

PAPER • OPEN ACCESS

Deep learning multimodal fNIRS and EEG signals for bimanual grip force decoding

To cite this article: Pablo Ortega and A Aldo Faisal 2021 *J. Neural Eng.* **18** 0460e6

View the [article online](#) for updates and enhancements.

You may also like

- [Decoding human mental states by whole-head EEG+fNIRS during category fluency task performance](#)

Ahmet Omurtag, Haleh Aghajani and Hasan Onur Keles

- [Imagined speech increases the hemodynamic response and functional connectivity of the dorsal motor cortex](#)

Xiaopeng Si, Sicheng Li, Shaoxin Xiang et al.

- [A hybrid BCI based on EEG and fNIRS signals improves the performance of decoding motor imagery of both force and speed of hand clenching](#)

Xuxian Yin, Baolei Xu, Changhao Jiang et al.

结果：

1. 多模态（脑电+肌氧）>单模态
2. cnnattn-lstm > cnnattn > linear
3. 特征处理效果 > 直接使用原始特征

MDN捕捉两只手的协方差特征
展示中间过程（时序、不同频率）显著性拓扑图
脑电信号不同频率分段处理分析

limitation:

1. 物理设备：多模态通道数量更多更好
2. 准备信号会影响动态响应
3. 动态响应时间过短

Journal of Neural Engineering



OPEN ACCESS

PAPER

Deep learning multimodal fNIRS and EEG signals for bimanual grip force decoding

RECEIVED
19 March 2021REVISED
26 July 2021ACCEPTED FOR PUBLICATION
4 August 2021PUBLISHED
31 August 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.

Pablo Ortega^{1,2} and A Aldo Faisal^{1,2,3,4,*} ¹ Brain and Behaviour Lab, Department of Bioengineering, Imperial College London, London SW7 2AZ, United Kingdom² Brain and Behaviour Lab, Department of Computing, Imperial College London, London SW7 2AZ, United Kingdom³ Data Science Institute, Imperial College London, London, United Kingdom⁴ MRC London Institute of Medical Sciences, SW7 2AZ London, United Kingdom

* Author to whom any correspondence should be addressed.

E-mail: a.faisal@imperial.ac.uk**Keywords:** fNIRS, EEG, brain-machine interface, deep learning, bimanual force decoding

Abstract

Objective. Non-invasive brain-machine interfaces (BMIs) offer an alternative, safe and accessible way to interact with the environment. To enable meaningful and stable physical interactions, BMIs need to decode forces. Although previously addressed in the unimanual case, controlling forces from both hands would enable BMI-users to perform a greater range of interactions. We here investigate the decoding of hand-specific forces. **Approach.** We maximise cortical information by using electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) and developing a deep-learning architecture with attention and residual layers (*cnnatt*) to improve their fusion. Our task required participants to generate hand-specific force profiles on which we trained and tested our deep-learning and linear decoders. **Main results.** The use of EEG and fNIRS improved the decoding of bimanual force and the deep-learning models outperformed the linear model. In both cases, the greatest gain in performance was due to the detection of force generation. In particular, the detection of forces was hand-specific and better for the right dominant hand and *cnnatt* was better at fusing EEG and fNIRS. Consequently, the study of *cnnatt* revealed that forces from each hand were differently encoded at the cortical level. **Cnnatt also revealed traces of the cortical activity being modulated by the level of force** which was not previously found using linear models. **Significance.** Our results can be applied to avoid hand-cross talk during hand force decoding to improve the robustness of BMI robotic devices. In particular, we improve the fusion of EEG and fNIRS signals and offer hand-specific interpretability of the encoded forces which are valuable during motor rehabilitation assessment.

1. Introduction

Invasive microelectrodes in BMI systems have provided a practical tool for the restoration of movements in individuals with paralysed limbs [1, 2]. In these BMIs, a motor command is decoded from cortical neurons and then sent to a device that stimulates an actual body part or to an artificial effector that supports or substitutes the real limb control. The spiking activity recorded with high-density multi-electrode arrays from the primary motor cortex (M1) provides the highest-resolution motor control brain signals in the space and time domains [3]. The invasive domain has also shown that force decoding can provide more

generalisable BMI decoders [4], especially in changing dynamic conditions [5], which become particularly important to enabling complex manipulations through the decoding of hand forces [3, 6–8].

