Cross-Subject Movement Decoding using Neural Latent Factor Analysis and Convolutional Networks

Daniel Katz, Roi Wayner
Deep Learning for Physiological Signals (3360209)
Faculty of Biomedical Engineering, Technion, Israel Institute of Technology
https://github.com/KanielDatz/cross-subject-movement-decoding

August 5, 2025

Abstract

Cross-subject neural decoding represents a critical challenge for practical brain-computer interface deployment due to substantial inter-subject variability in neural activity patterns. This study systematically evaluates multiple deep learning approaches for cross-subject movement classification using motor cortex recordings from two rhesus monkeys performing 12 distinct digit and wrist movements. We compare convolutional neural networks (CNNs) applied to raw neural data, channel-shuffled inputs, and latent factors extracted via Latent Factor Analysis via Dynamical Systems (LFADS). Cross-subject generalization proved challenging across most methods, with LFADS successfully generalizing for one movement only. We believe improving LFADS could achieve better latent space representation enabling broader future generalization. Throughout the study, we explored the data using several algorithms and clustering methods to investigate neural patterns. We demonstrated reconstruction of movement kinematics and EMG signals from neural activity, with attention analysis.

1 Introduction

Brain-computer interfaces (BCI) have shown remarkable progress in decoding movement intentions from neural activity [1, 2]. However, cross-subject generalization remains a significant barrier due to individual differences in electrode placement, cortical anatomy, and neural firing patterns that create substantial domain shifts limiting model transferability between subjects. Our case study aims to perform cross-subject generalization for finger movement prediction between two different monkeys from which we received previously recorded neuronal data [3]. We employed a cascading hierarchy of algorithms, starting with a simple convolutional network for movement classification. Understanding that cross-subject generalization requires invariance to input order, we progressed to an enhanced convolutional network capable of successfully handling shuffled inputs. However, when we tested this network on the second monkey, it led to complete failure. Consequently, we decided to transition to a more powerful tool - LFADS. LFADS (Latent Factor Analysis via Dynamical Systems) extracts low-dimensional neural dynamics that capture the underlying structure of motor cortical activity, providing a potentially more robust approach for cross-subject transfer learning [4, 5].

2 Methods

2.1 Data Collection and Preprocessing

Neural data were collected by M.H. Schieber [3] from two rhesus monkeys (*Chip* and *Gabby*) implanted with single-cell electrodes in the primary motor cortex (M1). Both subjects performed a cued movement task consisting of 12 distinct movements: flexion and extension of digits 1–5 and the wrist (labeled 1e, 1f, 2e, 2f, 3e, 3f, 4e, 4f, 5e, 5f, 6e, 6f). Neural spike trains were recorded and converted into binned spike counts for analysis. Notably, each recording session captured activity from a single neuron; thus, the dataset was synthetically constructed by aligning multiple trials of the same movement task around movement onset, with each channel representing a neuron recorded in a separate session.

The curated datasets contained 152 active motor cortex neurons for *Chip* and 50 active neurons for *Gabby*. Temporal windows were cropped to 800 ms ($\pm$400 ms around movement onset) to maintain consistent timing across sessions. This preprocessing ensured temporal alignment while preserving movement-related neural dynamics during the critical execution period.

2.2 Model Architectures

We evaluated four main models:

CNN on Raw Neural Data: A 1D convolutional neural network was trained directly on multichannel neural spike trains. The architecture consisted of temporal convolutional layers with kernel size 3 and 64 filters, followed by fully connected layers for 12-class movement classification.

CNN on Shuffled Channels: To test the hypothesis that spatial channel organization limits cross-subject transfer, we trained improved version CNN architectures on randomly shuffled channel inputs. This approach aimed to force the model to learn channel-invariant temporal features.

