Toward a generalizable deep CNN for neural drive estimation across muscles and participants

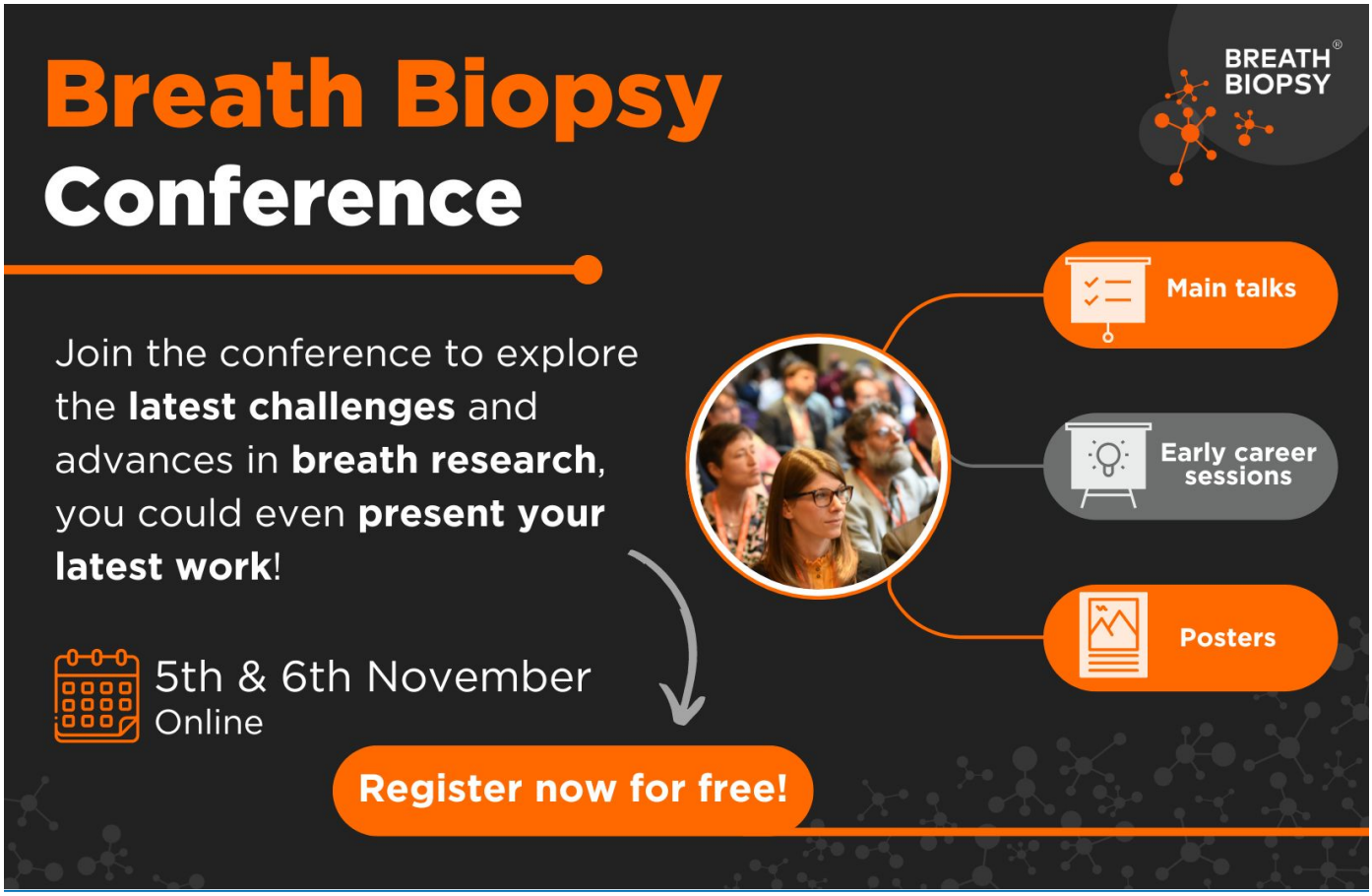
Yue Wen, Sangjoon Kim, Simon Avrillon, Jackson Levine, François Hug, José Pons

Journal of Neural Engineering

PAPER • OPEN ACCESS

Toward a generalizable deep CNN for neural drive estimation across muscles and participants

To cite this article: Yue Wen *et al* 2023 *J. Neural Eng.* **20** 016006



Journal of Neural Engineering

PAPER

Toward a generalizable deep CNN for neural drive estimation across muscles and participants

Yue Wen[1,5], **Sangjoon J Kim**[1,5], **Simon Avrillon**[1], **Jackson T Levine**[1], **François Hug**[2,3] and **José L Pons**[4,*]

[1] Legs and Walking Lab of Shirley Ryan AbilityLab and Feinberg School of Medicine, Northwestern University, Chicago, IL, United States of America
[2] Université Côte d'Azur, LAMHESS, Nice, France
[3] School of Biomedical Sciences, University of Queensland, Brisbane, QLD, Australia
[4] Legs and Walking Lab of Shirley Ryan AbilityLab, McCormick School of Engineering, and Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States of America
[5] These authors contributed equally to this work.
* Author to whom any correspondence should be addressed.

E-mail: jpons@sralab.org

Keywords: high-density electromyography, neural drive, deep CNN, human–machine interface

OPEN ACCESS

RECEIVED
22 July 2022

REVISED
14 December 2022

ACCEPTED FOR PUBLICATION
22 December 2022

PUBLISHED
18 January 2023

Abstract

Objective. High-density electromyography (HD-EMG) decomposition algorithms are used to identify individual motor unit (MU) spike trains, which collectively constitute the neural code of movements, to predict motor intent. This approach has advanced from offline to online decomposition, from isometric to dynamic contractions, leading to a wide range of neural-machine interface applications. However, current online methods need offline retraining when applied to the same muscle on a different day or to a different person, which limits their applications in a real-time neural-machine interface. We proposed a deep convolutional neural network (CNN) framework for neural drive estimation, which takes in frames of HD-EMG signals as input, extracts general spatiotemporal properties of MU action potentials, and outputs the number of spikes in each frame. The deep CNN can generalize its application without retraining to HD-EMG data recorded in separate sessions, muscles, or participants. *Approach.* We recorded HD-EMG signals from the vastus medialis and vastus lateralis muscles from five participants while they performed isometric contractions during two sessions separated by ∼20 months. We identified MU spike trains from HD-EMG signals using a convolutive blind source separation (BSS) method, and then used the cumulative spike train (CST) of these MUs and the HD-EMG signals to train and validate the deep CNN. *Main results.* On average, the correlation coefficients between CST from the BSS and that from deep CNN were 0.983 ± 0.006 for leave-one-out across-sessions-and-muscles validation and 0.989 ± 0.002 for leave-one-out across-participants validation. When trained with more than four datasets, the performance of deep CNN saturated at 0.984 ± 0.001 for cross validations across muscles, sessions, and participants. *Significance.* We can conclude that the deep CNN is generalizable across the aforementioned conditions without retraining. We could potentially generate a robust deep CNN to estimate neural drive to muscles for neural-machine interfaces.

