Predicting Human Grip Force from EEG, fNIRS and EMG Data

Oliver Shetler

Abstract
Brain computer interfacing (BCI) will become the bedrock of the next generation of computer user interfaces. Early commercial BCIs will use portable, non-invasive brain measurement techniques. One major problem for the development of wearable non-invasive BCIs is that existing light-weight measurement modalities have severe temporal and spatial resolution limitations. In general, non-invasive BCIs have only been successfully used for classification tasks with only a few target categories. The HYGRIP data set[1] was developed by a team of scientists at the Imperial College, London in order to facilitate the development of predictive models for continuous muscular response. This paper lays the ground-work for developing multimodal neural models by building a series of hierarchically integrated unimodal neural models, which can be integrated into multimodal models. Future work will leverage cloud APIs to build multimodal models and test whether they can improve on a benchmark LSTM model that uses force alone.

Keywords
Neural Data — Brain Computer Interfacing, BCI — Deep Learning

[1] *Department of Applied Mathematics and Statistics, CUNY—Hunter College, New York, NY*
*__Corresponding author__: oliver.shetler06@myhunter.cuny.edu.com

Contents

Introduction

Brain computer interfacing (BCI) will become the bedrock of the next generation of computer user interfaces. Early commercial BCIs will use portable, non-invasive brain measurement techniques (techniques that do not require the penetration of skin or skull). One major problem for the development of wearable non-invasive BCIs is that existing light-weight measurement modalities such as electroencephalography (EEG), magnetoencephalography (MEG), and functional near infrared spectroscopy (fNIRS) have severe spatial (EEG / MEG) or temporal (fNIRS) resolution limitations. As a consequence, extracting relevant neural data with one method alone can be difficult or impossible for certain tasks.

In general, non-invasive BCIs have only been successfully used for classification tasks with few distinctions—often between only two broad categories such as ”focused” and ”distracted.” One major open problem for non-invasive BCIs is the estimation of continuous quantities such as muscular motor signals, or spectrum of mental states. The development of continuous quantity predictive models will play an important role in the development of robotic prosthetics, exoskeletons and virtual / augmented reality user interfaces.

In recent years, researchers have begun to explore the possibility of improving non-invasive BCIs by using multiple measurement modalities in one system. For example, combining EEG and fNIRS can potentially compensate for the respectively low spatial and temporal resolutions of each modality alone. The Hybrid Dynamic Grip (HYGRIP) data set—discussed in more detail below—was constructed with this idea in mind. It includes EEG and fNIRS data from the scalp region above the motor cortices of several people who participated in a force-feedback gripping task.

This paper lays the groundwork for future work the multimodal modeling of the HYGRIP data set by developing a series of hierarchical unimodal models, and assessing their respective performances. Unfortunately, my machine is not equipped to merge the unimodal models into multimodal models. As a consequence, that project will have to be implemented on a paid cloud API.

1. Data

The HYGRIP data set[1] was developed by a team of scientists at the Imperial College, London in order to facilitate the development of predictive models for continuous muscular response. [1]. This open problem is of particular interest to researchers because it is a first step towards several useful applications of portable non-invasive BCIs. First, the anticipation of motor movements is a central problem for the development of useful exoskeletons for mobility assistance and strength augmentation. When an exoskeleton fails to anticipate a user's movements quickly enough, using the device can feel like moving through sand. Current exoskeletons generally use historical force-feedback data, but neural data could augment their performance. Moreover, motor-neural data could enable mobility impaired people to use exoskeletons, or other prostheses by transfer learning from neural-motor predictive models.

1.1 Motor Control Task

14 right handed participants[2] were asked to squeeze a hand grip 10 times for each trial at approximately 1Hz. Participants performed either 10 or 13 trials with each hand (total of 20 or 26). For each trial, a square-wave tone (on/squeeze 1.55s; off/relax 0.55s) was used to indicate when the participant should squeeze and when the participants should relax. Participants were asked to squeeze at between 25% and 50% of their maximum capacity, which was calibrated at the start of the sequence of trials with a single maximum strength squeeze. Visual feedback on force was given by a red circle on a computer screen. There was a solid red circle and a dotted white circle, both centered at the same point. When the squeeze force was inside the target range, the solid circle was the same size as the dotted circle. When the force was below the target range, the solid circle shrank, and when force exceeded the target range, the circle grew.