LFADS + Conv1D + fully connected head: The LFADS framework encodes high-dimensional neural spike trains into lower-dimensional latent factors that capture essential motor cortical dynamics while filtering subject-specific variations. Subsequently, the convolutional head extracts a set of vectors for classification from these latent representations, which are then fed into the fully connected layers that predict finger movements.

2.3 Attention Mechanisms and Interpretability

Attention mechanisms were incorporated into CNN models to provide interpretability and identify which neurons and time windows contribute most to movement classification. These attention maps allowed us to visualize model focus and compare attention patterns between raw and shuffled channel approaches.

2.4 Training Protocol and Evaluation

All software experiments were carried out on a Windows 11 workstation with a NVIDIA RTX 4090 GPU (24 GB VRAM), an Intel i7-13700 CPU, and 64 GB RAM, using Python and PyTorch. All models were trained on Chip's data and tested on Gabby's data to evaluate cross-subject generalization. Training used standard supervised learning with cross-entropy loss for movement classification. Model performance was evaluated using classification accuracy and confusion matrices. We also conducted single neuron analysis to identify individual neurons capable of movement classification.

3 Results

3.1 CNN Performance on Raw and Shuffled Data

Figure 1-Right shows the performance comparison demonstrates that the simple CNN model achieved high predictive accuracy on the test set for the same subject (91% on Chip). However, when we evaluated this model on the second monkey, we encountered complete failure. This poor cross-subject performance was observed both on raw input data and on shuffled input configurations.

Figure 1-Left demonstrates the performance of the improved CNN model designed to handle shuffled input. The results show that the model achieves good performance on the same subject (58% on Chip). However, when we tested it on the second monkey, it also failed completely. Consequently, we recognized the need to adopt a more aggressive approach to address the cross-subject generalization challenge.

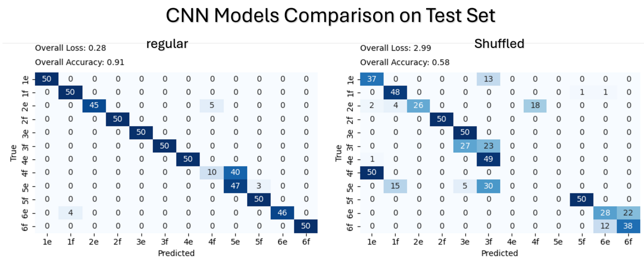

Figure 1: Regular(L) vs Shuffled(R) CNN models performance comparison.

3.2 Attention Analysis

The attention mechanism analysis revealed crucial insights into model behavior. Figure 2 shows attention maps for the standard channel CNN, demonstrating that specific neurons provide the majority of information for decision-making across multiple movements. This indicates that the network performs classification based primarily on these particular neurons, while the contribution of other neurons remains minimal. This model explains the inability of the shuffled model to generalize across subjects, as it most likely simply locates these predictive neurons rather than relying on deeper, more generalizable patterns for decision-making.

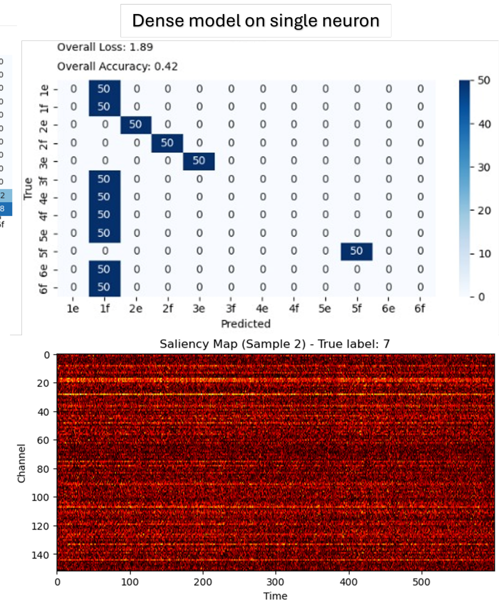

Figure 2: Top: A single neuron Dense model was trained for each of the high attention neurons - all demonstrated good performance on several distinct movements. Bottom: Attention map for regular channel CNN showing temporal channel importance across different time windows.

