects were provided with a 10 s duration to perform the MVC of the required direction and finger. Within the 10 s window, subjects could perform MVC at any time, but they were required to maintain the MVC for at least 2 seconds so that the average force during the steady period can be taken as the MVC value. A 30 s inter-trial resting period was provided to avoid the impact of muscle fatigue. During the experiment, the real time measured force data were presented on the screen as a visual feedback for subjects. We provided the acquired force segments corresponding to the 10 trials (5-channel time series for each segment, one finger per channel) in all our datasets. The MVC of each finger in both extension and flexion directions can be used to normalize the acquired force data.

3) One DoF Dataset: Subjects were instructed to perform 25-s duration, isometric contraction of each single finger. Subjects used the real time finger force as visual feedback to track a slowly force-varying "triangle" target trajectory, as shown in Fig. 3. The target force trajectory for the active finger ranged from 30% MVC flexion to 30% MVC extension. All other fingers were relaxed without co-contractions. Subjects performed 3 trials for each finger tracking the target force trajectory shown in Fig. 3. Our previous work [16] has demonstrated that a small size of data acquired from the same subject on a single day can capture most of the important neuromuscular information, and shows limited performance differences compared with a large data size using least squares based model estimation. For cross-subject validation of generalized control models, the dominant factor is the "data diversity" (the number of training subjects) rather than "data size" (the total signal duration from each subject). Accordingly, we assume that three trials of each DoF from each subject, with a relatively long duration of 25 s each trial, can support most research directions. Subjects were asked to re-perform a trial if any obvious co-contraction of other fingers was observed. A total of 30 trials (5 fingers $\times$ 3 trials) were included in the one DoF dataset. Subjects performed all thumb flexion and extension trials, then progressed in order to the index finger, middle finger, ring finger and little finger. After finishing one trial, subjects were provided with a resting period of self-selected duration. The target force trajectory only instructs subjects to vary their contraction efforts. Failure to perfectly track the force trajectory does not result in that trial being excluded.

4) N DoF Dataset: Subjects were instructed to perform 25-s duration, isometric contractions of a combination of multiple fingers. Subjects used the finger force as feedback to track up to five targets (one per finger), as shown in Fig. 4. The larger blue were targets which moved along the vertical screen direction. Subjects performed different slowly force-varying contraction efforts of up to five fingers to control the movement of smaller red triangles to track the targets. The target force trajectories ranged from 30% MVC flexion to 30% MVC extension. A total of 15 different contraction combinations of DoFs (fingers) were examined. In 10 combinations, all active fingers used the same force trajectory (same "triangular" trajectory as the 1 DoF dataset, with extension performed first), those fingers being: (1) thumb + index, (2) thumb + middle, (3) thumb + ring, (4) thumb + little, (5) index + middle, (6) thumb + index + middle, (7) index + middle + ring, (8) middle + ring + little, (9) index + middle + ring + little, and (10) all five fingers. In the other 5 combinations, pairs of fingers were active but with opposing "triangle-trajectory" extension vs. flexion effort with the first finger performing extension first and the second finger performing flexion first), those finger pairs being: (11) thumb + index, (12) thumb + middle, (13) thumb + ring, (14) thumb + little, and (15) index + middle. Subjects performed two trials for each combination of DoFs before they continued to the next one, following the order from combination (1) to combination (15) listed above. A total of 30 trials (15 different combinations of fingers $\times$ 2 trials) were included in the N DoF

dataset. After finishing one trial, subjects were provided with a resting period of self-selected duration. Because we aimed to acquire HD-sEMG signals with multiple DoFs to advance the development of multi-DoF prosthetic control, co-contractions of other fingers did not lead to the exclusion of trials.

5) Random Task Dataset: Subjects were allowed to randomly perform isometric contractions of muscles corresponding to any combination of fingers with any force trajectory. The force trajectories in all trials can be different. We aim to provide HD-sEMG data for multi-DoF prosthetic control in the most realistic scenario, where subjects do not need to follow any instructions. The real time measured force data corresponding to five fingers were presented on the screen. An example of random task was shown in Fig. A1 (in Appendix). A total of 5 trials were included in the random task dataset. Each contraction task was with a 25 s duration. A 5 s inter-trial resting period was provided to avoid the impact of muscle fatigue on sEMG properties.

III. Benchmark Analytics and Results

A. Data Preprocessing

The acquired HD-sEMG signals were first filtered with a 10–500 Hz 8-order Butterworth bandpass filter. A notch filter combination was then used to attenuate the power line interference at 50 Hz and all harmonic components up to 400 Hz. Force data were filtered by an 8-order 10 Hz low-pass Butterworth filter.

B. Pattern Recognition of Hand Gestures

1) Methods for Pattern Recognition:

(1.1) LDA-based Method

The first 0.25 s reaction time after each task onset was removed, leaving 0.75 s and 3.75 s signals for dynamic and maintenance tasks, respectively. The widely used EMG features of root mean square (RMS) [3], waveform length (WL) [9], zero crossing (ZC) [17], and slope sign change (SSC) [17], were extracted separately from each EMG channel. For ZC and SSC, a noise threshold (approximately 3% of the RMS during rest [17]) was used. For each feature, a 256-length feature vector was constructed, one value per channel. The four 256-length feature vectors were concatenated together to obtain a 1024-length feature vector. All feature vector were normalized to a mean of zero and a standard deviation of one.

Principal component analysis (PCA) [18] was applied to reduce the length the constructed feature vectors. For a classification task with N_f-length feature vectors and N_s training samples, the dimensionality D of the feature space is constrained by both $D \leq N_s - 1$ and $D \leq N_f$. For the pattern recognition of hand gestures in our work, $N_f = 1024 > N_s - 1$. Accordingly, we reduce the length of feature vectors to the maximal dimensionality of the feature space, $N_s - 1$. In our work, N_s varies with different subjects, depending on the number of correctly performed gestures. All PCA processed feature vectors were used to train and test a LDA [19] classifier.

