Fast & Accurate Classification of Task Stages in ECoG Generalizes to Continues Recordings

Amirhossein Khalilian-Gourtani (akg404@nyu.edu)
Neurology Department, New York University School on Medicine
New York, NY, USA

Yasamin Esmaeili (yasamin@nyu.edu)
Biomedical Engineering Department, New York University Tandon School of Engineering
Brooklyn, NY, USA

Adeen Flinker (adeen@nyu.edu)
Neurology Department, New York University School on Medicine
New York, NY, USA

Abstract

In this study, we explore the decoding of task-specific cognitive states and how they generalize to continuous neural recordings. Using electrocorticography (ECoG) data from eight neurosurgical patients performing various speech production tasks, we compare the performance of three time-series classification models (predicting perception, speech and rest). Results demonstrate that our proposed framework based on a mini-rocket model achieves highest accuracy and fastest inference time compared to the other models. Further, we assess the electrode importance for decoding and find a strong correlation between signal activity and decoding accuracy across sensory and motor regions. However, we do not find this relationship in other frontal regions. Our framework exhibits robust generalization capability across recording sessions, showcasing its potential for analyzing continuous neural recordings and deciphering cognitive states accurately.

Keywords: electrocorticography; classification; decoding; generalization; mini-rocket; speech perception and production.

Introduction

Human neuroscience research has undergone a notable transformation, gravitating towards the integration of unstructured, naturalistic, and continuous neural recordings alongside task-structured recordings to unveil deeper insights into brain dynamics. While this transition from structured task-based data allows for a more comprehensive understanding of neural activity, it comes at the cost of increased complexity and the need for laborious manual annotations, especially for motor tasks. Here we ask if learning from different cognitive states in task-based structures can help us identify relevant events in continuous neural recordings?

Methods

We use electrocorticography (ECoG) recordings from eight neurosurgical patients with electrodes sampling perisylvian and occipital regions while they perform speech production tasks varying in modality. The participants were instructed to complete three tasks to produce the same target words in response to certain auditory or visual stimuli: auditory repetition (i.e., repeat an auditory presented word), visual word reading (i.e., read out loud a visually presented written word), and picture naming (i.e., name a word based on a line drawing). For each task and trial the stimulus was randomly presented and participants produced the target word freely after which the next trial started (Fig. 1(A)). We extract neural activity (high gamma broadband: 70-150 Hz; z-scored across the entire recording session) during three event-related cognitive states of each task (rest, stimulus presentation, and speech production; five class labels in total). For each participant, we train a classifier to predict these states given the high-gamma neural activity. We extend the mini-rocket transform (Dempster, Schmidt, & Webb, 2021) to handle multivariate neural signals with unequal temporal lengths for feature extraction, followed by feature normalization and a logistic regression classifier (Fig. 1(A)). We compare mini-rocket against two other state-of-the-art feature extraction methods: HIVE-COTEv2 (Middlehurst et al., 2021) and random interval sampling with Catch22 (Lubba et al., 2019), using five fold stratified cross-validation.

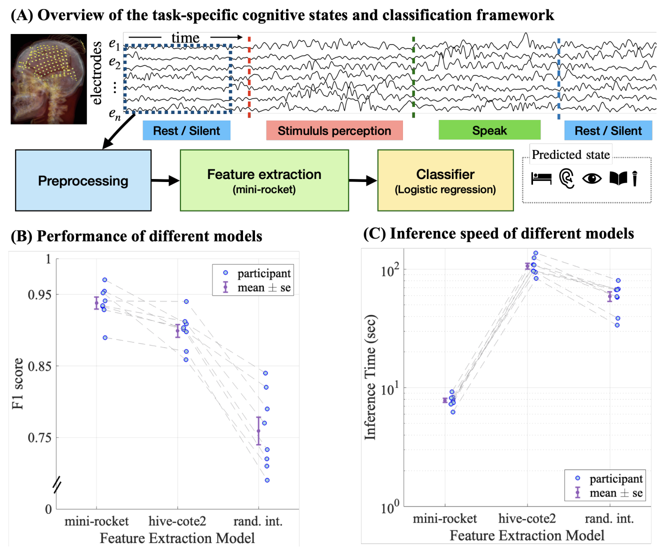

Figure 1: (A) Overview of the classification framework: Each task involves three main cognitive states: rest, stimulus perception, and articulation. A time-series classifier learns to classify these different states given the neural activity. (B,C) Three different feature extraction models are compared based on their classification performance and inference speed (mean and standard error of the mean shown over participants for each model).

