



## PAPER

## Generalized neural decoders for transfer learning across participants and recording modalities

## OPEN ACCESS

RECEIVED  
30 October 2020REVISED  
3 December 2020ACCEPTED FOR PUBLICATION  
8 January 2021PUBLISHED  
1 March 2021

Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Steven M Peterson<sup>1,2</sup> , Zoe Steine-Hanson<sup>3</sup> , Nathan Davis<sup>3</sup>, Rajesh P N Rao<sup>3,4,5</sup> and Bingni W Brunton<sup>1,2,\*</sup> <sup>1</sup> Department of Biology, University of Washington, Seattle, WA 98195, United States of America<sup>2</sup> eScience Institute, University of Washington, Seattle, WA 98195, United States of America<sup>3</sup> Paul G Allen School of Computer Science and Engineering, University of Washington, Seattle, WA 98195, United States of America<sup>4</sup> Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195, United States of America<sup>5</sup> Center for Neurotechnology, University of Washington, Seattle, WA 98195, United States of America

\* Author to whom correspondence should be addressed.

E-mail: [bbrunton@uw.edu](mailto:bbrunton@uw.edu)**Keywords:** neural decoding, deep learning, spectral power, transfer learningSupplementary material for this article is available [online](#)**Abstract**

*Objective.* Advances in neural decoding have enabled brain-computer interfaces to perform increasingly complex and clinically-relevant tasks. However, such decoders are often tailored to specific participants, days, and recording sites, limiting their practical long-term usage. Therefore, a fundamental challenge is to develop neural decoders that can robustly train on pooled, multi-participant data and generalize to new participants. *Approach.* We introduce a new decoder, HTNet, which uses a convolutional neural network with two innovations: (a) a Hilbert transform that computes spectral power at data-driven frequencies and (b) a layer that projects electrode-level data onto predefined brain regions. The projection layer critically enables applications with intracranial electrocorticography (ECoG), where electrode locations are not standardized and vary widely across participants. We trained HTNet to decode arm movements using pooled ECoG data from 11 of 12 participants and tested performance on unseen ECoG or electroencephalography (EEG) participants; these pretrained models were also subsequently fine-tuned to each test participant. *Main results.* HTNet outperformed state-of-the-art decoders when tested on unseen participants, even when a different recording modality was used. By fine-tuning these generalized HTNet decoders, we achieved performance approaching the best tailored decoders with as few as 50 ECoG or 20 EEG events. We were also able to interpret HTNet's trained weights and demonstrate its ability to extract physiologically-relevant features. *Significance.* By generalizing to new participants and recording modalities, robustly handling variations in electrode placement, and allowing participant-specific fine-tuning with minimal data, HTNet is applicable across a broader range of neural decoding applications compared to current state-of-the-art decoders.

**1. Introduction**

Brain-computer interfaces that interpret neural activity to control robotic or virtual devices have shown tremendous potential for assisting patients with neurological disabilities, including motor impairments, sensory deficits, and mood disorders [1–8]. At the same time, brain-computer interfaces offer new insights about the function of neural circuits, including how sensorimotor information is represented in the brain [9–11]. Advances in brain-computer

interfaces have been driven in part by improved neural decoding algorithms [12, 13]. However, it can be difficult to collect enough data to train decoders, especially given the non-stationary nature of the recorded signals, leading to decoders that generalize poorly to new data and require frequent re-calibrations [14–17]. Alternatively, *generalized neural decoders* can be trained by pooling data across multiple participants [18–21]. Such generalized decoders must be robust to inter-participant differences and capable of fine-tuning with only a few

training examples. By increasing decoder robustness and reducing the burden of repeated calibrations, generalized decoders have the potential to greatly enhance the practical long-term usage of brain-computer interfaces [22].

Frequency-domain techniques that extract spectral power features from time-domain recordings have long been shown to be useful in decoding neural population recordings. These techniques are especially well-suited to neural recordings such as intracranial electrocorticography (ECoG) and scalp electroencephalography (EEG) recordings, which contain oscillatory signals at specific frequency bands that correspond to different behaviors or neural phenomena [23–27]. In addition, relative spectral power patterns between a task and a baseline condition can be surprisingly similar across studies, even when measured with different neural recording modalities [28–33]. Such similarities motivate the use of spectral power features for generalized decoding. Because neural recordings are non-stationary, many decoders use instantaneous spectral power features, computed by band-pass filtering the data and then applying the Hilbert transform [17].

To make power spectral features directly comparable, inter-participant differences in electrode placement and frequency content must be addressed when developing generalized decoders. While EEG electrode coverage is typically standardized across participants, invasive ECoG electrode placement is clinically motivated and highly variable, making it difficult to align electrodes from one participant to the next [24, 34]. A similar cross-participant alignment issue occurs with EEG cortical dipoles following blind source separation [35]. To overcome these variable dipole locations, one successful approach has been to project EEG measures onto common brain regions using radial basis function interpolation [36, 37]. ECoG signals can similarly be projected from electrodes to common brain regions [38], but it has remained unclear how useful this method is for neural decoding. Even with aligned electrode placements, the frequency bands containing behaviorally-relevant spectral power can be highly variable across participants [38–40]. This variability is problematic when selecting band-pass filter cutoff frequencies prior to applying the Hilbert transform because researchers often rely on pre-existing knowledge of the neural signal, and traditional frequency bands may not apply to a particular task of interest. To decrease user bias, data-driven approaches have been proposed to analyze neural spectral power [41–43]. However, these techniques apply to frequency bands with distinct spectral power peaks and thus ignore other frequencies that might be useful for decoding. While a promising approach has been recently proposed [44], developing decoders that robustly handle variable frequency content remains an open problem.

We are motivated to address such cross-participant differences in frequency content by using convolutional neural networks, which generate data-driven features and can also be fine-tuned when presented with new data. Convolutional neural networks combine data-driven feature extraction with pattern recognition and have set the bar for state-of-the-art neural decoding performance for speech, motor imagery, and attention tasks [14, 45–49]. Importantly, trained convolutional neural networks can be fine-tuned to new data, a process made more efficient by freezing various network layers during re-training [50, 51]. Far from being ‘black-box’ models, convolutional layers perform data-driven temporal and spatial filtering, and careful analysis of trained convolutional weights can be used to interpret the spatiotemporal features that are key for decoding [52–54]. Our approach builds on EEGNet—a compact convolutional neural network decoder that can be trained on small datasets, provides interpretable model structure, and has outperformed other deep learning decoders [55].

In this paper, we present *HTNet*, a convolutional neural network architecture that decodes neural data with variable electrode placements using data-driven spectral features projected onto common brain regions (figure 1(A)). By training on data pooled across multiple participants, HTNet avoids the issue of neural data scarcity that often results in overfit decoders. We developed HTNet by augmenting EEGNet with custom Hilbert transform and projection onto brain region layers. The main advantage of our approach is its ability to generate data-driven features in the frequency domain that are spatially aligned across participants even when electrode placement is variable, enabling generalized decoding. To characterize its decoding performance, we trained HTNet to classify naturalistic arm movement vs. rest using ECoG recordings from a dataset containing simultaneous video and ECoG recordings from 12 patients being monitored before neurosurgery in a hospital [38]. We then tested these trained decoders’ ability to generalize to unseen ECoG participants or to unseen EEG participants performing a similar behavioral task [56]. In both cases, HTNet was able to decode these naturalistic arm movements and consistently outperformed other state-of-the-art decoders. Furthermore, we fine-tuned pretrained HTNet decoders with a few events from individual participants, and the resultant fine-tuned decoders approach the performance of the best tailored decoders with only 50 ECoG or 20 EEG events. Finally, we interpreted HTNet’s trained weights to understand how it generalizes and found that it primarily relied on physiologically-relevant features at low-frequencies (<20 Hz) near the motor cortex. Our findings demonstrate that HTNet is a generalized decoder that robustly handles inter-participant and inter-modality variability and

fine-tunes to new participants using minimal data. We believe our work advances the field of generalized neural decoders with an architecture that is robust, interpretable, and successful at generalizing across unseen participants and recording modalities.

The rest of this paper is structured as follows. We introduce the neural datasets and pre-processing steps used (sections 2.1–2.3). We then detail HTNet’s model structure, alternative decoders used for comparison, and the three generalization scenarios (sections 2.4–2.5). Next, we describe the methods used to tune decoder hyperparameters, analyze fine-tuning performance, and interpret trained model weights (sections 2.6–2.11). In section 3, we present our results for cross-participant and cross-modality decoding and fine-tuning. Finally, we discuss the implications of our findings along with the limitations of our approach (section 4).

## 2. Methods

### 2.1. Intracranial electrocorticography (ECoG) dataset

We obtained concurrent ECoG and video recordings from 12 human participants (8 males, 4 females) during continuous clinical epilepsy monitoring conducted at Harborview Medical Center in Seattle, WA. These recordings lasted  $7 \pm 2$  days per participant (mean  $\pm$  SD). Participants were aged  $29 \pm 8$  years old (mean  $\pm$  SD) and had electrodes implanted primarily in one hemisphere (5 right, 7 left). Our study was approved by the University of Washington Institutional Review Board for the protection of human participants, and all participants provided written informed consent.

Our decoding task was to classify upper-limb ‘move’ and ‘rest’ events of the arm contralateral to the implanted electrode hemisphere. We obtained non-concurrent move and rest events from video recordings via markerless pose tracking and automated state segmentation (see Singh *et al* [57] for further details). Move events correspond to wrist movement that occurred after at least 0.5 s of no movement, while rest events indicate no movement in either wrist for at least 3 s.

