

A generic neural network model to estimate populational neural activity for robust neural decoding

Rinku Roy, Feng Xu, Derek G. Kamper, Xiaogang Hu^{*}

Joint Department of Biomedical Engineering, University of North Carolina at Chapel Hill and North Carolina State University, USA

ARTICLE INFO

Keywords:

Neural networks
Hand function
Neural decoding
Motor unit
Finger force

ABSTRACT

Background: Robust and continuous neural decoding is crucial for reliable and intuitive neural-machine interactions. This study developed a novel generic neural network model that can continuously predict finger forces based on decoded populational motoneuron firing activities.

Method: We implemented convolutional neural networks (CNNs) to learn the mapping from high-density electromyogram (HD-EMG) signals of forearm muscles to populational motoneuron firing frequency. We first extracted the spatiotemporal features of EMG energy and frequency maps to improve learning efficiency, given that EMG signals are intrinsically stochastic. We then established a generic neural network model by training on the populational neuron firing activities of multiple participants. Using a regression model, we continuously predicted individual finger forces in real-time. We compared the force prediction performance with two state-of-the-art approaches: a neuron-decomposition method and a classic EMG-amplitude method.

Results: Our results showed that the generic CNN model outperformed the subject-specific neuron-decomposition method and the EMG-amplitude method, as demonstrated by a higher correlation coefficient between the measured and predicted forces, and a lower force prediction error. In addition, the CNN model revealed more stable force prediction performance over time.

Conclusions: Overall, our approach provides a generic and efficient continuous neural decoding approach for real-time and robust human-robot interactions.

1. Introduction

1.1. Background

Neural decoding of individual finger forces is a crucial part of the neural machine interface. It enables the control of assistive [1–3] or rehabilitative devices [4,5]. Surface electromyography (sEMG) is a noninvasive approach to drive such devices [5–8]. In conventional myoelectric control, pattern recognition algorithms uses time-domain [9] or frequency-domain features of sEMG [10] for discrete gesture classifications [11]. These approaches are relatively simple to implement. However, the classification output is limited to a few predefined gestures, and the generalizability to different users is also unsatisfactory. In parallel, EMG features such as EMG amplitude can be made proportional to the movement of robotic devices using regression approaches for continuous decoding [12–14]. However, sEMG signals are intrinsically unstable due to electrode shifts [15,16] or muscle fatigue [17,18], which requires frequent recalibration of model parameters.

An alternative approach is to perform neuronal population decoding based on decomposition of EMG signals [19,20], where the output is continuous neuron firings of a number of motor units (MUs) suitable for proportional control of assistive devices. Substantial progress has been made to extract the MU discharge information from EMG signals via different decomposition algorithms [21–28], and the performance tends to be more robust for continuous motor intent interpretation than EMG amplitude signals [29,30]. Nevertheless, one drawback of such approach is the inefficiency in calculating the composite MU firing activities. During the decomposition process, the spike trains corresponding to individual MUs are extracted and then additively merged in order to represent the neural population behavior. As a result, the information of individual MUs, obtained in the most time-consuming step of the decomposition algorithm, is lost, causing the decomposition-based decoding approach to be inefficient for real-time implementation.

Deep neural network models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have also been explored

^{*} Corresponding author. University of North Carolina at Chapel Hill, 10206B Mary Ellen Jones Building, Chapel Hill, NC, 27599, USA.

E-mail address: xiaogang@unc.edu (X. Hu).

for sEMG decoding. Although RNNs are suitable for time-series data, because of their lateral connections and optional memory cells [31], previous work showed that a two-layer CNN outperforms RNNs for onset detection of audio signal processing [32]. In addition, RNNs generally have a higher computational load compared to CNNs. Gesture recognition can be achieved using multi-stream CNNs and representation fusion based on sEMG features [33], and CNN outperforms traditional classifiers [34,35]. Moreover, CNN has been used for inter-subject recognition of hand gestures [36,37], which demonstrates the generalizability of CNN across subjects. Regression based decoding of wrist movements has also been explored recently using CNN [38] that maps EMG activities directly to continuous wrist joint kinematics. However, the model needs to be trained for individual subjects, because of substantial inter-subject variations in the specific relations between EMG features and motor output.

1.2. Research gap and focus of the current study

Despite the exciting progress in myoelectric based decoding, we still lack robust and efficient neural decoding algorithms that can continuously interface with assistive robotic hands at individual finger levels. To address this gap, we implemented a CNN-based algorithm to continuously decode the populational neuron firing activity in real-time based on HD-EMG signals. EMG signals are intrinsically stochastic, and high spatiotemporal variations have been observed. To address these variations, spectrum features are commonly computed prior to the neural network inference in neural signal processing [39,40]. Therefore, instead of directly using raw EMG signals as the CNN input, we first extracted two sets of EMG features, termed energy map (spatiotemporal features of EMG amplitude) and frequency spectrum map (spatiotemporal features of EMG frequency). High-level representations of these two features were then extracted by the convolutional layers and fused into a representation vector for the estimation of populational neuron discharge frequency. These hand-crafted feature calculations allowed us to accurately learn the EMG mapping to populational neuron firing activities in an efficient manner with a relative small dataset. We compared the performance of our neural decoding algorithm with two state-of-the-art neural decoding approaches: a MU decomposition method and a classic EMG amplitude method. One innovation of this approach is that we trained the model using motoneuron discharge frequency instead of direct force output. This allowed us to establish a robust and generic model that is generalizable to different human subjects, because EMG features have a more direct relation with neuron discharge activity than with force output. The main contribution of the study is that we provided a generic, efficient, and real-time neural network model for robust and continuous neural decoding, which allows intuitive human-robot interactions.

2. Methods

2.1. Human participants

Seven neurologically intact participants (age: 22–34) were recruited. All participants gave informed consent via protocols approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

2.2. Data acquisition

During data acquisition, the participants were seated in a height-adjustable chair in front of a desk where their forearms were put on a soft foam pad. A pair of stiff form pads were placed on both sides of the hands, holding the palm and the back of the hand to reduce unintended hand movement and torque transmission from the wrist.