We validated the LDA-based classification accuracy of dynamic tasks and maintenance tasks separately. Specifically, we employed a leave one out cross-validation strategy on data (either dynamic or maintenance task) acquired within a session from each subject. We first mixed all data samples (204 dynamic tasks or 68 maintenance tasks, if all gestures were correctly performed). Then we held out one testing sample, using all remaining samples for training. For testing set feature normalization, the mean value to subtract and the standard deviation to divide are taken as the corresponding values of the training set. Both the projection matrix of PCA and the parameters of LDA model were obtained using training data. The trained LDA model was then used to give the gesture label of the held out testing sample. The same procedure was performed for all data.

(1.2) Deep Learning-based Method

Compared with traditional machine learning methods, deep neural networks with a considerably higher number of parameters, require a larger training dataset to avoid overfitting. Accordingly, we further segmented the HD-sEMG signals using 250 ms windows with 125 ms window overlap. In this way, the training set were significantly augmented. A convolutional neural network (CNN) [20] was trained and tested using 250 ms signals. Since spectrogram is an efficient representation in deep learning based sEMG pattern recognition tasks [21], we transformed each obtained 250 ms signal to spectrogram via short-time Fourier transform (STFT) [22]. STFT was performed using Hamming window with a length of 256 data points and 50% overlap. The output of STFT for each 250 ms signal is with 3 time frames and 128 spectral bands (8 Hz for each band). Because the signals have been filtered bellow 500 Hz, only the first 64 spectral bands were retained. For each 250 ms signal, the representation $\mathcal{X} \in \mathbb{R}^{F \times N_R \times N_C}$ was fed into the CNN model, where $F = 3 \times 64 = 192$ denotes the length of vectors stacked by 3 time frames in all 64 spectral bands, $N_C \times N_R = 16 \times 16$ denotes the 16×16 electrode array reshaped by the 256 channels. We used Adam optimizer to update network weights. We used batch normalization to speed up convergence and avoid gradient vanishing problem. Dropout technique was applied to address overfitting issues. The CNN architecture and all hyper-parameters were presented in Table II and Table A2 (in Appendix), respectively. Given a testing sample, the segmented M 0.25 s signals were successively fed into the trained CNN model to obtain M 34-length score vectors $S_m = [s_1, s_2, ...s_p, ..., s_{34}]$, where $S_m, m \in \{1, 2, ..., M\}$, denotes the score vector of the m^{th} 250 ms signal window and $s_p, p \in \{1, 2, ..., 34\}$, denotes the probability that the gesture label of the corresponding 250 ms signal is p. The average of all obtained M score vectors was calculated. The index number corresponding to the maximal score value in the average score vector was the final gesture label.

To validate the deep learning based hand gesture classification, we pooled data from all subjects in one session (either session 1 or session 2) together to increase the data size, considering deep learning-based models rely on a large data size to train a large number of parameters. The 6-fold and 2-fold cross-validation was used for dynamic and maintenance tasks, respectively. To make the data in all folds balanced across all gestures, in each fold, we randomly allocated 1 out of 6 dynamic samples or 1 out of 2 maintenance samples of all 34 gestures from all subjects. Overall, each fold included almost

TABLE II: Network architecture

Layers	Components	In/Out
Conv 1	Conv ($w = 3, s = 1$) LReLU MP ($w = 3, s = 2$)	320/32
Conv 2	Conv ($w = 3, s = 1$) LReLu MP ($w = 3, s = 2$)	32/64
FC 1	FC ($nn = 576$) LRelu	64/1
FC 2	FC ($nn = N_{gesture} = 34$) LRelu	1/1
Softmax	Softmax operation	1/1

*Conv, LReLU, MP, FC denote convolution, leaky ReLU, max pooling, and fully connected layers, respectively. Variables w, s, nn denote the kernel width, convolution stride, and neuron numbers, respectively. $N_{gesture} = 34$ represents the number of gestures. The depth of input and output features is summarized in the rightmost column

the same number of signal samples for each gesture. Each time we employed data in one fold as testing set and all other folds as training set. Note that, different with LDA-based hand gesture classification, we used data from all subjects together to train our deep learning model to avoid overfitting of such a large number of parameters. By contrast, in all other following validations, data acquired in different sessions from different subjects were processed separately. Cross-day and cross-subject variation of performance was not taken into consideration, as we only provide benchmark results.

2) Results:

LDA-based classification results of dynamic and maintenance tasks for the 34 gestures are shown in Table III. Average classification accuracies of 96.86% and 93.80% were achieved for the dynamic and maintenance tasks, respectively. Deep learning based classification results are shown in Table IV. Average accuracies of 88.96% and 89.84% were achieved for the dynamic and maintenance tasks, respectively. Such classification accuracies are generally considered high with a ∼2.94% random assignment chance level for a 34-class classification problem. For LDA-based method, our tests were performed with training and testing data acquired in the same session. Future studies using our dataset can further improve the classification accuracy in both within-session and cross-session (both cross-day and cross-subject) validation using advanced signal processing and machine learning techniques. Deep learning based hand gesture classification has attracted increasing attention in recent years. So far, no consensus exists on the optimal network architecture. Future studies can develop more effective architectures by making modifications on ours in the toolbox.

C. One-DoF Dataset

1) Feature Extraction and Model Description:

EMG amplitude of each channel was estimated by computing the RMS [17] of each 40 contiguous samples (19.5 ms). WL, ZC and SSC [17] of each 40 continuous samples were also extracted, as described previously. The average force of force samples corresponding to each 19.5 ms window was also calculated. EMG-force regression was performed on the downsampled features and force values using training data.

TABLE III: LDA-based classification accuracy (%) of 34 hand gestures. The reported value for each subject is the average of cross validations and two sessions.