1. Introduction

Surface electromyography (EMG) has wide applications for neural-machine interfacing because of its non-invasive access to the neural drive to muscles, i.e. the net output of all the motor neurons that innervate the muscle. Various control methods, including pattern recognition, regression models, and musculoskeletal models, have been used to control wearable robots (e.g. prostheses and exoskeleton) using surface EMG [1–4]. Essentially, the neural drive to muscles is composed of discrete spike signals, which are then amplified and transformed to continuous motor unit action potential (MUAP) signals

by the innervated muscle. When measured at the surface of the skin, the MUAP signals from all activated MUs are overlapped and temporal-spatially filtered by the tissues between muscle fibers and the skin. Therefore, surface EMG is a rough estimation and is not linearly related to the neural drive to muscles due to the superimposition and cancellation of MUAPs [5].

To have a more accurate estimation of the neural drive to muscles, several MU identification techniques have been developed, allowing the decomposition of high-density surface EMG (HD-EMG) into MU spike trains. Holobar and Zazula combined a convolution kernel compensation (CKC) method and a gradient descent algorithm to extract multiple MU spike trains simultaneously [6, 7]. Alternatively, Chen and Zhou proposed a progressive FastICA peel-off method for fast and accurate HD-EMG decomposition [8]. Both methods were validated through a two-source validation protocol [6, 9], and they demonstrated a high agreement in MU identification [10]. Recently, Negro *et al* proposed a convolutive blind source separation (BSS) method for MU identification using multi-channel invasive and non-invasive EMG signals by combining FastICA and CKC algorithms [11]. The feasibility of these methods for HD-EMG decomposition has been extensively validated with different muscles (from upper limb and lower limb), with different clinical populations [12–14], and for different purposes (pathology studies, interfacing with assistive devices).

Recently, the neural drive to muscles has been used to extract motor intent for human–machine interfaces. Joint torque was estimated using the cumulative spike train (CST) [15], a MU twitch model [16], and a MU-specific image-based deep convolutional neural network (CNN) [17]. Joint kinematics were also estimated using MU discharge times [18] with forward kinematics [19], the first principal component of MU firing rates [20], and the CST [21]. Furthermore, these studies demonstrated the improved reliability and accuracy of force and kinematic estimations over traditional EMG-amplitude based approaches. To apply these estimations and achieve real-time human–machine interfacing, efforts have been made to push MU identification to an online platform by applying separation matrices, initialized offline, to real-time HD-EMG data [22, 23]. Additionally, MU identification of dynamic movements has recently been achieved during tasks such as flexion and extension of the elbow and gait [24–26]. Both online and dynamic methods were based on template matching of MUAP, using features stored in a separation matrix that were periodically updated to account for MUAP changes due to muscle fatigue and changes in muscle length.

Although the periodical update strategy could account for slow and gradual MUAP changes, the separation matrix approach is limited to a fixed number of MUs after offline initialization, which restricts its application to the same muscle within the same session. The number of recruited MUs during a contraction is different for different muscles and different participants, and the amplitude and duration of the MUAP are affected by the diameter of the muscle fibers [27, 28]. Therefore, the MUAP shapes are also different across muscles and participants. In addition, when HD-EMG signals are recorded on the surface of the skin, all MUAPs features are influenced by the relative position between the muscle fiber and the electrodes, as well as the tissue in between [29]. Moreover, the MUAP shapes from the same muscle could also change when recorded from a different day or session. Therefore, the separation matrix approach always requires offline initialization for each application to different muscles, different participants, or the same muscle in a different session. A generalizable approach for neural drive estimation across the aforementioned conditions would dramatically reduce the initialization efforts and simplify real-time human–robot interfacing.

The purpose of this study is to validate the generalizability of a deep CNN framework for neural drive estimation across sessions, muscles, and participants. A deep CNN is chosen because it demonstrated success (a) in many complex image pattern recognition problems regardless of the orientation and distortion of the objects [30, 31]; (b) in EMG and electroencephalography analysis as well as MU identification using HD-EMG [32, 33]. In addition, we have demonstrated that the deep CNN is feasible for neural drive estimation across different contraction intensities and joint angles in our previous study [34]. As the deep CNN is a non-linear complex feature extraction approach, it could handle more challenging situations and store much more information than a linear separation matrix. In this study, we first identified the MU spike trains from HD-EMG signals using BSS and then used both HD-EMG and corresponding CST to train the deep CNN. We validated the generalizability of the deep CNN for within participant variations (i.e. across muscle-session variations), across participants variations, and across muscle-session-participant variations.

2. Method

2.1. Experiment

2.1.1. Participants

Five healthy participants were enrolled in this study (30 ± 6 years, height 183 ± 6 cm, body mass 74 ± 8 kg). All participants had no history of lower leg injury within the past 6 months. The institutional research ethics committee 'Comité de protection des personnes Ile de France XI' approved this study (CPP-MIP-013), and all procedures were in accordance with

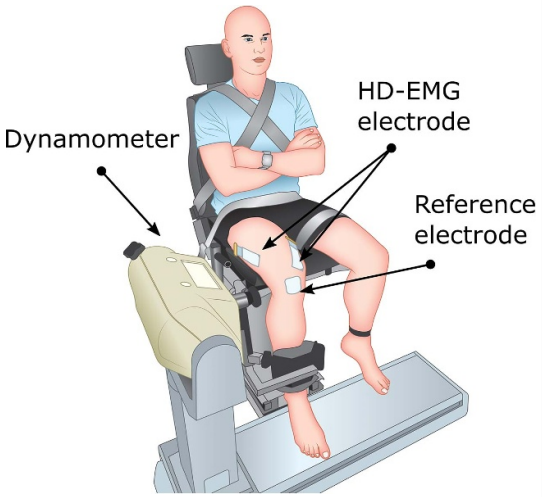


Figure 1. Experimental setup. Participants were seated on a dynamometer (Biodex System 3 Pro, Biodex Medical, Shirley, NY) with their knee joint positioned at 80°. HD-EMG signals were recorded from the vastus lateralis (VL) and the vastus medialis (VM) muscles while participants performed isometric contractions.

the Declaration of Helsinki. All participants provided their written informed consent prior to participation in the study. Note that the experimental data were a subset of a previously published dataset [35].

2.1.2. Experimental protocol

The experimental protocol included two experiment sessions (i.e. S1 and S2) separated by ∼20 months (±1 month). During each session, participants were seated on a dynamometer (Biodex System 3 Pro, Biodex Medical, Shirley, NY) with their torso immobilized with two inextensible straps that crossed the torso. Their hip joint was positioned at 90°, where 0° is the neutral position; their knee joint was positioned at 80°, where 0° is full extension (figure 1). After a series of warm-up contractions, participants performed three maximal voluntary (isometric) contractions for 3–5 s with two minutes of rest in between. Then, participants performed three submaximal isometric contractions by tracking a target displayed on a screen corresponding to 25% of their peak torque. The target followed a trapezoidal trajectory with a 10 s linear ramp-up, a 10 s plateau, and a 10 s linear ramp-down phase. Each contraction was followed by 30 s of rest. Similarly, on the second visit, participants performed three trapezoidal contractions at 25% of their peak torque with a 5 s linear ramp-up, a 20 s plateau, and a 5 s ramp-down phase. Each contraction was separated by a 60 s period of rest. Torque signals were collected using the same Quattrocento system used for HD-EMG acquisition (in the following section).