We performed ECoG data processing using custom MNE-Python scripts [58]. We first removed median DC drift and high-amplitude discontinuities. Each participant’s ECoG data was then band-pass filtered (1–200 Hz), notch filtered, and re-referenced to the common median across electrodes. We also removed noisy electrodes based on abnormal standard deviation ( $> 5$  IQR) or kurtosis ( $> 10$  IQR). Next, we generated 10 s ECoG segments centered around each ‘move’ and ‘rest’ event. ECoG segments with missing data or large artifacts were removed based on abnormal spectral power density. See Peterson *et al* [38] for further ECoG pre-processing details.

We then downsampled to 250 Hz and trimmed segments to 2 s centered around each event. For every participant, we balanced the number of move and rest segments within each recording day, resulting in  $1155 \pm 568$  events per participant (mean  $\pm$  SD).

Electrode positions were localized using the Fieldtrip toolbox in Matlab [59, 60]. This process involved co-registering preoperative MRI and post-operative CT scans, manually selecting electrodes in 3D space, and warping electrode positions into Montreal Neurological Institute (MNI) space [61]. Using this common MNI coordinate system enabled us to directly compare electrode positions between ECoG participants (see figure 1(C)).

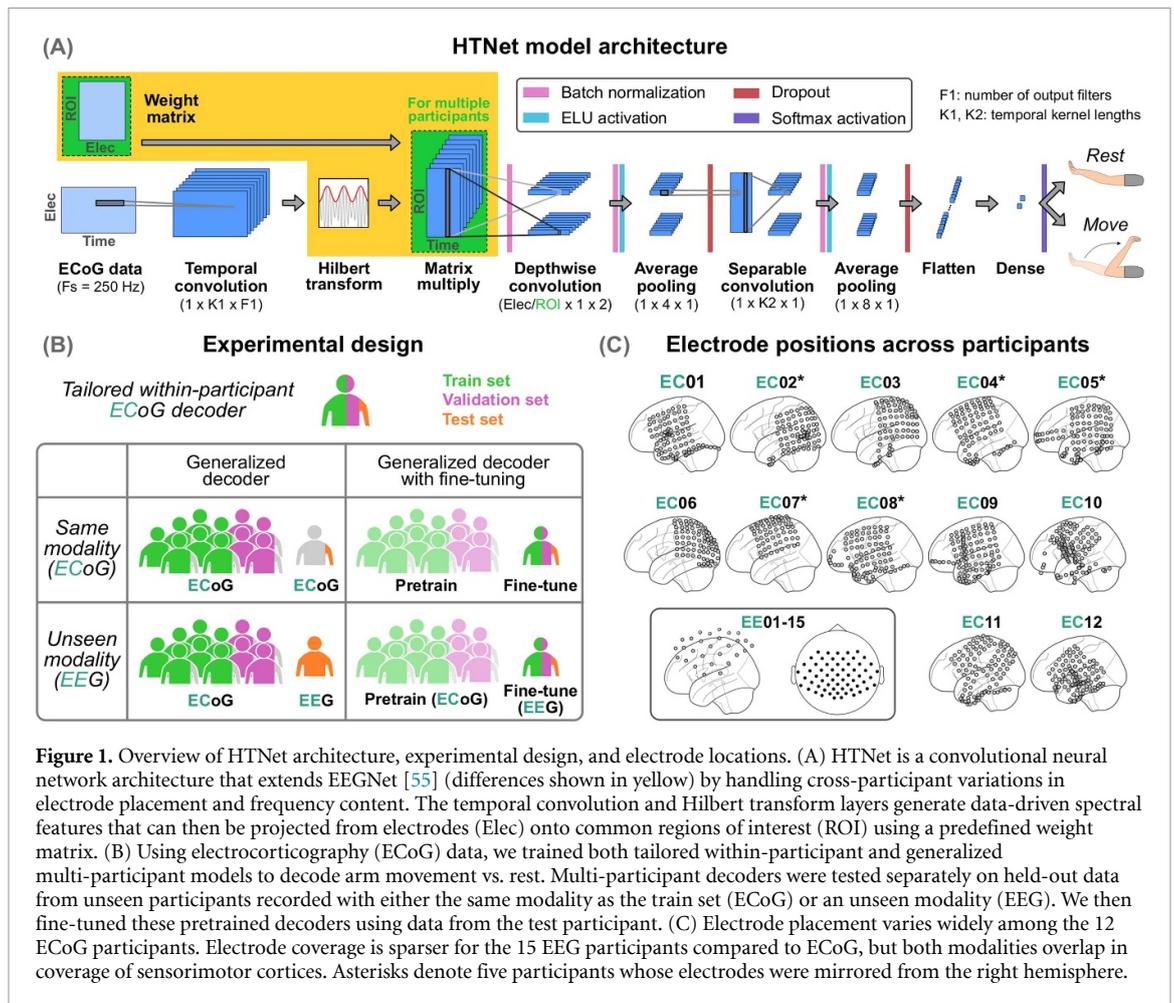
### 2.2. Comparative cross-modal dataset (EEG)

To test decoder generalizability across recording modalities, we used a publicly available EEG dataset of 15 human participants performing cued right elbow flexion movements [56]. Participants were aged  $27 \pm 5$  years old (mean  $\pm$  SD) and performed 60 movement and 60 rest trials each, resulting in 120 total events. Because only cue onset times were available, we determined the onset of each movement event by thresholding the hand’s radial displacement after it was cued to move.

EEG data were recorded at 512 Hz, notch filtered at 50 Hz, referenced to right mastoid, and band-pass filtered between 0.01 and 200 Hz. We pre-processed the data by average referencing, 1 Hz high-pass filtering, resampling to 250 Hz, and generating 2 s segments centered around each event. Each participant had 61 EEG electrodes, whose MNI positions were estimated using the 10-5 system template from Fieldtrip [60, 62].

### 2.3. Computing projection matrices

We accounted for variations in electrode placement by mapping electrode positions onto common brain regions based on distance as described below. To increase electrode overlap among ECoG participants, we mirrored all right hemisphere electrode positions onto the left hemisphere. Using EEGLAB and Matlab, we mapped from electrode positions to small, predefined brain regions by computing radial basis function kernel distances between each electrode and brain region (2 cm full-width at half-maximum) [36, 63, 64]. This projection procedure was performed separately for each participant. We projected to regions within sensorimotor areas, as defined by the AAL atlas [65] (precentral, postcentral, and inferior parietal), in order to limit the projected data size. We then normalized these distance values for every region so that each region’s values across all electrodes summed to one. These normalized distances created a projection matrix for each participant, which we later used to estimate the activity at each common region of interest by performing a weighted average of electrode-level data.



## 2.4. HTNet architecture

HTNet builds upon EEGNet [55], a compact convolutional neural network developed using Python 3.8 and Tensorflow 2.2. EEGNet has three convolution layers: (a) a one-dimensional convolution analogous to temporal band-pass filtering, (b) a depthwise convolution to perform spatial filtering, and (c) a separable convolution to identify temporal patterns across the previous filters. For HTNet, we added a Hilbert transform layer after this initial temporal convolution to compute relevant spectral power features using a data-driven filter-Hilbert analog (see figure 1(A)). We then added a matrix multiplication layer to project electrode-level spectral power onto common brain regions of interest, using the pre-computed weight matrices described in the previous section. Note that the matrix multiplication layer was not necessary when the same participant was used for training and testing, as the electrodes remain consistent. It is worth noting that the time complexity of the Hilbert transform and matrix multiplication layers are  $O(n \log n)$  and  $O(n^3)$ , respectively, where  $n$  represents the number of samples per segment in the input neural data [66]. All other HTNet layers were the same as EEGNet.

## 2.5. Data division and cross-validation

We compared HTNet decoding performance against EEGNet, random forest, and minimum distance decoders. The minimum distance decoder used Riemannian mean and distance values for classification [67, 68]. We assessed decoder performance during three scenarios (figure 1(B)): (a) testing on an untrained recording day for the same ECoG participant (*tailored decoder*), (b) testing on an untrained ECoG participant (*same modality*), and (c) testing on participants from the EEG dataset after training only on the ECoG dataset (*unseen modality*). Note that we used the same trained, multi-participant ECoG decoders for same and unseen modality conditions. Additionally, we projected data onto common regions of interest for all decoders in order to enable reasonable decoding for the same and unseen modality conditions. For all scenarios, we performed 36 pseudo-random selections (folds) of the training and validation datasets, such that each of the 12 ECoG participants was the test participant three times. We used each ECoG participant's last recording day as the test set (orange in figure 1(B)) and excluded it from all training and validation sets. All training, validation, and test sets were balanced with equal numbers of move and rest events. We used nonparametric

statistics to test for significant effects of decoder type on test accuracy (Friedman test,  $p < 0.05$ ) and significant pairwise differences among decoders (Wilcoxon signed-rank test with false discovery rate correction [69]).

## 2.6. Hyperparameter tuning

We performed hyperparameter tuning to identify optimal values for each decoder. We tuned HTNet and EEGNet simultaneously using six hyperparameters: temporal kernel length, separable kernel length, temporal filter count, dropout rate, dropout type, and model type (HTNet or EEGNet). For the random forest decoder, we tuned two hyperparameters: maximum depth and number of estimators. The minimum distance decoder had no tunable hyperparameters.