Subject	Dynamic	Maintenance
1	95.53	94.79
2	98.77	99.25
3	99.23	97.73
4	98.52	97.03
5	90.19	85.83
6	99.26	97.74
7	96.57	94.12
8	95.57	91.18
9	97.55	88.24
10	99.75	93.94
11	97.72	90.74
12	95.31	96.29
13	97.28	94.81
14	91.81	94.12
15	98.26	90.23
16	96.50	87.12
17	98.01	97.06
18	95.82	93.32
19	99.75	97.73
20	95.82	94.85
Average	96.86±2.49	93.80±3.80

TABLE IV: Deep learning based classification accuracy (%) of 34 hand gestures. The reported values are the average of cross validations and two sessions.

Dynamic	Maintenance
88.96±1.76	89.84±1.91

The first 2 s and the last 2 s of the data were removed to account for filter startup and tail transients [1].

We trained linear finite impulse response models (to provide dynamics) of the form:

$$Force[i] = \sum_{q=0}^{Q} \sum_{m=1}^{M} \theta_{q,m} \cdot x_m[i-q] \quad (1)$$

where i is the decimated sample index; $Q = 20$ (390 ms) is the number of time lags, similar to our previous work [17]; and M is the number of utilized features. In this work, a 1024-length feature vector was extracted (4 features×256 channels). PCA [18] was applied to reduce the dimensionality of the 1024-length feature vector to 200. Accordingly, $M = 200$ in our work. Optimal model parameters were obtained using linear least squares, with singular values of the design matrix discarded if their ratio to the largest singular value was less than a tolerance threshold (selected as 0.05).

2) Validation Methodologies:

To validate the performance of EMG-force regression for the one DoF dataset, we employed a leave one out cross-validation strategy. For each finger per subject, we used 2 trials for PCA-based feature reduction and training the regression model. This model was then tested on the third trial. The same procedure was repeated until all trials of all fingers have been used as testing trial. Note that for each finger (DoF), force values in extension and flexion directions were estimated using the same model, with positive and negative force values representing extension and flexion, respectively. Root mean square error (RMSE) between the estimated force and the ground truth force for each test trial was used for performance evaluation.

[1]In real-time scenarios, all processing procedures would be performed using causal filters, with no need to remove any tail transients.

TABLE V: RMSE (%MVC) of EMG-force regression for one DoF dataset. The reported value for each subject is the average of cross validations and two sessions.

Subject	Thumb	Index	Middle	Ring	Little
1	4.80	5.42	4.38	4.19	6.01
2	5.39	3.77	2.91	3.62	3.92
3	4.27	5.97	4.55	5.96	3.85
4	3.91	4.31	4.93	5.03	4.52
5	6.09	7.78	6.37	6.84	5.89
6	6.54	5.67	8.82	3.85	5.84
7	5.35	4.98	4.55	3.03	3.83
8	7.25	5.72	5.01	3.74	7.82
9	9.53	6.37	4.61	3.71	7.47
10	8.71	4.59	5.58	3.04	9.68
11	5.31	4.83	5.10	5.97	6.98
12	5.44	3.43	7.53	8.20	14.46
13	5.37	6.29	5.23	5.94	7.35
14	10.33	6.03	7.87	7.94	8.56
15	7.77	6.57	5.21	4.79	7.94
16	4.34	4.64	8.46	4.19	5.70
17	6.73	4.80	4.05	5.53	7.47
18	8.40	5.34	8.40	9.30	4.93
19	8.00	7.99	8.06	5.96	11.09
20	5.13	6.57	7.97	6.35	9.60
Average	6.43 $\pm$1.84	5.55 $\pm$1.19	5.98 $\pm$1.78	5.36 $\pm$1.77	7.15 $\pm$2.67
Grand Average	6.09$\pm$1.98				

TABLE VI: RMSE (%MVC) of EMG-force regression for N DoF dataset. The reported value for each subject is the average of cross validations and two sessions.

Subject	Thumb	Index	Middle	Ring	Little
1	5.84	7.28	4.02	4.01	6.05
2	5.36	3.89	3.22	4.14	5.43
3	6.21	5.05	4.29	3.83	3.54
4	3.78	4.28	4.57	3.56	3.67
5	6.42	5.46	4.51	4.92	4.84
6	6.67	6.48	5.28	4.73	5.72
7	8.00	6.67	4.02	5.47	5.43
8	5.64	5.00	3.11	3.93	3.95
9	8.60	5.81	5.05	3.74	4.89
10	9.81	9.40	6.00	4.54	8.13
11	7.57	7.09	4.85	5.50	7.40
12	5.39	5.62	6.02	4.33	12.45
13	5.77	5.63	5.79	5.29	5.95
14	11.38	10.54	6.94	15.23	19.43
15	6.76	6.21	7.17	4.66	6.55
16	5.15	6.13	6.45	4.44	5.33
17	7.45	7.43	7.24	5.02	7.02
18	11.30	8.48	9.75	14.93	10.05
19	7.05	8.65	7.55	5.01	7.53
20	7.90	7.52	7.29	7.01	8.39
Average	7.10 $\pm$1.99	6.63 $\pm$1.70	5.66 $\pm$1.68	5.71 $\pm$3.30	7.09 $\pm$3.63
Grand Average	6.44$\pm$2.62				

3) Results:

The summary RMSE results of EMG-force regression in the one DoF dataset is presented in Table V. The average RMSE values for the thumb, index, middle, ring and little fingers are 6.43% MVC, 5.55% MVC, 5.98% MVC, 5.36% MVC and 7.15% MVC, respectively. The average RMSE of all fingers is 6.09% MVC. Representative time series of the ground truth and corresponding estimated force trajectories of the one DoF dataset are presented in Fig.A2 (in Appendix).

D. N-DoF Dataset

1) Feature Extraction and Model Description:

All data were preprocessed as in the one-DoF case, except that we trained five models to estimate the force values of five fingers separately, using the same features extracted from HD-sEMG but different ground truth force data from 5 fingers. Each of the five models gives the force estimation of one finger. The five models together achieve concurrent force estimation of all five fingers.

2) Validation Methodologies:

To validate the performance of EMG-force regression for the N DoF dataset, we employed a 2-fold cross-validation strategy. A total of 15 DoF combinations were examined. In each fold, one trial of each DoF combination was included, used for PCA-based feature reduction and training the regression model. This model was then tested on the other fold. Root mean square error (RMSE) between the estimated force and the ground truth force for each trial was calculated and used for performance evaluation.