2.1.3. Data acquisition

During all contraction trials, HD-EMG signals were recorded from the vastus lateralis (VL) and the vastus medialis (VM). Two-dimensional adhesive grids of 64 electrodes (13 × 5 gold-coated electrodes with one electrode absent on a corner; interelectrode distance: 8 mm (ELSCH064NM2, OT Bioelettronica, Italy)) were placed over the VL and VM muscles aligned parallel to the direction of the muscle fascicles. Prior to placing the grids, the skin was shaved and cleaned with an abrasive pad and alcohol. Semi-disposable bi-adhesive foam layers (SpesMedica, Battipaglia, Italy) were used to attach the adhesive grids on the skin. These foam layers were equipped with cavities such that conductive paste (SpesMedica, Battipaglia, Italy) could fill the cavities to make skin-electrode contact. Reference electrodes (Kendall Medi-Trace, Canada) were placed over the patella. The multi-channel acquisition system (EMG-Quattrocento; 400-channel EMG amplifier, OT Bioelettronica, Italy) was used to record EMG signals in monopolar mode, which was then band-pass filtered (10–900 Hz) and digitized at a sampling rate of 2048 Hz.

2.1.4. HD-EMG decomposition

The HD-EMG signals from all channels were band-pass filtered offline (Butterworth 2nd order, 20–500 Hz) and were visually inspected to remove noisy channels with low signal-to-noise ratio or motion artifacts. After pre-processing the data, the HD-EMG signals were decomposed into MU spike trains using the BSS method by Negro *et al* [11]. The BSS method is a general framework for MU decomposition of multi-channel invasive and non-invasive EMG signals by combining fastICA method [8] to identify the MU spike trains and CKC method [6] to further refine the MU spike trains, and it has been extensively validated using experimental and simulated signals. Then, all MU spike trains were visually inspected and manually edited to remove any false positives (FPs; labeled artifact) or false negatives (non-labeled discharges) by an experienced operator. Only the MUs that exhibited a pulse-to-noise ratio (PNR) greater than 30 dB were retained for further analysis. This threshold ensured a sensitivity higher than 90% and a FP rate lower than 2% [29].

2.2. Neural drive estimation using deep CNN

2.2.1. Deep CNN for neural drive estimation

We have previously developed a deep CNN framework to directly estimate the neural drive in the form of the CST [34]. The deep CNN identified the number of MU spikes in a given window of HD-EMG signals. The deep CNN took the window of HD-EMG signals as its input where the width of the window was M (the number of HD-EMG channels) and the length of the window was W (e.g. 40 data points from each channel when window size equals 40). A sliding window approach, which is commonly used in offline EMG analysis methods [15, 23, 36], was used to segment the data with the increment of the window defined as

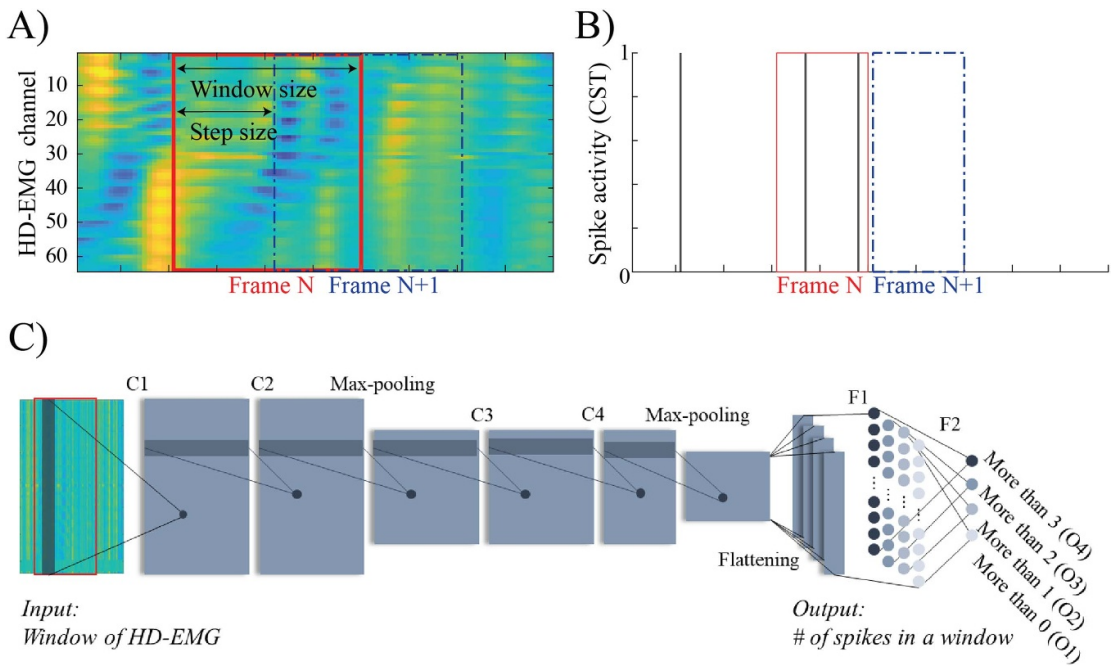


Figure 2. (A) The input of the deep CNN: the HD-EMG signals were segmented into windows using a sliding window method with a window size of 40 and a step size of 20, generating an overlap of 20 data points between windows. (B) The output of the deep CNN: the CST decomposed using BSS was also segmented into windows using a similar sliding window method with a window size of 20 and a step size of 20, generating no overlap between windows to avoid double counting. (C) Deep CNN framework for neural drive estimation in form of CST.

the step size. Then the deep CNN estimated the number of MU spikes within a given window as its output, and if no spikes was identified, the output was set to zero for all nodes (see figure 2). A node was set equal to one for every spike identified within a window (e.g. for two spikes identified, the first two nodes were set to one and the rest to zero).

The full structural design of the deep CNN is presented in [34]. In short, the deep CNN consists of six parametric layers (four convolutional layers and two fully connected layers) and three non-parametric layers (two max pooling layers and one flatten layer) between the input and output layers. The first three convolutional layers (C1–C3) had 128 parallel kernel filters with a kernel size of 3 and the last convolutional layer (C4) had 64 kernel filters to reduce the parameters of the fully connected layers. The two fully connected layers (F1 and F2) had 256 nodes and 1 node, respectively. Max-pooling with size of 2 was applied to C2 and C4 to reduce the dimension of intermediate features and a dropout rate of 50% was applied to C2, C4, and F2 layers to prevent overfitting. A sigmoid activation function, ranging from 0 to 1, was used for the output layer and a rectified linear unit activation function was used for all other layers. For the training of the deep CNN, we used a 'binar_crossentropy' as the loss function considering that the MU activity is a binary state (i.e. 0 or 1). The network weights were updated and optimized using RMSprop optimizer with a base learning rate of 0.001 and mini batches of 64. The training process was terminated if it reached 100 epochs or the training loss had no improvement for 50 consecutive epochs. Previously, we have demonstrated the ability to generalize this deep CNN across different contraction intensities, contraction target profiles (e.g. trapezoidal vs sinusoidal), and muscle geometries (e.g. for different joint angles). Here we aim to demonstrate the ability to generalize across sessions and muscles, and across participants. (The training and validation datasets are available at [37]. The Python implementation of the deep CNN can be found at https://github.com/ywen3/dcnn_mu_decomp.)