1.2 Multimodal Sensors

EEG and fNIRS detectors were symmetrically placed over the left and right motor cortices of each participant. EMG data were collected from 4 electrodes placed on the left or right arm. Force data were collected using a handle with a grip force sensor. Two potential confounders were recorded as well. First, breathing was recorded using a chest-band sensor (breathing rate can impact overall blood oxygenation, measured by fNIRS). Second, electrooculography (EOG)[3] data were recorded to measure eye blinks and other facial movements that might confound the EEG signals.

1.3 Data Preparation

Force. The force data was sampled at a rate of 50 Hz in the published data set. Each force signal was band-pass filtered

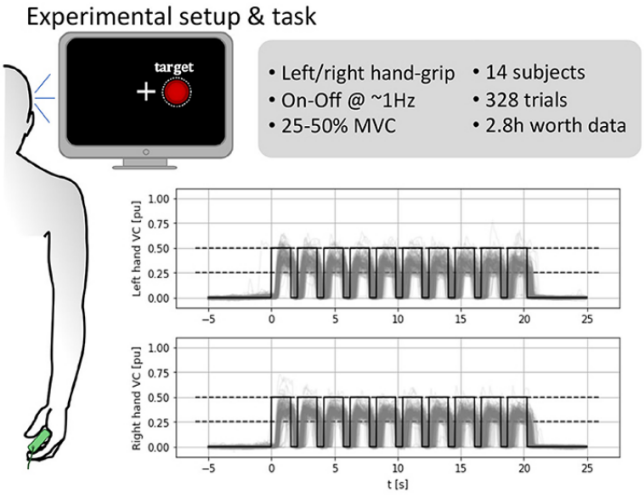

Figure 1. The graphs are of the force variable for trials for left and right hands overlayed with the target range (dotted lines) and the tone (portrayed as a square wave with 0 and 0.5 as its levels). The cartoon portrays a subject holding the gripper while attending to the feedback on the screen.

at frequencies between 10^{-4} and 9 Hz using a second order elliptical filter. Each signal was then divided up into 30-second trials and band-passed filtered again for frequencies above 10^{-3}. Then, the resulting signal was normalized against the grip calibration by dividing all the trial data by the peak calibration grip force value.[1] The resulting data were all within the unit interval, and could be interpreted as a fraction of an individual's grip strength.[4]

Finally, each trial was further divided into 3-second epochs. For each trial, the first 9 epochs were retained as co-variates, and epochs 2-10 were retained as a response variable.

fNIRS. The fNIRS data were sampled at 12.5 Hz in the published data set. The optical intensity, $\widehat{I}_{ij}^{\lambda}$ for each wavelength, λ, was low-pass filtered for frequencies below 0.25 Hz with a 7 th order elliptical filter. Changes in optical densities for each wavelength, $\Delta\mathrm{OD}_{ij}^{\lambda}(t)$, were found with the following formula,

$$\Delta\mathrm{OD}_{ij}^{\lambda}(t) = -\log\left(\widehat{\mathrm{I}}_{ij}^{\lambda}(t)/\bar{\mathrm{I}}_{ij}^{\lambda}\right)$$

where i and j are the indices of valid pairs of sensors (i) and detectors (j), and t is the time and $\bar{\mathrm{I}}_{ij}^{\lambda}$ is the average of the optical intensity 1 second before to the "start" event. Changes in oxygenated and deoxygenated hemoglobin concentrations (ΔHbO and ΔHbR, respectively) were computed solving the modified Beer-Lambert law[3][1],

$$\Delta\mathrm{OD}_{ij}^{\lambda} = \mathrm{L}_{ij}^{\lambda}\mathrm{DPF}^{\lambda}\left(\varepsilon_{\mathrm{HbR}}^{\lambda}\Delta\mathrm{HbR} + \varepsilon_{\mathrm{HbO}}^{\lambda}\Delta\mathrm{HbO}\right)$$

[1] At the time of this draft, the data have only been used to classify the handedness of trials via the fNIRS sub-data-set.[2]

[2] The handedness of each participant was confirmed using the Edinburgh inventory.[**?**]

[3] EOG are electrodes placed on the forehead over the eyes.

[4] The above procedure was conducted using the code that came with the data.