To measure the individual finger extension forces, the index, middle, ring, and pinky fingers were separately secured to four miniature load

cells (SM-200 N, Interface), sampling at 1000 Hz. In the meantime, an 8×20 HD-EMG electrode array was attached to each participant's forearm to record the activities of the extensor digitorum communis (EDC) muscle (Fig. 1). The electrodes were 3 mm in diameter and evenly distributed on the electrode array at a distance of 10 mm. The EMG signals were amplified with a gain of 1000 and band-pass filtered from 10 to 900 Hz. The EMG signals were sampled at 2048 Hz using EMG-USB2+ (OT Bioelettronica). The participants were asked to follow a predefined pseudorandom trajectory ranging from 0% to 40% maximum voluntary contraction (MVC) for each finger for 5 min. Both the target trajectory and real-time feedback of the actual force were displayed to the participants. The recordings for each participant of the index, middle, and ring fingers were conducted in a random order with a 2-min interval between trials to alleviate muscle fatigue. The pinky finger was excluded due to inevitable co-contractions with other fingers [41, 42] and overlapped muscle activation distribution [43].

2.3. MU discharge time calculation using FastICA

As a training target for the CNN model, fast independent component analysis (FastICA)-based decomposition algorithm [44] was used to extract MU discharge activities. Briefly, the discharge timings of the MUs were recognized as a convolutive mixture of delta function series in EMG signals with additive white noise [24]. After the EMG signals were extended and whitened, the decomposition procedure extracted individual spike trains using the FastICA algorithm based on iterations of the fixed point algorithm [45,46]. The output of each iteration of the fixed-point algorithm was the source signal and the separation vector. A binary spike train of a single MU was further derived from the source signal based on a binary classification based on the *Kmeans++* algorithm. The derived binary spike trains were shortlisted using a modified 'Silhouette distance [47]' measurement [44] by quantifying the classification quality of the source signals. The retained spike trains were further merged into a single composite train representing the populational firing frequency of the MU pool. The populational firing frequency was then normalized and categorized into target classes ranging from 0 to 15 for the neural network training, which was a tradeoff between the resolution of the estimation and the convergence of the neural network with limited data. 0 corresponded to the lowest firing rate, and 15 corresponded to the highest firing rate. During the normalization, the populational firing frequency was divided by the maximum firing frequency, and then scaled up to 15.

2.4. Force estimation using CNN

Feature calculation for CNN model: The 160-channel HD-EMG recordings were first segmented into a sequence of 64-sample (31.25 ms) windows with a step of 16-sample (7.81 ms). Two types of features (Frequency and Energy maps) were estimated from it (Fig. 2). Frequency spectrum map (F_{spec}) was calculated using Fourier transform (FT) within a window for each channel. A Hanning window was applied before the FT. The spectrum vector was $(window - 2)/2 = 33$ in height. 160 spectrum vectors (one per channel) were calculated within each window, leading to a 33×160 (height $\times$ width) frequency map. Energy map (F_{engy}) was calculated based on the mean square (MS) value of all the samples of each channel within the window. The dimension of the energy map was 8×20 (height $\times$ width), matching the spatial distribution of the electrode grid as shown in Fig. 1.

CNN Model: We implemented a neural network model (detailed in Table 1) with two parallel convolutional neural networks before the feature fusion and fully connected layers for the output of the populational discharge rate. The two EMG feature maps were fed to two separate pathways (Fig. 2). To provide the neural network with more contextual information for the estimation of populational neuronal firings, we combined the energy maps and frequency maps from five consecutive windows. This resulted in the dimensions of the two

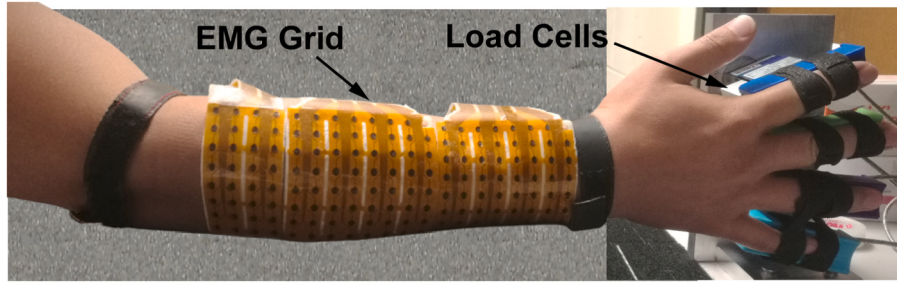


Fig. 1. Experimental setup for HD-EMG and force signal acquisition.

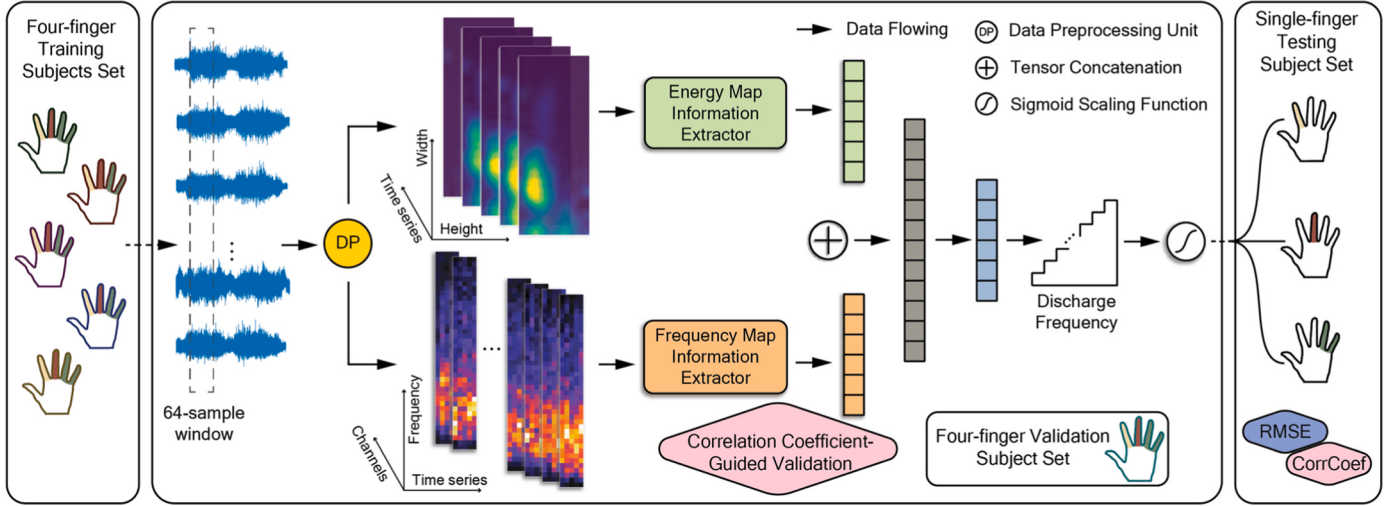


Fig. 2. Diagram of the proposed CNN model. DP in yellow circle represents the data processing module. Modules with diamond shapes denotes the evaluation metrics used, where ‘CorrCoef’ denotes the correlation coefficient evaluation and ‘RMSE’ denotes the root mean square error evaluation. Data from different participants are represented in different color of the hand profiles. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Convolutional neural network structure.