2.2.2. Data preparation

The HD-EMG signals were visually inspected, and channels with low signal-to-noise ratio or motion artifacts were assigned the mean values of the signals from surrounding channels. Additionally, we aligned the spike indices to the center of the MUAP. First, the MUAP shape was extracted using the spike indices of each MU. Then, spike indices were adjusted to align with the center of the MUAP. After pre-processing, we used the same sliding window method as in [34] to segment the HD-EMG signals and CST into small windows. For the HD-EMG windowing, we used a window length of 40 data points and sliding step of 20 data points with an overlap of 20 data points

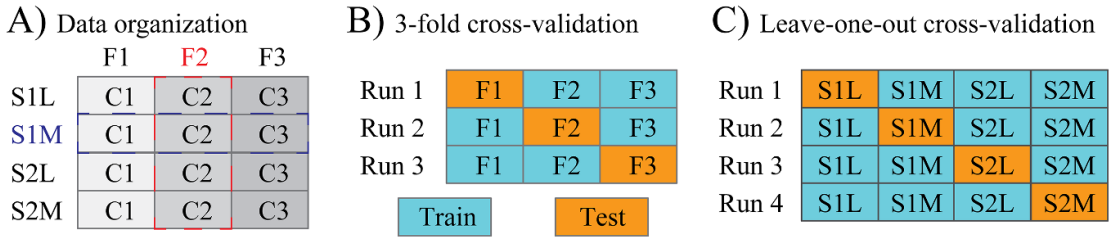


Figure 3. Data organization and validation methods: (A) Data organization. The experimental data was organized in two ways: (1) Fold 1 to Fold 3 represented the three contractions of each dataset (S1L, S1M, S2L, and S2M); (2) S1L, S1M, S2L and S2M represented the four dataset from two muscles in two separated sessions; S1 and S2 represent the first and second session, respectively; L and M represent the VL and VM muscles, respectively. (B) The standard three-fold cross-validation (3CV). For 'Run 1', two folds (F2 and F3) were used for training and one fold (F1) was used for validation. Three runs were performed to get the averaged performance. (C) The leave-one-out cross-validation (LOOCV). For 'Run 1', three datasets (S1M, S2L, and S2M) were used for training and one dataset (S1L) was left out for validation. Four runs were performed to get the averaged performance.

(figure 2(A)). For the CST windowing, we used a window length of 20 data points and sliding step of 20 data points without overlapping to avoid double counting (figure 2(B)). The CST frame and HD-EMG frame were aligned in the center. The summation of the spikes in each CST frame was used to label the corresponding HD-EMG frame. In other words, the effective HD-EMG signals were the 20 data points (approximately 10 ms) in the middle of the frame, but we provided extra 10 data points on each side to improve the estimation performance. Based on our previous study [34], we limited the number of outputs of the deep CNN to 4. If there was one spike in the window, the first node was set to one; if there were two spikes in the window, the first two nodes were set to 1. However, if there were more than four spikes in the window, we set all four nodes to 1. Therefore, we might lose some of the spikes in the training and testing data, but this is the best choice when comparing the output of the deep CNN with the full CST extracted from the BSS [34].

2.2.3. Training and validation of deep CNN

For within-participant validation, we validated the generalizability of the deep CNN in estimating the neural drive across two factors (i.e. muscle and session) using four datasets (i.e. S1L and S1M from session one, S2L and S2M from session two; L refers to VL muscle, and M refers to VM muscle) from each participant. Since the performance of the deep CNN heavily relies on the training and validation datasets, we organized the data into two formats: First, we grouped data from three contractions of each muscle-session condition to obtain one fold of data; Second, we grouped each contraction from four muscle-session conditions to create one fold of data (figure 3(A)). Accordingly, we performed two types of cross-validations (CV): (a) a standard three-fold CV, where each fold included data from one contraction from each dataset (figure 3(B)) and two folds were used for training and one fold was used for validation; (b) a leave-one-out cross-validation (LOOCV), where three datasets were used for training and one dataset was left out for validation (figure 3(C)); the three-fold CV is to validate the capacity of the deep CNN when it has full access to the datasets. The LOOCV is to validate the performance of the deep CNN when trained and tested with different datasets. We performed two types of CVs for each participant.

For across-participant validation, we also performed three-fold CV and LOOCV to investigate the generalizability of the deep CNN for neural drive estimation across participants using five datasets (one from each participant). For the three-fold CV, each fold included data from five contractions, one from each participant; two folds were used to train the deep CNN and the remaining fold was used to validate the deep CNN. For LOOCV validation, the deep CNN was trained with data from four participants and validated with data from the left-out participant. We performed the aforementioned two types of CVs for each muscle in each session.

In addition, we validated the generalizability of the deep CNN across all conditions regardless of the muscles, sessions, and participants. Among 20 datasets (5P*2S*2M), we randomly selected a subset of data for training and used the remaining datasets for testing. Here we also investigated how the training data size (i.e. the number of datasets included for training) affected the generalizability of the deep CNN. First, we trained the deep CNN with each one of the 20 datasets and tested with the remaining 19 datasets. Here, we ran 20 iterations for all possible combinations as a baseline performance. Then, we trained the deep CNN with combinations of multiple datasets (up to six datasets) and tested with the remaining datasets. For each training data size condition, we ran five iterations, each with a randomly selected fixed number of datasets for training, to get representative results.

Table 1. The number of MUs and number of spikes for all participants and experimental conditions.

	Session 1				Session 2			
	VL		VM		VL		VM	
	# of MUs	# of spikes	# of MUs	# of spikes	# of MUs	# of spikes	# of MUs	# of spikes
P1	36	18290	20	8827	30	17581	16	10500
P2	31	16148	18	9553	37	22384	19	13029
P3	22	8363	9	3455	38	23286	13	7392
P4	19	11322	15	9321	33	19459	14	9198
P5	32	24376	16	12886	34	24290	15	9996

P1–P5: Participant 1–5; VL: vastus lateralis muscle, VM: vastus medialis muscle.

2.3. Performance evaluation

The outputs of the deep CNN were summed to form the CST, which was then smoothed using a 400-ms Hanning window to estimate the neural drive [38]. We calculated the correlation coefficient and the normalized root-mean-square error (nRMSE) between the smoothed CST from manually edited MU spike trains from BSS and that from the deep CNN to evaluate the accuracy of deep CNN in neural drive estimation. Specifically, the nRMSE was defined as

$$\begin{aligned} \mathrm{RMSE} &= \sqrt{\frac{\sum_{n=1}^{N}(\hat{x}_n - x_n)^2}{N}} \\ \mathrm{nRMSE} &= \frac{RMSE}{x_{\max} - x_{\min}} \times 100\%, \end{aligned} \tag{1}$$

where $\hat{x}_n$ and x_n are the nth sample of the smoothed CST from deep CNN and the smoothed CST from BSS, respectively. $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of the smoothed CST from BSS, and N is the number of samples.

The deep CNN could generate FP errors during the neural drive estimation, i.e. identifying spike activities that were not identified by the BSS at very low muscle activation. Therefore, we included 2 s data before and after the neural drive identified by the BSS when calculating the correlation coefficient and nRMSE.