3) Results:

The summary RMSE results of EMG-force regression in the N DoF dataset is presented in Table VI. The average RMSE values for the thumb, index, middle, ring and little fingers are 7.10% MVC, 6.63% MVC, 5.66% MVC, 5.71% MVC and 7.09% MVC, respectively. The average RMSE of all fingers is 6.44% MVC.

E. Random Dataset

1) Feature Extraction and Model Description:

All data were preprocessed as in the N-DoF case (five models were used to estimate force values of five fingers).

2) Validation Methodologies:

To validate the performance of EMG-force regression for the random task dataset, we employed a leave one out cross-validation strategy. We used 4 trials for PCA-based feature reduction and training the regression model. This model was then tested on the fifth trial. The same procedure was performed with each trial held out. Root mean square error (RMSE) between the estimated force and the ground truth force for each trial was calculated and used for performance evaluation.

3) Results:

The summary RMSE results of EMG-force regression in the random dataset is presented in Table VII. The average RMSE values for the thumb, index, middle, ring and little fingers are 8.03% MVC, 8.63% MVC, 7.59% MVC, 7.23% MVC and 11.39% MVC, respectively. The average RMSE of all fingers is 8.57% MVC.

F. Decomposition of HD-sEMG into MU Spike Trains

1) Procedures of ICA-based HD-sEMG Decomposition:

A series of studies by Chen et al. [5], [23]–[25] demonstrated the excellent performance of using fastICA [26] for HD-sEMG decomposition. Negro et al. [6] also applied a combination of fastICA and convolution kernel compensation (CKC) [27] to decompose HD-sEMG into MU spike trains. To avoid the convergence of ICA algorithms to the same MU, either an automatic peel-off operation in [25] or an orthogonalization operation in [6] was required. In our implementation, the latter one was applied. The details of procedures to decompose sEMG in our work can be found in [6]. We present the main steps of sEMG decomposition briefly. All of the following processing steps were performed separately for extensor and flexor muscle array data (128 channels for each muscle).

TABLE VII: RMSE (%MVC) of EMG-force regression for random task dataset. The reported value for each subject is the average of cross validations and two sessions.

Subject	Thumb	Index	Middle	Ring	Little
1	8.86	10.81	6.65	6.16	8.60
2	6.40	4.30	3.82	4.45	5.98
3	5.25	5.39	6.68	3.69	3.49
4	4.43	5.42	5.59	6.33	4.47
5	6.65	8.23	9.70	7.28	8.24
6	6.28	10.78	5.87	5.78	6.56
7	4.79	5.38	3.57	4.71	7.03
8	4.56	6.27	5.26	5.15	7.92
9	13.56	9.93	9.71	12.26	10.74
10	8.89	8.75	7.19	5.59	8.37
11	6.60	7.93	3.89	5.02	8.45
12	7.32	7.35	6.58	5.88	50.64
13	13.62	9.54	8.57	10.26	9.28
14	8.52	10.69	8.44	15.76	20.99
15	9.80	14.25	11.58	7.35	18.95
16	7.86	10.02	12.92	7.93	9.47
17	8.81	10.10	7.76	8.80	11.32
18	11.20	10.36	9.68	10.40	12.64
19	6.34	7.17	7.95	4.61	5.25
20	10.78	9.93	10.44	7.26	9.43
Average	8.03 $\pm$2.72	8.63 $\pm$2.48	7.59 $\pm$2.58	7.23 $\pm$3.01	11.39 $\pm$10.19
Grand Average	8.57$\pm$5.27				

(1) Stack 4 copies of raw sEMG signals in each channel to extend the number of channels from 256 to 1024 [6]. In each copy, we progressively added one more sample delay to the original signals.

(2) Whiten the extended sEMG through eigenvalue decomposition.

(3) Perform fastICA on the whitened sEMG signals to obtain the sources corresponding to different MUs.

(4) Identify MU spike train (discharge timings) of each individual MU through peak detection and k-means clustering. Silhouette distance values (SIL), as the indicators of the consistency within clusters of data, were calculated. Only MUs with a SIL higher than a threshold (0.6 in our work) were retained for further analysis.

(5) Remove duplicate MUs. The ICA algorithm may converge repetitively to the same MU or its delayed copies due to limitations of the algorithm itself and/or the extension operation in step *(1)*. If several MU spike trains share more than 50% synchronized firing events within a $\pm$ 1 ms match window after compensating for delay, retain the one with the highest SIL [28].

2) Validation Methodologies:

To validate the ICA-based HD-sEMG decomposition code in our open access toolbox on the one DoF dataset, we employed the three metrics:

(1) The number of decomposed MUs. A common limitation of most current HD-sEMG decomposition algorithms is that only superficial and large MUs can be identified [29]. The number of decomposed MUs varies with different segments of HD-sEMG signals. A larger number of decomposed MUs could provide more information.

(2) Average SIL of decomposed MUs. SIL measures the degree of separation of the MU spike trains from noise (both the background noise and other potential source signals) [28]. A higher average SIL represents better decomposition performance.

(3) Pearson correlation coefficient [30] between the ground truth force and the fitted force of MU spike trains. Decomposition of each trial was performed separately. We then calculated the firing frequency of the spike trains of each MU in each contiguous 19.5 ms window (40 sample points). This firing frequency was then filtered by an 8-order 10 Hz low-pass Butterworth filter. All of the processed firing frequencies from a trial were fit to the ground truth force, using linear regression to determine one optimal scaling gain per firing frequency/MU. A higher Pearson correlation coefficient represents better decomposition performance.