Understanding the representation and decoding of forces from M1 can offer a more intuitive BMI control, exploiting motor signals already present in the cortex [7]. However, there is a considerable need for non-invasive approaches. Bio-compatibility issues and the complexity of the implantation surgeries required in invasive techniques limit the long-term adoption of invasive BMI technologies in paralysed and amputee end-users [9]. For the same reasons, invasive approaches are not suitable for users with

short term disabilities (e.g. after a stroke or concussion) or with no disabilities at all, including consumer applications [10].

Here, we pick up the challenge to decode continuous force generation by fusing non-invasive cortical signals in the form of co-located EEG and fNIRS targeting the bilateral sensorimotor cortices. EEG and fNIRS measure the cortical neuronal activity using very different physical and physiological phenomena. EEG captures the sub-second electrical potential changes of hundreds of cortical neurons while fNIRS measures the changes in optical absorption by oxy and deoxy-haemoglobin after neuronal metabolic demands. Previous studies have shown that EEG and fNIRS provide complementary information in the decoding of discrete characteristics of executed [11] or imagined force [12, 13] in a trial-by-trial basis. Despite the benefits introduced by these multimodal approaches, EEG [14, 15] or fNIRS [16] have only been used separately to decode force on a continuous-time basis, reflecting the current lack of methods that can efficiently fuse these non-invasive modalities [17].

In this context, deep learning (DL)—a flexible machine learning framework for non-linear modelling and automatic feature extraction—can introduce further advantages in non-invasive BMI force decoding. It has been shown that DL is a powerful tool for data driven feature extraction in EEG [18] and fNIRS [19, 20]. In this work, we explore bimanual continuous force reconstruction from multimodal (fNIRS and EEG) non-invasive cortical signals with DL. We first examine if multimodal signals (fNIRS and EEG) provide an advantage in bimanual continuous force decoding compared to each separate modality. For comparison, we used the linear approach in [14] to do this. Secondly, we explore DL techniques to fuse these multimodal signals with different neurophysiological origins.

In particular, our bimanual task allows us to specifically answer questions about bimanual force decoding in the continuous case. Namely, we evaluate three decoding performance aspects: (1) force trajectory reconstruction, i.e. how well a method can reconstruct bimanual force trajectories from non-invasive signals; (2) force detection performance, i.e. how well a model can detect if the hands apply force; and (3) hand disentanglement, which measures how accurate a model is at distinguishing which hand is producing more force, i.e. is more active or contracted at a given time.

Previous studies suggested that non-invasive brain signals are not linearly modulated with force [14, 15]. In this sense, the use of our proposed non-linear methods raises exciting questions on the non-linear modulation of brain-signals with continuously graded force. We hypothesise that the modulation of non-invasive brain signals with force might be better captured non-linearly as in the invasive case

[8, 21]. Hence, we interpret our non-linear approach in search of non-linear modulation signatures of non-invasive brain signals.