Input Feature		$F_{engy} \in \mathbb{R}^{T \times H \times W}$	$F_{spec} \in \mathbb{R}^{N \times M \times T}$
CNN module		Energy map information extractor	Frequency spectrum map information extractor
Parallel CNN structure	L#1	Conv-10@2 × 6	Conv-160@6 × 2
	2	Maxpool	Maxpool
	3	Conv-20@2 × 4	Conv-320@6 × 2
	4	Maxpool	Maxpool
	5	Conv-40@2 × 2	Conv-640@3 × 3
	6	Maxpool	Maxpool
	7	Conv-80@2 × 2	Conv-1280@2 × 2
	8	Maxpool	Maxpool
	9	FC-256	FC-1024
	10	–	FC-512
	11	FC-256	
	12	Dropout	
	13	FC-16 (Output discharge frequency)	

Note: The number after ‘Conv’ denotes the number of convolutional kernels in each convolutional layer, while the numbers after ‘@’ denotes the shape of the kernel in that layer. The ReLU activation function was added after each fully connected layer and each convolutional layer before Maxpool. The number after ‘FC’ denotes the number of neurons in the fully connected layer. ‘L#’ denotes layer number. Number of total trainable parameters in the proposed convolutional neural network structure: 8,345,886.

features to be $F_{engy} \in \mathbb{R}^{T \times H \times W}$ and $F_{spec} \in \mathbb{R}^{N \times M \times T}$. We processed the dimension arrangement of the two features in different ways so that the neural network can inspect the data in multiple aspects. Specifically,

Table 2
Details of the CNN parameters.

# of trials	# of trials for training	# of trials for validation	# trials for Testing	Learning Rate
210	150	30	30	3e-4

F_{engy} had a height (H) and width (W) of eight pixels and 20 pixels respectively, covering the EDC muscle by the electrode array. The third dimension of F_{engy} was the depth in time (T), where $T = 5$ consecutive windows. Similarly, F_{spec} had a height (M) of 33 frequencies, and a width (T) of 5 consecutive windows, because the spectrum vectors from five windows were concatenated over time to form one frequency spectrum map. The third dimension of F_{spec} was the depth of the number of channels (N), where $N = 160$ EMG channels (see Table 2).

The upper pathway of the network (as in Fig. 1) extracted the information from the energy map feature. Convolutional kernels in each layer were able to explore the shape of the activated area and the change in the EMG amplitude in each channel. The depth dimension provided the information of changes in five energy maps varied over time. The lower pathway of the network extracted the information from the frequency spectrum feature. The height and width of the kernels in the lower CNN pathway covered the frequencies and time of the frequency map respectively, which enabled the exploration of the changes in frequencies at different time scales. As neuron firing spikes could arrive at different timings among channels, the CNN kernel explored the temporal-frequency relation among channels. By passing through four

convolutional layers, the high-level representation of the energy and frequency spectrum map feature was extracted. Both high-level representations were viewed as representation vectors, which were passed through one or two local fully connected layers, and concatenated into a fusion representation vector. Dropout ($p = 0.5$) [48] was applied before the output layer to prevent overfitting. The estimated discharge frequency was obtained at the output layer.

Model Training and Validation: We randomly chose one participant as validation and a different participant as testing out of the seven participants. The training data were from the five remaining participants. In the training session, the preprocessed features from the five participants were shuffled both in time and participant before fed to the neural network. The shuffle was made among the features along different timestamps in the trials and in different subjects. In the evaluation session, including the validation and the final testing, the features were fed to the model in their original chronological order, generating a new estimation every 16 samples (7.81 ms). The training of the neural network was optimized by the Adam optimizer [49] at a learning rate of $3e-4$.

In the validation session, only the correlation coefficient between the predicted force and the measured force was used. The validation was performed every 500 iterations by evaluating the performance of the model on three 75-s segments randomly chosen from three 5-min index, middle, and ring finger trials of the validating participant. The number of iterations of the final model selected was determined by the highest arithmetic mean of the three correlation coefficients among all validation sessions. After the training session has reached the maximum number of iterations, the neural network model at the number of iterations with highest correlation coefficient value was used to determine the network parameters. The network training session took 8 h, and the procedure was repeated seven times, until all the participants were used as a testing participant.

Force Estimation: The estimated populational firing frequency was smoothed by a moving average filter of 0.5-s windows with 0.1-s moving steps. A linear regression model initialized with the first 25 s of data was used to predict the forces based on the firing frequency. The correlation coefficient and root mean square error (RMSE) between the predicted force and the measured force were used in the testing session to evaluate the force estimation performance. Note that the RMSE was not used during the validation session to finalize the network model parameters.

2.5. Force estimation using FastICA

Although MU decomposition was performed offline for neural network training, MU firing activities needed to be obtained in real-time for force estimation. We then performed online FastICA decomposition based on an earlier study [44]. A multivariate linear regression model [50] was initialized with the first 25-s of the data, and was applied to the subsequent MU firing activities for force estimation.

2.6. Force estimation using EMG amplitude

The root mean square (RMS) values were calculated using the same windowing method for each channel and averaged across all channels. A linear regression model was also initialized with the first 25 s of the data. To reduce computation time and limit the influence of channels with low signal-to-noise ratio, only 60 channels with the highest RMS values within the first 25 s were used for both the FastICA and EMG amplitude methods [51,52].

2.7. Statistical measure

Repeated measures analysis of variance (ANOVA) was performed on the dependent variables. A significance level α was set as 0.05. A post hoc analysis was conducted using Bonferroni adjustment when necessary. As the correlation coefficient values were close to 1 and had an

upper limit of 1, arcsine square root transformation [53] was used prior to ANOVA to ensure normality of residuals.

3. Results

The implemented neural network is computationally efficient with inference time of 4.26 ms (Table 3), including the time to compute the frequency spectrum map and the energy map. This delay is well below the acceptable loop delay (100–150 ms) for real-time control of assistive devices [54].

