

# Data Augmentation and Deep Learning Techniques to Improve Interface Noise Tolerance of Myoelectric Pattern Recognition Controllers

Yuni Teh (✉ [yuni@u.northwestern.edu](mailto:yuni@u.northwestern.edu))

Northwestern University

Levi J. Hargrove

Shirley Ryan AbilityLab

---

## Research Article

**Keywords:** Myoelectric control, pattern recognition, prosthetics, data augmentation, deep learning, convolutional neural networks

**Posted Date:** February 4th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1318759/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Clinically available myoelectric PR controllers deteriorate under conditions that generate interface noise, such as electrode liftoff or wire failure. Previous solutions relied on additional processing steps like signal denoising and controller adaptation to mitigate these negative effects. However, there are no clinically practical controllers that are inherently robust to interface noise. This paper investigated the use of data-driven methods to build clinically practical, noise tolerant PR control strategies for transradial prostheses.

**Methods:** We developed a data augmentation protocol to increase training data variability and two deep neural networks (MLP and CNN) to compress redundancies across a six-channel EMG array to a latent space that is robust to noise. Using data from fourteen intact limb subjects and seven below-elbow amputee subjects, we evaluated five control strategies: an LDA classifier (LDA), an existing adaptive LDA classifier that recalibrates after discarding noisy EMG channels (LDA-), an LDA classifier trained with an augmented data set (LDA+), an LDA classifier trained with MLP-aligned latent variables of an augmented data set (MLP-LDA), and an LDA classifier trained with CNN-aligned latent variables of an augmented data set (CNN-LDA). We computed the classification accuracies on clean signals and signals that had interface noise in up to four channels.

**Results:** Compared to the baseline LDA method, LDA+ and MLP-LDA significantly increased classification accuracy on noisy data but decreased accuracy on clean data. On the other hand, LDA- and CNN-LDA improved noise tolerance while maintaining high performance with clean data.

**Conclusion:** The CNN-LDA control strategy is robust to multichannel interface noise and can improve the reliability and usability of PR-based upper limb prostheses.

## Background

Major upper limb amputations cause impairments that inhibit basic activities of daily living such as eating and dressing [1]. Myoelectric prostheses have the potential to restore some lost functionality, thereby helping users to regain their independence and improve their quality of life. Commercial devices employ two main types of control systems: amplitude-based direct control and pattern recognition (PR) control.

Direct control uses electromyographic (EMG) signals measured from agonist-antagonist muscle pairs to control prosthesis movements [2]. Each muscle pair activates one degree-of-freedom (DOF); thus, the dexterity of a direct controlled prosthesis is limited by the number of distinct muscle sites that can be individually contracted. Although co-contractions can be used to toggle between DOFs, this renders the control method less intuitive. Furthermore, while direct control is reliable, it requires manual tuning to accommodate the unique muscle anatomies of each user.

PR control methods measure EMG signals from an array of electrodes and learn the patterns of muscle activity that correspond to intended movements [3–5]. Typically, PR controllers use a classifier to classify descriptive features extracted from windowed EMG signals. Use of linear discriminant analysis (LDA) classifiers is common as they are computationally simple to train and implement. PR control has been shown to improve functional outcomes for prosthesis users, provided the EMG interface is stable and the signals are repeatable [6].

However, regular usage of myoelectric prosthesis gives rise to various sources of signal disturbances that degrade classification accuracy [7–12]. Changes in residual limb volume, limb position, and socket loading can cause electrodes to intermittently lose contact with the skin. Signal abnormalities stemming from electrode or wire failures also occur with prolonged prosthesis use [8,13,14]. Interface noise in just one channel is often detrimental to the accuracy of PR control strategies, rendering the device unusable until it can be recalibrated [8,9].

Previous studies have proposed several approaches to resolve the effects of interface noise. For example, signal processing algorithms can be used to denoise affected channels before classification [7,15–21]. The generalizability of these methods is limited, however, as most focus on filtering out periodic noise (eg. electrical noise) and do not address the effects of intermittent noise signals [22]. Another approach uses control strategies that adjust their classifier parameters to adapt to changes in EMG signals [23–25]. Of note is a fast-retraining LDA method that detects and removes noisy channels before recalibrating its weights [9]. Though they can increase classification accuracy, these adaptive control methods require additional processing steps during classification.

One promising solution exploits the signal redundancies across EMG channels to build a classifier that is inherently robust to interface noise. Since surface electrodes measure diffuse muscle activity, an array of surface EMG signals contains overlapping neural information. These redundant signals can therefore be compressed into a low-dimensional manifold that retains discriminative features and is less sensitive to input disturbances [26–29]. Using this concept, prior works have employed spatial filtering, linear factorization, and data fusion to build robust classifiers [30–32]. However, these examples used setups that are not yet clinically practical (eg. high-definition EMG arrays).

To train an accurate PR controller, it is essential that the training data typifies signals in real scenarios. For example, performing dynamic arm movements during training data collection instead of maintaining a static position significantly improves LDA classification performance across different limb positions [33,34]. By increasing training data variability, classifiers are encouraged learn discriminative features that are consistent across sources of variance, thus preventing overfitting. However, it is a challenge to physically collect enough data to sufficiently represent realistic scenarios. To alleviate the burden of extensive training data collection, data augmentation can be used to artificially introduce variability in a systematic manner [35–37].

As the variance of the training data increases and the data become more complex, simple linear classifiers may not be equipped to adequately model hidden structures. Deep learning models are known

for their ability to learn complex nonlinear relationships within large datasets. Specifically, deep encoders compress high-dimensional inputs into low-dimensional latent subspaces [38]. A desirable characteristic of these manifolds is that large disturbances in the input have minimal effects on their latent representations. Depending on the model's loss function, the latent space can be optimized for a specific objective, such as reconstructing or classifying the input.

Convolutional neural networks (CNNs) are another useful deep learning tool [39,40]. CNNs are commonly used for image processing applications where the relative locations of pixels are crucial to the underlying structure. While sequential neural networks use flattened vectors as inputs, CNNs preserve spatial relationships by allowing multidimensional input matrices. Thus, CNNs are well-equipped to disentangle EMG signal redundancies, which are dependent on electrode locations [37,41–44].

The objective of this study was to develop a clinically feasible, noise-tolerant myoelectric PR controller that classifies hand and wrist movements. To that end, we explored the use of data augmentation and deep learning techniques to uncover a latent subspace in which movement classes are separable and robust to interface noise. We trained five control strategies and evaluated their performances on normal EMG data and EMG data that contained up to four channels of interface noise. These strategies included a conventional LDA algorithm, an adaptive LDA algorithm, an LDA algorithm trained with an augmented data set, an LDA algorithm that classifies latent EMG variables computed by a multilayer perceptron, and an LDA algorithm that classifies latent EMG variables computed by a CNN. We hypothesized that the CNN control strategy would be the most accurate non-adaptive classifier across all noise conditions because it preserves spatial dependencies and contains nonlinearities.