3) Results:

The average number of decomposed MUs, average SIL, and Pearson correlation coefficient between the ground truth force and the force estimated from the MU spike trains were shown in Table VIII. On average, 32.57 MUs were obtained through the ICA-based HD-sEMG decomposition algorithm. The average SIL value is 0.7351. An average Pearson correlation coefficient of 0.8611 was achieved. An example MU discharge plot, showing the ground truth force and the force estimated from MU spike trains were shown in Fig. A3 in Appendix. The optimal linear combination of the filtered discharge frequencies of all MUs showed a similar trend with the ground truth finger force, indicating the decomposed source signals obtained via HD-sEMG decomposition are physiologically significance. These results were obtained with a 0.6 SIL threshold. We also reported the decomposition results vs. different predefined SIL threshold (0.7, 0.8 and 0.9) in Table A4 in Appendix. The selection of SIL threshold varies in different studies, depending on their research purposes. We present decomposition results with different SIL thresholds to provide a baseline for comparisons in different studies. These results support that the ICA-based HD-sEMG decomposition algorithm in our toolbox was properly implemented and the dataset has good signal quality.

G. Summary of our Hyser Dataset and Toolbox

All available data segments in our dataset and important functions in our toolbox were presented in Table A3 and Table Table A5 in Appendix, respectively. Gesture label data were saved in "*.txt" files with comma-separated values format. All force trajectory waveforms and EMG signal waveforms were saved in waveform database (WFDB) format, with one "*.dat" file storing all 16-bit signed type quantitized values, and one "*.hea" file storing the scaling factors. To fully implement our toolbox, users still need a MATLAB license. Instructions on how to use the codes are given in the first few lines of each function in our toolbox.

IV. Discussion

A. Limitations of our Dataset and Analysis

For LDA-based analysis of the PR dataset, we did not use a sliding window to extract features, because the sliding window technique yielded an extremely large feature dimensionality. For example, a 3.75 s signal duration of maintenance tasks combined with a 250 ms window length and a 125 ms sliding step would result in a 29696-length feature vector (29 windows

TABLE VIII: Summary results of HD-sEMG decomposition of all one DoF trials of both experimental sessions from each subject. For the average number of MUs, the reported results are the number of MUs in extensor muscle + that in flexor muscle. The reported value for each subject is the average of all trials in two sessions.

Subject	Average Number of MUs	Average SIL	Pearson Correlation Coefficient
1	18.20+4.27	0.7352	0.7983
2	12.23+7.07	0.7237	0.8328
3	13.67+9.37	0.7130	0.8882
4	12.60+14.87	0.7554	0.8869
5	17.73+28.17	0.7180	0.8866
6	19.03+11.93	0.7026	0.9028
7	10.73+3.77	0.7057	0.8266
8	19.20+9.03	0.7091	0.8848
9	20.13+11.33	0.7759	0.8876
10	21.07+21.73	0.7580	0.9128
11	21.20+17.60	0.7256	0.8775
12	15.10+16.13	0.7333	0.8367
13	19.13+14.10	0.7382	0.8653
14	15.90+19.23	0.7281	0.8536
15	26.50+11.70	0.7634	0.8695
16	20.03+7.20	0.7300	0.8912
17	11.37+22.50	0.7507	0.8225
18	36.63+18.63	0.7395	0.8828
19	16.30+8.03	0.7299	0.7850
20	18.90+29.10	0.7664	0.8301
Average	18.28 ($\pm$5.82) + 14.29 ($\pm$7.32)	0.7351$\pm$0.0210	0.8611$\pm$0.0358

$\times$ 4 features $\times$ 256 channels). Such a large number of features make subsequent analysis, e.g. singular value decomposition of the large covariance matrix in PCA, difficult to implement, due to both the high computational complexity and the high memory requirements. Many previously proposed techniques, e.g., sliding window, need to be re-considered in the context of HD-sEMG to better address the challenges from the large number of channels, together with the huge advantages. Our toolbox provides the functions for tuning the width and step length of sliding windows, supporting investigation of window selection in future studies.

Analysis of data from each subject in each session was performed separately, with training and testing sets drawn from the same session. We selected this analysis protocol because our main purpose of analysis is to provide benchmark results for dataset users to verify the good signal quality of our dataset and the proper implementation of algorithms in our toolbox. In practical scenarios, cross-session validation is preferred because the trained model should be robust when applied to a new subject or the same subject on a second day. So far, most studies in the literature validated their methods with training and testing data acquired in the same session from the same subject. Cross-session validation of hand gesture classification methods have attracted increasing attention in recent years [3]. Validating advanced algorithms in a more realistic cross-session protocol is highly encouraged for future studies. Such more rigorous validation may lead to performance degeneration compared with previously reported classification results. In this case, transfer learning algorithms such as transfer component analysis (TCA) [31] and correlation-based data weighting (COR-W) [16] are promising to tackle this issue. Additionally, electrode arrays cannot be replaced on exactly the same position day-to-day, so the issue of electrode shift needs to be investigated in future studies. Wu et al. proposed a data augmentation algorithm [32], which is promising to solve electrode shift problems in gesture recognition tasks. Although no data from patients were provided in our dataset and data from each subject were not acquired across a large number of days (two limitations of our dataset), data from 20 intact subjects acquired on two separate days can still largely fulfill the demands of research on cross-day and cross-subject control of neural interfaces. Training models with data from only one day, tested on data from a second day is also favorable for practical applications, to achieve a low-cost training model.

Another limitation of our dataset is the lack of ground truth of MU spike trains. The best way to validate decomposition results is two-source validation [24]. However, we acquired HD-sEMG from the entire forearm muscle groups, covering a large area. Common intramuscular EMG can only detected a limited area of muscle. Accordingly, applying two-source validation for all spike trains is not feasible in our work. The reliability of HD-sEMG decomposition was validated via indirect metrics, such as the SIL. SIL has been demonstrated to show a high positive correlation with the accuracy of HD-sEMG decomposition [33]. Employing SIL to assess decomposition accuracy is reasonable.