Fig. 3 shows an exemplar trial of real-time and continuous force predictions using the three methods. The measured force was considered as the ground truth for subsequent evaluations. The EMG amplitude method was the least accurate one among the three methods. This is more evident at the end of the force estimation due to a drift of the EMG amplitude. The real-time prediction by CNN and FastICA was more stable over time than the EMG amplitude method.

Fig. 4 illustrates the overall results of correlation coefficients and mean square error (RMSE) of the three methods. The CNN model revealed the best performance, with the highest correlation coefficients and the lowest RMSE. The difference in the correlation coefficients was significant ($F(2,12) = 19.655$, $p < 0.001$) across the methods. The post hoc comparison showed that the correlation coefficients of the CNN model (0.92) were significantly higher ($p < 0.01$) than that of the FastICA method (0.86), and that of the EMG amplitude-based method ($p < 0.01$). The difference in the RMSE results was also significant ($F(2,12) = 6.715$, $p = 0.011$). The RMSE results were 5.0% (CNN), 5.6% (FastICA method), and 7.6% (EMG amplitude method), respectively.

Fig. 5 (A) shows the correlation coefficients grouped by different fingers. The two-way (*method* $\times$ *finger*) ANOVA indicated that both the *method* ($F(2,12) = 19.655$, $p < 0.001$) and the *finger* ($F(2,12) = 9.147$, $p = 0.004$) was significant, with no interaction. The post hoc comparison revealed that the CNN model was significantly higher than the FastICA method ($p < 0.01$) and the EMG amplitude-based method ($p < 0.01$). Significant differences among fingers only appeared between the index and ring fingers ($p < 0.05$).

Fig. 5 (B) depicts the RMSE grouped by different fingers. The ANOVA showed that both *method* ($F(2,12) = 6.715$, $p = 0.035$) and *finger* ($F(2,12) = 8.126$, $p = 0.006$) had significant effects on the RMSE, with no significant interaction. The post hoc revealed a significant difference among fingers between the index and ring fingers ($p < 0.05$).

We also quantified the stability of the force estimation over time. The first 25-s segment of each 5-min trial was used to initialize the linear regression for all the three methods, and thus was not included for the analysis. The remaining 275-s of data were divided into four segments, resulting in four 68.75-s segments, where the correlation coefficients and RMSE were calculated separately for each segment.

Fig. 6 illustrates the correlation coefficients over different time segments. The two-way (*method* $\times$ *segment*) ANOVA showed a significant interaction ($F(6,36) = 7.967$, $p < 0.001$). The post hoc comparison revealed that there were no significant differences in correlation coefficients among the three methods in segment 2 ($p > 0.05$), but the CNN model was higher than the FastICA method from segments 3 to 5 (11%, $p < 0.05$; 14.1%, $p < 0.01$; 19.4%, $p < 0.001$). The CNN model was also

Table 3
Computation time.

Approach	CNN	FastICA-online	FastICA-offline	EMG
Time	4.26 ms (62.5 ms)	60 ms (1 s)	15 s (1 s)	0.38 ms (0.5 s)

Note: The table denotes time in the format: calculation time (required window length of the HD-EMG segment for computation). The computation time was based on a PC with Intel Core i7-6700@3.4 GHz, NVIDIA GTX 1080, and 24 GB of memory.

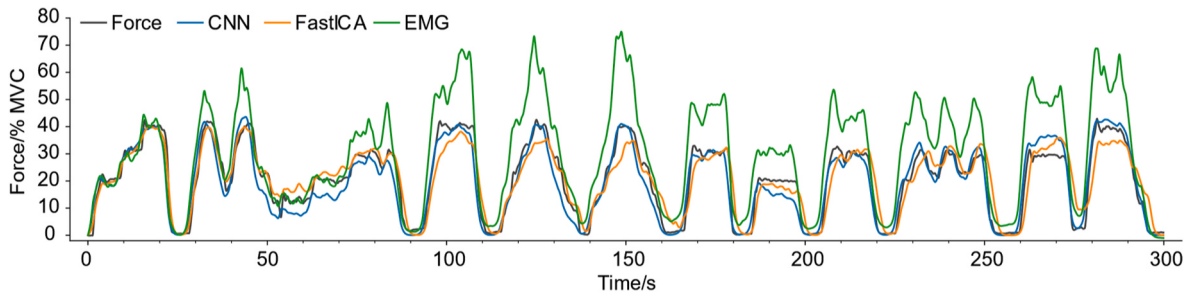


Fig. 3. Example of force prediction of individual fingers in real-time using the CNN model, the FastICA method, and the EMG amplitude method.

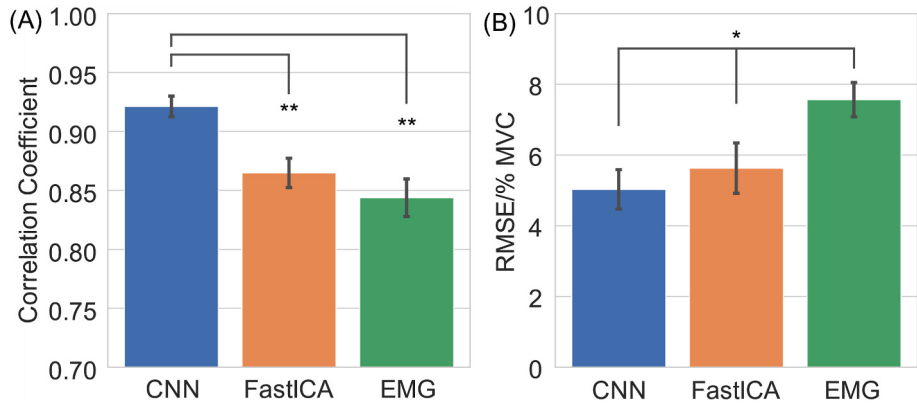


Fig. 4. (A) Overall correlation coefficient results of the three methods averaged across all fingers. (B) Overall RMSE results of the three methods across all fingers. The error bar represents standard error. *, $p < 0.05$. **, $p < 0.01$.