We tuned hyperparameters separately for the tailored decoder and same modality conditions, using the Optuna toolbox [70]. We ran 25 random forest parameter selections (trials) and 100 HTNet/EEGNet trials for each condition. Performance was measured using validation accuracy, averaged over 36 folds for tailored decoding or 12 folds for same modality. We sampled from parameter space using Optuna's tree-structured Parzen estimator [71, 72], which selects optimal parameter values based on the performance during previous trials.

Overall, we found that hyperparameter selections minimally affected decoder performance (see figure S1 (available online at [stacks.iop.org/JNE/18/026014/mmedia](https://stacks.iop.org/JNE/18/026014/mmedia))). Still, we selected hyperparameter values from the trial with the highest validation accuracy (table S1) for each condition to ensure optimal decoder performance.

## 2.7. Fine-tuning decoder performance to the test participant

In addition to testing generalizability, we assessed how much a generalized HTNet decoder improves when re-trained using data from the test participant, a process known as *fine-tuning*. Fine-tuning is a transfer learning technique where some layers of the pre-trained model are 'frozen' and not adjusted during re-training, reducing the number of parameters to fit [73]. We fine-tuned our pretrained same and unseen modality decoders using a portion of the test participant's data. Additionally, we fine-tuned each HTNet convolutional layer separately and all layers together, resulting in four re-trained models per fold (see figure 3(A)). When separately tuning each convolutional layer, we also re-trained the nearby batch normalization layers, as shown in figure 3(A), which notably boosted performance. We tested for significant differences among these four fine-tuning models using Wilcoxon signed-rank tests with false discovery rate correction.

During fine-tuning, we also varied the amount of training/validation data available in order to

assess its impact on fine-tuning performance. For all variations, the test set remained fixed as the last recording day for each ECoG participant and 30 randomly-selected events for every EEG participant. Using the remaining test participant data, we selected four amounts of training/validation data: 17% training/8% validation, 33% training/17% validation, 50% training/25% validation, and 67% training/33% validation. We then linearly modelled the relationship between test accuracy and the logarithm of the number of training events for each fine-tuning model. All training, validation, and test sets contained equal numbers of move and rest events. Note that the number of events used for fine-tuning differs across ECoG participants because the total number of events varies. In addition, we trained randomly-initialized HTNet decoders with the same training/validation data used for fine-tuning in order to compare tailored decoding with our fine-tuned models.

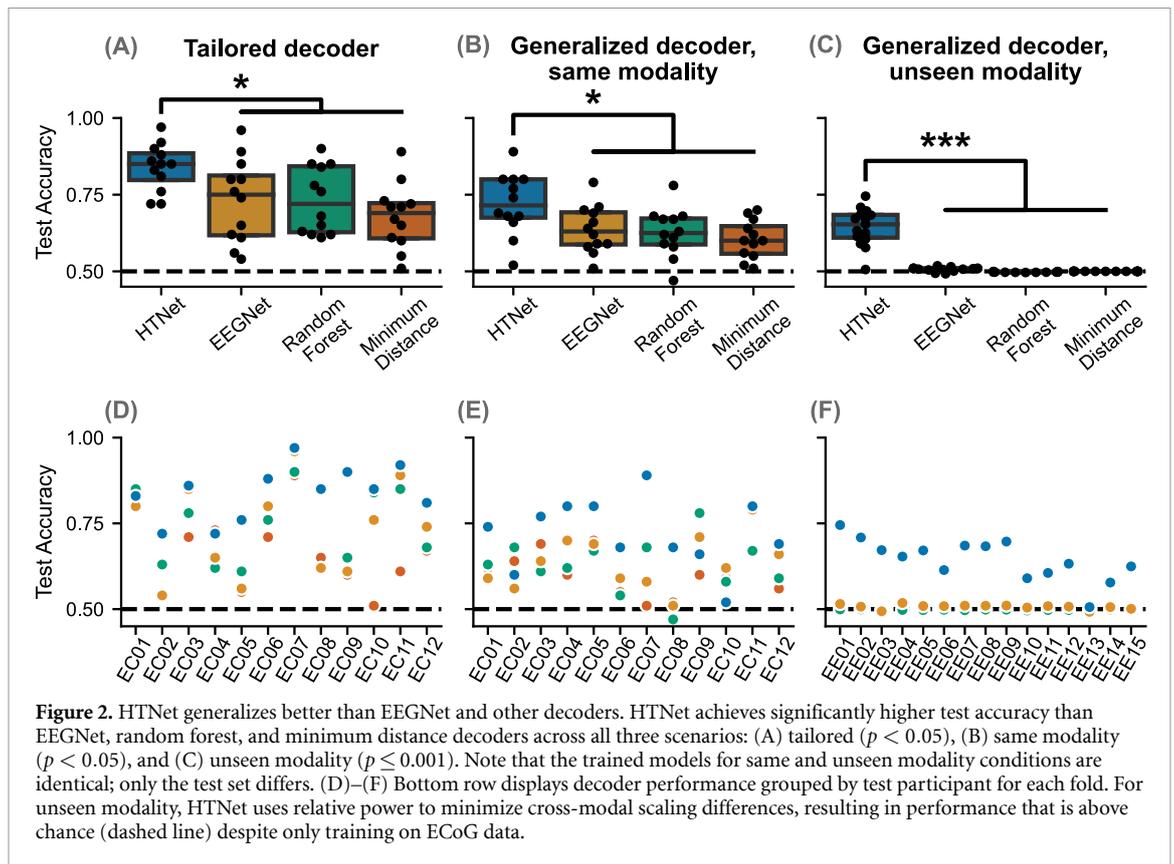
## 2.8. Comparing performance of HTNet spectral measures

Although we primarily used HTNet to generate data-driven spectral power features, HTNet can be easily adapted to generate other spectral measures that may boost decoding performance. These spectral measures still use the Hilbert transform, which can be used to find instantaneous power, phase, or frequency [17]. For the same three conditions, we tested four spectral measures: (a) log-transform of one plus power, (b) relative power, (c) unwrapped instantaneous phase, and (d) instantaneous frequency, also known as frequency sliding [74]. We tested for significant differences in test accuracy when HTNet decoders implemented one of these four measures or spectral power, using Wilcoxon signed-rank tests with false discovery rate correction.

When computing relative power, we first took the log-transform of one plus power and then subtracted out the average power from  $-1$  to  $-0.5$  s before each event. This procedure is analogous to baseline-subtraction of spectral power and should be robust to large-scale spectral power variations across days, participants, and recording modalities. Because of its potential robustness to large differences in signal scaling, relative power was also used in HTNet decoders that we used to compare with other decoder types, but only for the unseen modality condition.

## 2.9. Interpreting model weights

Like EEGNet, HTNet's first two convolutional layers are interpretable and can indicate the spatiotemporal features used for decoding. To analyze the temporal features, we fed a white noise signal into HTNet's trained temporal convolution and computed the frequency response magnitude using Welch's method, similar to a Bode plot [75]. We then averaged frequency responses across temporal filters and folds. To



**Figure 2.** HTNet generalizes better than EEGNet and other decoders. HTNet achieves significantly higher test accuracy than EEGNet, random forest, and minimum distance decoders across all three scenarios: (A) tailored ( $p < 0.05$ ), (B) same modality ( $p < 0.05$ ), and (C) unseen modality ( $p \leq 0.001$ ). Note that the trained models for same and unseen modality conditions are identical; only the test set differs. (D)–(F) Bottom row displays decoder performance grouped by test participant for each fold. For unseen modality, HTNet uses relative power to minimize cross-modal scaling differences, resulting in performance that is above chance (dashed line) despite only training on ECoG data.

determine important spatial features, we computed the absolute value of HTNet’s trained depthwise convolution weights [55], averaged across filters. This process generated one weight for each brain region of interest. We then scaled the maximum value per fold to one and averaged across folds.

While the temporal frequency response shows which frequencies were used for decoding, it does not show what the activity looks like at these frequencies. To visualize such activity, we computed the difference in log spectral power between move and rest conditions, projected onto a region near the motor cortex (figure 4(B) far left region in the second row from the top). We took the difference between the average move and average rest log spectrograms for each participant and then averaged the resulting differences across participants for both the ECoG and EEG datasets.

## 2.10. Effect of training participants on performance

We also assessed how many training participants are needed for improved decoder performance. Increasing the number of training participants should improve decoder performance on an unseen test participant, but the improvement from adding new participants will likely diminish as more training participants are added. We varied the number of training participants from 1 to 10, always using one validation participant. Participants were pseudo-randomly selected across 36 folds such that each participant was the test participant three times. We linearly

modelled the relationship between test accuracy and the logarithm of the number of training participants, identifying decoders with significant nonzero trends ( $p < 0.05$ , two-tailed  $t$ -test with false discovery rate correction).

## 2.11. Effect of electrode overlap on performance

Because electrode locations varied among ECoG participants, we also tested if higher decoding performance corresponded to increased electrode overlap between same modality training and test participants. We estimated electrode overlap between training and test participants using a custom fraction overlap metric that allowed us to combine multiple participants from the training set. For each sensorimotor region, we summed unnormalized projection matrix weights across electrodes to estimate how many electrodes were nearby. We then identified regions of high electrode coverage for each participant by thresholding summations greater than 0.07. Because these regions are common across participants, we could average these summation values across training participants prior to thresholding. Next, we divided the number of thresholded regions common to both training and test participants by the number of thresholded regions in the training set to obtain fraction overlap. A fraction overlap of 1.0 indicates that the test participant’s thresholded regions include all thresholded regions from the training participants. We linearly modelled the relationship between fraction overlap and test accuracy across folds and identified decoders with significant nonzero slope

( $p < 0.05$ , two-tailed  $t$ -test with false discovery rate correction [69]).