B. Possible Research Directions in Future Studies Which Might Benefit from our Dataset and Toolbox

Establishing a new dataset for research purposes is time-consuming, thus our dataset can save a huge amount of time for researchers. Research directions which might benefit from our dataset and toolbox include:

1) HD-sEMG-based neuroprosthetic control. In previous studies, both macroscopic features extracted from global sEMG [3] and microscopic features extracted from MU spike trains obtained via decomposition have been used as the input of control models [2]. Our dataset and toolbox can be used to develop neuroprostheses based on both pattern recognition and proportional control, using both macroscopic and microscopic features. Using our dataset, generalized neural interface techniques can also be developed for intact users to manipulate mobile devices in Internet of Things (IoT) applications. Additionally, deep learning-based control models have attracted increasing attention in recent years. However, the large number of parameters used in deep learning models necessitates large training data sets, which can be a limitation (as compared to conventional processing methods) in the neural interface field..

2) Compression of HD-sEMG signals. HD-sEMG acquires signals from a large number of channels, greatly increasing the burdens of data storage and transmission in tele-rehabilitation applications. Several unique properties of HD-sEMG, such as the similarity between neighboring channels, may facilitate new solutions for multi-channel sEMG compression. To-date, investigations of HD-sEMG signal compression are very scarce in the literature.

3) Signal quality assessment of HD-sEMG. In many applications, sEMG measurement needs to be achieved in an unsupervised way. In this case, low-quality signals may disproportionately reduce the robustness of systems. By designing a signal quality descriptor, we can exclude low-quality channels from the analysis procedure, or set the system

(neuroprostheses or health monitoring systems) to an idle state if signal quality is lower than a predefined threshold. In the context of signal quality assessment of sEMG, the specific properties of large HD-sEMG arrays may provide new challenges and opportunities.

4) Neuromuscular physiology studies. Neuromuscular physiology studies highly rely on decoding the discharge activities of MUs via non-invasive sEMG measurement. The ICA-based HD-sEMG decomposition algorithm in our toolbox can contribute to extending the body of knowledge in neuromuscular physiology.

5) Neuromuscular biometrics decoded from HD-sEMG for user authentication or identification. Our recent study [34] has demonstrated excellent performance using HD-sEMG as a new cancelable biometric trait (validated on the Hyser PR dataset), due to the individually-unique characteristics of HD-sEMG signals. Our dataset provide HD-sEMG under diverse hand gestures and muscle contraction efforts, which can be used to investigate HD-sEMG-based biometrics. Data acquired on different days can also support the evaluation of cross-day variation of HD-sEMG biometrics.

V. Conclusion

In this work, we provide an open access HD-sEMG dataset, signal processing toolbox, and benchmark application results for neural interface studies. The dataset and toolbox support a diversity of research directions. Signal analysis of the provided data using the open access toolbox demonstrated the good signal quality of our dataset and proper implementation of the algorithms in our toolbox. To facilitate the dataset and toolbox for research purposes, the data and code in our toolbox were made as clear as possible. Our HD-sEMG dataset and processing toolbox are available online via the website[2].

Appendix A
Supplementary Figures and Tables on Description of Data, Methods and Results in our Validation

TABLE A1: Detailed subject information.

Subject	Age (Years)	Gender	Height (cm)	Weight (kg)	Intervel Between Sessions (Days)
1	32	Male	173	78	7
2	24	Male	170	71	7
3	22	Male	177	59	3
4	21	Male	175	75	4
5	22	Male	191	95	5
6	22	Female	162	50	3
7	22	Female	168	60	5
8	26	Male	168	61	17
9	30	Male	175	80	4
10	30	Male	175	80	18
11	26	Female	170	56	25
12	22	Female	158	52	10
13	23	Male	183	86	8
14	29	Female	158	50	4
15	32	Male	178	75	11
16	34	Male	170	87.5	22
17	23	Female	158	47	4
18	27	Female	160	55	7
19	31	Female	165	53	3
20	33	Male	174	70	3

[2]The website is available immediately after the review process is finished

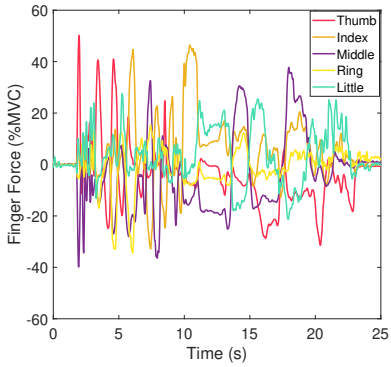

Fig. A1: Example of random task dataset.

TABLE A2: Summary of used hyper-parameters in CNN

Parameter	Value
Batch Size	512
Learning Rate	10^{-3}
Dropout Probability	0.2
Adam α	0.9
Adam β	0.999
L2 Penalty	10^{-4}
Training iterations	100

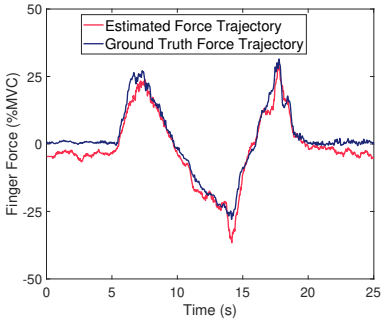

Fig. A2: Example waveforms of ground truth and estimated finger force trajectories (subject 2, session 1, middle finger, 2^{nd} trial of 1 DoF dataset). RMSE of the example trial was 3.56%MVC

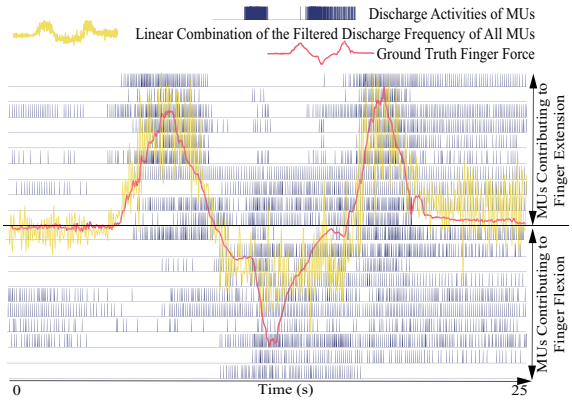

Fig. A3: Example MU discharge plot (subject 1, session 2, middle finger, 2^{nd} trial of 1 DoF dataset). The Pearson correlation coefficient of the presented trial is 0.8562.