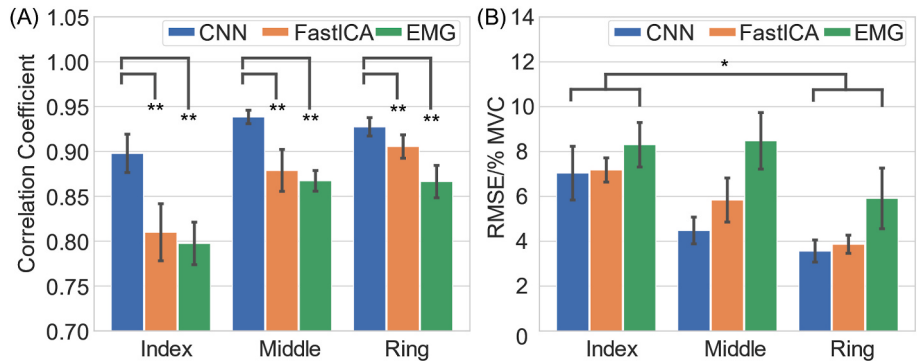


Fig. 5. (A) The Correlation coefficient results of different fingers across all subjects. (B) The RMSE results of different fingers across all subjects. The error bar represents standard error. *, $p < 0.05$. **, $p < 0.01$.

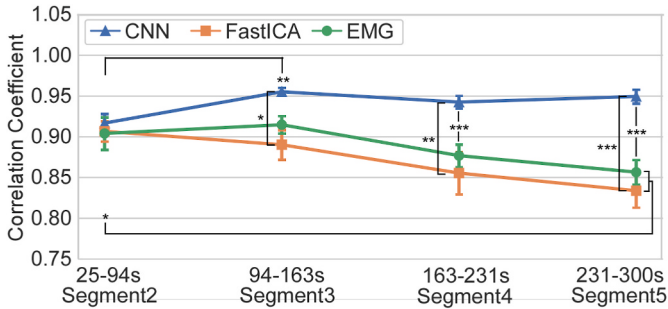


Fig. 6. The correlation coefficient results across all subjects over four segments. Time stamps in the label are rounded to the nearest integer. The error bar represents standard error. *, $p < 0.05$. **, $p < 0.01$. ***, $p < 0.001$.

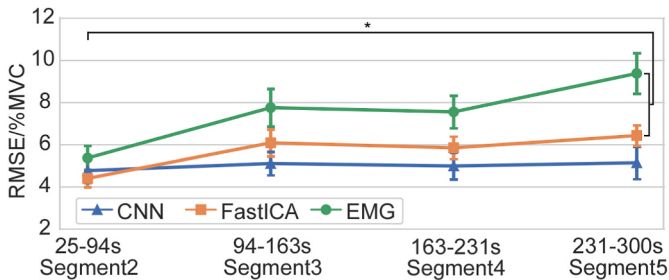


Fig. 7. The RMSE results across all subjects over four segments. Timestamps in the label are rounded to the nearest integer. The error bar represents standard error. *, $p < 0.05$.

higher than the EMG amplitude-based method from segments 4 to 5 (11.8%, $p < 0.001$; 16.3%, $p < 0.001$).

Fig. 7 shows the RMSE over different time segments. The ANOVA showed a significant interaction ($F(6,36) = 4.819$, $p = 0.001$). The differences among the three methods were not significant in segment 2 ($p > 0.05$). There were no significant differences among four segments of the CNN model in RMSE (all $p > 0.05$), indicating a relatively stable and robust estimation performance. In contrast, the RMSE of the FastICA method and the EMG amplitude method increased significantly from segment 2 to segment 5 ($p < 0.05$).

4. Discussion

In this study, we performed populational neural decoding in real-time based on CNNs without obtaining individual motor unit information as in spike sorting. We first extracted two types of features from the HD-EMG signals, energy map feature and frequency spectrum map feature. These features were learned by the CNNs to extract high-level representations, and were used to predict the populational neuronal firing frequency. The estimated force output from the CNN model was compared with two state-of-the-art methods: FastICA-based decomposition and conventional EMG-amplitude-based approaches. We found that the CNN model outperformed the FastICA method and the EMG amplitude method as indicated by a higher correlation coefficient, a lower RMSE, and a more robust prediction performance over time. The results can provide a generic, computationally efficient, and real-time neural network model for continuous neural decoding. The outcomes can facilitate intuitive and robust human-machine interactions. Our CNN based neural decoding method has several advantages compared with existing approaches.

Decoding accuracy: The CNN model is more accurate even though it is trained on a dataset from a pool of subjects, while the FastICA and EMG methods only performed subject-specific decoding. The current decomposition algorithms [19,55] are relatively accurate but with certain levels of decomposition errors. We developed several processes to facilitate the CNN model training to improve decoding performance. First, the offline decomposition method was used to obtain the firing frequency for the training. Second, we filtered the results of the offline decomposition algorithm by evaluating the performance of correlation coefficients with force. If the correlation coefficient was lower than 0.70 for a single fragment, all the decomposition results in the data fragment would not be used as the CNN training target. Lastly, the validation session served as a guidance during the training process. As a result, the chosen number of iterations with the highest validation performance has the learned parameters from the most accurate firing events.

Computational Efficiency: The proposed approach is also considered efficient regarding the computational load. The calculation to sort individual spike trains was not needed, and the direct output was populational firing frequency, making the algorithm more straightforward and applicable in real-time. The decomposition approach first extracts individual MU spike trains, and the spike trains are then additively merged to predict motor output. In the process, the information of individual MUs is lost and thus is considered inefficient. The output of the CNN model was the populational firing frequency, which can be directly used for multiple finger force prediction tasks. Besides, FastICA method required 1 s of data to perform the separation matrix multiplication, and the EMG amplitude method also needed 0.5-s of data for reliable predictions. In contrast, the CNN method only needed two contextual windows in advance for prediction. The algorithm was capable of calculating the prediction procedure starting from feature preprocessing and output discharge frequency within the interval between two predictions. As a result, the algorithm was practical for real-time estimation. Given the accuracy and robustness over time in the force estimation, the generic model is efficient because one can directly apply the trained model to new subjects, without recalibration of the model parameters frequently.

The feasibility of neural decoding without spike sorting has been demonstrated mainly in intracortical recordings [56,57]. In the spike sorting approach for intracortical recordings, the human intervention, which is lengthy and repetitive, is required to manually identify the waveform snippet categories. In addition, the shape of the waveforms from neuronal sources may change over time, and the categories may not be distinguishable to each other anymore. An alternative approach is to perform adaptive thresholding to estimate the collective activities from a population of neurons, which possesses long-term stability and robustness against noise on individual neurons. The conceptual framework of our study is similar to decoding without spike sorting, which estimates the populational neuronal activity from EMG recordings with a mixture of signals from multiple sources.