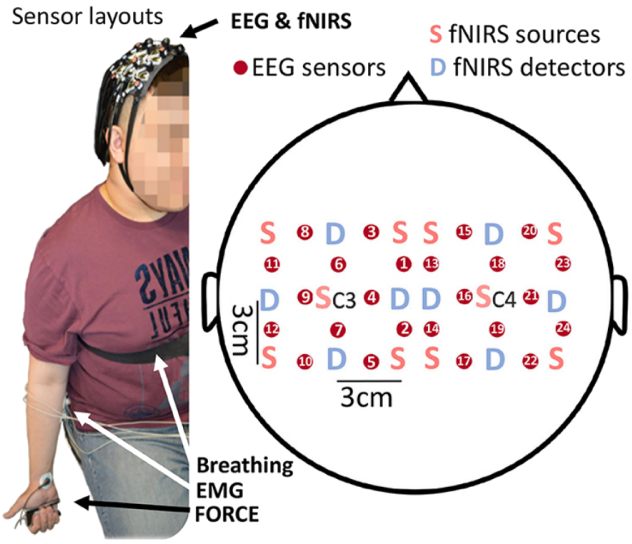

Figure 2. EEG and fNIRS sensors were placed on each participant's scalp near the C3 and C4 regions according to the 10-20 system, in order to access the left and right motor cortices. The EEG, fNIRS EMG, breathing and Force detectors can be seen on a subject to the left.

Blood oxygenation (HbO) increased during the task, and decreased after, showing a gross correspondence between the fNIRS data, and the gripping task.[5]

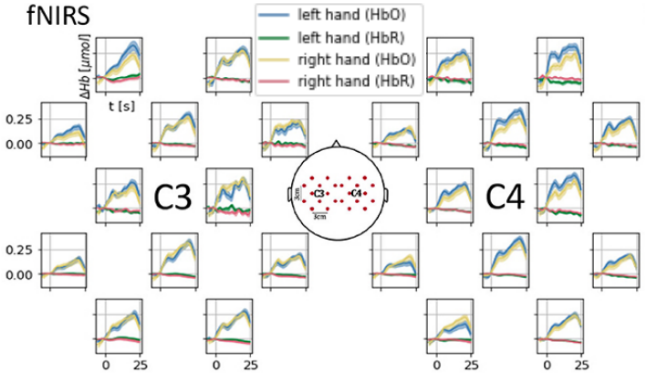

Figure 3. Blood Oxygenation (HbO) and Deoxygenation (HbR) for each trial are plotted overlayed and plotted against the 30 second trial interval ($t = -10$ to $t = 30$).

The fNIRS trials were each further sub-divided into 9 3-second long epochs for use in predictive models (the first 9 epochs were taken from the data after it was divided into 10 epochs).

EEG The EEG data were sampled at a rate of 125 Hz in the published data set. Notch filters were used at the frequency of the mains, 50 Hz, and at the frequency of the fNIRS array, 12.5 Hz. An Independent Components Analysis was conducted and used to remove components correlated with EOG data. Components exceeding an individual correlation value of 0.3 were removed; only one component met the criteria. The procedure was then repeated and components correlated with EMG data were also removed. The removal criteria was stricter for EMG. Any component with a component with a correlation above 10^{-4} was rejected.[1] The data were divided into 30 second epochs.[6] The data were then further sub-divided into 3-second epochs (for each epoch, the first 9 were retained). Then each epoch was further sub-divided into 10 time-steps for input into Long Short Term Memory (LSTM) in deep learning models.

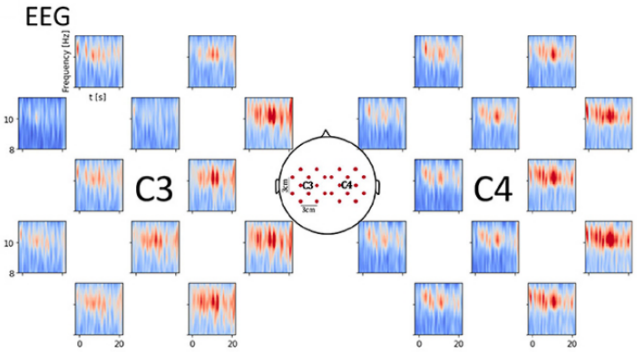

Figure 4. Spectral plots of aggregated EEG data from each channel are shown in place according to the 10-20 system placements.

EMG The EMG signals were high-pass filtered with a 17th order Butterworth filter of above 110 Hz. Then, the Hilbert envelope was computed, and the mean power was taken. The published data set was sampled at a rate of 125Hz. EMG signals were broken into 30 second long trial intervals using the code provided by the authors. Then the trial intervals were further broken into 3 second epochs, the first 9 of which were retained for model building. Then, the three second epochs were broken into 10 time steps for input into LSTM blocks.