3.3 LFADS-Based Approaches

LFADS extracted low-dimensional latent factors that captured movement-related neural dynamics while reducing noise and inter-trial variability. We evaluated both 15-factor and 100-factor models to understand the impact of latent dimensionality on cross-subject transfer.

Figure 3 shows 3D PCA visualization of LFADS factors from Chip, The analysis reveals a clear separation in the latent space between different movements, as well as consistent representations of the same movement across different trials for the same monkey.

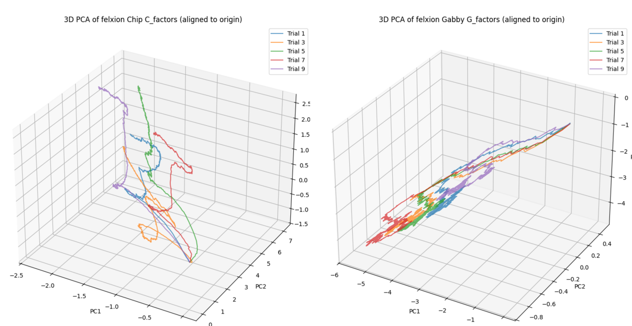

Figure 3: 3D PCA visualization of LFADS factors from Chip (Left) and Gabby (Right) showing movement-specific clustering in latent space, demonstrating the ability to capture structured neural dynamics.

However, when examining the latent space representations of the same movement across two different monkeys, a significant discrepancy emerges. This difference may hinder the model's ability to generalize across subjects, as ideally, the latent space should capture only the essential features of the movement - features that are expected to be similar, if not identical, for the same movement regardless of the individual performing it. (Figure 4)

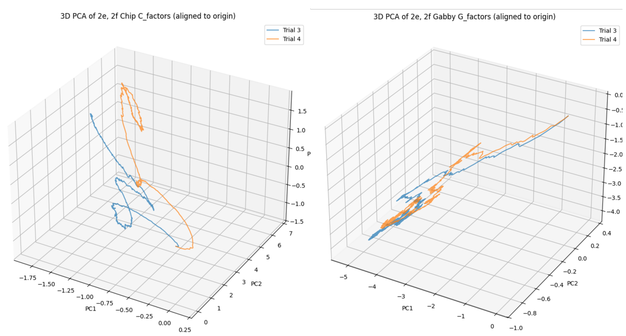

Figure 4: 3D PCA of LFADS factors for same-digit movements showing inconsistent clustering patterns