## Methods

The following experiment was approved by the Northwestern Institutional Review Board. Fourteen individuals with intact limbs (ITL) and seven individuals with below-elbow amputations (AMP, Table 1) participated in this study after providing written informed consent. Due to partial data loss, results from one ITL participant and one AMP participant were excluded from the final analysis.

**Table 1**

### **Amputee Subject Demographics**

Subject	Age	Gender	Time Since Amputation	Level Of Amputation	DOFs Controlled
AMP1	73	M	32 years	Transradial	3DOF
AMP2	33	M	5 years	Wrist disarticulation	3DOF
AMP3	65	M	6 years	Transradial	2DOF
AMP4	56	M	40 years	Transradial	2DOF
AMP5	48	M	11 months	Transradial	2DOF
AMP6	19	M	10 months	Transradial	2DOF

## Experimental Setup

For ITL participants (Fig. 1a), six channels of EMG signals were collected using dry stainless-steel bipolar electrodes (Motion Control Inc.) that were embedded in an adjustable armband. The electrodes were equally spaced around the subject’s right arm, with the reference electrode positioned just distal to the olecranon. An HTC Vive tracker was attached to the dorsal side of the armband and used to track the participant’s limb position. Participants also wore an orthosis around the wrist and hand to promote isometric contractions that would more closely resemble amputee contractions. Finally, a 400g weight was attached to the distal end of the orthosis to simulate the weight of a prosthesis.

Due to the unique size and anatomy of each residual limb, dry electrode setups that are not specifically customized for an amputee are prone to electrode liftoff. Hence, we used wet electrodes for AMP participants to prevent unwanted interface noise (Fig. 1b). Six channels of EMG signals were collected using adhesive Ag/AgCl bipolar electrodes (Bio-Medical Instruments) that were secured under a silicone liner. The electrodes were equally spaced around the subject’s residual limb and the reference electrode was placed just distal to the olecranon. An adjustable lightweight frame was fastened around the residual limb and lengthened to match the subject’s intact limb length. A 400g weight was attached to the distal end of the frame to simulate the weight of a prosthesis.

## Data Collection Protocol

All data collection was conducted in an HTC Vive virtual reality environment (Fig. 1c). Each participant collected a training data set and a test data set during one experimental session. EMG signals were sampled at a rate of 1 kHz, band-pass filtered between 70-300 Hz, and segmented into 200 ms windows in 25 ms increments. In addition to a hardware gain of 2 and a software gain of 1000, there were channel-specific software gains that were customized for each subject. These channel gains were calculated by scaling the signals in each channel to span the output range of -5V to 5V. Although the channel gains were calculated using the training data set alone, they were applied to the training and test data sets.

To collect training data, the subject performed hand and wrist gestures while moving their arm around the workspace. This simple training protocol has been shown to achieve high real-time performance [33,34]. All ITL subjects and two AMP subjects completed seven gestures (rest, wrist flexion/extension, wrist pronation/supination, hand open/close), corresponding to a 3DOF controller. Based on clinician input and to minimize fatigue, the remaining four AMP participants completed five gestures (rest, wrist pronation/supination, hand open/close), corresponding to a 2DOF controller. Each gesture was held for 2.5 seconds and repeated five times, resulting in 12.5 seconds (500 overlapping windows) of clean training examples per gesture.

To collect test data, the subject performed the trained hand and wrist gestures in four limb positions (Fig. 2a). Each gesture was held for 2.5 seconds and repeated five times. Therefore, each participant had 50 seconds (2000 overlapping windows) of clean test data for each gesture.

## Offline Analyses

After EMG data collection, all further analyses were conducted offline on a Windows 10 laptop computer with 16GB RAM, an Intel Core i7-9850H CPU at 2.60GHz, and a 4GB NVIDIA Quadro T1000 GPU. These analyses included training data augmentation, training five control strategies, testing those strategies, and statistical evaluations.

## Training Data Augmentation

We constructed an augmented training data set by systematically corrupting up to four channels in copies of the original raw training signals. The augmented data contained one clean copy of the original data and eight noisy copies. Within the eight noisy copies, there were two copies of data that had one noisy EMG channel, two copies that had two noisy channels, two copies that had three noisy channels, and two copies that had four noisy channels. We evenly distributed 12 types of synthetic noise into all possible channel combinations. These synthetic noise types included flatlining, in which the signal was completely attenuated to 0V, five levels of Gaussian noise centered at 0V ( $\sigma = 1,2,3,4,5V$ ), five levels of 60 Hz noise (amplitude = 1,2,3,4,5V), and a randomized mixture of all noise types.

## Control Strategies

Five control strategies were trained and evaluated in this study. Before we trained the controllers, four time-domain features were extracted from the training data sets: mean absolute value, waveform length, zero crossings, and slope sign changes.

## *Traditional LDA Classifiers*

Three control methods were based on the traditional LDA classifier algorithm.

1. Baseline LDA (LDA) – To act as the baseline model, an LDA classifier was trained with the original training data set. This algorithm is used in most clinically available PR systems and therefore demonstrates what current prosthesis users experience.
2. Augmented LDA (LDA+) – To investigate how data augmentation affects the reliability of a standard LDA algorithm, we trained an LDA classifier with the augmented training data set.
3. Adaptive LDA (LDA-) – We implemented an existing fast-retraining LDA classifier that circumvented signal disturbances by adjusting its LDA weights after removing noisy EMG channels [9]. First, we trained an LDA classifier with the original training data set and stored the class mean and covariance matrices. During classification, we omitted the elements corresponding to noisy EMG channels from the mean and covariance matrices and recalculated the LDA weights. Then, we removed the noisy EMG channels from the classifier inputs and used the new LDA weights to classify the remaining signals. In practice, this control strategy requires a fault detector to detect noisy signals. However, we excluded this step and instead assumed that a perfect fault detector was used. Thus, the LDA- classifier shows the best-case scenario for an adaptive LDA control system. When there were no noisy channels, the LDA- classifier was identical to the baseline LDA classifier.

## ***Neural Network-Aligned Classifiers***

The remaining two control strategies comprised two stages: a latent encoder network that aligned the EMG inputs to a low-dimensional manifold and an LDA classifier that classified these latent variables (Fig. 2d). One control strategy used a multilayer perceptron network (MLP-LDA) while the other used a convolutional neural network (CNN-LDA). Both models were implemented using Keras 2.3.1 with the Tensorflow backend.

We used five-fold cross-validation with the augmented training data set to tune each model's hyperparameters. To avoid overlapping training and validation data, each fold corresponded to one gesture repetition. After the hyperparameters were determined, we trained the final models with the entire augmented training data set.

Both networks were trained using the Adam optimization algorithm with a learning rate of 0.001 and mini-batch gradient descent with 30 training epochs and a batch size of 128. To accelerate training time, we used a minmax scaler to standardize the input features between [0,1] and applied batch normalization after each hidden layer.