2.4. Statistical analyses

A two-factor ANOVA (factor 1: validation methods; factor 2: data sources) was performed to test whether the performance of the deep CNN was validation method dependent or data source dependent. Tukey's honestly significant difference procedure was used to test the statistical significance between conditions. Here, data sources refer to datasets from each one of the five participants for the within-participant validation and datasets from each one of the four session-muscle combinations for across-participant validation. On one hand, if the deep CNN performed well in three-fold CV method but significantly worse in LOOCV method, this could indicate that the knowledge might not be transferable from one participant to another participant or from one muscle to another muscle. On the other hand, if the deep CNN performed well with data from one participant/muscle but much worse with data from another participant/muscle, this could indicate that the data quality has a significant effect on the performance of the deep CNN. Additionally, a Pearson correlation test was performed to quantify the relationship between the generalizability of the deep CNN and the number of MUs to determine the Pearson correlation coefficient (r) and its associated P value ($\alpha = .05$). All analyses were conducted using custom-written scripts in Matlab 2020.

3. Results

3.1. Number of MUs and spikes identified using BSS

The number of MUs and total number of spikes from five participants across four experimental conditions are reported in table 1. The number of MUs and the number of spikes varied across experimental conditions (table 1). For session 1, across participants, the average number of MUs identified using BSS was 28 ± 7 for VL muscle and 16 ± 4 for VM muscle; the number of spikes was $15\,700 \pm 6229$ for VL muscle and 8808 ± 3395 for VM muscle. For session 2, the number of MUs identified using BSS was 34 ± 3 for VL muscle and 15 ± 2 for VM muscle; the number of spikes was $21\,400 \pm 2794$ for VL muscle and $10\,023 \pm 2053$ for VM muscle. Similarly, the number of MUs and the number of spikes varied across participants (table 1). Averaged across all four experimental conditions, P2 was the highest in both the number of MUs (26 ± 9) and the number of spikes ($15\,279 \pm 5449$), while P3 was the lowest in both the number of MUs (20 ± 13) and the number of spikes ($10\,624 \pm 8704$).

3.2. Generalizability of the deep CNN within participant

For the within-participant validation, the correlation coefficient was 0.988 ± 0.003 (mean $\pm$ std across five participants) for three-fold CV and 0.983 ± 0.006 for LOOCV (figure 4(A)). The nRMSE was $5.19 \pm 0.53\%$ for three-fold CV and $6.17 \pm 0.86\%$ for LOOCV

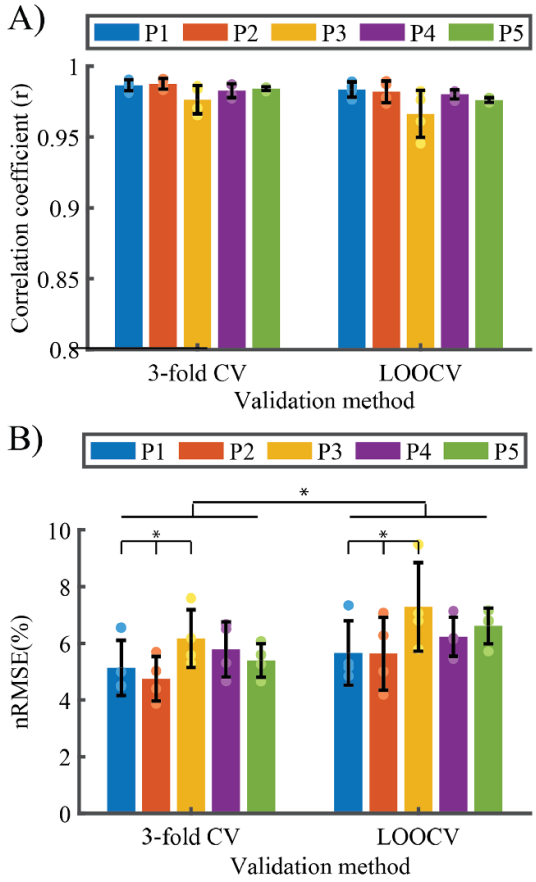


Figure 4. (A) The correlation coefficient between results of the BSS and deep CNN, (B) the nRMSE between results of the BSS and deep CNN from the within-participant cross validation. The validation performed using both three-fold CV and LOOCV for five participants. The error bar represents the standard deviation from multiple runs.

(figure 4(B)). Figure 5 shows that the neural drive from BSS and that from deep CNN are highly correlated ($r = 0.987 \pm 0.005$; $nRMSE = 5.36 \pm 1.22\%$) for all four muscle-session combinations from one representative participant in the LOOCV. Regardless of the validation method, the deep CNN trained and validated with data from P2 generated the best performance with a correlation coefficient of 0.989 ± 0.003 and a nRMSE of $4.94 \pm 0.60\%$; when trained with data from P3, the deep CNN performed the worst resulting in a correlation coefficient of 0.978 ± 0.008 and a nRMSE of $6.66 \pm 1.15\%$. For the correlation coefficient, the ANOVA detected no significant difference between data sources ($p = 0.20$), no significant difference between validation methods ($p = 0.29$), and no significant interaction between validation methods and data sources ($p = 0.99$). For nRMSE, the ANOVA detected significant differences between data sources ($p < 0.01$), significant differences between validation methods ($p < 0.01$), but no significant interaction between validation methods and data sources ($p = 0.77$).

3.3. Generalizability of the deep CNN across participants

For across-participant validation, the correlation coefficient was 0.987 ± 0.003 (mean $\pm$ std across 4 muscle-session combinations) for three-fold CV and 0.989 ± 0.002 for LOOCV (figure 6(A)). The nRMSE was $5.37 \pm 0.61\%$ for three-fold CV and $4.95 \pm 0.57\%$ for LOOCV (figure 6(B)). Regardless of the validation method, the deep CNN performed the best when trained and validated with data from S2L with a resulting correlation coefficient of 0.991 ± 0.001 (mean $\pm$ std across two validation methods) and nRMSE of $4.61 \pm 0.23\%$ (figure 6). The deep CNN performed the worst when trained and validated with data from S2M. The resulting correlation coefficient was 0.985 ± 0.001 and nRMSE was $5.81 \pm 0.32\%$. For both correlation coefficient and nRMSE, the ANOVA detected a significant difference between data sources (r: $p < 0.02$; nRMSE: $p < 0.01$), no significant difference between validation methods (r: $p = 0.15$; nRMSE: $p = 0.12$), and no significant interaction between validation methods and data sources (r: $p = 0.99$; nRMSE: $p = 0.99$).

3.4. Generalizability of the deep CNN across all conditions

The generalizability of the deep CNN across all conditions increased as more datasets were included in the training data (figure 7). Trained with only one dataset, the correlation coefficient was 0.977 ± 0.007 (mean and standard deviation) across 20 runs, ranging from 0.965 to 0.982 (MIX1 in figure 7). Trained with four datasets, the correlation coefficient was 0.984 ± 0.001 (mean and standard deviation) across 5 runs, ranging from 0.982 to 0.985 (MIX4 in figure 7). As shown in figure 8, the number of MUs in the training data has a significant effect on the average correlation coefficient across all testing conditions ($r = 0.60$ and $p < 0.01$).