Generalizability: The trained CNN model has high generalizability. The estimation of the finger forces or joint angle from the EMG signals is not straightforward, which requires a subject-dependent transformation from the neural populational firing activity. However, the relation between EMG signals and the neuron firing rate is direct. Our study has demonstrated that the EDC activation patterns learned by the CNN model from different subjects maintain inter-subject similarity, and thus the approach can be applied to a larger group of new users with minimum adjustment in the model parameters, even though the placement of the HD-EMG electrode array was not consistent across individuals or optimized for the CNN model. Moreover, the output of the CNN model was populational decoding of neuronal activities instead of force output directly. Therefore, it is possible that this approach can be applied directly to other decoding tasks such as prediction of finger joint angles, which can be derived based on the populational coding.

The CNN model trained in our study is a generic model that is viable for potential data from new subjects. The validation session not only guided the performance of discharge frequency for finger-specific predictions but also examined the expected inter-subject performance. The data used for validation were from the sixth participant, which was different from the training participants and the final testing participant. Therefore, the trained CNN model under this process was considered applicable for unknown new subjects. A previous study [58,59] showed that, via a similar MU decomposition method, the derived discharge frequency could be used to predict the finger joint angle with a second-order polynomial regression, providing high prediction accuracy. Because the presented CNN model was trained on the composite discharge trains, the algorithm could be used directly for joint kinematic prediction. On the other hand, if the CNN model was trained on the measured force, the extra layers of transformation from firing events to the force output would make it unable to generalize neither to the joint angle prediction task nor in a cross-subject manner. However, there is high heterogeneity in clinical populations, due to differences in the severity of impairment. In such circumstances, an initial fine-tuning of the model parameters may be needed for specific users. Further study is needed to investigate the model generalizability in clinical populations.

Robust Decoding: We also observed stable performance in the CNN decoder over time as shown in Figs. 6 and 7. In contrast, the EMG amplitude approach showed progressive decline of performance, mainly due to variations in EMG signals over time, and a more moderate decline was observed in the FastICA approach. With weight-sharing kernels moving from left to right and from top to bottom, the CNN worked well in detecting the distribution or shape of the muscle activation even with small displacement of the 2D electrode array during electrode placement. The Maxpool layers further improved this image translation invariance property for stable decoding performance. It was achieved by reporting the maximum output within a rectangular neighborhood in the representation map from the previous layer, which improved the robustness of the output representation relatively to small translations of the input.

Limitations and Future Directions: Although the CNN model was relatively accurate regarding single finger force prediction, the current model has not been evaluated on concurrent prediction of multi-finger

forces. Current assistive and rehabilitation robotic hands have fully articulated fingers, and it is necessary to independently drive individual digits. In addition, only isometric finger forces were decoded in the current study. Given that the decoding is performed in real-time after the model is trained, in future studies, we plan to evaluate this approach on finger force and finger joint angle prediction interchangeably using the same model. Lastly, the actual evaluation of robotic control was not performed in the current study. We plan to use the decoded output to continuously control a physical robotic hand when performing dexterous finger movement tasks in real-time.

5. Conclusion

This study presented a populational neural decoding scheme on HD-EMG signals using CNN. The decoding was performed without spike sorting and can be performed in real-time for individual finger force prediction. Two spatiotemporal related EMG energy and frequency features were extracted, and were fed to the CNN model for populational discharge frequency estimation. The CNN model demonstrated high generalizability to data from new subjects by training the CNN model on populational neuronal firing activities. The results of the force estimation evaluation demonstrated that the CNN model also showed robust performance over time compared with other decoding methods, and this can avoid frequent model recalibration during long-term use. The outcomes of our study can provide a generic and efficient CNN model for continuous and real-time neural decoding, which can lead to intuitive and robust human-machine interactions.

Declaration of competing interest

The authors have no financial relationships that may cause a conflict of interest.