Results

We evaluate the decoding performance of task-specific cognitive states for the three different models on a hold-out set (Fig. 1(B,C)). As shown, mini-rocket model, achieves higher accuracy with faster inference time (a: 94% f: 94% t: 7.8s) compared with HIVE-COTEv2 (a: 91% f: 89% t: 106.7s) and random interval Catch22 (a: 76% f: 76% t: 58.9s). Figure 1(B) illustrates the capacity of our adapted mini-rocket model to consistently outperform other models across participants. Leveraging the convolutional architecture, the mini-rocket transform exhibits this high accuracy while significantly reducing inference time by an order of magnitude (Fig. 1(C)). Consequently, the mini-rocket model emerges as a promising candidate for scenarios where we analyze an stream of continues data. We focus the rest of our experiments on the architecture with mini-rocket model.

To quantify the importance of an electrode for decoding, we use the permutation importance test for feature evaluation (Breiman, 2001). The importance score of an electrode is calculated as follows. For the model trained for each participant, we first evaluate the baseline accuracy on a hold-out set. Next, for any given electrode, we permute its signal from the hold-out set and evaluate the accuracy again. The permutation importance score is defined as the mean difference between the baseline accuracy and the accuracy from permuting the electrode signal after 10 repetition of each trial in the hold-out set. We compare the importance score to the high-gamma signal activity of each electrode as in Figure 2.

We observe many electrodes with high importance scores across perisylvian cortex. However, many electrodes with high signal activity do not necessarily exhibit a high importance score (compare Fig. 2(A) and (B)). To quantify the relation between the importance score and signal activity we directly compare and correlate the two values across five regions of interest: superior temporal, pre-central, post-central, inferior frontal, and middle frontal gyri (Fig. 2(C-G), respectively). For electrodes in speech sensory and motor regions we observe a positive relation between signal activity and importance score; electrodes with higher signal activity are contributing more to decoding accuracy. However, in frontal regions, especially middle frontal gyrus, this relation is not significant. In fact, we observe a set of electrodes in middle and inferior frontal gyri with low signal activity that achieve a high importance score. This finding suggests that low values of signal activity can still lead to discernible feature patterns, especially after the mini-rocket convolution transformation. This highlights the significance of investigating decoding importance compared with signal activity alone when investigating mechanisms underlying speech perception and production.

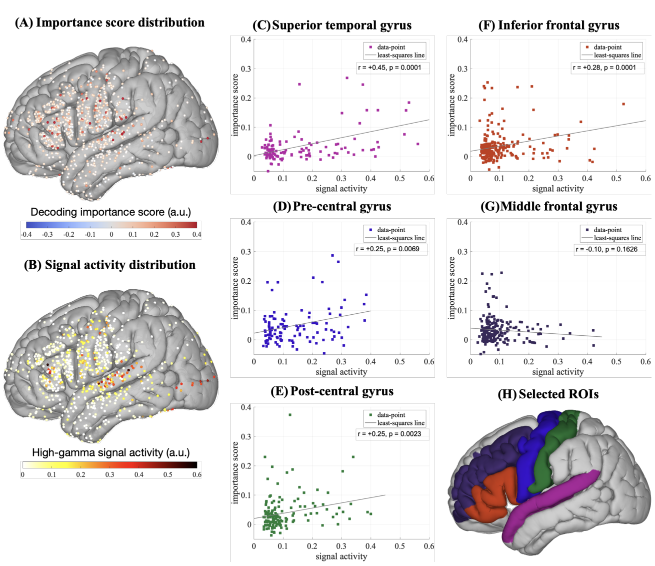

Figure 2: The spatial distribution of decoding importance score (A) and high-gamma signal activity (B) are shown across participants on a normal brain. Importance score plotted against signal activity for five regions of interest (C-G). Each point represents an electrode in a selected ROI (color-codes shown in H; region assignment for each electrode is based on individual participant's anatomy) and the lines show the least-squares line fit (Pearson-correlation and corresponding p-value are shown in boxes for each plot).