4. Discussion

This study aimed to validate the feasibility of using deep CNN for neural drive estimation from HD-EMG and to investigate the generalizability of this approach across two muscles, two sessions, and five participants. In principle, the deep CNN can learn the general features of MUAPs from a pool of MUs and estimate the neural drive from HD-EMG in an online manner. We recorded HD-EMG signals from the VL and VM muscles during two different sessions and decomposed the HD-EMG signals into corresponding MU spike trains using a BSS method. Then, the HD-EMG signals and the CST (summation of all MU spike trains) were segmented into data frames for the training and validation of the deep CNN. We performed three-fold CV and LOOCV for within-participant and across-participant situations.

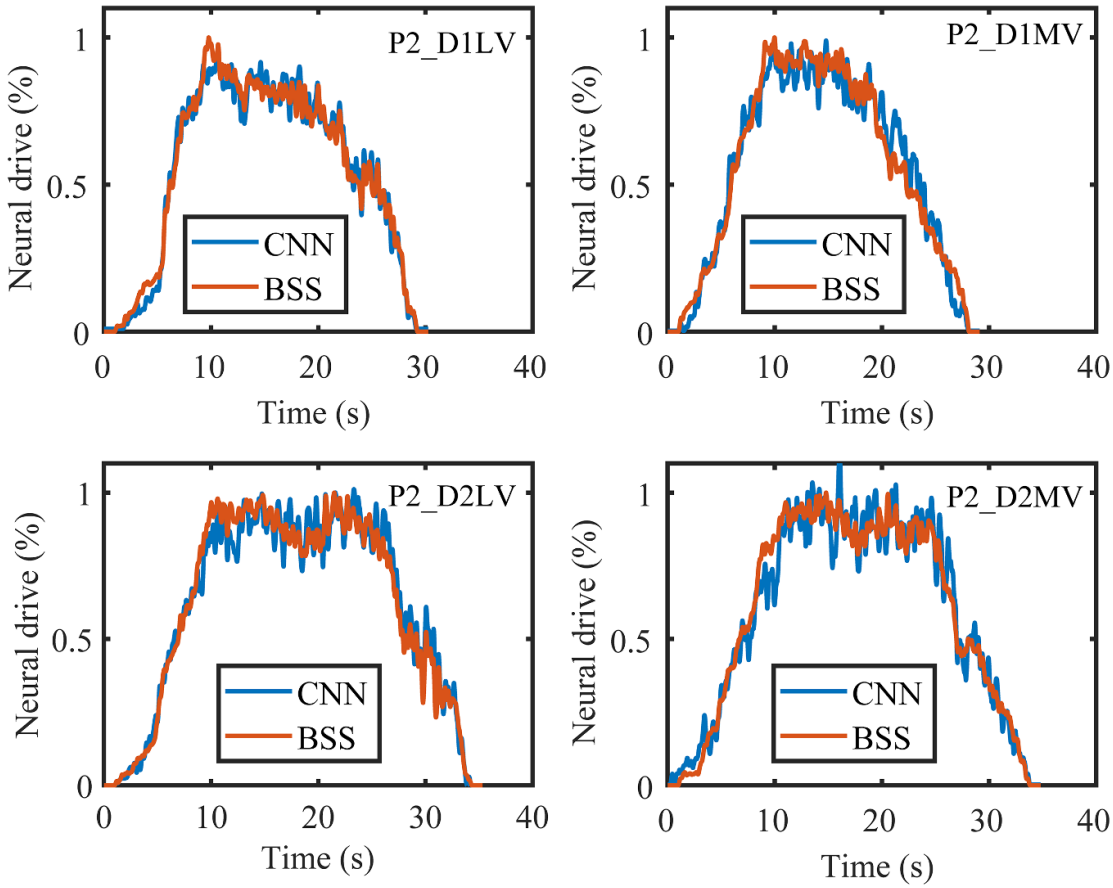


Figure 5. Comparison between the neural drive estimated using BSS and the neural drive estimated using the deep CNN for within-participant LOOCV. The deep CNN was trained with three muscle-session conditions and validated with the remaining one. Here are the results from participant 2.

Without retraining, the deep CNN is generalizable across different muscles and participants for long-term neural drive estimation. More important, in LOOCV, the deep CNN achieved similar performance as in three-fold CV, indicating it is generalizable to new dataset with comparable performance. With a higher number of MUs included in the training data, the deep CNN is more generalizable to different muscles, sessions, and participants.

For data analysis, the ramp-up and ramp-down segments were included to take into account the performance in estimating the neural drive during linear changes in the force generation profile, which is typical in evaluating neural-machine interface [15–17]. The 2 s period before and after neural drive identified by BSS at very low muscle activation was used to account for potential FPs in estimating the neural drive. This is a trade-off to assess the FPs without inflation of the correlation coefficient with a long non-active period. The deep CNN performed reasonably well compared to a previous study on intramuscular EMG decomposition, which reported correlation coefficients of 0.95 ± 0.04 between estimated CST using an online method and offline manually edited CST using EMGlab decomposition software [39]. A few studies first estimated the neural drive using online methods and then used a linear regression model to estimate joint force using neural drive, and they reported correlation coefficients ranging from 0.83 to 0.91 [36, 40]. We reference these studies but cannot make fair comparison between our results and theirs because of differences in selection of muscles and experimental setups.

The performance of the deep CNN varied slightly for each dataset (e.g. a lower correlation coefficient for participant 3) in figures 4 and 6. We speculate that the performance of the deep CNN was related to the number of MUs and the number of spikes in the training data. In table 1, we observed that (a) participant 3 has fewer MUs and spikes than other participants and (b) the VM muscle has fewer MUs and spikes than VL muscle for both sessions across all participants. These observations are aligned with the performance decreases with certain datasets. This is also supported by the generalizability validation of the deep CNN across all conditions. The number of MUs was significantly positively correlated with the averaged correlation coefficient across all testing datasets (figure 8). It worth investigate the effects of the number of MUs on the neural drive estimation, for example, training the deep CNN with data that includes a large number of MUs and testing the deep CNN with data that

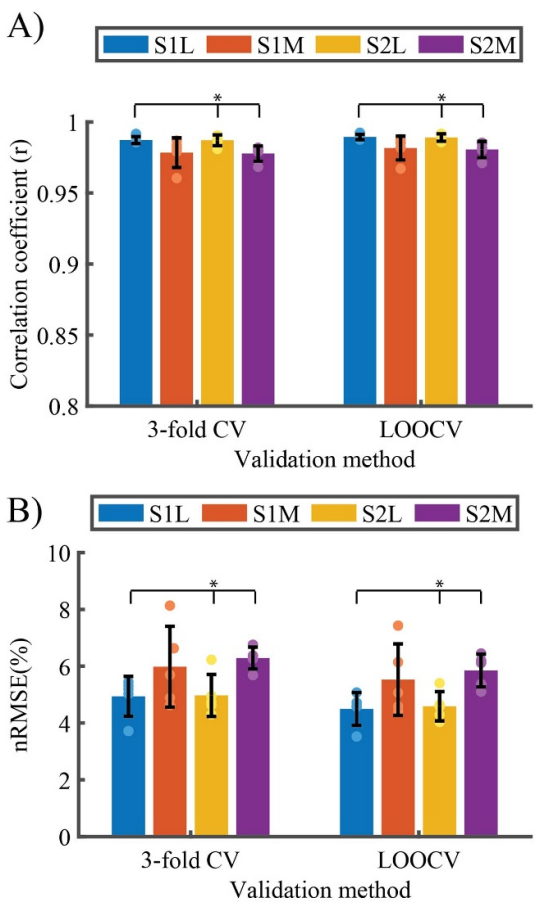


Figure 6. The correlation coefficient (A) and the nRMSE (B) of the deep CNN from the across-participant cross validation. The validation was performed using both three-fold CV and LOOCV for four muscle-session conditions. The error bar represents the standard deviation from multiple runs.