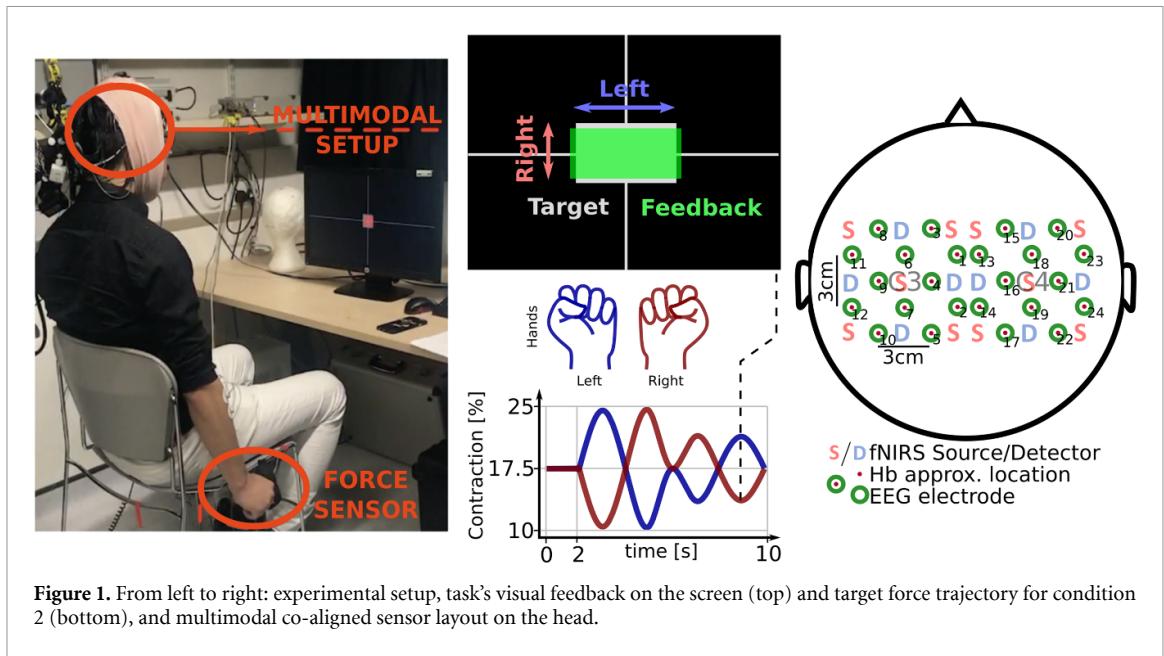
2. Methods

2.1. Task and participants

Nine ($N = 9$) healthy and right-handed participants were asked to perform a bimanual isometric contraction task. Handedness was confirmed by the Edinburgh inventory [22] for all participants. The Imperial College Research Ethics Committee approved all procedures, and all participants gave their written informed consent. The experiment complied with the Declaration of Helsinki for human experimentation and national and applicable international data protection rules.

We designed our task to study muscle-like and high-level characteristics of force control and their non-invasive decoding. To do so, we provided the force profile with two characteristics (figure 2(A) rightmost). First, both hands were either contracted or relaxed at the same time. This allows us to distinguish between a relaxed and contracted state representing a high-level or coarse force control state. Second, once contracted, the hands had to track five force profiles (figure 2(A) left to right). The first profile consisted of a constant contraction with both hands and was not used in this study. The remaining profiles (2 to 5) required each hand to track a different dynamic force target (dashed lines in figure 2(A)). Since each hand tracked a different target simultaneously, this part of the task represents muscle-like characteristics of force generation. Namely, these muscle-like characteristics are specifically associated with the group of muscles controlling each hand. In summary, the task's high-level and muscle-like characteristics allow us to differentiate between relaxed and contracted bimanual states. Additionally, since the dynamic part was different for each hand, the task introduces the challenge of decoding simultaneous and hand-specific hand grip generation. In particular, the different way each hand is engaged in dynamic force tracking gives a better representation of continuous control of force. Altogether, our task represents a controlled way of studying force generation during bimanual interactions.

The participants received visual feedback on the desired contraction trajectory that they were instructed to follow with each hand (figure 1). Except for condition 1, in which both hands performed the same force profile, the remaining 2 to 5 conditions required contraction trajectories that were specific to each hand. These target force trajectories (dashed lines in figure 2(A) overlying the average behaviour) required each hand to generate a different contraction. All conditions lasted 10 s, and all participants did 30 trials per condition, except two participants who were tired to perform the last block. The order of the conditions



was randomised, and each condition was performed in blocks of 30 trials. For conditions 2–5, we set the highest level of contraction to 25% MVC and the lowest to 10% MVC. The average of the contraction trajectory during the 10 s corresponded to the same contraction level of condition 1, i.e. 17.5% MVC.