4. Multilayer perceptron-aligned LDA (MLP-LDA) - We trained a fully connected five-layer neural network (Fig. 2b) to take in a 24 by 1 EMG feature vector and output a 4 by 1 latent feature vector  $z$  and a predicted gesture label  $y$ .

The first four hidden layers aligned the EMG inputs to the latent space. We applied ReLU activation functions after the first three layers and a linear activation function after the fourth layer. Based on our cross-validation results, we found that classification accuracy improved as the dimensionality of the latent space increased but began to plateau after a dimensionality of 4. Thus, we set the latent dimension to 4. We also regularized the weights of the fourth layer with L1 regularization ( $\lambda = 10e-5$ ) to encourage sparsity and improve generalization.

The last hidden layer in the MLP was a linear classifier that used a softmax activation function to classify the latent features. The network was trained to minimize the categorical cross entropy loss between the predicted class and the ground truth, thus optimizing linear separability between movement classes in the latent space.

Since neural network classifiers are prone to overfitting, we trained an LDA classifier with the latent features of the augmented data set and used it in tandem with the MLP network to form the MLP-LDA control strategy (Fig. 2d). During classification, the EMG input vectors were passed through the MLP encoder to compute the latent features  $z$ , which were then fed to the LDA classifier to obtain gesture predictions.

In total, the MLP had 1267 trainable parameters.

5. Convolutional neural network-aligned LDA (CNN-LDA) - We trained a CNN (Fig. 2c) with the same objectives as the MLP: to output a 4 by 1 EMG latent feature vector  $z$  and a predicted gesture label  $y$ .

While the inputs for the previous control strategies were 24 by 1 feature vectors, the CNN input was a 6 by 4 feature matrix, corresponding to the 6 EMG channels and 4 time-domain features. This enabled the 2-dimensional convolutional layers to exploit the spatial relationships between EMG channels and learn more robust latent representations.

The first five hidden layers served as the encoder, starting with two convolutional layers with ReLU activation functions. Then, we flattened the output of the convolutional layers before passing it through two sequential layers with a ReLU and a linear activation function, respectively. Thus, the latent encoder modules of the CNN and MLP each had three ReLU functions and one linear function. Like the MLP, the layer preceding the bottleneck was regularized with L1 regularization ( $\lambda = 10e-5$ ) and the latent space had a dimensionality of 4.

The last layer of the CNN classified the latent feature vector  $z$  using a softmax activation function. The CNN was trained to minimize the categorical cross entropy loss between the predicted class and the ground truth, once again to encourage linear separability between the class latent representations.

Finally, we trained an LDA classifier with the augmented training data set after it was aligned by the CNN. The CNN-LDA control strategy (Fig. 2d) used the CNN encoder to compute latent variables  $z$  which were then classified with the LDA classifier.

In total, the CNN had 12999 trainable parameters.

## Evaluation

To evaluate control performance and robustness, we calculated the offline classification accuracies of each control strategy on clean and noisy test data. Since it was impractical and challenging to introduce interface noise in a controlled manner during data collection, we constructed noisy test data offline by fusing the original test raw signals with examples from a real noise database.

### *Real Noise Database*

The effects of four noise types were investigated in this study: broken wires, moving broken wires, contact artifacts, and loose electrodes. A database containing 25 seconds (1000 overlapping windows) of each type was collected from one ITL subject (Fig. 3). Since the housing of the armband prevented access to individual electrodes, this database was recorded using the wet electrode setup. Although all six channels were recorded, only the affected noisy channel was stored in the database.

To simulate the broken wire and moving broken wire conditions, one wire was cut at the connection point between the wire and the electrode. For the broken wire condition, the subject maintained a 90-degree angle at the elbow throughout data collection. For the moving broken wire condition, the subject moved their arm around freely in a workspace that contained sources of electrical noise, such as monitors and laptops. For the contact artifact condition, the electrode was tapped approximately every 200 ms. Finally, for the loose electrode condition, the electrode was peeled off and gently shifted around the surface of the subject's skin throughout signal recording.

### *Fusion of Test Signals and Real Noise*

We constructed four noisy test sets, each containing a distinct number of noisy EMG channels (1 to 4 channels). Each noisy set started as a copy of the clean test raw signals. We then systematically superimposed pseudorandomized samples from the real noise database onto the copy, ensuring that all combinations of affected channels and noise types were equally represented. To maintain signal amplification consistency, the subject-specific channel gains were applied to the noise windows according to the channels with which they were being fused. Signals were then truncated to stay within the output range of  $[-5V, 5V]$ . Finally, we extracted the four time-domain features from the noisy test signals.

## Statistical Analyses

The statistical analyses were conducted separately for ITL and AMP populations. We used linear mixed effects models to evaluate the statistical effects of each control algorithm with respect to the baseline LDA method. Initially, we fit a model with classification accuracy as the response variable, the control strategy (LDA, LDA+, LDA-, MLP-LDA, and CNN-LDA), number of noisy electrodes (0-4), and their interactions as fixed factors, and the subject identifier as a random factor. Statistical significance was judged based on a significance level of  $\alpha = 0.05$ . After observing that all interaction factors were statistically significant ( $p < 0.001$ ), the data were separated by the number of noisy electrodes. These data sets were used to fit five new models that each had the control strategy as a fixed factor and subject identifier as a random factor. We used the Bonferroni method to correct for multiple comparisons.

## Results

The classification accuracies of the five control strategies and summary of the statistical models are shown in Fig. 4.

### Interface noise degrades LDA accuracy

Under standard conditions, the baseline LDA classifier decoded gestures with average accuracies of  $79.92 \pm 1.14\%$  and  $78.10 \pm 1.66\%$  for ITL and AMP participants respectively. When noise was present in just one channel, the accuracies dropped to  $61.55 \pm 0.98\%$  and  $41.74 \pm 1.62\%$ , demonstrating that a minor change in input signals can have a large impact on control performance, particularly for AMP subjects. As the number of corrupted channels increased, the accuracy continued to decrease.

### Data augmentation may increase noise tolerance, but at a cost

For the AMP population, augmenting the training data set with synthetic noise increased the robustness of an LDA classifier. Compared to the 36.36% drop in baseline LDA accuracy between the noiseless and single channel conditions, the LDA- accuracy decreased by only 4.19%. Consequently, the LDA- control strategy significantly outperformed the baseline LDA control strategy for all noisy conditions ( $p < 0.001$ ). However, with an accuracy of  $59.12 \pm 1.79\%$ , the LDA- algorithm was also significantly worse at classifying clean signals compared to the baseline LDA algorithm ( $p < 0.001$ ).

The trends for the LDA- classifier were different in ITL subjects. Its outcomes were significantly worse compared to those of the baseline LDA classifier for the noiseless and single channel noise conditions ( $p < 0.001$ ), but not significantly different for the remaining noisy conditions.