TABLE A3: All available signal segments.

Dataset	Folder Name	File Name in Each Folder	Description
pr_dataset (Size: 37.1 GB)	subject***i***_session***j***	***taskType_sigType***_sample***k***.dat ***taskType_sigType***_sample***k***.hea label_***taskType***.txt	***i***$\in$ {'01','02', ..., '20'} represents the subject index; ***j***$\in$ {'1','2'} represents the session index; ***taskType***$\in$ {'dynamic','maintenance'} represents the two tasks of each gesture; ***sigType***$\in$ {'raw','preprocess'} represents raw and preprocessed EMG segments, respectively; ***k***$\in$ {'1','2',...,'N_s'} represents the segment index for each task.
mvc_dataset (Size: 7.8 GB)	subject***i***_session***j***	mvc_***sigType***_finger***u***_***direction***.dat mvc_***sigType***_finger***u***_***direction***.hea	***sigType***$\in$ {'raw','preprocess','force'} represents signal segments of raw EMG, preprocessed EMG and ground truth force, respectively; ***u***$\in$ {'1','2',...,'5'} represents contractions of thumb, index, middle, ring and little finger, respectively; ***direction***$\in$ {'extension','flexion'} represents the two contraction directions; ***i,j*** are defined the same as PR dataset.
1dof_dataset (Size: 29.3 GB)	subject***i***_session***j***	1dof_***sigType***_finger***u***_sample***k***.dat 1dof_***sigType***_finger***u***_sample***k***.hea	***sigType,i,j,k,u*** are defined the same as PR and MVC datasets.
ndof_dataset (Size: 58.6 GB)	subject***i***_session***j***	ndof_***sigType***_combination***u***_sample***k***.dat ndof_***sigType***_combination***u***_sample***k***.hea	***u***$\in$ {'1','2',...,'15'} represents the index number of 15 finger combinations (as listed in the text); ***sigType,i,j,k*** are defined the same as PR and MVC datasets.
random_dataset (Size: 9.8 GB)	subject***i***_session***j***	random_***sigType***_sample***k***.dat random_***sigType***_sample***k***.hea	***sigType,i,j,k*** are defined the same as PR and MVC datasets.

*For files with a file name "label_***taskType**.txt", the stored data are the gesture labels of all segments (either dynamic or maintenance tasks, depending on the value of ***taskType***) in PR dataset, formatted as $1 \times N_s$ comma-separated values (one value per segment). For all other data, the waveforms (either ground truth force trajectory waveforms or HD-sEMG signal waveforms) were stored in WFDB format (one "*.dat" file storing all 16-bit signed type quantitized values, and one "*.hea" file storing the scaling factors). Data can be loaded to Matlab using "load_pr.m", "load_1dof.m", "load_ndof.m" and "load_random.m" functions in our toolbox.

TABLE A4: Summary results of HD-sEMG decomposition thresholded by SIL. Results are the average of trials in the one DoF dataset.

SIL Threshold	Number of MUs	Average SIL	Pearson Correlation Coefficient
0.7	9.13 ($\pm$3.63) + 8.20 ($\pm$5.31)	0.8200$\pm$0.0142	0.7411$\pm$0.0707
0.8	4.70 ($\pm$2.17) + 4.89 ($\pm$3.53)	0.8836$\pm$0.0099	0.5950$\pm$0.0866
0.9	1.87 ($\pm$1.14) + 2.32 ($\pm$1.86)	0.9391$\pm$0.0058	0.4076$\pm$0.1101

TABLE A5: Important functions in our toolbox.

Function Name	Description
main_pr.m	Full LDA-based PR benchmark analysis
main_pr_cnn.m	Full CNN-based PR benchmark analysis
main_1dof.m	Full 1-DoF benchmark analysis
main_ndof.m	Full N-DoF benchmark analysis
main_random.m	Full random benchmark analysis
main_decomposition.m	Full decomposition
main_decomposition_validate.m	Validate full decomposition
demo_pr.m	Partial LDA-based PR benchmark analysis
demo_pr_cnn.m	Partial CNN-based PR benchmark analysis
demo_1dof.m	Partial 1-DoF benchmark analysis
demo_ndof.m	Partial N-DoF benchmark analysis
demo_random.m	Partial random benchmark analysis
demo_decomposition.m	Partial decomposition
load_pr.m	Load PR data of 1 subject & session
load_1dof.m	Load 1-DoF data of 1 subject & session
load_ndof.m	Load N-DoF data of 1 subject & session
load_random.m	Load Random data of 1 subject & session
get_mvc.m	Estimate MVC of each finger and contraction direction.
get_rms.m	Extract RMS features
get_wl.m	Extract WL features
get_zc.m	Extract ZC features
get_ssc.m	Extract SSC features
feature_normalize.m	Normalize features (with or without PCA)
e_sifir_trn.m	Train an EMG-force model
e_sifir_tst.m	Test an EMG-force model
FastICA.m	Perform ICA for HD-sEMG decomposition
SILCal.m	Calculate SIL value of decomposed sources

*Functions "main_pr.m", "main_pr_cnn.m", "main_1dof.m", "main_ndof.m", "main_random.m", "main_decomposition.m" and "main_decomposition_validate.m" run the corresponding benchmark analysis methods as described above on data of all subjects and sessions, one by one. These functions yield the global results listed in this report but are time-consuming. Functions "demo_pr.m", "demo_pr_cnn.m", "demo_1dof.m", "demo_ndof.m", "demo_random.m" and "demo_decomposition.m" run the same methods on a small proportion of data, with a fast running speed, which can be used to quickly understand the implementation of our toolbox. Also note that function "main_decomposition.m" can be used to obtain the decomposition sources of HD-sEMG, and "main_decomposition_validate.m" calculates the number of decomposed MUs, SIL values and Pearson correlation coefficients listed in Table VIII.