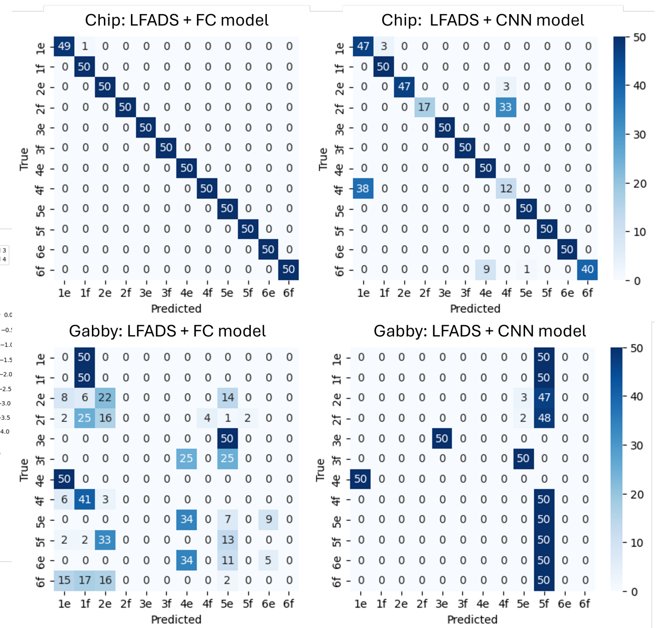

Figure 5: LFADS + Conv1D model results

As seen in Fig. 5 for Chip, both the LFADS + Fully Connected (FC) and LFADS + Convolutional Neural Network (CNN) models achieved near-perfect classification across all 12 movements. In contrast, performance on Gabby was significantly worse. The FC model correctly classified several movements but exhibited widespread misclassification overall, with predictions frequently assigned to incorrect labels. The CNN model failed entirely on Gabby, collapsing to predict nearly all examples as a single movement class. Surprisingly, one additional class was correctly identified alongside the dominant predicted label.

4 Discussion

Our results demonstrate that convolutional neural network–based classifiers, whether operating on raw multichannel spike trains or shuffled-channel inputs, achieve strong within-subject performance but fail to generalize across animals. The standard CNN appears to rely heavily on subject-specific spatial patterns—i.e., individual neurons with high predictive value—while the shuffled-channel variant, despite encouraging channel invariance, still collapses when faced with a new subject. LFADS, by extracting low-dimensional latent dynamics, markedly improves robustness within Chip's data and reveals clear movement-specific clustering in latent space. However, the same-digit PCA plots across monkeys highlight a pronounced misalignment of latent representations between subjects, which impedes reliable cross-subject decoding. The fact that one movement could be classified successfully suggests that there is some shared latent structure, but this is insufficient for broad transfer.

Moreover, our dataset—comprised of only 152 neurons for Chip and 50 for Gabby—may be too limited for LFADS to fully capture the rich dynamics underlying motor control. Prior work by M. Churchland has C. Pandarinath leveraged multi-electrode arrays recording from several hundred neurons to achieve more stable latent factor extraction and improved generalization [4, 6]. To address this, future studies should consider expanding the recorded neural population, collecting additional trials, and optimizing LFADS training (e.g., longer training schedules, finer hyperparameter sweeps) to better harness latent dynamics. Finally, to bridge the cross-subject gap, domain-adaptation techniques—such as adversarial alignment or contrastive learning—and mathematical transformations like canonical correlation analysis (CCA) should be investigated to explicitly align latent spaces across subjects and emphasize shared movement dynamics.

5 Conclusion

LFADS-based models achieved strong within-subject decoding but limited cross-subject transfer. Future work should: Expand neural recordings (hundreds of neurons, more trials). Refine LFADS training (longer schedules, hyperparameter sweeps). Align latent spaces across subjects (domain adaptation, CCA)

These steps will help harmonize representations and boost subject-invariant decoding.

References

[1] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.

[2] B. Wodlinger, J. E. Downey, E. C. Tyler-Kabara, A. B. Schwartz, M. L. Boninger, and J. L. Collinger, "Ten-dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations," *Journal of neural engineering*, vol. 12, no. 1, p. 016011, 2015.

[3] M. H. Schieber, "Limited functional grouping of neurons in the motor cortex hand area during individuated finger movements: a cluster analysis," *Journal of neurophysiology*, vol. 82, no. 6, pp. 3488–3505, 1999.

[4] C. Pandarinath, D. J. O'Shea, J. Collins, R. Jozefowicz, S. D. Stavisky, J. C. Kao, E. M. Trautmann, M. T. Kaufman, S. I. Ryu, L. R. Hochberg *et al.*, "Inferring single-trial neural population dynamics using sequential auto-encoders," *Nature methods*, vol. 15, no. 10, pp. 805–815, 2018.

[5] M. M. Churchland and K. V. Shenoy, "Preparatory activity and the expansive null-space," *Nature Reviews Neuroscience*, vol. 25, pp. 213–236, March 2024, accepted: 26 January 2024, Published: 05 March 2024, Issue Date: April 2024.

[6] M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy, "Neural population dynamics during reaching," 2012, data collected by M.H. Schieber.