### Neural network models outperform non-adaptive LDA algorithms

Generally, the neural network-aligned methods improved overall outcomes compared to all non-adaptive LDA methods. Across all test sets, MLP-LDA had average accuracies of 64.84% (ITL) and 58.88% (AMP), which were 18.72% and 7.73% higher than LDA+ accuracies. MLP-LDA also improved classification of noisy signals by 15.77% (ITL) and 25.22% (AMP) compared to the baseline LDA classifier. However, there were statistically significant drops in accuracy on clean signals (ITL: 6.13%, AMP: 6.18%). Thus, at best, MLP-LDA had a  $73.78 \pm 0.99\%$  accuracy for ITL subjects and a  $71.92 \pm 1.72\%$  accuracy for AMP subjects.

In contrast, the CNN-LDA control strategy significantly improved classification of noisy EMG signals ( $p < 0.001$ ) without decreasing accuracy on clean EMG signals ( $p = 1.000$ ). CNN-LDA classified normal signals with an accuracy of  $80.25 \pm 1.21\%$  (ITL) and  $78.91 \pm 1.89\%$  (AMP), exhibiting the best performances across all five control strategies. Unsurprisingly, these accuracies decreased as noise was introduced into the system. However, CNN-LDA scored  $65.52 \pm 0.90\%$  (ITL) and  $53.49 \pm 1.71\%$  (AMP) with four noisy channels, meaning that at its worst, it still performed better than baseline LDA did with only one noisy channel. Therefore, CNN-LDA was the most accurate and robust non-adaptive control strategy.

## CNN-LDA eliminates need for controller adaptation

Overall, CNN-LDA and LDA- achieved similar performances. For AMP participants, the decrease in accuracy from LDA- to CNN-LDA for noisy data ranged from 1.61% for single-channel noise to 7.20% for four-channel noise. For ITL subjects, CNN-LDA accuracies surpassed LDA-, with improvements ranging from 0.79% for single-channel noise and 6.87% for four channel noise. Although we did not statistically compare these differences, it is unlikely that they would cause significant clinical impact. Thus, CNN-LDA was functionally equivalent to an adaptive LDA control system with a perfect fault detector.

## LDA components illustrate performance differences

To visualize the underlying mechanisms of the non-adaptive classifiers, we plotted the first three LDA components of the clean training data, clean test data, and noisy test data for subject AMP5 (Fig. 5). For baseline LDA and LDA+, we used the projection matrices from the trained LDA classifiers to reduce the high-dimensional input features. For MLP-LDA and CNN-LDA, we aligned the input feature data to their latent manifolds through the MLP and CNN encoders before applying the projection matrices from their corresponding LDA classifiers.

CNN-LDA was the only control strategy that maintained separable clusters across all three data sets, explaining why it was able to effectively classify both clean and noisy signals. In contrast, LDA and MLP-LDA clusters lost their separability when noise was introduced, thus depicting their lower noise tolerance. The LDA+ clean test set clusters did not match the clean training set clusters; therefore, the decision boundaries computed from the training set were not able to delineate the gestures in clean test data set.

# Neural network methods require longer processing times

In Table 2, we show the processing times of each control method to assess their practicality in clinical settings. Notably, the initial training processing times for the neural network methods, which included data augmentation time, were substantially longer than those for the traditional LDA methods. Likewise, the classification processing times were longer for the neural network methods. However, at ~5 ms, this was still shorter than the EMG window sampling increments (25 ms), indicating that a real-time implementation is plausible.

**Table 2**

**Controller Processing Times**

	<b>Training Processing Time</b>	<b>Classification Processing Time</b>
<b>LDA</b>	0.0259 s $\pm$ 1.37 ms	0.00659 ms $\pm$ 64.4 ns
<b>LDA+</b>	1.22 s $\pm$ 48.6 ms	0.00665 ms $\pm$ 68.8 ns
<b>LDA-</b>	0.0259 s $\pm$ 1.37 ms	0.328 ms $\pm$ 4.02 us
<b>MLP-LDA</b>	51.3 s $\pm$ 2.82 s	5.54 ms $\pm$ 38.1 us
<b>CNN-LDA</b>	47.8 s $\pm$ 3.12 s	5.75 ms $\pm$ 38.7 us

## Discussion

Clinical PR-based myoelectric control systems often encounter error-producing signal noise stemming from EMG interface instabilities. Previous works have reduced these errors with additional processing steps such as signal filtering, noise detection, and controller recalibration or adaptation. To our knowledge, this is the first study that presents a PR control strategy that is inherently resistant to various types of multichannel interface noise. Using examples of real interface noise, we evaluated the reliability of LDA-based classifiers that employed data augmentation and deep learning techniques.

In general, we found that the adverse effects of interface noise were more severe in the amputee population than the intact limb population. One potential reason is that intact limb muscle contractions often have larger amplitudes compared to residual limb contractions. Therefore, noisy intact limb EMG signals have higher signal-to-noise ratios (SNR), leading to better classification rates.

The high SNR in intact limb participants may also explain why augmenting the training data set with artificial noise did not improve overall outcomes for the augmented LDA classifier. The augmented data set contained various levels of signal corruption that spanned the output range of the EMG channels. Consequently, it had a wide range of SNRs as opposed to the relatively high SNRs of the noisy test set. Thus, the augmented LDA classifier was trained on distributions that were not representative of the testing data, causing it to perform poorly.

Our results also showed that manifold alignment facilitated by neural networks produced impactful performance gains. This suggests that nonlinear transformations are crucial to extracting useful discriminative structures within EMG signals. Furthermore, the CNN outperformed the multilayer perceptron, emphasizing that the spatial relationships between channels are valuable and should be preserved and leveraged for better control. For the intact limb population, the CNN-LDA method was also more accurate than the adaptive LDA method. This exemplified an important advantage of the CNN-LDA method: it retains discriminative features in the noisy channels that would normally be discarded by recalibration algorithms. Notably, CNN-LDA was able to classify signals with real interface noise even though the augmented training data set only contained synthetic noise, highlighting the generalizability of the deep learning model.

Ultimately, the CNN-LDA control strategy is an attractive solution to the interface noise problem, as it was the only non-adaptive method that improved noise tolerance without reducing accuracy on normal signals. In reality, disturbances would most likely occur in one or two EMG channels. For these cases, CNN-LDA obtained classification accuracies (ITL: >75.19%, AMP: >68.83%) that would suggest effective real-time control.

The clinical implications of the CNN-LDA controller are that patients would be able to maintain good control of their prosthesis across common scenarios that produce intermittent noise (eg. electrode liftoff, contact artifacts) or continuous noise (eg. broken wires). Moreover, it does not require a noise detector, which can be inaccurate, or recalibration, which adds a processing step. This method is highly beneficial for amputee users, as their low EMG SNR renders their clinical LDA systems unusable in the presence of noise.

However, there are some practical limitations to the CNN-LDA controller. For example, its training and execution processing time is slower than traditional LDA methods. Also, the memory requirements of the model may hinder its implementation on a prosthesis microcontroller. Lastly, it is difficult to quickly retrain a black box model such as the CNN-LDA. If the user wanted to recalibrate a single gesture, the entire backpropagation process would have to be repeated.