### 3. Results

Here, we show that our approach to movement decoding, HTNet, is generalizable and tunable, learning common patterns from the training data that transfer to unseen participants and recording modalities. Through a series of systematic experiments, we use decoders tailored to each participant—by far the most common approach to training decoders—as a standard to which the performance of our generalized decoders are compared (figure 1(B)). In particular, our training data are from 12 ECoG participants during uninstructed, naturalistic arm movements (figure 1(C), [57, 38]); our test data are then either one ECoG participant withheld from the training set, or participants from an entirely independent EEG dataset. HTNet consistently outperformed other decoders, and fine-tuning pre-trained HTNet decoders with a small number of the unseen participant's events yielded decoders that approached the performance of tailored decoders trained on many more events. Further, we show that HTNet works by extracting physiologically-relevant spectral features from the data.

#### 3.1. Decoder generalization

HTNet consistently outperformed other decoders in all conditions, including tailored within-participant decoding, generalized decoding to unseen participants in the same modality, and generalized decoding to unseen modality participants (figure 2). For each condition, we found a significant effect of decoder type on test accuracy (tailored:  $p = 3.94e-4$ , same modality:  $p = 0.034$ , unseen modality:  $p = 1.95e-8$  respectively; Friedman test). For tailored decoding, HTNet achieved test accuracy of  $84\% \pm 8\%$  (mean  $\pm$  SD), which was significantly higher than EEGNet ( $73\% \pm 14\%$ ,  $p = 0.003$ ), random forest ( $73\% \pm 11\%$ ,  $p = 0.004$ ), and minimum distance decoders ( $68\% \pm 11\%$ ,  $p = 0.010$ ; Wilcoxon signed-rank test with false discovery rate correction). For same modality, HTNet test accuracy of  $72\% \pm 10\%$  was again significantly increased compared to EEGNet ( $64\% \pm 8\%$ ,  $p = 0.042$ ), random forest ( $63\% \pm 8\%$ ,  $p = 0.037$ ), and minimum distance decoders ( $60\% \pm 6\%$ ,  $p = 0.021$ ). No other significant pairwise differences were found for either condition.

In the unseen modality condition, HTNet was the only decoder to generalize above chance. All pairwise comparisons were statistically significant ( $p < 0.05$ ) in the unseen modality condition, due to low variability in cross-participant test accuracy for each non-HTNet decoder. However, HTNet's unseen modality test accuracy was, on average,  $\sim 15\%$  higher than

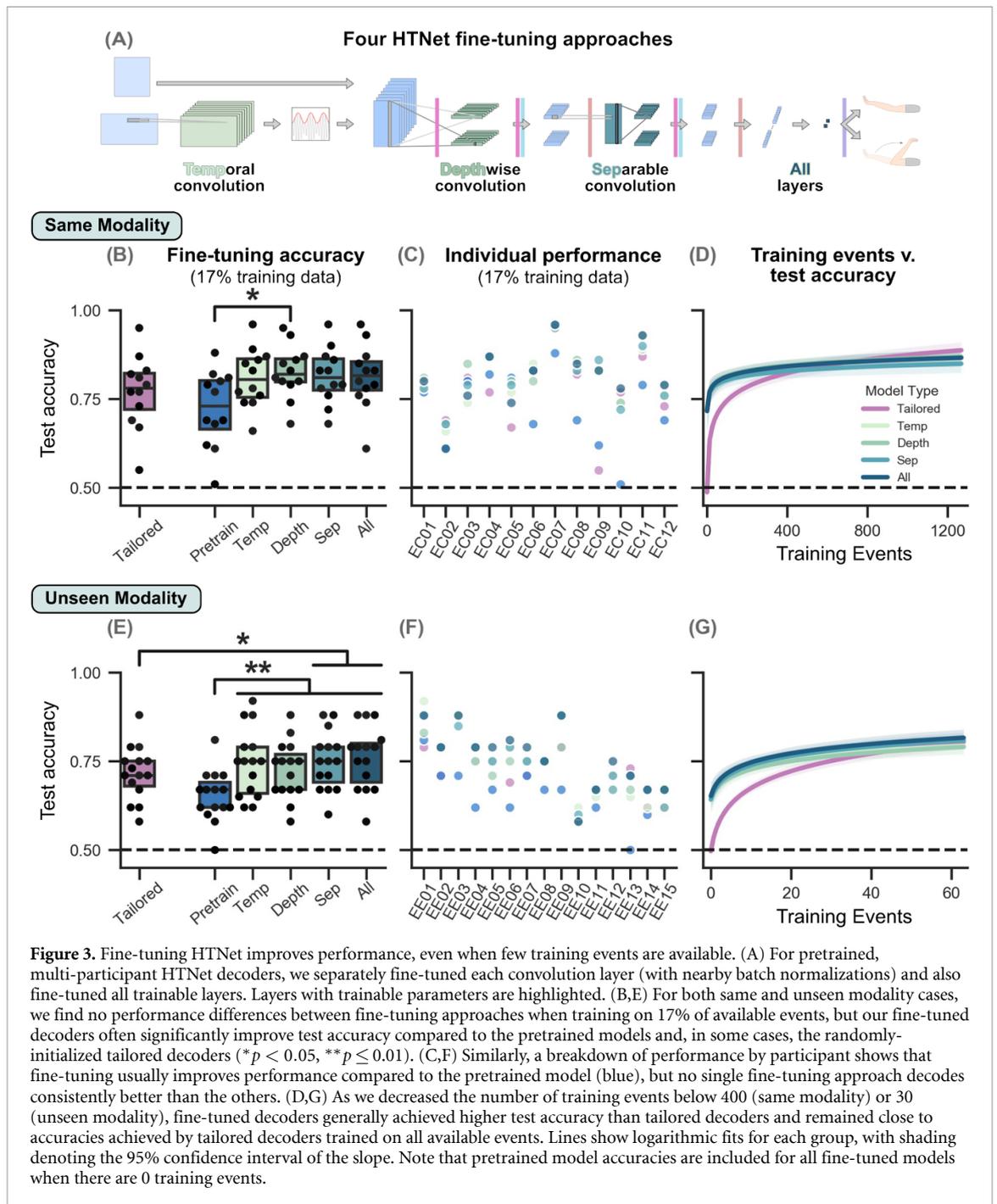
all other decoders and the only decoder to perform well-above chance (50%). Therefore, we only show the pairwise statistical results in figure 2(C) between HTNet and each of the other three decoders. Average computational time during training is shown for each decoder type in tables S2 and S3. In addition to spectral power, we also developed HTNet models that decoded using instantaneous phase and frequency features (figure S2 and tables S4 and S5). Briefly, HTNet with phase performed worse than spectral power for tailored and same modality conditions, while using instantaneous frequency resulted in similar performance to HTNet with spectral power.

In addition to outperforming the other decoders on average, we show in figures 2(D)–(F) that HTNet was the single best arm movement decoder for almost every individual participant in all conditions. HTNet performance was also consistently well-above random chance for all but two participants (EC10 same modality and EE13), much more than any other decoder. HTNet's consistently high accuracy demonstrates its ability to robustly generalize to a variety of participants with differences in electrode placement and signal quality.

#### 3.2. Fine-tuning generalized decoders

By fine-tuning these pretrained HTNet decoders using as few as 50 events from the held-out participant, HTNet can approach performances of the best tailored decoders trained using all available participant data. We tested fine-tuning each convolutional layer separately as well as re-training all layers of the network, as shown in figure 3(A). We found no significant differences in performance between our fine-tuning approaches, but we did find significant improvements in test accuracy when comparing fine-tuned decoders to pretrained decoders or tailored decoders trained on the same data, even when training on only 17% of the test participant's available events. For same modality (figure 3(B)), fine-tuning just the depthwise convolution significantly increased test accuracy compared to the pretrained models ( $p = 0.049$ ; Wilcoxon signed-rank test with false discovery rate correction). In the unseen modality condition (figure 3(E)), all fine-tuning approaches significantly increased test accuracy compared to the pretrained models (temporal convolution:  $p = 0.0054$ , depthwise convolution:  $p = 0.0094$ , separable convolution:  $p = 0.0054$ , all layers:  $p = 0.0054$ ). Additionally, fine-tuning the separable convolution and all HTNet layers resulted in significantly higher test accuracy than tailored HTNet decoders ( $p = 0.021$  and  $p = 0.015$ , respectively). No other comparisons were statistically significant.