Two seconds before the trial started a grey square appeared on the screen indicating the subjects to get ready. After this an acoustic signal indicated the start of the trial, and for the first 2 s of conditions 2–5 the target showed a flat 17.5% MVC that the subjects aimed to reach before the dynamic part of the task started. A randomised resting period uniformly distributed between 15 and 21 s followed each trial to avoid phasic constructive interference of systemic artefacts, e.g. Mayer waves and breathing, in the brain responses. The refreshing of the feedback on the screen was set to 100 Hz.

2.1.1. Brain and force signals

Twenty four ($N = 24$) co-aligned EEG and fNIRS channels covered the bilateral sensorimotor cortex and provided the brain signals from which to decode the force signals generated with each hand. EEG and fNIRS co-alignment is achieved by arranging the probe layout in such a way that for each fNIRS source and detector an EEG electrode is placed in between them (see figure 1).

A NIRScout system (NIRx Medizintechnik GmbH, Berlin, Germany) was used to record fNIRS signals. Twelve channels per hemisphere ($N = 12/\text{hemisphere}$, 10 sources and 8 detectors in total) sampling at 12.5 Hz were used. We used an ActiChamp amplifier (BrainProducts, Berlin, Germany) operating at 4 kHz (BrainVision software, v1.20.0801) to record EEG. We first downsampled EEG to 500 Hz (with anti-aliasing filtering) and then

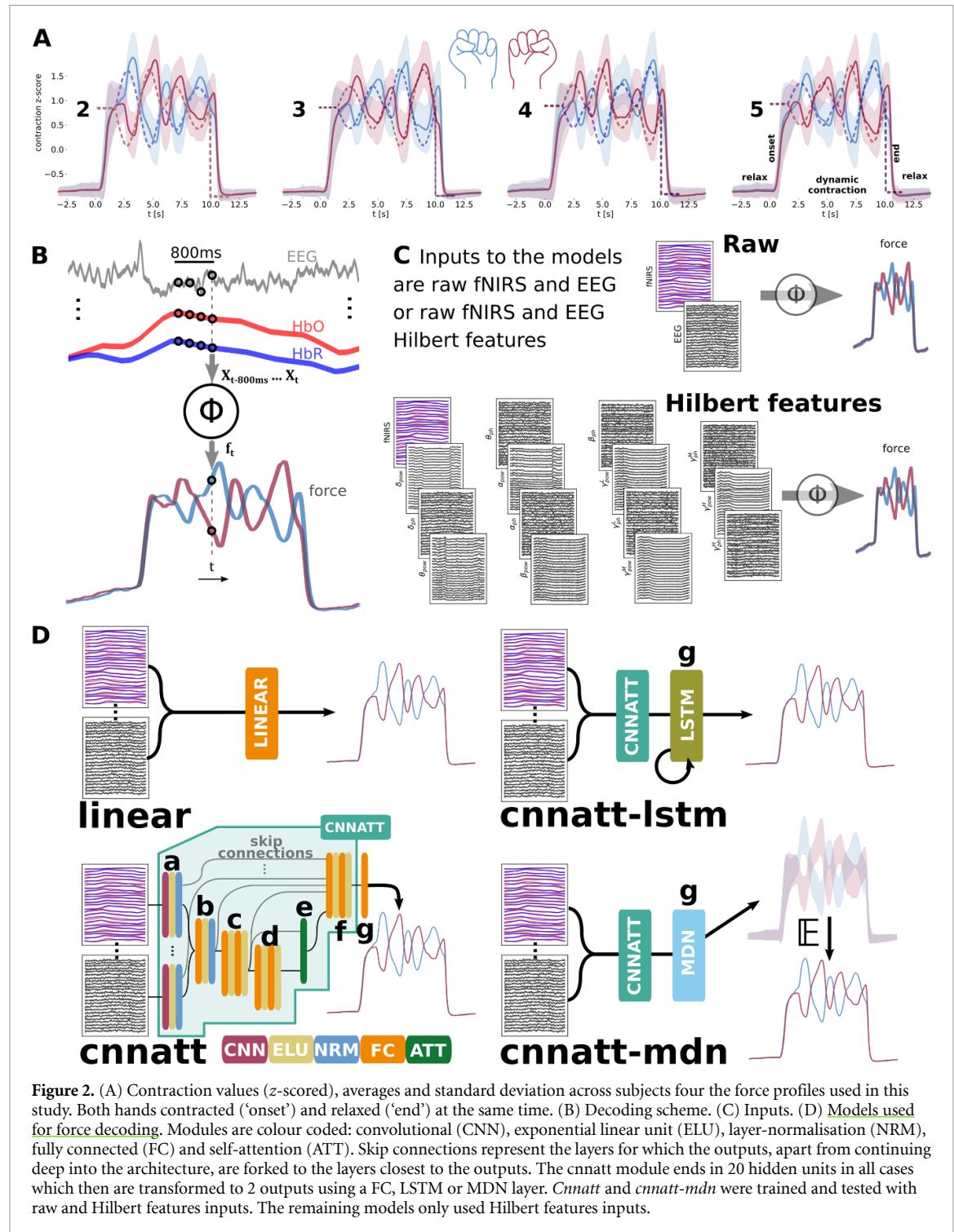
applied Notch filters at the mains (50 Hz) and fNIRS (12.5 Hz) frequencies and their harmonics. We high-pass filtered the EEG above 1 Hz (5th order Butterworth filter) EEG. Finally, we automatically rejected the independent components extracted from the EEG having a correlation with electrooculography (EOG) (independently recorded using the same ActiChamp amplifier) above 0.3 in absolute value (in practice, no more than one component was rejected per trial).