## Limitations and Future Work

The main limitation of this study is that the controllers were evaluated offline. Real-time experiments should be conducted to investigate the effects of user adaptation and provide a more realistic evaluation of prosthesis controllability. Additionally, while commercial systems typically include eight EMG channels, the experimental setup only used six EMG channels. The implication is that each noisy signal has a greater influence on classification accuracy. Thus, we expect the performance to be better with clinical setups that have eight EMG channels.

Four amputee subjects only completed enough gestures for a 2DOF controller instead of 3DOF controller. These algorithms were also limited to sequential control but would be more impactful if extended to

simultaneous control. To facilitate practical implementation of the CNN-LDA, we plan to further minimize the network architecture and adjust hyperparameters to balance controller performance and processing time. In addition to conducting real-time experiments with a physical prosthesis, it would be useful to investigate the controller's robustness to donning/doffing and its long-term stability.

## **Conclusion**

In this study, we investigated the use of training data augmentation and deep learning models to achieve reliable myoelectric control of hand and wrist gestures. Our results showed that a CNN-based control strategy can learn latent representations of muscle patterns that are linearly separable and robust to multichannel interface noise. The performance gains exhibited by the CNN-LDA are a significant improvement on clinical myoelectric PR controllers and motivate further development.

## **Abbreviations**

PR: Pattern recognition, EMG: Electromyography, DOF: Degrees of freedom, ITL: Intact limb participants, AMP: Amputee participants, LDA: Linear discriminant analysis, MLP: Multilayer perceptron, CNN: convolutional neural network, SNR: Signal-to-noise ratio

## **Declarations**

## **Ethics approval and consent to participate**

This study was approved by the Northwestern Institutional Review Board.

## **Consent for publication**

All participants gave written informed consent.

## **Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## **Competing interests**

The authors declare that they have no competing interests.

## **Funding**

This work was supported by the National Institutes of Health NIH R01HD094861.

## Authors' contributions

YT and LJH conceived this work. YT designed the protocols, developed the experimental devices and software, performed data collection and analyses, and prepared the manuscript. LJH supervised and provided critical feedback throughout all steps.

## Acknowledgements

We would like to acknowledge Kristi Turner and Laura Miller for their assistance during data collection.

## References

1. Cordella F, Ciancio AL, Sacchetti R, Davalli A, Cutti AG, Guglielmelli E, et al. Literature review on needs of upper limb prosthesis users. *Front Neurosci.* 2016;10(MAY):1–14.
2. Williams TW, Meier RH, Atkins D. Control of powered upper extremity prostheses. *Funct Restor adults Child with Up Extrem amputation.* 2004;207–24.
3. Hudgins B, Parker P, Scott RN. A new strategy for multifunction myoelectric control. *IEEE Trans Biomed Eng.* 1993;40(1):5541–8.
4. Englehart K, Hudgins B. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans Biomed Eng.* 2003;50(7):848–54.
5. Parker P, Englehart K, Hudgins B. Myoelectric signal processing for control of powered limb prostheses. *J Electromyogr Kinesiol.* 2006;16(6):541–8.
6. Hargrove LJ, Lock BA, Simon AM. Pattern recognition control outperforms conventional myoelectric control in upper limb patients with targeted muscle reinnervation. In: *35th Annu Int Conf IEEE Eng Med Biol Soc.* 2013. p. 1599–602.
7. De Luca CJ, Gilmore LD, Kuznetsov M, Roy SH. Filtering the surface EMG signal movement artifact and baseline noise contamination. *J Biomech.* 2010;1573–9.
8. Simon AM, Lock BA, Stubblefield KA. Patient training for functional use of pattern recognition-controlled prostheses. *J Prosthetics Orthot.* 2012;24(2):56–64.
9. Zhang X, Huang H. A real-time, practical sensor fault-tolerant module for robust EMG pattern recognition. *J Neuroeng Rehabil.* 2015;12(1):1–16.
10. Kyranou I, Vijayakumar S, Erden MS. Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses. *Front Neurorobot.* 2018;12(September):1–22.
11. Scheme EJ, Englehart K. Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. *J Rehabil Res Dev.* 2011;48(6):643–60.

12. Chowdhury RH, Reaz MBI, Bin Mohd Ali MA, Bakar AAA, Chellappan K, Chang TG. Surface electromyography signal processing and classification techniques. *Sensors (Switzerland)*. 2013;13(9):12431–66.
13. Hargrove L, Miller L, Turner K, Kuiken T. Control within a virtual environment is correlated to functional outcomes when using a physical prosthesis. *J Neuroeng Rehabil*. 2018;15(Suppl 1).
14. Miller LA, Turner K, Simon AM. Data logging during pattern recognition calibration as a remote diagnostic tool. In: *Myoelectric Control Symp*. 2020.
15. Ortolan RL, Mori RN, Pereira RR, Cabral CMN, Pereira JC, Cliquet A. Evaluation of adaptive/nonadaptive filtering and wavelet transform techniques for noise reduction in EMG mobile acquisition equipment. *IEEE Trans Neural Syst Rehabil Eng*. 2003;11(1):60–9.
16. Reaz MBI, Hussain MS, Mohd-Yasin F. Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biol Proced Online*. 2006;8(1):11–35.
17. Phinyomark A, Limsakul C, Phukpattaranont P. A comparative study of wavelet denoising for multifunction myoelectric control. In: *Int Conf Comput Autom Eng*. IEEE; 2009. p. 21–5.
18. Fraser GD, Chan ADC, Green JR, Abser N, Maclsaac D. CleanEMG - Power line interference estimation in sEMG using an adaptive least squares algorithm. In: *33rd Annu Int Conf IEEE Eng Med Biol Soc*. IEEE; 2011. p. 7941–4.
19. Rehbaum H, Farina D. Adaptive common average filtering for myocontrol applications. *Med Biol Eng Comput*. 2015;53(2):179–86.
20. Powar OS, Chemmangat K, Figarado S. A novel pre-processing procedure for enhanced feature extraction and characterization of electromyogram signals. *Biomed Signal Process Control*. 2018;42:277–86.
21. Maier J, Naber A, Ortiz-Catalan M. Improved prosthetic control based on myoelectric pattern recognition via wavelet-based de-noising. *IEEE Trans Neural Syst Rehabil Eng*. 2018;26(2):506–14.
22. Phinyomark A, Campbell E, Scheme E. Surface electromyography (EMG) signal processing, classification, and practical considerations. In: *Biomedical Signal Processing*. Springer Singapore; 2020. p. 3–29.
23. Sensinger JW, Lock BA, Kuiken TA. Adaptive pattern recognition of myoelectric signals: Exploration of conceptual framework and practical algorithms. *IEEE Trans Neural Syst Rehabil Eng*. 2009;17(3):270–8.
24. Tommasi T, Orabona F, Castellini C, Caputo B. Improving control of dexterous hand prostheses using adaptive learning. *IEEE Trans Robot*. 2013;29(1):207–19.
25. Stachaczyk M, Farokh Atashzar S, Farina D. Adaptive spatial filtering of high-density EMG for reducing the influence of noise and artefacts in Mmyoelectric control. *IEEE Trans Neural Syst Rehabil Eng*. 2020;28(7):1511–7.
26. Gazzoni M, Farina D, Merletti R. A new method for the extraction and classification of single motor unit action potentials from surface EMG signals. *J Neurosci Methods*. 2004;136(2):165–77.