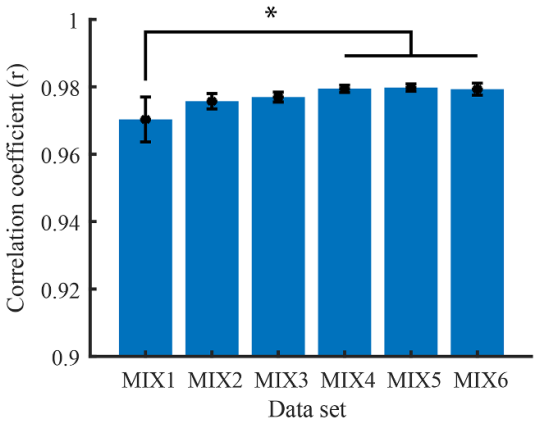


Figure 7. The generalizability of deep CNN across all training conditions with different numbers of datasets. The error bar represents the standard deviation across multiple runs in each condition. One-way ANOVA test identified significant difference ($p < 0.01$) between MIX1 and MIX4, MIX5, MIX6.

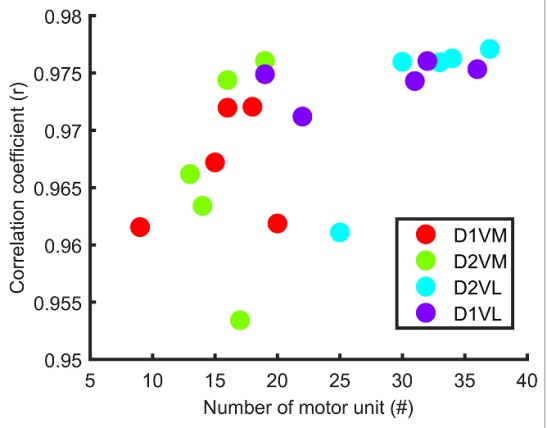


Figure 8. Relationship between the generalizability of deep CNN and the number of MUs in the training data. The training data included only one dataset (i.e. MIX1). There are 20 data points in total, including five participants each with four muscle-session conditions.

includes a small number of MUs, and vice versa. In addition, to make the results more generalizable, we will train the deep CNN with benchmark HD-EMG data and validate with our experimental data in our future study.

In this study, we performed a few pre-processing procedures, such as visual inspection of the HD-EMG signals to remove noisy channels, that need to be automated for online application. Current efforts have been made to automate removal of noisy channels using techniques such as independent component analysis [41], spatial similarity via normalized mutual information between electrodes [42], and outlier detection via local distance-based outlier factor [43]. To this end, we could easily integrate one of these methods to our framework or include a calibration step to remove noisy channels. After automatic removing of noisy channels, we will investigate the robustness of the deep CNN method and how the number of missing channels affects its performance, as this will determine to what extend the deep CNN can reliably estimate the neural drive in real applications.

Additionally, we have only investigated the generalization across two muscles (the VL and VM muscles), two sessions, and five participants. VL and VM muscles have very similar functions in terms of actuating the knee, which may have led to a high cross-correlation between the two muscle activities. In our future work, we will test the generalization across more muscles with higher independence and greater architectural differences (e.g. fusiform vs. pennate muscle) across additional participants, and we will validate the generalizability of our method for neural drive estimation during non-isometric contractions. Currently, MU identification requires manual-edits by an operator, which is time consuming and labor intensive. If simulated data (i.e. simulated HD-EMG signals) can be used to train a deep CNN that can be generalizable for different muscles, this can greatly improve the practicality of this approach to reduce the calibration process.

5. Conclusion

We proposed a general deep CNN framework to estimate the neural drive in the form of CST from HD-EMG signals across different muscles, sessions, and participants. Our study demonstrated that (a) the deep CNN is generalizable across two muscles and five participants, i.e. estimating the neural drive for conditions that are not included in the training data, and (b) the deep CNN is scalable to capture complex features of variant MUAP shapes when trained with combined data that are recorded from different muscles and participants. Compared with the commonly used offline neural drive extraction approach (i.e. BSS), the proposed deep CNN framework could accurately estimate the neural drive using HD-EMG signals. Moreover, the proposed deep CNN is generalizable to different muscles and participants without retraining and could run in a real-time manner as was done in a previous study [32]. Therefore, the proposed deep CNN framework is a promising candidate to extract neural drive for real-time human–machine interface for assistive technology (e.g. exoskeletons and prostheses).

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.6084/m9.figshare.21685418.v1.

Acknowledgment

This work was supported by National Science Foundation/National Robotics Initiative (Grant No. 2024488).

ORCID iDs

Yue Wen https://orcid.org/0000-0001-5297-6230
Sangjoon J Kim https://orcid.org/0000-0002-3245-2752
Simon Avrillon https://orcid.org/0000-0002-2226-3528
Jackson T Levine https://orcid.org/0000-0002-1597-0136
François Hug https://orcid.org/0000-0002-6432-558X
José L Pons https://orcid.org/0000-0003-0265-0181