We used two ($N = 2$) grip force transducers (1 kHz, PowerLab 4/25 T, ADInstruments, Castle Hill, Australia) to record the force generated by each hand simultaneously. We first resampled the force signals to 250 Hz and then band-pass filtered between 0.01 and 10 Hz (3rd order Butterworth filter). Force measures (originally in Volts) were divided by the recorded maximum voluntary contractions (MVCs) values (in Volts) recorded at the beginning of the experiment to compute contraction values (cont. = Force[V]/MVC[V]).

Before the decoding, we further processed the fNIRS and EEG signals. The fNIRS optical intensities ($I(t)$) were low-pass filtered below 0.25 Hz (7th order elliptical filter) and changes in optical densities per wavelength, $\Delta OD_{ij}^\lambda(t)$, were obtained as $\Delta OD_{ij}^\lambda(t) = \log \frac{(I(t))}{\sum_{t \in \Delta t} I(t))}$ where Δt is the duration of the trial. Finally, we extracted HbO and HbR concentrations along time using the modified Beer–Lambert law [23]. We used this HbO and HbR time series as input to the models without further feature extraction.

In contrast, we separated the EEG processing in two different streams, one generating ‘raw’ EEG measures and the other extracting time-courses of the power and phases of different EEG bands or, as we refer to them, ‘Hilbert features’, i.e. the amplitude envelope and instantaneous phase extracted using the Hilbert transform (figure 2(A)). These

幅度包络
瞬时相位



preprocessing streams allowed us to compare our algorithms' decoding capabilities when using 'raw' EEG data or its spectral content per frequency band ('Hilbert features'). We resampled all signals to 80 Hz for the 'raw' EEG stream. Otherwise, in order to extract 'Hilbert features' the EEG was decomposed in the power and phases of the following frequency bands as in [14]: delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), low-gamma (30–50 Hz), mid-gamma (70–110 Hz), high-gamma (130–200). Finally using the Hilbert transform [24] we computed

the amplitude envelope and the instantaneous phase in each frequency band, which we refer to as the 'EEG Hilbert features'. In this case, we resampled all signals and the newly computed EEG 'Hilbert features' to 12.5 Hz to have an equivalent number of samples as the HbO and HbR signals and to decrease the input dimensionality in the models.