27. Jiang N, Englehart K, Parker P. Extracting simultaneous and proportional neural control information for multiple-DOF prostheses from the surface electromyographic signal. *IEEE Trans Biomed Eng.* 2009;56(4):1070–80.
28. Ison M, Artemiadis P. The role of muscle synergies in myoelectric control: Trends and challenges for simultaneous multifunction control. *J Neural Eng.* 2014;11(5).
29. Teh Y, Hargrove LJ. Using latent representations of muscle activation patterns to mitigate myoelectric interface noise. In: *10th Int IEEE/EMBS Conf Neural Eng.* 2021. p. 1148–51.
30. Hahne JM, Graimann B, Muller KR. Spatial filtering for robust myoelectric control. *IEEE Trans Biomed Eng.* 2012;59(5):1436–43.
31. Muceli S, Jiang N, Farina D. Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms. *IEEE Trans Neural Syst Rehabil Eng.* 2014;22(3):623–33.
32. López NM, di Sciascio F, Soria CM, Valentinuzzi ME. Robust EMG sensing system based on data fusion for myoelectric control of a robotic arm. *Biomed Eng Online.* 2009;8:1–13.
33. Fougner A, Scheme EJ, Chan ADC, Englehart K, Stavadahl Ø. Resolving the limb position effect in myoelectric pattern recognition. *IEEE Trans Neural Syst Rehabil Eng.* 2011;19(6):644–51.
34. Teh Y, Hargrove LJ. Understanding limb position and external load effects on real-time pattern recognition control in amputees. *IEEE Trans Neural Syst Rehabil Eng.* 2020;28(7):1605–13.
35. Hu B, Simon AM, Hargrove L. Deep generative models with data augmentation to learn robust representations of movement intention for powered leg prostheses. *IEEE Trans Med Robot Bionics.* 2019;1(4):267–78.
36. Yang D, Liu H, Member S. An EMG-based deep learning approach for multi-DOF wrist movement decoding. *IEEE Trans Ind Electron.* 2021;0046(52075114).
37. Luo T, Zhang X, We L, Chen X, Chen X, Chen X. Convolutional neural network with data augmentation for robust myoelectric control. In: *IEEE Int Conf Comput Intell Virtual Environ Meas Syst Appl.* 2019.
38. Lecun Y. *Modeles connexionistes de l'apprentissage.* Universit´e de Paris VI; 1987.
39. Lecun Y, Bengio Y. Convolutional networks for images, speech, and time-Series. *Handb Brain Theory Neural Networks.* 1995;3361(10):63–6.
40. Holden D, Saito J, Komura T, Joyce T. Learning motion manifolds with convolutional autoencoders. *SIGGRAPH Asia 2015 Tech Briefs, SA 2015.* 2015;1–4.
41. Atzori M, Cognolato M, Müller H. Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands. *Front Neurobot.* 2016;10(SEP):1–10.
42. Geng W, Du Y, Jin W, Wei W, Hu Y, Li J. Gesture recognition by instantaneous surface EMG images. *Sci Rep.* 2016;6(October):6–13.
43. Ameri A, Akhaee MA, Scheme E, Englehart K. Regression convolutional neural network for improved simultaneous EMG control. *J Neural Eng.* 2019;16(3).

44. Yang W, Yang D, Liu Y, Liu H. Decoding simultaneous multi-DOF wrist movements from raw EMG signals using a convolutional neural network. IEEE Trans Human-Machine Syst. 2019;49(5):411–20.