References

[1] Kuiken T A, Li G, Lock B A, Lipschutz R D, Miller L A, Stubblefield K A and Englehart K B 2009 Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms *JAMA* **301** 619–28
[2] Huang H, Kuiken T A and Lipschutz R D 2009 A strategy for identifying locomotion modes using surface electromyography *IEEE Trans. Biomed. Eng.* **56** 65–73
[3] Lenzi T, De, Rossi S M, Vitiello N and Carrozza M C 2012 Intention-based EMG control for powered exoskeletons *IEEE Trans. Biomed. Eng.* **59** 2180–90
[4] Gordon K E and Ferris D P 2007 Learning to walk with a robotic ankle exoskeleton *J. Biomech.* **40** 2636–44
[5] McGill K C 2004 Surface electromyogram signal modelling *Med. Biol. Eng. Comput.* **42** 446–54
[6] Holobar A and Zazula D 2007 Multichannel blind source separation using convolution kernel compensation *IEEE Trans. Signal Process.* **55** 4487–96
[7] Holobar A and Zazula D 2007 Gradient convolution kernel compensation applied to surface electromyograms *Int. Conf. on Independent Component Analysis and Signal Separation* pp 617–24
[8] Chen M and Zhou P 2016 A novel framework based on FastICA for high density surface EMG decomposition *IEEE Trans. Neural Syst. Rehabil. Eng.* **24** 117–27
[9] Chen M, Zhang X, Lu Z, Li X and Zhou P 2018 Two-source validation of progressive FastICA peel-off for automatic surface EMG decomposition in human first dorsal interosseous muscle *Int. J. Neural Syst.* **28** 1850019
[10] Chen M, Holobar A, Zhang X and Zhou P 2016 Progressive FastICA peel-off and convolution kernel compensation demonstrate high agreement for high density surface EMG decomposition *Neural Plast.* **2016** 3489540
[11] Negro F, Muceli S, Castronovo A M, Holobar A and Farina D 2016 Multi-channel intramuscular and surface EMG decomposition by convolutive blind source separation *J. Neural Eng.* **13** 026027
[12] Li X, Holobar A, Gazzoni M, Merletti R, Rymer W Z and Zhou P 2015 Examination of poststroke alteration in motor unit firing behavior using high-density surface EMG decomposition *IEEE Trans. Biomed. Eng.* **62** 1242–52
[13] Gallego J A, Dideriksen J L, Holobar A, Ibáñez J, Pons J L, Louis E D, Rocon E and Farina D 2015 Influence of common synaptic input to motor neurons on the neural drive to muscle in essential tremor *J. Neurophysiol.* **113** 182–91
[14] Holobar A, Glaser V, Gallego J A, Dideriksen J L and Farina D 2012 Non-invasive characterization of motor unit behaviour in pathological tremor *J. Neural Eng.* **9** 056011
[15] Dai C, Cao Y and Hu X 2019 Prediction of individual finger forces based on decoded motoneuron activities *Ann. Biomed. Eng.* **47** 1357–68
[16] Zhang X, Zhu G, Chen M, Chen X, Chen X and Zhou P 2020 Muscle force estimation based on neural drive information from individual motor units *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** 3148–57
[17] Yu Y, Chen C, Sheng X and Zhu X 2020 Wrist torque estimation via electromyographic motor unit decomposition and image reconstruction *IEEE J. Biomed. Health Inform.* **25** 2557–66
[18] Kapelner T, Vujaklija I, Jiang N, Negro F, Aszmann O C, Principe J and Farina D 2019 Predicting wrist kinematics from motor unit discharge timings for the control of active prostheses *J. Neuroeng. Rehabil.* **16** 47
[19] Kapelner T, Sartori M, Negro F and Farina D 2020 Neuro-musculoskeletal mapping for man-machine interfacing *Sci. Rep.* **10** 5834
[20] Chen C, Chai G, Guo W, Sheng X, Farina D and Zhu X 2019 Prediction of finger kinematics from discharge timings of motor units: implications for intuitive control of myoelectric prostheses *J. Neural Eng.* **16** 026005
[21] Dai C and Hu X 2020 Finger joint angle estimation based on motoneuron discharge activities *IEEE J. Biomed. Health Inform.* **24** 760–7
[22] Glaser V, Holobar A and Zazula D 2013 Real-time motor unit identification from high-density surface EMG *IEEE Trans. Neural Syst. Rehabil. Eng.* **21** 949–58
[23] Zheng Y and Hu X 2019 Real-time isometric finger extension force estimation based on motor unit discharge information *J. Neural Eng.* **16** 066006

[24] De Luca C J, Chang S S, Roy S H, Kline J C and Nawab S H 2015 Decomposition of surface EMG signals from cyclic dynamic contractions *J. Neurophysiol.* **113** 1941–51
[25] Glaser V and Holobar A 2019 Motor unit identification from high-density surface electromyograms in repeated dynamic muscle contractions *IEEE Trans. Neural Syst. Rehabil. Eng.* **27** 66–75
[26] Chen C, Ma S, Sheng X, Farina D and Zhu X 2020 Adaptive real-time identification of motor unit discharges from non-stationary high-density surface electromyographic signals *IEEE Trans. Biomed. Eng.* **67** 3501–9
[27] Howard J E, McGill K C and Dorfman L J 1988 Age effects on properties of motor unit action potentials: ADEMG analysis *Ann. Neurol.* **24** 207–13
[28] Howard J E, McGill K C and Dorfman L J 1988 Properties of motor unit action potentials recorded with concentric and monopolar needle electrodes: ADEMG analysis *Muscle Nerve* **11** 1051–5
[29] Holobar A and Farina D 2014 Blind source identification from the multichannel surface electromyogram *Physiol. Meas.* **35** R143–65
[30] Ahlawat S, Choudhary A, Nayyar A, Singh S and Yoon B 2020 Improved handwritten digit recognition using convolutional neural networks (CNN) *Sensors* **20** 3344
[31] He K, Zhang X, Ren S and Sun J 2015 Spatial pyramid pooling in deep convolutional networks for visual recognition *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 1904–16
[32] Wen Y, Avrillon S, Hernandez-Pavon J C, Kim S J, Hug F and Pons J L 2021 A convolutional neural network to identify motor units from high-density surface electromyography signals in real time *J. Neural Eng.* **18** 056003
[33] Clarke A K, Atashzar S F, Vecchio A D, Barsakcioglu D, Muceli S, Bentley P, Urh F, Holobar A and Farina D 2021 Deep learning for robust decomposition of high-density surface EMG signals *IEEE Trans. Biomed. Eng.* **68** 526–34
[34] Wen Y, Kim S J, Avrillon S, Levine J, Hug F and Pons J 2022 A deep CNN framework for neural drive estimation from HD-EMG across contraction intensities and joint angles *IEEE Trans. Neural Syst. Rehabil. Eng.* **30** 2950–9
[35] Avrillon S, Del Vecchio A, Farina D, Pons J, Vogel C, Umehara J and Hug F 2021 Individual differences in the neural strategies to control the lateral and medial head of the quadriceps during a mechanically constrained task *J. Appl. Physiol.* **130** 269–81
[36] Xu F, Zheng Y and Hu X 2020 Real-time finger force prediction via parallel convolutional neural networks: a preliminary study *Annual Int. Conf. IEEE Engineering in Medicine and Biology Society* pp 3126–9
[37] Wen Y, Kim S J, avrillon S, Levine J, Hug F and Pons J 2022 HD-EMG signals and corresponding neural drive signals for training and validation of a deep CNN for neural drive estimation across muscles and participants figshare *Dataset.* (https://doi.org/10.6084/m9.figshare.21685418.v1)
[38] De Luca C J, LeFever R S, McCue M P and Xenakis A P 1982 Control scheme governing concurrently active human motor units during voluntary contractions *J. Physiol.* **329** 129–42
[39] Karimimehr S, Marateb H R, Muceli S, Mansourian M, Mañanas M A and Farina D 2017 A real-time method for decoding the neural drive to muscles using single-channel intra-muscular EMG recordings *Int. J. Neural Syst.* **27** 1750025
[40] Dai C, Zheng Y and Hu X 2018 Estimation of muscle force based on neural drive in a hemispheric stroke survivor *Frontiers Neurol.* **9** 187
[41] Zheng Y and Hu X 2019 Interference removal from electromyography based on independent component analysis *IEEE Trans. Neural Syst. Rehabil. Eng.* **27** 887–94
[42] Bingham A, Jelfs B, Arjunan S and Kumar D 2018 Identifying noisy electrodes in high density surface electromyography recordings through analysis of spatial similarities *2018 40th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE) pp 2325–8
[43] Marateb H, Rojas-Martínez M, Mansourian M, Merletti R and Villanueva M A M 2012 Outlier detection in high-density surface electromyographic signals *Med. Biol. Eng. Comput.* **50** 79–89