REFERENCES

[1] C. G. McDonald *et al.*, "A Myoelectric Control Interface for Upper-Limb Robotic Rehabilitation Following Spinal Cord Injury," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 4, Apr. 2020.

[2] C. Dai and X. Hu, "Finger Joint Angle Estimation Based on Motoneuron Discharge Activities," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 3, pp. 760–767, Mar. 2020.

[3] A. Phinyomark *et al.*, "EMG feature evaluation for improving myoelectric pattern recognition robustness," *Expert Syst. Appl.*, vol. 40, no. 12, pp. 4832–4840, Sep. 2013.

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

[5] M. Chen and P. Zhou, "A Novel Framework Based on FastICA for High Density Surface EMG Decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 117–127, Jan. 2016.

[6] F. Negro *et al.*, "Multi-channel intramuscular and surface EMG decomposition by convolutive blind source separation." *J. Neural Eng.*, vol. 13, no. 2, pp. 026 027–026 027, Apr. 2016.

[7] D. Farina *et al.*, "Principles of Motor Unit Physiology Evolve With Advances in Technology," *Physiology*, vol. 31, no. 2, pp. 83–94, Feb. 2016.

[8] X. Li *et al.*, "Examination of Poststroke Alteration in Motor Unit Firing Behavior Using High-Density Surface EMG Decomposition," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 5, pp. 1242–1252, May 2015.

[9] X. Jiang *et al.*, "Neuromuscular Password-based User Authentication," *IEEE Trans. Ind. Inform.*, pp. 1–1, 2020.

[10] M. Atzori *et al.*, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, no. 1, p. 140053, Dec. 2014.

[11] S. Pizzolato *et al.*, "Comparison of six electromyography acquisition setups on hand movement classification tasks," *PLoS One*, vol. 12, no. 10, p. e0186132, Oct. 2017.

[12] Y. Du *et al.*, "Surface EMG-Based Inter-Session Gesture Recognition Enhanced by Deep Domain Adaptation," *Sensors*, vol. 17, no. 3, p. 458, Mar. 2017.

[13] C. Amma *et al.*, "Advancing Muscle-Computer Interfaces with High-Density Electromyography," in *Proceedings of the 33rd Annual ACM*

Conference on Human Factors in Computing Systems, ser. CHI '15. New York, NY, USA: Association for Computing Machinery, Apr. 2015, pp. 929–938.

[14] A. Matran-Fernandez *et al.*, "SEEDS, simultaneous recordings of high-density EMG and finger joint angles during multiple hand movements," *Sci. Data*, vol. 6, no. 1, pp. 1–10, Sep. 2019.

[15] D. Yang, "HIT Simultaneous Control EMG Dataset (HIT-SIMCO)," http://homepage.hit.edu.cn/ydp.

[16] X. Jiang *et al.*, "Data Management for Transfer Learning Approaches to Elbow EMG-Torque Modeling," *IEEE Trans. Biomed. Eng.*, 2021, accepted.

[17] C. Dai *et al.*, "Comparison of Constant-Posture Force-Varying EMG-Force Dynamic Models About the Elbow," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 9, pp. 1529–1538, Sep. 2017.

[18] D. Zhang *et al.*, "A comparative study on PCA and LDA based EMG pattern recognition for anthropomorphic robotic hand," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. Hong Kong, China: IEEE, May 2014, pp. 4850–4855.

[19] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Ann. Eugenics*, vol. 7, no. 2, pp. 179–188, Sep. 1936.

[20] A. Krizhevsky *et al.*, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[21] U. Côté-Allard *et al.*, "Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019.

[22] L. Durak and O. Arikan, "Short-time Fourier transform: two fundamental properties and an optimal implementation," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1231–1242, May 2003.

[23] M. Chen *et al.*, "Progressive FastICA Peel-Off and Convolution Kernel Compensation Demonstrate High Agreement for High Density Surface EMG Decomposition," *Neural. Plast.*, vol. 2016, p. e3489540, Aug. 2016, publisher: Hindawi.

[24] M. Chen *et al.*, "A Novel Validation Approach for High-Density Surface EMG Decomposition in Motor Neuron Disease," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 6, pp. 1161–1168, Jun. 2018.

[25] M. Chen *et al.*, "Automatic Implementation of Progressive FastICA Peel-Off for High Density Surface EMG Decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 144–152, Jan. 2018.

[26] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Netw.*, vol. 13, no. 4, pp. 411–430, Jun. 2000.

[27] A. Holobar and D. Zazula, "Multichannel Blind Source Separation Using Convolution Kernel Compensation," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4487–4496, Sep. 2007.

[28] C. Dai and X. Hu, "Independent component analysis based algorithms for high-density electromyogram decomposition: Experimental evaluation of upper extremity muscles," *Comput. Biol. Med.*, vol. 108, pp. 42–48, May 2019.

[29] D. Farina *et al.*, "Decoding the neural drive to muscles from the surface electromyogram," *Clin. Neurophysiol.*, vol. 121, no. 10, pp. 1616–1623, Oct. 2010.

[30] J. Benesty *et al.*, "On the Importance of the Pearson Correlation Coefficient in Noise Reduction," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 4, pp. 757–765, May 2008.

[31] X. Jiang *et al.*, "Transfer Component Analysis to Reduce Individual Difference of EEG Characteristics for Automated Seizure Detection," in *2019 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Oct. 2019, pp. 1–4, iSSN: 2163-4025.

[32] L. Wu *et al.*, "Improved High-density Myoelectric Pattern Recognition Control Against Electrode Shift Using Data Augmentation and Dilated Convolutional Neural Network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, pp. 1–1, 2020.

[33] C. Dai *et al.*, "Prediction of Individual Finger Forces Based on Decoded Motoneuron Activities," *Ann. Biomed. Eng.*, vol. 47, no. 6, pp. 1357–1368, Jun. 2019.

[34] X. Jiang *et al.*, "Enhancing IoT Security via Cancelable HD-sEMG-based Biometric Authentication Password, Encoded by Gesture," *IEEE Internet of Things Journal*, pp. 1–1, 2021.