## Figures

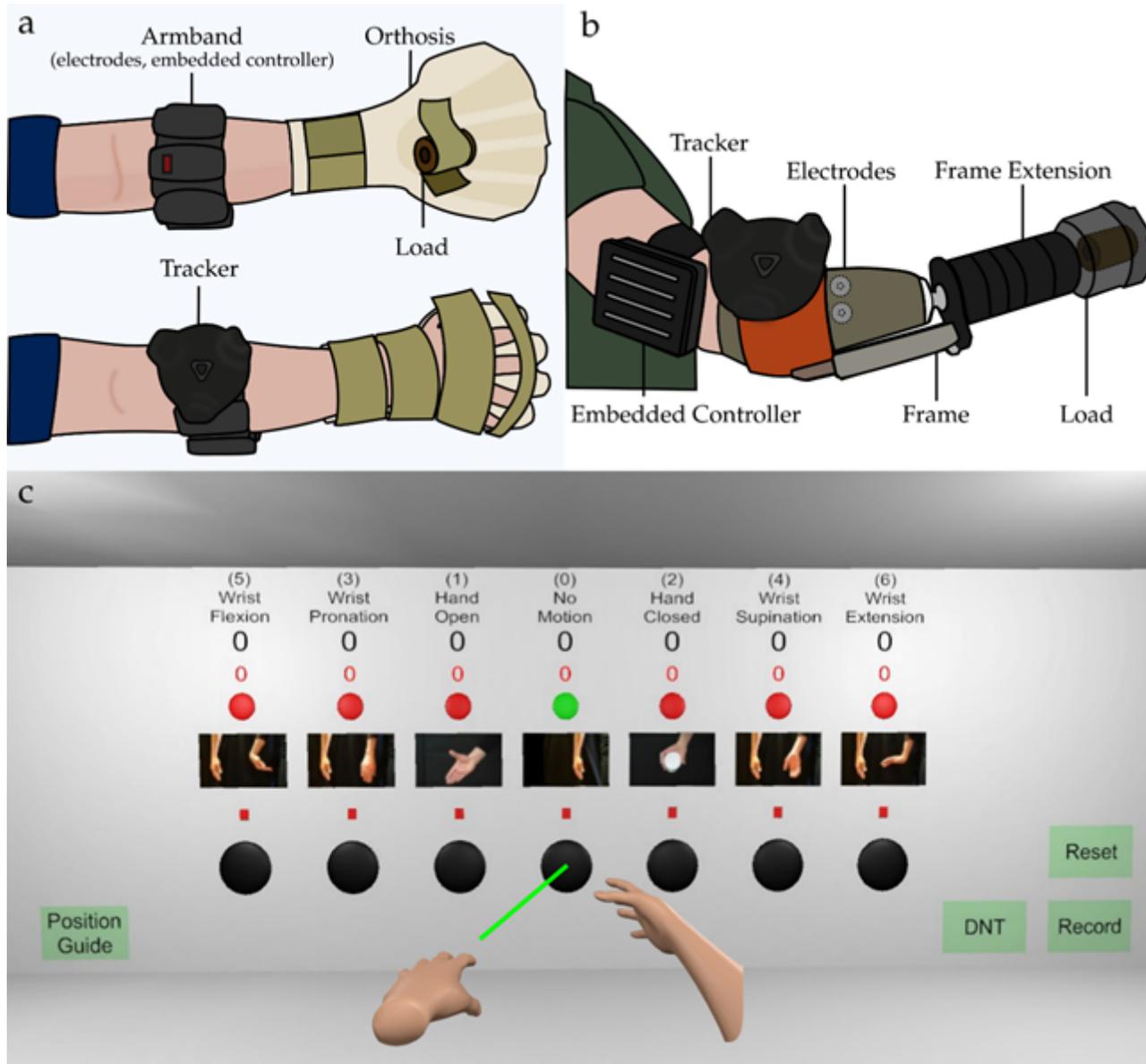
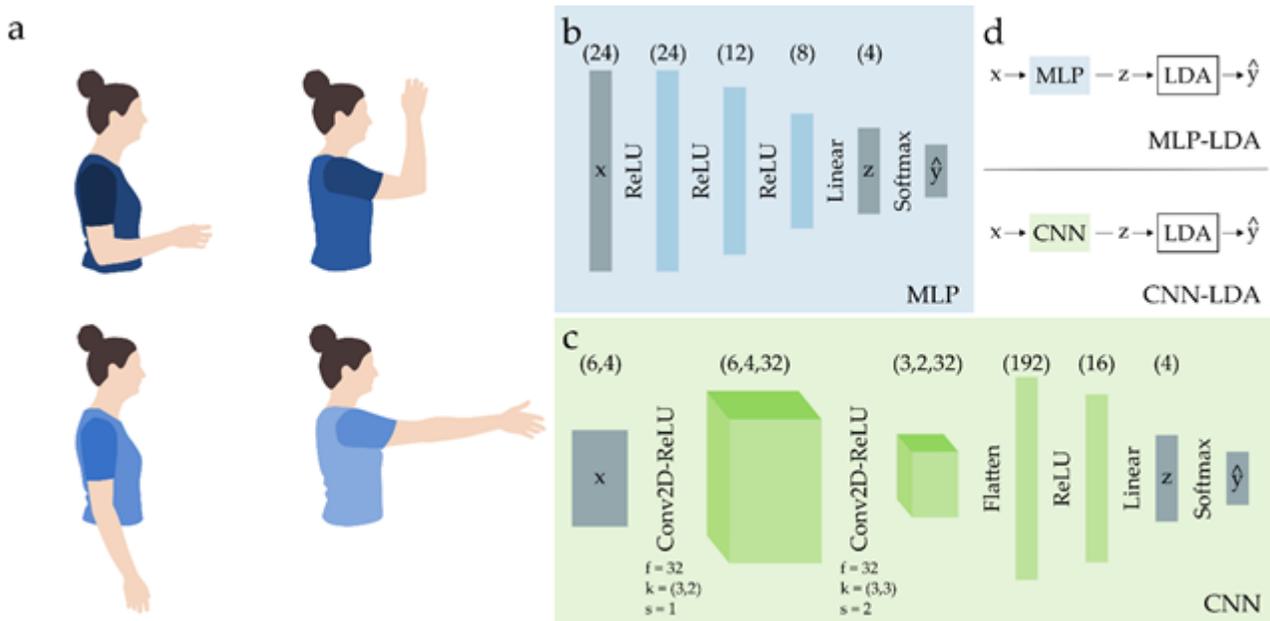


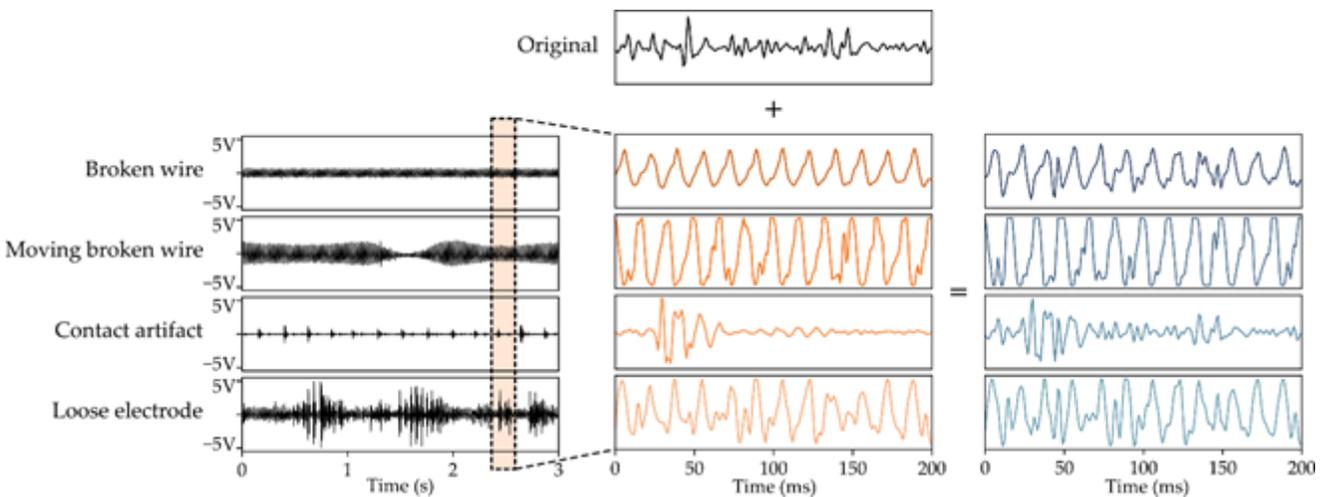
Figure 1

Experimental setups for (a) intact limb participants and (b) amputee participants. (c) Virtual reality environment used for data collection.