2. Analysis

2.1 Data Set

The data were divided into a training set, a validation set and a test set, segregated by subject. In order to reduce the chances of over-fitting, the divisions were made without overlap between subjects. Subjects A-J were used to generate the test set, which was 2016 epochs long. Subjects K and L were used to generate the validation set, which was 468 epochs long. Subjects M and N were used to generate the test set which was also 486 epochs long. The data for each modality in each data set was prepared according to the process specified in the respective subsection of 1.3.

The goal of each model was to predict 3 seconds of force data from the previous 3 seconds of whichever covariates were used.

[5]The filtering and plotting of the trials was done using code provided by the developers of the HYGRIP data set[1]

[6]The above procedure was conducted using the code provided by the authors of the data set.

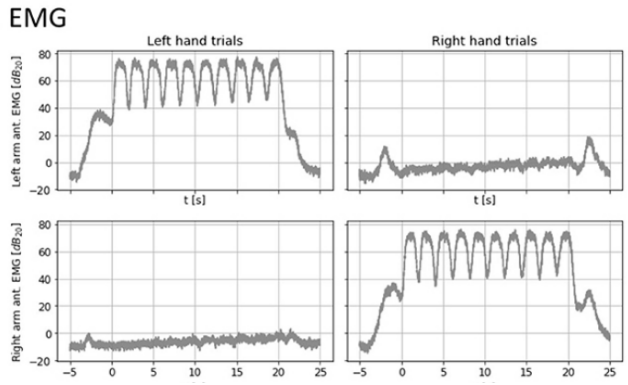

Figure 5. Decibles of EMG power density from each trial are overlayed and are shown against time. The time $t = 0$ denotes the beginning of the active gripping task. They were computed using the Hilbert envelope of the signal.

2.2 Benchmark Model

As a benchmark, an LSTM model with one hidden layer was used on force history to predict future force data. This model was considered to be a reasonable bench mark because it is a simple time-series, with no neural data. This is a simple version of the sort of model a modern exoskeleton might use to anticipate various pressures on its actuators, without neural data.

Force LSTM The Force LSTM model took had a test MSE of 0.01058, lower than the other models tested.

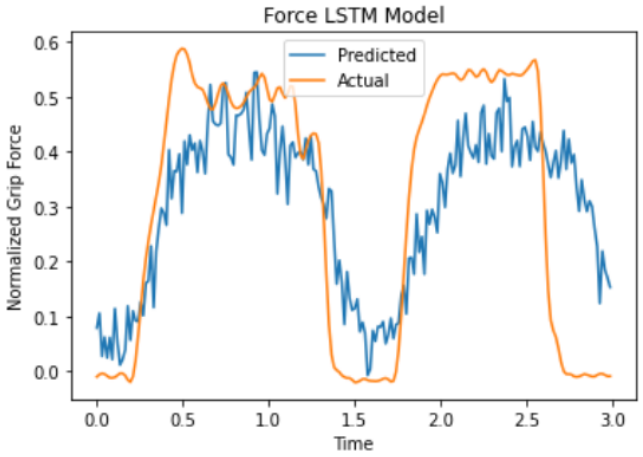

Figure 6. The force LSTM model predictions plotted next to the .

2.3 Hierarchical Models

A series of models were built, trained and merged into increasingly complex hierarchical models. First, each channel of each modality (force, fNIRS, EEG and EMG) was fed into an appropriate shallow single channel model. Each channel model was trained for 30 epochs, each of which consisted of 504 data points (there was redundant sampling, so over-fitting had to be addressed separately). After training, each channel model was locked to reduce the chances of aggregate over-fitting. (Additionally, regularization was employed deeper into the model). Next, the top hidden layer of each channel model was concatinated with the top hidden layer of other channel models to create unimodal hierarchical models (discussed more below). These models were each trained for 10 epochs with the same epoch size, and then frozen.

All layers used the rectified linear unit activation function. All models were trained using the Nadam optimize on the Mean Squared Error.

2.3.1 Channel Models

fNIRS The fNIRS epochs were 50 sample points long. Each channel had an input for oxygenated and de-oxygenated blood signals. Each of these 48 inputs was connected to a 50 unit long Dense layer. These two hidden layers were then concatenated, and fed to another Dense layer with 50 neurons. The model was then fed to a 200 unit long Dense output layer. Each of the 24 channel models was trained against the forward-time-displaced force signal.