References

- [1] P.K. Artemiadis, K.J. Kyriakopoulos, EMG-based Position and Force Estimates in Coupled Human-Robot Systems: towards EMG-Controlled Exoskeletons, *Springer Tracts Adv. Robot.*, 2009, pp. 241–250, https://doi.org/10.1007/978-3-642-00196-3_29.
- [2] C. Castellini, P. Artemiadis, M. Wininger, A. Ajoudani, M. Alimusaj, A. Bicchi, B. Caputo, W. Craelius, S. Dosen, K. Englehart, D. Farina, A. Gijssels, S.B. Godfrey, L. Hargrove, M. Ison, T. Kuiken, M. Marković, P.M. Pilarski, R. Rupp, E. Scheme, Proceedings of the first workshop on Peripheral Machine Interfaces: going beyond traditional surface electromyography, *Front. Neurobot.* 8 (2014) 22.
- [3] Z. Tang, H. Yu, H. Yang, L. Zhang, L. Zhang, Effect of velocity and acceleration in joint angle estimation for an EMG-Based upper-limb exoskeleton control, *Comput. Biol. Med.* 141 (2022), 105156, <https://doi.org/10.1016/j.compbmed.2021.105156>.
- [4] C.S. Klein, S. Li, X. Hu, X. Li, Editorial: electromyography (EMG) techniques for the assessment and rehabilitation of motor impairment following stroke, *Front. Neurol.* 9 (2018) 1122, <https://doi.org/10.3389/fneur.2018.01122>.
- [5] J. Stein, K. Narendran, J. McBean, K. Krebs, R. Hughes, Electromyography-controlled exoskeletal upper-limb-powered orthosis for exercise training after stroke, *Am. J. Phys. Med. Rehabil.* 86 (2007) 255–261, <https://doi.org/10.1097/PHM.0b013e3180383cc5>.
- [6] K. Englehart, B. Hudgins, A. Robust, real-time control scheme for multifunction myoelectric control, *IEEE Trans. Biomed. Eng.* 50 (2003) 848–854, <https://doi.org/10.1109/TBME.2003.813539>.
- [7] F.H.Y. Chan, Y.S. Yang, F.K. Lam, Y.T. Zhang, P.A. Parker, Fuzzy EMG classification for prosthesis control, *IEEE Trans. Rehabil. Eng.* 8 (2000) 305–311, <https://doi.org/10.1109/86.867872>.
- [8] X. Hu, A.K. Suresh, W.Z. Rymer, N.L. Suresh, Assessing altered motor unit recruitment patterns in paretic muscles of stroke survivors using surface electromyography, *J. Neural. Eng.* 12 (2015), 066001, <https://doi.org/10.1088/1741-2560/12/6/066001>.
- [9] T.A. Kuiken, G. Li, B.A. Lock, R.D. Lipschutz, L.A. Miller, K.A. Stubblefield, K. B. Englehart, Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms, *JAMA, J. Am. Med. Assoc.* 301 (2009) 619–628, <https://doi.org/10.1001/jama.2009.116>.
- [10] K. Englehart, B. Hudgins, P.A. Parker, M. Stevenson, Classification of the myoelectric signal using time-frequency based representations, *Med. Eng. Phys.* 21 (1999) 431–438, [https://doi.org/10.1016/S1350-4533\(99\)00066-1](https://doi.org/10.1016/S1350-4533(99)00066-1).
- [11] S.A. Raurale, J. McAllister, J.M. Del Rincon, Real-time embedded EMG signal analysis for wrist-hand pose identification, *IEEE Trans. Signal Process.* 68 (2020) 2713–2723, <https://doi.org/10.1109/TSP.2020.2985299>.
- [12] E.A. Clancy, N. Hogan, Probability density of the surface electromyogram and its relation to amplitude detectors, *IEEE Trans. Biomed. Eng.* 46 (1999) 730–739, <https://doi.org/10.1109/10.764949>.
- [13] A. Ameri, E.J. Scheme, E.N. Kamavuako, K.B. Englehart, P.A. Parker, Real-time, simultaneous myoelectric control using force and position-based training paradigms, *IEEE Trans. Biomed. Eng.* 61 (2014) 279–287, <https://doi.org/10.1109/TBME.2013.2281595>.
- [14] C. Dai, X. Hu, Extracting and classifying spatial muscle activation patterns in forearm flexor muscles, *Int. J. Neural Syst.* 29 (2019), 1850025, <https://doi.org/10.1142/S0129065718500259>.
- [15] I. Kyranou, S. Vijayakumar, M.S. Erden, Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses, *Front. Neurobot.* 12 (2018) 58, <https://doi.org/10.3389/fnbot.2018.00058>.
- [16] D. Špúlák, R. Čmejla, R. Bačáková, B. Kračmar, L. Satrapová, P. Novotný, Muscle activity detection in electromyograms recorded during periodic movements, *Comput. Biol. Med.* 47 (2014) 93–103, <https://doi.org/10.1016/j.compbmed.2014.01.013>.
- [17] T. Moritani, M. Muro, A. Nagata, Intramuscular and surface electromyogram changes during muscle fatigue, *J. Appl. Physiol.* 60 (1986) 1179–1185, <https://doi.org/10.1152/jappl.1986.60.4.1179>.
- [18] Y. Zheng, X. Hu, Interference removal from electromyography based on independent component analysis, *IEEE Trans. Neural Syst. Rehabil. Eng.* 27 (2019) 887–894, <https://doi.org/10.1109/TNSRE.2019.2910387>.
- [19] Y. Zheng, X. Hu, Real-time isometric finger extension force estimation based on motor unit discharge information, *J. Neural. Eng.* 16 (2019), 066006, <https://doi.org/10.1007/s10439-020-02557-2>.
- [20] R. Merletti, A. Botter, A. Troiano, E. Merlo, M.A. Minetto, Technology and instrumentation for detection and conditioning of the surface electromyographic signal: state of the art, *Clin. Biomech.* 24 (2009) 122–134.
- [21] A. Holobar, D. Zazula, Multichannel blind source separation using convolution kernel compensation, *IEEE Trans. Signal Process.* 55 (2007) 4487–4496.
- [22] R.S. LeFever, C.J. De Luca, A procedure for decomposing the myoelectric signal into its constituent action potentials—Part I: technique, theory, and implementation, *IEEE Trans. Biomed. Eng. BME-* 29 (1982) 149–157, <https://doi.org/10.1109/TBME.1982.324881>.
- [23] Y. Ning, X. Zhu, S. Zhu, Y. Zhang, Surface EMG decomposition based on K-means clustering and convolution kernel compensation, *IEEE J. Biomed. Heal. Informatics.* 19 (2015) 471–477, <https://doi.org/10.1109/JBHI.2014.2328497>.
- [24] M. Chen, P. Zhou, A novel framework based on FastICA for high density surface EMG decomposition, *IEEE Trans. Neural Syst. Rehabil. Eng.* 24 (2016) 117–127.
- [25] M.D. Twardowski, S.H. Roy, Z. Li, P. Contessa, G. De Luca, J.C. Kline, Motor unit drive: a neural interface for real-time upper limb prosthetic control, *J. Neural. Eng.* 16 (2019) 16012, <https://doi.org/10.1088/1741-2552/aaeb0f>.
- [26] V. Glaser, A. Holobar, D. Zazula, Real-time motor unit identification from high-density surface EMG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 21 (2013) 949–958, <https://doi.org/10.1109/TNSRE.2013.2247631>.
- [27] Y. Ning, N. Dias, X. Li, J. Jie, J. Li, Y. Zhang, Improve computational efficiency and estimation accuracy of multi-channel surface EMG decomposition via dimensionality reduction, *Comput. Biol. Med.* 112 (2019), 103372, <https://doi.org/10.1016/j.compbmed.2019.103372>.
- [28] Y. Zheng, X. Hu, Adaptive real-time decomposition of electromyogram during sustained muscle activation: a simulation study, *IEEE Trans. Biomed. Eng.* 69 (2022) 645–653, <https://doi.org/10.1109/TBME.2021.3102947>.
- [29] Y. Zheng, X. Hu, Concurrent prediction of finger forces based on source separation and classification of neuron discharge information, *Int. J. Neural Syst.* 31 (2021), 2150010, <https://doi.org/10.1142/S0129065721500106>.
- [30] C. Dai, Y. Cao, X. Hu, Prediction of individual finger forces based on decoded motoneuron activities, *Ann. Biomed. Eng.* 47 (2019) 1357–1368, <https://doi.org/10.1007/s10439-019-02240-1>.
- [31] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (1997) 1735–1780.
- [32] J. Schlüter, S. Böck, Improved musical onset detection with convolutional neural networks, in: ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., 2014, pp. 6979–6983, <https://doi.org/10.1109/ICASSP.2014.6854953>.
- [33] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, W. Geng, A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface, *Pattern Recogn. Lett.* 119 (2019) 131–138, <https://doi.org/10.1016/j.patrec.2017.12.005>.
- [34] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, J. Li, Gesture recognition by instantaneous surface EMG images, *Sci. Rep.* 6 (2016) 36571, <https://doi.org/10.1038/srep36571>.
- [35] A.E. Olsson, A. Björkman, C. Antfolk, Automatic discovery of resource-restricted Convolutional Neural Network topologies for myoelectric pattern recognition, *Comput. Biol. Med.* 120 (2020), 103723, <https://doi.org/10.1016/j.compbmed.2020.103723>.
- [36] K.T. Kim, C. Guan, S.W. Lee, A subject-transfer framework based on single-trial EMG analysis using convolutional neural networks, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (2020) 94–103, <https://doi.org/10.1109/TNSRE.2019.2946625>.
- [37] Q. Wang, X. Wang, Deep convolutional neural network for decoding EMG for human computer interaction, in: 2020 11th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf., 2020, pp. 554–557, <https://doi.org/10.1109/UUEMCON51285.2020.9298064>.
- [38] A. Ameri, M.A. Akhaee, E. Scheme, K. Englehart, Regression convolutional neural network for improved simultaneous EMG control, *J. Neural. Eng.* 16 (2019) 36015, <https://doi.org/10.1088/1741-2552/ab0e2e>.