To test the generalization capability of our framework to future recordings, one participant repeated the tasks for an extra separate recording session. The mini-rocket classifier trained on one session generalizes well when tested on the other (Fig. 3(A)). We note that the confusion between the two visual perceptions is expected due to no coverage of occipital cortex in this participant. We then applied the classifier trained on recording session 1 to a stream of continuous neural recordings from session 2. Figure 3(B) shows that the mini-rocket model can pinpoint the event windows with high accuracy (vs. annotated event onsets). Notably, the model correctly classifies the transitions between different states with varying time lengths.

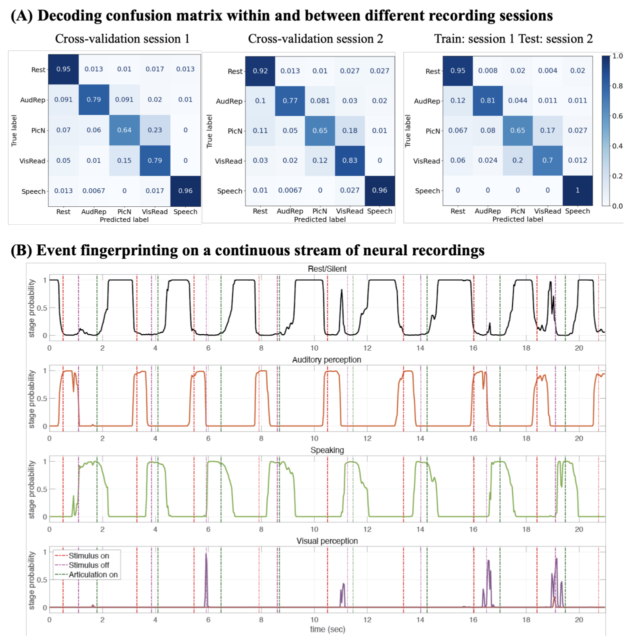

Figure 3: (A) Decoding confusion matrices shown for training and testing on matching and different recording sessions for one participant. The mini-rocket classifier trained on one session generalizes well when tested on the data from a different recording session. (B) Event fingerprinting performed using the mini-rocket model on a continuous stream of neural recordings from auditory repetition is compared to human-annotated event onsets. Each row in (B) plots the probability of being at a given cognitive state over time (from top to bottom: resting, auditory perception, speech production, and visual perception from both visual tasks overlaid). Dotted lines show the human-annotated event onsets in time.

Conclusion

We evaluate classifiers trained on task-structured data for event detection in continuous neural activity streams. We report a fast and accurate classifier that generalizes well to continuous recordings. When scaled up, our model enables annotation and investigation of complex neural activity dynamics in naturalistic scenarios.

Acknowledgment

This work is supported by the National Institute of Health NINDS 1R01NS109367, 1R01NS115929, and NIDCD R01DC018805 grants.

References

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Dempster, A., Schmidt, D. F., & Webb, G. I. (2021). Minirocket: A very fast (almost) deterministic transform for time series classification. In *Proceedings of the 27th acm sigkdd conference on knowledge discovery & data mining* (pp. 248–257).

Lubba, C. H., Sethi, S. S., Knaute, P., Schultz, S. R., Fulcher, B. D., & Jones, N. S. (2019). catch22: CAnonical Time-series CHaracteristics: Selected through highly comparative time-series analysis. *Data Mining and Knowledge Discovery*, *33*(6), 1821–1852.

Middlehurst, M., Large, J., Flynn, M., Lines, J., Bostrom, A., & Bagnall, A. (2021). HIVE-COTE 2.0: a new meta ensemble for time series classification. *Machine Learning*, *110*(11), 3211–3243.