Similarly, decoder performance for each participant shows that, in general, fine-tuning HTNet increases test accuracy. All tested fine-tuning approaches yielded similar performance (figures 3(C)

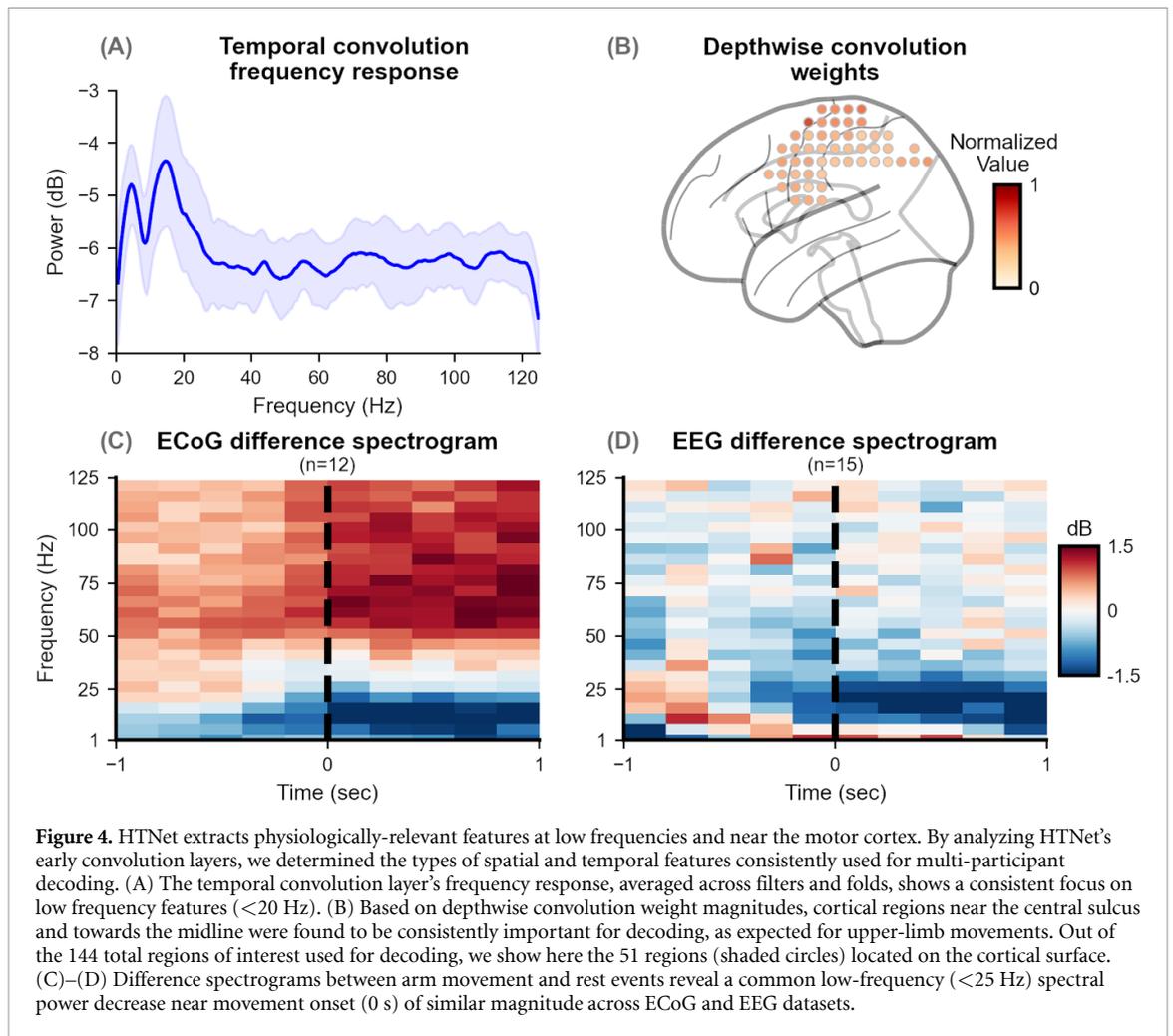


and (F)). In addition, computation times are consistent across fine-tuning approaches (tables S6, S7). By varying the number of events used (figures 3(D) and (G)), we find that fine-tuning approaches outperform randomly-initialized, tailored decoders when fewer than  $\sim 400$  ECoG or  $\sim 30$  EEG events from the test participant are available for training (see figure S3 for separated plots of each fine-tuning approach). In addition, fine-tuned HTNet decoders approach performances of the best tailored decoders with as few as 50 ECoG or 20 EEG events available for training. Finally, we demonstrated that we can fine-tune HTNet decoders to ECoG events after pretraining on EEG events (figure S4).

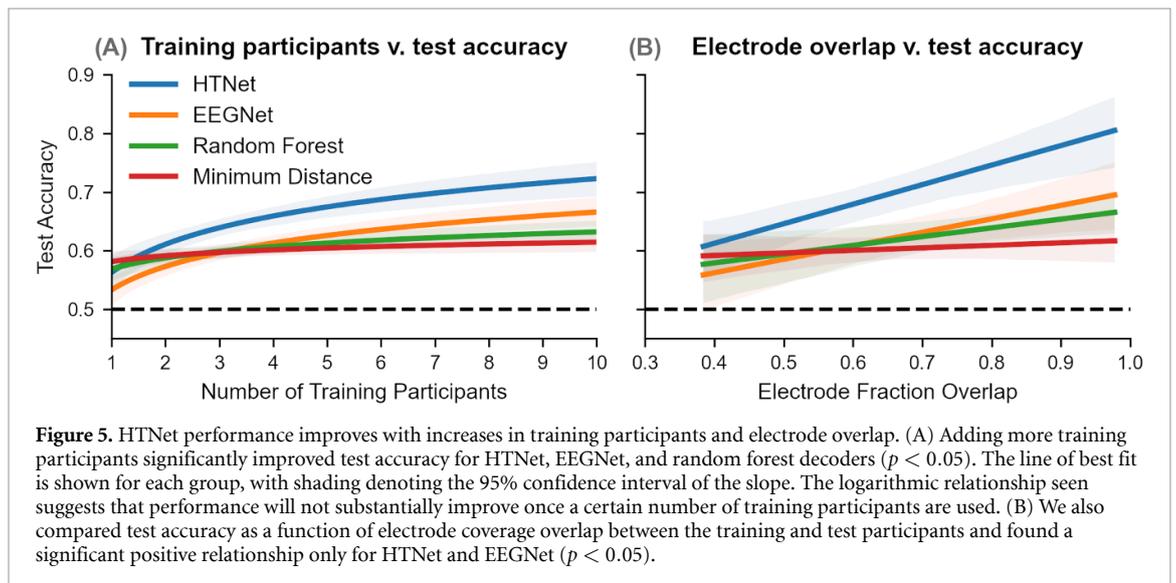
### 3.3. Interpreting network computations

Trained HTNet models achieve generalized decoding by extracting physiologically-relevant features, specifically at low frequencies and near the motor cortex (figure 4). HTNet's temporal and depthwise convolutional layers are interpretable, so we can probe these trained layers to determine the spatiotemporal features most often used for decoding, just like with EEGNet [55]. Here, we analyzed the trained layers of generalized HTNet decoders that were used for the same and unseen modality conditions.

Across decoders, we find a consistent emphasis on low-frequency ( $< 20$  Hz) temporal features



**Figure 4.** HTNet extracts physiologically-relevant features at low frequencies and near the motor cortex. By analyzing HTNet's early convolution layers, we determined the types of spatial and temporal features consistently used for multi-participant decoding. (A) The temporal convolution layer's frequency response, averaged across filters and folds, shows a consistent focus on low frequency features (<20 Hz). (B) Based on depthwise convolution weight magnitudes, cortical regions near the central sulcus and towards the midline were found to be consistently important for decoding, as expected for upper-limb movements. Out of the 144 total regions of interest used for decoding, we show here the 51 regions (shaded circles) located on the cortical surface. (C)–(D) Difference spectrograms between arm movement and rest events reveal a common low-frequency (<25 Hz) spectral power decrease near movement onset (0 s) of similar magnitude across ECoG and EEG datasets.



**Figure 5.** HTNet performance improves with increases in training participants and electrode overlap. (A) Adding more training participants significantly improved test accuracy for HTNet, EEGNet, and random forest decoders ( $p < 0.05$ ). The line of best fit is shown for each group, with shading denoting the 95% confidence interval of the slope. The logarithmic relationship seen suggests that performance will not substantially improve once a certain number of training participants are used. (B) We also compared test accuracy as a function of electrode coverage overlap between the training and test participants and found a significant positive relationship only for HTNet and EEGNet ( $p < 0.05$ ).

(figure 4(A)) when analyzing the average frequency response over temporal convolution filters. These decoders also frequently focused on cortical regions near the motor cortex (towards the central sulcus and midline), based on trained depthwise convolution weights (figure 4(B)). When we take the difference in spectral power between arm movement

and rest events near these motor cortical regions, we find, consistent with previous ECoG and EEG studies [76, 77], that low-frequency spectral power decreases during movement onset, with a similar magnitude for both ECoG and EEG data. This consistent magnitude in low-frequency power between recording modalities likely explains why

our trained HTNet models generalized to the EEG participants prior to any fine-tuning. It is important to note that the increase in ECoG high-frequency (>50 Hz) power, a pattern commonly linked to active movement [78, 79], was not present in the EEG dataset, likely due to temporal filtering by the skull [80].

Another crucial factor in HTNet performance is the spatial overlap of electrodes between the training and test participants, as well as how many participants were used for training (figure 5). We observe significant positive trends between electrode fraction overlap and test accuracy for HTNet (0.33 slope,  $p = 0.005$ ;  $t$ -test with false discovery rate correction) and EEGNet (0.23,  $p = 0.025$ ). We also find significant logarithmic relationships between the number of training participants and test accuracy for HTNet (0.07 slope,  $p = 2.77e-7$ ), EEGNet (0.06,  $p = 2.77e-7$ , and random forest (0.03,  $p = 0.003$ ). We initially tried a linear fit, but found that logarithmically scaling the number of training events resulted in a better fit. This logarithmic trend suggests that, at a certain point, adding more participants will not noticeably improve decoder performance. The number of training participants also affects computation time and number of epochs used during training (figure S5). Our findings suggest that HTNet is best able to incorporate information when many training participants are used, especially those with similar electrode placement to the test participant.