We also recorded the scalp skin's hemodynamic activity on the forehead using a NONIN 8000 R (Tilburg, The Netherlands) in 17 of the sessions. The skin hemodynamics reflect the variations of

haemoglobin due to the heart and breathing activity [25, 26] but do not contain brain hemodynamic responses [25]. We used scalp hemodynamics to discard that pulse and breathing were not predictive of force by training a linear decoding model to predict force from skin hemodynamics and comparing its decoding performance to a similar linear model trained on fNIRS. After preprocessing, we extracted epochs from 4 s before the ‘Go’ instruction to 14 s after. Data is available in [27].

2.2. Decoders

We used six different methods to compare their ability to decode bimanual force from non-invasive brain signals. In some cases, the decoders’ main difference was the kind of EEG (raw or Hilbert features) that we used as input. Remaining signals, fNIRS and force had always undergone the same processing independently of the EEG. The only exception was the frequency at which we resampled the signals to align all measures in time. We built the models with increasing complexity. From less to more complex the used models (figure 2(D)) were: (1) A *linear* model, (2) a convolutional neural network with residual modules and self-attention using ‘Hilbert EEG features’ (*cnnatt*) or ‘raw’ EEG (*cnnatt-raw*); (3) we then included an long-short term memory (LSTM) network layer to *cnnatt*’s output (20 hidden units) and only tested with ‘Hilbert envelope features’ (*cnnatt-lstm*) since EEG raw features delivered a high sampling frequency that considerably slowed down their training times and reduced the performance; (4) and, finally, we added an **mixture density network (MDN)** layer to *cnnatt*’s output with ‘raw EEG’ or ‘Hilbert EEG features’ (*cnnatt-mdn-raw* and *cnnatt-mdn*, respectively).

As shown in figure 2(D), from input to output the *cnnatt* module consists of: (a) A 1D convolutional layer convolving the time axis with a 4 units kernel and specific for each input type. In particular, for fNIRS, this kernel also convolved HbO and HbR together, hence the kernel had a 2×4 shape. After the convolution, there was an ELU non-linearity and a normalisation layer. (b) All the outputs from the convolutional layer were concatenated both skipped to the last layer and passed through a FC layer with output size 20, followed by an ELU unit and a normalisation layer. (c) Two FC layers with output size 20 and ELU units followed. The output was skipped to the last layer, acted as the query values ($\mathbf{Q} \in \mathcal{R}^{1 \times 20}$) of the attention module and was sent to the next layer. (d) A similar layer as in (c) but with 400 output size was applied, reshaped and sent to the attention mechanism acting as context ($\mathbf{C} \in \mathcal{R}^{20 \times 20}$). (e) The self-attention mechanism was implemented as:

$$\mathbf{y}_{out}(\mathbf{Q}, \mathbf{C}) = \text{softmax}(\mathbf{Q}\mathbf{W}\mathbf{C}) \cdot \mathbf{Q},$$

with $\mathbf{W} \in \mathcal{R}^{20 \times 20}$ a learned matrix of weights and · the element-wise multiplication. (f) The attention

module output of size 20, is passed to the last module along with all intermediate values that skipped to this layer composed of two FC layers with output size 20 and one ELU unit and a Tanh unit for the output to respect the symmetry of the output as it was standardised between -1 and +1. (g) The last layer depended on the module attached and always had output size 2, i.e. a FC layer for *cnnatt*, an LSTM for *cnnatt-lstm* and a MDN for *cnnatt-mdn*.

Every decoding scheme used 800 ms of brain signal history to decode the bimanual force, except *cnnatt-lstm*, which in theory can have more memory. Although most BMI decoders exploiting electrical signals use memory depths below 500 ms, we increased the memory to 800 ms. This way, we aimed to better capture the slower fNIRS signals without excessively increasing the input dimensionality. All inputs (EEG and fNIRS) and outputs (force) were z-scored using the channel-wise average and standard deviation values computed using the training data.

The decoding can be expressed as $\mathbf{f}_t = \phi(\mathbf{X}_{t-800ms}, \dots, \mathbf{X}_t)$ where \mathbf{f}_t represents the vector of bimanual left (L) and right (R) forces $\mathbf{f}_t = [f_{L,t}, f_{R,t}]^T$ and \mathbf{X}_t the matrix of fNIRS and EEG features (either ‘raw’ or ‘Hilbert features’) at time t .