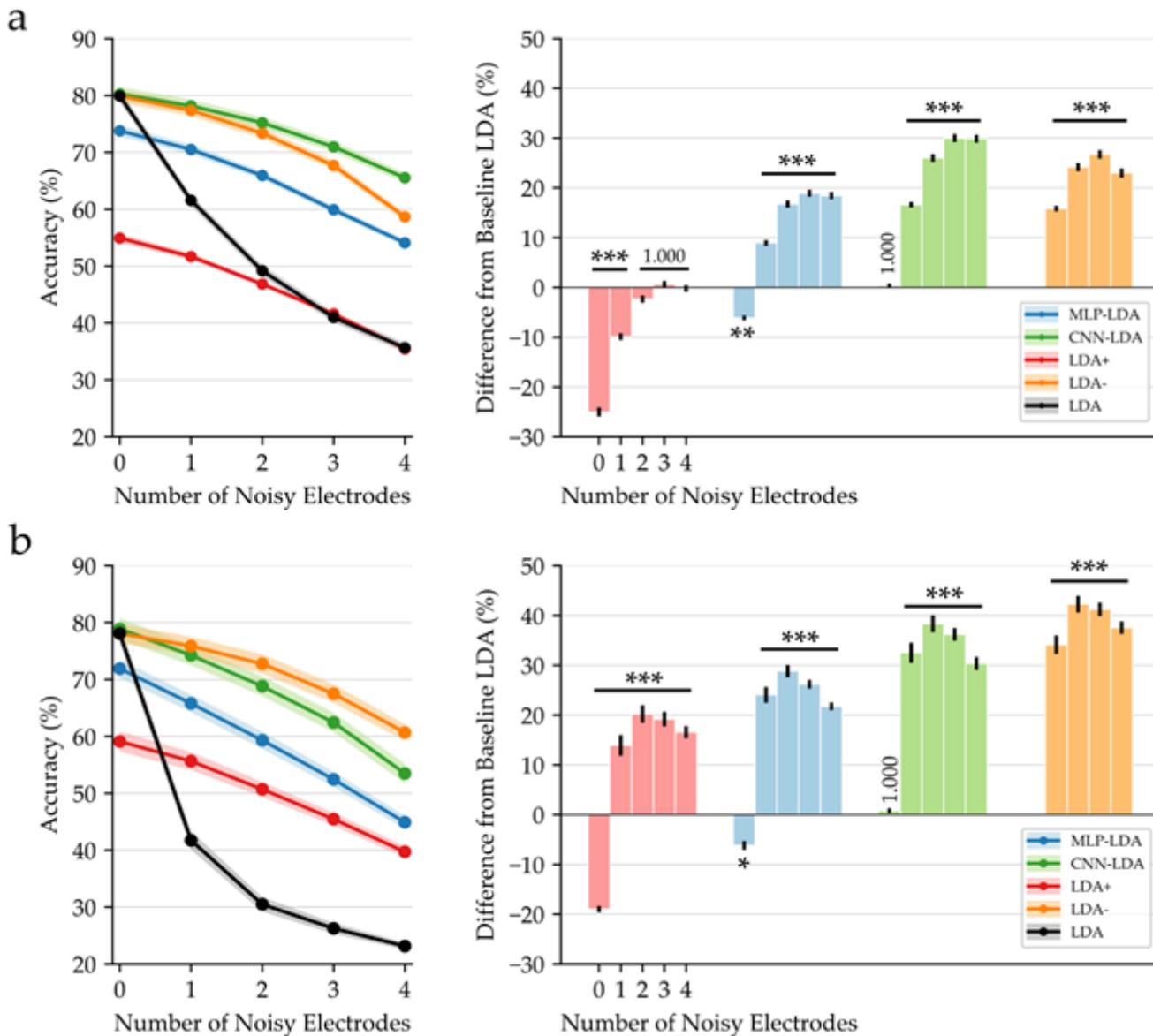
**Figure 2**

(a) To increase test data variability, participants completed wrist and hand gestures in four arm positions. Network architectures for (b) MLP and (c) CNN. (d) The MLP-LDA and CNN-LDA control strategies used neural networks to compute EMG latent features that were classified by LDA classifiers.



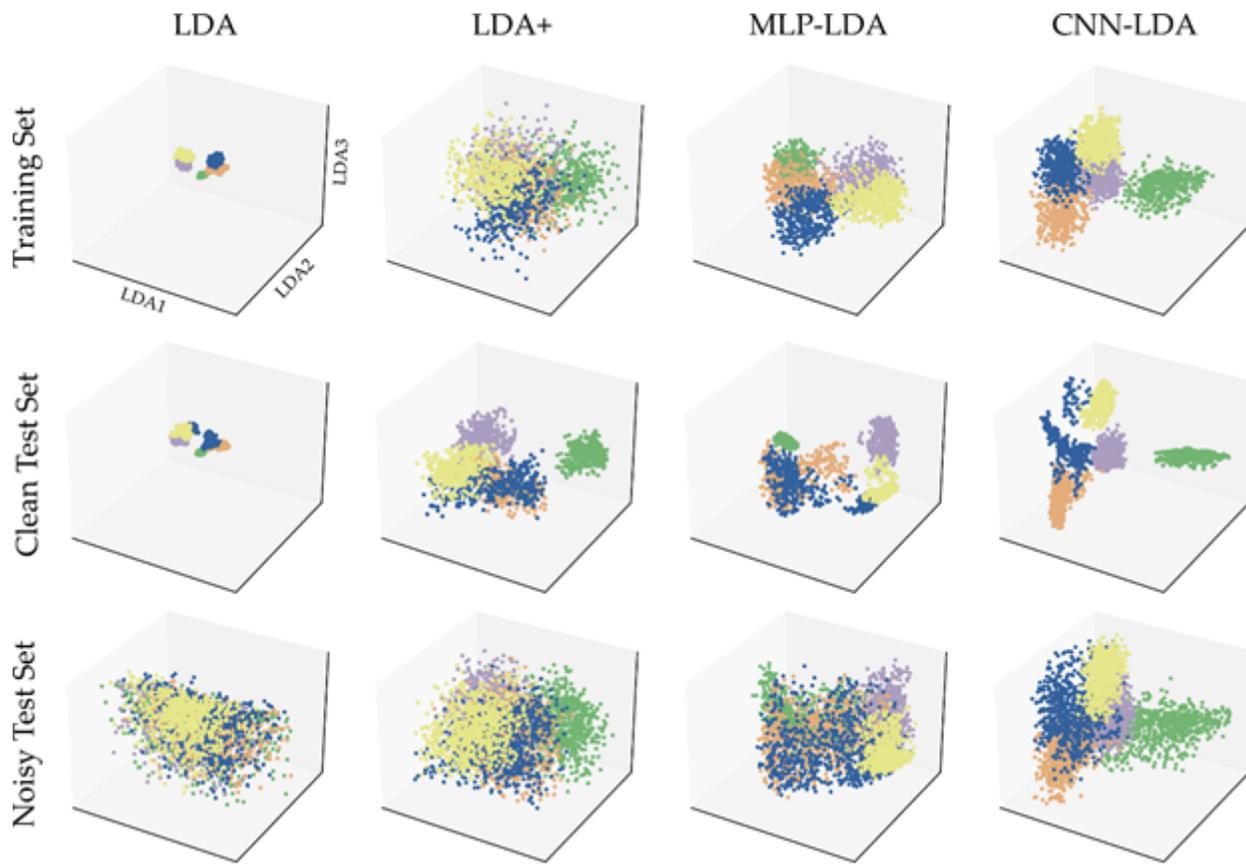
**Figure 3**

Examples from the real noise database fused with a clean EMG signal.



**Figure 4**

Average classification accuracies and differences from baseline LDA accuracies for (a) intact limb participants and (b) amputee participants.



**Figure 5**

Three-dimensional latent representations of AMP5's training and test data sets using LDA components.