#### 4. Discussion

We demonstrated that HTNet can outperform state-of-the-art neural decoders when generalizing to new participants, even when a different recording modality is used. Developed as an extension of EEGNet [55], our key contributions to the neural network architecture are the addition of a Hilbert transform layer and a weight matrix to project individual electrode locations onto common brain regions. These trained HTNet decoders can be fine-tuned to a new test participant with fewer than 100 of the test participant's events and still decode almost as well as tailored decoders that have been trained on substantially more events. To achieve this performance, HTNet decoders consistently extract physiologically-relevant features, as revealed by our analysis of trained decoder weights.

To our knowledge, HTNet is the first decoder that can generalize and transfer its learning across both ECoG participants and different recording modalities. Previous studies have implemented decoders that can transfer across different EEG devices [14, 46, 81, 82] or leverage data from concurrent recording modalities [83, 84], but none of these decoders have demonstrated the ability to generalize to an entirely different recording modality. Yet, cross-modal generalization can be hampered by the limitations of each recording modality. For example,

the high-frequency power increase seen with ECoG is not present in the EEG dataset (figures 4(C) and (D)), demonstrating the infeasibility of transferring from EEG to ECoG when high-frequency features are necessary to decode well. As for generalizing across participants, many decoders can do this with EEG data [14], including EEGNet [55], but development of analogous ECoG decoders has been hindered by the high variation in electrode placement across ECoG patients. Similarly, several studies have implemented fine-tuning when decoding EEG or ECoG data [21, 50, 85–88], but these decoders were trained on either one participant or non-brain data, instead of multiple ECoG participants. By training on multiple participants' data, HTNet can generalize effectively to unseen participants and avoid overfitting to any one participant.

When we compared performance across different types of decoders, HTNet consistently outperformed EEGNet, even in the tailored condition. These performance differences between HTNet and EEGNet arise solely from HTNet's Hilbert transform layer because for all decoders, we projected onto common brain regions when training on multiple participants. We used the Hilbert transform to compute spectral power, which unlike the time-domain signal does not include a phase component. When we used just this phase component for decoding, we found that HTNet performance substantially worsened compared to using spectral power (see figure S2). This difference in accuracy suggests that decoding directly from the time-domain signal is suboptimal for our specific decoding task because the phase component is less informative than spectral power. Based on this insight, we would expect EEGNet to perform as well as or better than HTNet when decoding neural data with highly informative phase, such as event-related potentials [89, 90], depending on whether HTNet uses phase or spectral power to decode.

More broadly, HTNet demonstrates the value of integrating computational models, such as deep learning, with insights from neural signal processing. Fusing computational methods with scientific insight has inspired novel solutions that leverage the strengths of both approaches [45, 53, 91]. An alternative to our approach would have been to develop an end-to-end neural network model that simply learns to compute spectral power or project electrode-level signals onto a common grid. However, incorporating explicit transformations to spectral power and common cortical regions minimized the number of trainable parameters (and hence the amount of training data needed) and, like EEGNet, kept HTNet's initial layers interpretable. Furthermore, we were able to apply insights from previous ECoG/EEG research by explicitly baseline-subtracting spectral power within HTNet in order to generalize from ECoG to unseen EEG data [92, 93].

On the other hand, deep learning models can generate data-driven features that may be computationally expensive to obtain using other available methods. While many data-driven spatial filtering methods are available [17, 94], identifying frequencies with relevant spectral power often requires either a brute force search or applying techniques such as wavelet convolution to compute power at several frequencies, increasing the size of an already high-dimensional dataset [95, 96]. In contrast, HTNet converges quickly and provides a low-dimensional feature representation in the spectral domain. We believe that further embedding of neural signal processing into data-driven methods such as deep learning will continue to enhance the robustness and generalizability of future neural decoders.

Our study has several important limitations to consider. First, we classified two event types (arm movement versus rest), which is substantially less than the types of complex behaviors present in many neural decoding paradigms. Nonetheless, HTNet's architecture allows for decoding more than two types of events, and EEGNet has been shown to perform well when decoding among four types of behavior [55]. Still, decoding more complex behaviors than the current task may impact HTNet's generalizability. We chose ECoG and EEG datasets from a task with only two event types because the electrode positions were in a common coordinate system (MNI), which was essential when projecting data onto common brain regions for multi-participant decoding. Our analysis also may have been limited by projecting only onto 144 sensorimotor brain regions, which we chose based on the decoding task. For real-world decoders, the most informative regions may not be clear, requiring data-driven region-selection approaches to avoid high memory usage and slow computation times [37, 38]. Additionally, we could have constrained HTNet's temporal convolution layer to learn more meaningful narrow-band filters [97]. Finally, HTNet itself is limited by the assumption that spatiotemporal features are consistent across participants and the use of traditional spectral measures that often conflate periodic and aperiodic components of neural activity [98].

We are currently exploring extensions of HTNet for a variety of applications such as cross-frequency coupling [99, 100], long-term state decoding [6], cross-task decoding [101], and data-driven regression [102, 103]. In addition, other decoding measures could be substituted for the Hilbert transform, including non-Fourier methods [104, 105], and more complex interpolation schemes could be used to generate the projection matrix by incorporating participant-specific cortical anatomy [106, 107]. Besides ECoG and EEG, HTNet may also be useful for generalizing across participants with stereotactic EEG or local field potential recordings [108, 109]. Overall, HTNet provides a useful decoding framework

applicable across a variety of tasks and overcomes important obstacles towards developing robust, generalized neural decoders that can be fine-tuned with minimal data.

## Code and data availability

Our HTNet code is publicly available at: [https://github.com/BruntonUWBio/HTNet\\_generalized\\_decoding](https://github.com/BruntonUWBio/HTNet_generalized_decoding). The code in this repository can be used in conjunction with publicly available ECoG ([https://figshare.com/projects/Generalized\\_neural\\_decoders\\_for\\_transfer\\_learning\\_across\\_participants\\_and\\_recording\\_modalities/90287](https://figshare.com/projects/Generalized_neural_decoders_for_transfer_learning_across_participants_and_recording_modalities/90287)) and EEG [56] datasets to generate all of the main findings and figures from our study.

## Acknowledgments

We thank Satpreet Singh, Ellie Strandquist, Nancy Wang, Pierre Karashchuk, and Kurt Weaver for discussions and help with data analysis and model design, along with neurosurgeons Jeffrey G Ojemann, Andrew Ko, and the excellent staff at Harborview Hospital Neurosurgery department for their care of the patients during their monitoring. This research was supported by funding from the Defense Advanced Research Projects Agency (FA8750-18-2-0259), the National Science Foundation (1630178 and EEC-1028725), the Washington Research Foundation, and the Weill Neurohub.

## ORCID iDs

Steven M Peterson  <https://orcid.org/0000-0003-0782-5788>

Zoe Steine-Hanson  <https://orcid.org/0000-0001-9290-2885>

Rajesh P N Rao  <https://orcid.org/0000-0003-0682-8952>

Bingni W Brunton  <https://orcid.org/0000-0002-4831-3466>

## References

- [1] Ganzer P D *et al* 2020 Restoring the sense of touch using a sensorimotor demultiplexing neural interface *Cell* **181** 763–73.e12
- [2] Miller K J, Hermes D and Staff N P 2020 The current state of electrocorticography-based brain–computer interfaces *Neurosurgical Focus* **49** E2
- [3] Volkova K, Lebedev M A, Kaplan A and Ossadtchi A 2019 Decoding movement from electrocorticographic activity: a review *Frontiers Neuroinform.* **13** 74
- [4] Niketeghad S and Pouratian N 2019 Brain machine interfaces for vision restoration: the current state of cortical visual prosthetics *Neurotherapeutics* **16** 134–43
- [5] Martin S, Millán Jedel R, Knight R T and Pasley B N 2019 The use of intracranial recordings to decode human language: challenges and opportunities *Brain Lang.* **193** 73–83