- [39] E. Campbell, J.A.D. Cameron, E. Scheme, Feasibility of data-driven EMG signal generation using a deep generative model, in: 2020 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE, 2020, pp. 3755–3758, <https://doi.org/10.1109/EMBC44109.2020.9176072>.
- [40] R.D. Flint, M.C. Tate, K. Li, J.W. Templer, J.M. Rosenow, C. Pandarinath, M. W. Slutzky, The representation of finger movement and force in human motor and premotor cortices, *ENeuro* 7 (2020) 1–15, <https://doi.org/10.1523/ENEURO.0063-20.2020>.
- [41] W.S. Yu, H. Van Duinen, S.C. Gandevia, Limits to the control of the human thumb and fingers in flexion and extension, *J. Neurophysiol.* 103 (2010) 278–289, <https://doi.org/10.1152/jn.00797.2009>.
- [42] C. Dai, Y. Zheng, X. Hu, Estimation of muscle force based on neural drive in a hemispheric stroke survivor, *Front. Neurol.* 9 (2018) 187, <https://doi.org/10.3389/fneur.2018.00187>.
- [43] X. Hu, N.L. Suresh, C. Xue, W.Z. Rymer, Extracting extensor digitorum communis activation patterns using high-density surface electromyography, *Front. Physiol.* 6 (2015) 279, <https://doi.org/10.3389/fphys.2015.00279>.
- [44] Y. Zheng, X. Hu, Real-time isometric finger extension force estimation based on motor unit discharge information, *J. Neural. Eng.* 16 (2019), 066006, <https://doi.org/10.1088/1741-2552/ab2c55>.
- [45] C. Dai, X. Hu, Independent component analysis based algorithms for high-density electromyogram decomposition: systematic evaluation through simulation, *Comput. Biol. Med.* 109 (2019) 171–181, <https://doi.org/10.1016/j.combiomed.2019.04.033>.
- [46] F. Negro, S. Muceli, A.M. Castronovo, A. Holobar, D. Farina, Multi-channel intramuscular and surface EMG decomposition by convolutive blind source separation, *J. Neural. Eng.* 13 (2016), 26027, <https://doi.org/10.1088/1741-2560/13/2/026027>.
- [47] P.J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *J. Comput. Appl. Math.* 20 (1987) 53–65, [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [48] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* 15 (2014) 1929–1958.
- [49] D.P. Kingma, J.L. Ba, Adam: a method for stochastic optimization, in: 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., 2015.
- [50] L.J. Bain, *Applied Regression Analysis*, John Wiley & Sons, 1967, <https://doi.org/10.1080/00401706.1967.10490452>.
- [51] Y. Zheng, X. Hu, Concurrent estimation of finger flexion and extension forces using motoneuron discharge information, *IEEE Trans. Biomed. Eng.* 68 (2021) 1638–1645, <https://doi.org/10.1109/TBME.2021.3056930>.
- [52] Y. Zheng, X. Hu, Dexterous force estimation during finger flexion and extension using motor unit discharge information, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS. 2020-July, 2020, pp. 3130–3133, <https://doi.org/10.1109/EMBC44109.2020.9175236>.
- [53] D.I. Warton, F.K.C. Hui, The arcsine is asinine: the analysis of proportions in ecology, *Ecology* 92 (2011) 3–10, <https://doi.org/10.1890/10-0340.1>.
- [54] J.E. Downey, L. Brane, R.A. Gaunt, E.C. Tyler-Kabara, M.L. Boninger, J. L. Collinger, Motor cortical activity changes during neuroprosthetic-controlled object interaction, *Sci. Rep.* 7 (2017) 1–10, <https://doi.org/10.1038/s41598-017-17222-3>.
- [55] C. Dai, X. Hu, Independent component analysis based algorithms for high-density electromyogram decomposition: experimental evaluation of upper extremity muscles, *Comput. Biol. Med.* 108 (2019) 42–48, <https://doi.org/10.1016/j.combiomed.2019.03.009>.
- [56] E.M. Trautmann, S.D. Stavisky, S. Lahiri, K.C. Ames, M.T. Kaufman, D.J. O'Shea, S. Vyas, X. Sun, S.I. Ryu, S. Ganguli, K. V. Shenoy, Accurate estimation of neural population dynamics without spike sorting, *Neuron* 103 (2019) 292–308, <https://doi.org/10.1016/j.neuron.2019.05.003>, e4.
- [57] B.P. Christie, D.M. Tat, Z.T. Irwin, V. Gilja, P. Nuyujukian, J.D. Foster, S.I. Ryu, K. V. Shenoy, D.E. Thompson, C.A. Chestek, Comparison of spike sorting and thresholding of voltage waveforms for intracortical brain-machine interface performance, *J. Neural. Eng.* 12 (2015) 16009, <https://doi.org/10.1088/1741-2560/12/1/016009>.
- [58] C. Dai, X. Hu, Finger joint angle estimation based on motoneuron discharge activities, *IEEE J. Biomed. Heal. Informatics.* 24 (2020) 760–767, <https://doi.org/10.1109/JBHI.2019.2926307>.
- [59] C. Dai, Y. Cao, X. Hu, Estimation of finger joint angle based on neural drive extracted from high-density electromyography, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS. 2018-July, 2018, pp. 4820–4823, <https://doi.org/10.1109/EMBC.2018.8513152>.