Table 1. fNIRS Channel Model Evaluation Performance

Model		
Modality	Channel	MSE
fNIRS	1	0.03956
fNIRS	2	0.03919
fNIRS	3	0.04020
fNIRS	4	0.04065
fNIRS	5	0.03984
fNIRS	6	0.04050
fNIRS	6	0.03951
fNIRS	8	0.03977
fNIRS	9	0.03879
fNIRS	10	0.04123
fNIRS	11	0.03921
fNIRS	12	0.04007
fNIRS	13	0.02627
fNIRS	14	0.02631
fNIRS	15	0.02629
fNIRS	16	0.02625
fNIRS	17	0.02633
fNIRS	18	0.02632
fNIRS	19	0.02629
fNIRS	20	0.02631
fNIRS	21	0.02627
fNIRS	22	0.02625
fNIRS	23	0.02632
fNIRS	24	0.02624

EEG The EEG epochs consisted of 10 time steps, each 200 sample points long. Each of the 24 EEG channel was fed into an LSTM layer with 200 output dimensions. This hidden layer was fed to a 200 unit long Dense layer. Each channel

model was trained against the forward-time-displaced force signal.

Table 2. EEG Channel Evaluation Performance

Model		
Modality	Channel	MSE
EEG	1	0.03137
EEG	2	0.03592
EEG	3	0.03033
EEG	4	0.02940
EEG	5	0.02919
EEG	6	0.02967
EEG	7	0.02836
EEG	8	0.02845
EEG	9	0.03482
EEG	10	0.03177
EEG	11	0.02859
EEG	12	0.02858
EEG	13	0.03015
EEG	14	0.03183
EEG	15	0.03129
EEG	16	0.03212
EEG	17	0.03099
EEG	18	0.02895
EEG	19	0.03020
EEG	20	0.02948
EEG	21	0.02905
EEG	22	0.03142
EEG	23	0.02990
EEG	24	0.02937

EMG The EMG epochs consisted of 10 time steps, each 200 sample points long. Each of the 4 EMG channel was fed into an LSTM layer with 200 output dimensions. This hidden layer was fed to a 200 unit long Dense layer. Each channel model was trained against the forward-time-displaced force signal.

Table 3. EMG Channel Model Evaluation Performance

Model		
Modality	Channel	MSE
EMG	1	0.02656
EMG	2	0.02549
EMG	3	0.02535
EMG	4	0.02554

2.3.2 Unimodal Models

fNIRS C3 (Left Hemisphere) The twelve top hidden layers in C3 (channels 1-12) were concatenated (dim = 600) and fed into a 400 dimension Dense layer. This was fed into a 200 dimension Dense output layer.

fNIRS C4 (Right Hemisphere) The C4 fNIRS model was structured the same as the C3 model, but channels 13-24 were fed into it.

EEG C3 The twelve channels in the C3 region were broken into four groups of 3 (1-3, 4-6, 6-8, and 8-12). The top hidden layer from each channel model was concatenated with its respective group members into 600 dimension layers. Each of these was fed into a 200 dimension Dense layer, and then to a 200 dimension output layer. The sub-models were trained for 4 epochs. Then, the top hidden layers of the 4 sub-models were concatenated into a 800 dimension layer and fed to a 200 dimension Dense layer. This was fed to a batch normalization layer, and this was fed to a 400 dimension Dense hidden layer. This was fed to a 200 dimensional output layer and trained for 10 epochs.

EEG C4 The C4 EEG model was constructed in the same fashion as the C3 model, except with channels 13-24.

EMG The hidden layers from each of the 4 EMG channels were concatenated into a 2000 dimension layer. This was fed into a 1000 dimension Dense layer, which was fed to a batch normalization layer, which was fed to a 400 dimension Dense layer. This was fed to a 200 dimensional output layer and trained for

Table 4. Unimodal Model Evaluation Performance

Model		
Modality	Region	MSE
fNIRS	C3	0.02639
fNIRS	C4	0.02632
EEG	C3	0.02748
EEG	C4	0.02841
EMG	NA	0.0291

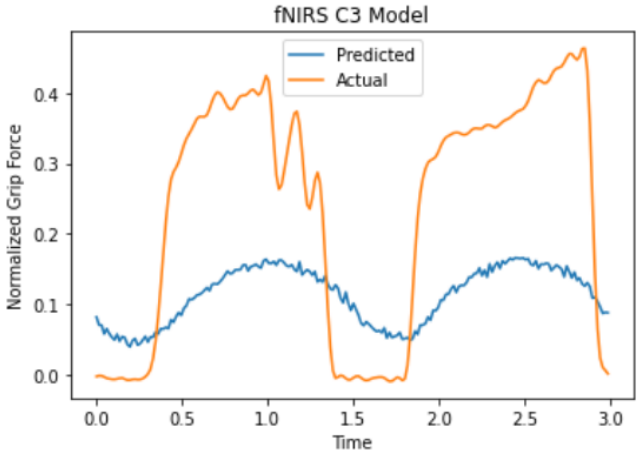

Figure 7. An fNIRS C3 model prediction against the actual force.