- [6] Sani O G, Yang Y, Lee M B, Dawes H E, Chang E F and Shanechi M M 2018 Mood variations decoded from multi-site intracranial human brain activity *Nat. Biotechnol.* **36** 954–61
- [7] Wang W et al 2013 An electrocorticographic brain interface in an individual with tetraplegia *PLoS One* **8** e55344
- [8] Leuthardt E C, Schalk G, Wolpaw J R, Ojemann J G and Moran D W 2004 A brain–computer interface using electrocorticographic signals in humans *J. Neural Eng.* **1** 63–71
- [9] Degenhart A D, Bishop W E, Oby E R, Tyler-Kabara E C, Chase S M, Batista A P and Byron M Y 2020 Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity *Nat. Biomed. Eng.* **4** 672–85
- [10] Oby E R, Hennig J A, Batista A P, Yu Byron M and Chase S M 2020 Intracortical brain–machine interfaces *Neural Engineering* (Berlin: Springer) pp 185–221
- [11] Collinger J L, Gaunt R A and Schwartz A B 2018 Progress towards restoring upper limb movement and sensation through intracortical brain–computer interfaces *Curr. Opin. Biomed. Eng.* **8** 84–92
- [12] Gu X, Cao Z, Jolfaei A, Xu P, Wu D, Jung T-P, and Lin C-T 2020 EEG-based brain–computer interfaces (BCIS): a survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications (arXiv: 2001.11337)
- [13] Rao R P N 2013 *Brain-Computer Interfacing: An Introduction* (Cambridge: Cambridge University Press)
- [14] Wu D, Xu Y and Lu B-L 2020 Transfer learning for EEG-based brain–computer interfaces: a review of progress made since 2016 *IEEE Trans. Cogn. Devel. Syst.* **1**
- [15] Van Erp J, Lotte F and Tangermann M 2012 Brain–computer interfaces: beyond medical applications *Computer* **45** 26–34
- [16] Huang G, Liu G, Meng J, Zhang D and Zhu X 2010 Model based generalization analysis of common spatial pattern in brain computer interfaces *Cogn. Neurodyn.* **4** 217–23
- [17] Cohen M X 2014 *Analyzing Neural Time Series Data: Theory and Practice* (Cambridge, MA: MIT Press)
- [18] Tan C, Sun F, Kong T, Zhang W, Yang C and Liu C 2018 A survey on deep transfer learning *ICANN 2018* (Berlin: Springer) pp 270–79
- [19] Volker M, Schirrmester R T, Fiederer L D J, Burgard W and Ball T 2018 Deep transfer learning for error decoding from non-invasive EEG 2018 6th Int. Conf. Brain–Computer Interface (BCI) (IEEE) pp 1–6
- [20] Zubarev I, Zetter R, Halme H-L and Parkkonen L 2019 Adaptive neural network classifier for decoding meg signals *Neuroimage* **197** 425–34
- [21] Azab A M, Mihaylova L, Ang K K and Arvaneh M 2019 Weighted transfer learning for improving motor imagery-based brain–computer interface *IEEE Trans. Neural Sys. Rehabil. Eng.* **27** 1352–9
- [22] Brunner C et al 2015 Bnci horizon 2020: towards a roadmap for the BCI community *Brain-Computer Interfaces* **2** 1–10
- [23] Miller K J 2019 A library of human electrocorticographic data and analyses *Nat. Hum. Behav.* **3** 1225–35
- [24] Parvizi J and Kastner S 2018 Promises and limitations of human intracranial electroencephalography *Nat. Neurosci.* **21** 474–83
- [25] Takaura K, Tsuchiya N and Fujii N 2016 Frequency-dependent spatiotemporal profiles of visual responses recorded with subdural ECoG electrodes in awake monkeys: differences between high- and low-frequency activity *NeuroImage* **124** 557–72
- [26] Gunduz A, Brunner P, Daitch A, Leuthardt E C, Ritaccio A L, Pesaran B and Schalk G 2011 Neural correlates of visual-spatial attention in electrocorticographic signals in humans *Front. Hum. Neurosci.* **5** 89–99
- [27] Pistohl T, Ball T, Schulze-Bonhage A, Aertsen A and Mehring C 2008 Prediction of arm movement trajectories from ECoG-recordings in humans *J. Neurosci. Methods* **167** 105–14
- [28] Martin A B, Yang X, Saalman Y B, Wang L, Shestyuk A, Lin J J, Parvizi J, Knight R T and Kastner S 2019 Temporal dynamics and response modulation across the human visual system in a spatial attention task: an ECoG study *J. Neurosci.* **39** 333–52
- [29] Peterson S M and Ferris D P 2018 Differentiation in theta and beta electrocortical activity between visual and physical perturbations to walking and standing balance *eNeuro* **5** 4
- [30] Hell F, Taylor P C J, Mehrkens J H and Bötzel K 2018 Subthalamic stimulation, oscillatory activity and connectivity reveal functional role of STN and network mechanisms during decision making under conflict *Neuroimage* **171** 222–33
- [31] Jiang J, Bailey K and Xiao X 2018 Midfrontal theta and posterior parietal alpha band oscillations support conflict resolution in a masked affective priming task *Front. Hum. Neurosci.* **12** 175
- [32] Zavala B, Tan H, Ashkan K, Foltyniec T, Limousin P, Zrinzo L, Zaghoul K and Brown P 2016 Human subthalamic nucleus–medial frontal cortex theta phase coherence is involved in conflict and error related cortical monitoring *Neuroimage* **137** 178–87
- [33] Fujioka T, Trainor L J, Large E W and Ross B 2012 Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations *J. Neurosci.* **32** 1791–802
- [34] Schalk G and Leuthardt E C 2011 Brain–computer interfaces using electrocorticographic signals *IEEE Rev. Biomed. Eng.* **4** 140–54
- [35] Onton J, Westerfield M, Townsend J and Makeig S 2006 Imaging human EEG dynamics using independent component analysis *Neurosci. Biobehav. Rev.* **30** 808–22
- [36] Bigdely-Shamlo N, Mullen T, Kreutz-Delgado K and Makeig S 2013 Measure projection analysis: a probabilistic approach to EEG source comparison and multi-subject inference *NeuroImage* **72** 287–303
- [37] Peterson S M, Rios E and Ferris D P 2018 Transient visual perturbations boost short-term balance learning in virtual reality by modulating electrocortical activity *J. Neurophysiol.* **120** 1998–2010
- [38] Peterson S M, Singh S H, Wang N X R, Rao R P N and Brunton B W 2020 Behavioral and neural variability of naturalistic arm movements (<https://doi.org/10.1101/2020.04.17.047357>)
- [39] Saha S and Baumert M 2019 Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: a review *Frontiers Comput. Neurosci.* **13** 87
- [40] Krumpke T, Baumgaertner K, Rosenstiel W and Spüler M 2017 Non-stationarity and inter-subject variability of EEG characteristics in the context of BCI development 7th Graz Brain-Computer Interface Conference p 48
- [41] Corcoran A W, Alday P M, Schlesewsky M and Bornkessel-Schlesewsky I 2018 Toward a reliable, automated method of individual alpha frequency (IAF) quantification *Psychophysiology* **55** e13064
- [42] Adam A, Shapiaz M I, Tumari M Z M, Mohamad M S and Mubin M 2014 Feature selection and classifier parameters estimation for EEG signals peak detection using particle swarm optimization *Sci. World J.* **2014** 973063
- [43] Chiang A K I, Rennie C J, Robinson P A, Roberts J A, Rigozzi M K, Whitehouse R W, Hamilton R J and Gordon E 2008 Automated characterization of multiple alpha peaks in multi-site electroencephalograms *J. Neurosci. Methods* **168** 396–411
- [44] Cohen M X 2020 A data-driven method to identify frequency boundaries in multichannel electrophysiology data *J. Neurosci. Methods* **347** 108949

- [45] Anumanchipalli G K, Chartier J and Chang E F 2019 Speech synthesis from neural decoding of spoken sentences *Nature* **568** 493–8
- [46] Xu L, Xu M, Ke Y, An X, Liu S and Ming D 2020 Cross-dataset variability problem in EEG decoding with deep learning *Frontiers Hum. Neurosci.* **14** 103
- [47] Ming Y, Ding W, Pelusi D, Wu D, Wang Y-K, Prasad M and Lin C-T 2019 Subject adaptation network for EEG data analysis *Appl. Soft Comput.* **84** 105689
- [48] Roy Y, Banville H, Albuquerque I, Gramfort A, Falk T H and Faubert J 2019 Deep learning-based electroencephalography analysis: a systematic review *J. Neural Eng.* **16** 051001
- [49] Zhang X, Yao L, Wang X, Monaghan J, Mcalpine D, and Zhang Y 2019 A survey on deep learning based brain computer interface: recent advances and new frontiers (arXiv: 1905.04149)
- [50] Makin J G, Moses D A and Chang E F 2020 Machine translation of cortical activity to text with an encoder–decoder framework *Nat. Neurosci.* **23** 575–82
- [51] Behncke J, Schirrmeyer R T, Volker M, Hammer J, Marusic P, Schulze-Bonhage A, Burgard W and Ball T 2018 Cross-paradigm pretraining of convolutional networks improves intracranial EEG decoding 2018 *IEEE Int. Conf. Systems, Man and Cybernetics (SMC)* (IEEE) pp 1046–53
- [52] Sakhavi S, Guan C and Yan S 2018 Learning temporal information for brain-computer interface using convolutional neural networks *IEEE Trans. Neural Netw. Learn. Syst.* **29** 5619–29
- [53] Schirrmeyer R T, Springenberg J T, Fiederer L D J, Glasstetter M, Eggenberger K, Tangermann M, Hutter F, Burgard W and Ball T 2017 Deep learning with convolutional neural networks for EEG decoding and visualization *Hum. Brain Mapp.* **38** 5391–420
- [54] Bashivan P, Rish I, Yeasin M and Codella N 2015 Learning representations from EEG with deep recurrent-convolutional neural networks (arXiv: 1511.06448)
- [55] Lawhern V J, Solon A J, Waytowich N R, Gordon S M, Hung C P and Lance B J 2018 Eegnet: a compact convolutional neural network for EEG-based brain–computer interfaces *J. Neural Eng.* **15** 056013
- [56] Ofner P, Schwarz A, Pereira J and Müller-Putz G R 2017 Upper limb movements can be decoded from the time-domain of low-frequency EEG *PLoS One* **12** 1–24
- [57] Singh S H, Peterson S M, Rao R P N and Brunton B W 2020 Towards naturalistic human neuroscience and neuroengineering: behavior mining in long-term video and neural recordings (arXiv: 2001.08349)
- [58] Gramfort A et al 2013 MEG and EEG data analysis with MNE-python *Frontiers Neurosci.* **7** 267
- [59] Stolk A et al 2018 Integrated analysis of anatomical and electrophysiological human intracranial data *Nat. Protoc.* **13** 1699–723
- [60] Oostenveld R, Fries P, Maris E and Schoffelen J-M 2011 Fieldtrip: open source software for advanced analysis of MEG, EEG and invasive electrophysiological data *Comput. Intell. Neurosci.* **2011** 156869
- [61] Fonov V, Evans A C, Botteron K, Robert Almli C, McKinstry R C and Louis Collins D 2011 Unbiased average age-appropriate atlases for pediatric studies *NeuroImage* **54** 313–27
- [62] Oostenveld R and Praamstra P 2001 The five percent electrode system for high-resolution EEG and ERP measurements *Clin. Neurophysiol.* **112** 713–9
- [63] Vert J-P, Tsuda K and Schölkopf B 2004 A Primer on Kernel Methods *Kernel Methods Comput. Biol.* (Cambridge, MA: The MIT Press) 35–70
- [64] Delorme A and Makeig S 2004 Eeglab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis *J. Neurosci. Methods* **134** 9–21
- [65] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B and Joliot M 2002 Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain *NeuroImage* **15** 273–89
- [66] Van Loan C 1992 *Computational Frameworks for the Fast Fourier Transform* (Philadelphia, PA: SIAM)
- [67] Congedo M, Barachant A and Bhatia R 2017 Riemannian geometry for EEG-based brain–computer interfaces; a primer and a review *Brain-Computer Interfaces* **4** 155–74
- [68] Yger F, Berar M and Lotte F 2016 Riemannian approaches in brain–computer interfaces: a review *IEEE Trans. Neural Sys. Rehabil. Eng.* **25** 1753–62
- [69] Benjamini Y and Hochberg Y 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing *J. R. Stat. Soc. Ser. B (Methodological)* **57** 289–300
- [70] Akiba T, Sano S, Yanase T, Ohta T and Koyama M 2019 Optuna: a next-generation hyperparameter optimization framework *Proc. 25th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining* 2623–631
- [71] Bergstra J, Yamins D and Cox D 2013 Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures *Int. Conf. Machine Learning* pp 115–23 (<https://proceedings.mlr.press/v28/bergstra13.html>)
- [72] Bergstra J S, Bardenet Remi, Bengio Y and Kégl Bazs 2011 Algorithms for hyper-parameter optimization *Adv. Neural Inf. Process. Syst.* pp 2546–54
- [73] Yosinski J, Clune J, Bengio Y and Lipson H 2014 How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* pp 3320–8
- [74] Cohen M X 07 2014 Fluctuations in oscillation frequency control spike timing and coordinate neural networks *J. Neurosci.* **34** 8988
- [75] Rao Yarlagadda R K 2010 *Analog and Digital Signals and Systems* vol 1 (Berlin: Springer)
- [76] Chung J W, Ofori E, Misra G, Hess C W and Vaillancourt D E 2017 Beta-band activity and connectivity in sensorimotor and parietal cortex are important for accurate motor performance *NeuroImage* **144** 164–73
- [77] Miller K J, Leuthardt E C, Schalk G, Rao R P N, Anderson N R, Moran D W, Miller J W and Ojemann J G 2007 Spectral changes in cortical surface potentials during motor movement *J. Neurosci.* **27** 2424–32
- [78] Nowak M, Zich C and Stagg C J 2018 Motor cortical gamma oscillations: what have we learnt and where are we headed? *Curr. Behav. Neurosci. Rep.* **5** 136–42
- [79] Başar E, Başar-Eroglu C, Karakaş S and Schürmann M 2001 Gamma, alpha, delta and theta oscillations govern cognitive processes *Int. J. Psychophysiol.* **39** 241–8
- [80] Srinivasan R, Tucker D M and Murias M 1998 Estimating the spatial Nyquist of the human EEG *Behav. Res. Methods Instrum. Comput.* **30** 8–19
- [81] Wu D, Lawhern V J, David Hairston W and Lance B J 2016 Switching EEG headsets made easy: reducing offline calibration effort using active weighted adaptation regularization *IEEE Trans. Neural Syst. Rehabil. Eng.* **24** 1125–37
- [82] Nakanishi M, Wang Y-T, Wei C-S, Chiang K-J and Jung T-P 2020 Facilitating calibration in high-speed BCI spellers via leveraging cross-device shared latent responses *IEEE Trans. Biomed. Eng.* **67** 1105–13
- [83] Muraskin J, Brown T R, Walz J M, Tu T, Conroy B, Goldman R I and Sajda P 2018 A multimodal encoding model applied to imaging decision-related neural cascades in the human brain *NeuroImage* **180** 211–22
- [84] Fatima S and Kamboh A M 2017 Decoding brain cognitive activity across subjects using multimodal m/eeg neuroimaging 2017 *39th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE) pp 3224–7

- [85] Elango V, Patel A N, Miller K J and Gilja V 2017 Sequence transfer learning for neural decoding (<https://doi.org/10.1101/210732>)
- [86] Wang R, Chen X, Khalilian-Gourtani A, Chen Z, Yu L, Flinker A and Wang Y 2020 Stimulus speech decoding from human cortex with generative adversarial network transfer learning 2020 IEEE 17th Int. Symp. Biomedical Imaging (ISBI) (IEEE) pp 390–4
- [87] Desai S A, Tcheng T and Morrell M 2019 Transfer-learning for differentiating epileptic patients who respond to treatment based on chronic ambulatory ECoG data 2019 9th Int. IEEE/Conf. Neural Engineering (NER) (IEEE) pp 1–4
- [88] Uran A, Van Gemeren C, van Diepen R, Chavarrriaga R and Millán Je del R 2019 Applying transfer learning to deep learned models for EEG analysis (arXiv: 1907.01332)
- [89] Aydarkhanov R A, Uscumlic M, Chavarrriaga R, Gheorge L and Millan J D R 2020 Spatial covariance improves BCI performance for late ERPs components with high temporal variability *J. Neural Eng.* (<https://doi.org/10.1088/1741-2552/ab95eb>)
- [90] Kappenman E S and Luck S J 2016 Best practices for event-related potential research in clinical populations *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging* **1** 110–5
- [91] Xie Z, Schwartz O and Prasad A 2018 Decoding of finger trajectory from ECoG using deep learning *J. Neural Eng.* **15** 036009
- [92] Hu Li, Xiao P, Zhang Z G, Mouraux Ae and Iannetti G D 2014 Single-trial time–frequency analysis of electrocortical signals: baseline correction and beyond *Neuroimage* **84** 876–87
- [93] Parra L C, Spence C D, Gerson A D and Sajda P 2005 Recipes for the linear analysis of EEG *Neuroimage* **28** 326–41
- [94] Makeig S, Bell A J, Jung T-P and Sejnowski T J 1996 Independent component analysis of electroencephalographic data *Advances in Neural Information Processing Systems* pp 145–51
- [95] Lee H K and Choi Y-S 2019 Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain–computer interface *Entropy* **21** 1199
- [96] Tayeb Z, Fedjaev J, Ghaboosi N, Richter C, Everding L, Qu X, Wu Y, Cheng G and Conradt Jorg 2019 Validating deep neural networks for online decoding of motor imagery movements from EEG signals *Sensors* **19** 210
- [97] Ravanelli M and Bengio Y 2018 Interpretable convolutional filters with sincnet (arXiv:1811.09725)
- [98] Donoghue T et al 2020 Parameterizing neural power spectra into periodic and aperiodic components *Nat. Neurosci.* **23** 1655–65
- [99] van der Meij R, Kahana M and Maris E 2012 Phase-amplitude coupling in human electrocorticography is spatially distributed and phase diverse *J. Neurosci.* **32** 111–23
- [100] Canolty R T and Knight R T 2010 The functional role of cross-frequency coupling *Trends Cogn. Sci.* **14** 506–15
- [101] He H and Wu D 2020 Different set domain adaptation for brain– interfaces: a label alignment approach *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** 1091–108
- [102] Cui Y, Xu Y and Wu. D 2019 EEG-based driver drowsiness estimation using feature weighted episodic training *IEEE Trans. Neural Syst. Rehabil. Eng.* **27** 2263–73
- [103] Wu D, Lance B J, Lawhern V J, Gordon S, Jung T-P and Lin C-T 2017 EEG-based user reaction time estimation using Riemannian geometry features *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** 2157–68
- [104] Cole S and Voytek B 2019 Cycle-by-cycle analysis of neural oscillations *J. Neurophysiol.* **122** 849–61
- [105] Cole S R and Voytek B 2017 Brain oscillations and the importance of waveform shape *Trends Cogn. Sci.* **21** 137–49
- [106] Owen L L W, Muntianu T A, Heusser A C, Daly P M, Scangos K W and Manning J R 2020 A Gaussian process model of human electrocorticographic data *Cerebral Cortex* (<https://doi.org/10.1093/cercor/bhaa115>)
- [107] Vermaas M, Piastra M C, Oostendorp T F, Ramsey N F and Tiesinga P H E 2020 Femfuns: a volume conduction modeling pipeline that includes resistive, capacitive or dispersive tissue and electrodes *Neuroinformatics* **18** 569–80
- [108] Herff C, Krusienski D J and Kubben P 2020 The potential of stereotactic-EEG for brain–computer interfaces: current progress and future directions *Frontiers Neurosci.* **14** 123
- [109] Jackson A and Hall T M 2016 Decoding local field potentials for neural interfaces *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** 1705–14