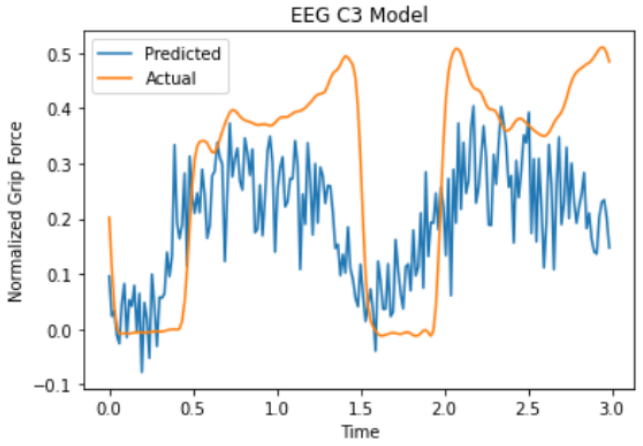

Figure 8. An EEG C3 model prediction against the actual force.

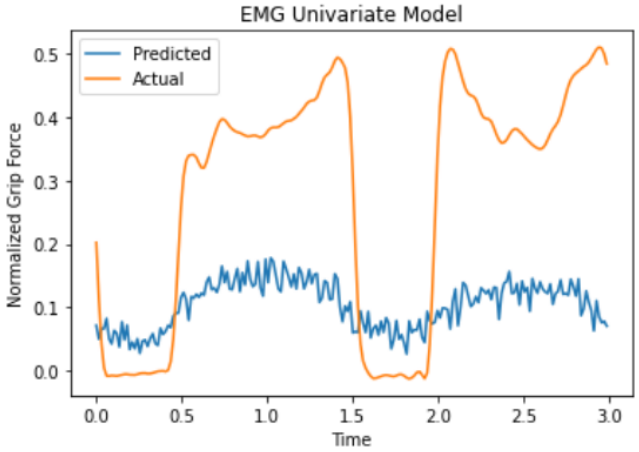

Figure 9. An EMG prediction against the actual force.

3. Conclusion

The baseline model exceeded the performances of all the univariate neural models, and channel models. However, all the unimodal models were able to identify the correct peaks and troughs in the force data.

It remains to be seen whether multimodal models will outperform simple non-neural models. However, these results suggest that data about continuous motor response is contained within all the various neural modalities. This raises the possibility that combining them will result in stronger predictions than using simple time-series methods. Combining EEG and fNIRS data looks especially promising because fNIRS produced less variable predictions while EEG produced more accurate mean predictions (see Fig 7. and Fig 8.). Unfortunately, due to constraints on available computing power, this question will remain un-answered for the time being.

It should be emphasized that while these results might seem under-whelming in another domain, the fact that models from each of the neural modalities was able to produce models that realistically identified the peaks and troughs is highly encouraging, and not commonly seen. There is a strong possibility that the hierarchical approach taken here can be refined into a viable method for predicting continuous motor signals from non-invasive neural data, especially with rapidly improving equipment.

References

[1] Pablo Ortega, Tong Zhao, and A. Aldo Faisal. HYGRIP: Full-Stack Characterization of Neurobehavioral Signals (fNIRS, EEG, EMG, Force, and Breathing) During a Bimanual Grip Force Control Task. 14.

[2] Pablo Ortega and Aldo Faisal. HemCNN: Deep Learning enables decoding of fNIRS cortical signals in hand grip motor tasks.

[3] M. Cope, D. T. Delpy, E. O. R. Reynolds, S. Wray, J. Wyatt, and P. van der Zee. Methods of Quantitating Cerebral Near Infrared Spectroscopy Data. In Masaji Mochizuki, Carl R. Honig, Tomiyasu Koyama, Thomas K. Goldstick, and Duane F. Bruley, editors, *Oxygen Transport to Tissue X*, Advances in Experimental Medicine and Biology, pages 183–189. Springer US.