For the linear model in particular, the previous expression takes the form $\mathbf{f}_t = \mathbf{W}\mathbf{X}_{t-800ms\dots t}$ where \mathbf{W} represents a matrix of weights with shape $(2, N \cdot M \cdot \text{samples})$ with N representing the number of features per channels (16 in total, 2 for HbR and HbO and 14 for the instantaneous phase and the amplitude envelope extracted with the Hilbert transform on each of the 7 EEG bands), M the number of channels (48 in total, 24 for EEG and 24 for fNIRS), and ‘samples’, the number of samples in an 800 ms windows sampled at 12.5 Hz (i.e. 10 samples); $\mathbf{X}_{t-800ms\dots t}$ represents a matrix of concatenated EEG Hilbert features and oxygenated haemoglobin (HbO) and deoxygenated haemoglobin (HbR) concentrations as specified for the second dimension of \mathbf{W} . The same decoding approach was used by [14] with some differences. Their approach used the Fourier transform to extract only the power content, not the phases, of each EEG band except for delta. For the delta band, they used the filtered EEG signal in the respective band. Additionally, they used a 250 ms window and did not include fNIRS.

Previous invasive BMI research suggests that exploiting the task’s temporal structure can be beneficial in the decoding of force [8, 28]. We introduce recurrences in the model using an LSTM layer designed to capture temporal force dependencies. In particular, for the LSTM architecture (*cnnatt-lstm*) the memory-depth can be learned in a more flexible way, emphasised in the expression via the parameter τ . At the same time, LSTM allows the modelling to depend on previous decoding outputs $\hat{\mathbf{f}}$ via recurrence, $\mathbf{f}_t = \phi(\mathbf{X}_{t-\tau}, \dots, \mathbf{X}_t, \hat{\mathbf{f}}_{t-\tau}, \dots, \hat{\mathbf{f}}_{t-1})$.

Finally, we added an MDN layer to *cnnatt* since invasive BMI evidence shows the importance of including the variance and covariance structure in the modelling [8, 28]. The covariance structure becomes particularly relevant in our decoding task since we aim to decode two hands simultaneously. Indeed, hand activities are correlated during the relaxation, onset and end of force generation, and non-correlated or anti-correlated during the dynamic contraction. These covariance characteristics of the task can be explicitly modelled using and better captured using MDN layers

In particular, we adopted the approach of [29] and explicitly modelled the bimanual force's covariance structure through a mixture of bivariate Gaussians using a maximum-likelihood loss during training. Hence, the MDN models (*cnnatt-mdn* and *cnnatt-mdn-raw*) represent the conditional probability of the force generated by the two hands given the 800 ms history of brain signals. The MDN were implemented as the weighted sum of bivariate Gaussians, $p(\mathbf{f}_t | \mathbf{X}_{t-800\text{ms}}, \dots, \mathbf{X}_t) = \sum_{j=1}^M \pi_t^j \mathcal{N}(\mathbf{f}_t | \boldsymbol{\mu}_t^j, \boldsymbol{\sigma}_t^j, \rho_t^j)$, with each bivariate Gaussian \mathcal{N} modelled as,

$$\mathcal{N}(\mathbf{f}_t | \boldsymbol{\mu}_t, \boldsymbol{\sigma}_t, \rho_t) = \frac{1}{2\pi\sigma_{t,L}\sigma_{t,R}\sqrt{1-\rho_t^2}} \times \exp\left[\frac{-Z_t}{2(1-\rho_t^2)}\right],$$

and

$$Z_t = \frac{(f_{t,L} - \mu_{t,L})^2}{\sigma_{t,L}^2} + \frac{(f_{t,R} - \mu_{t,R})^2}{\sigma_{t,R}^2} + \frac{\rho(f_{t,L} - \mu_{t,L})(f_{t,R} - \mu_{t,R})}{\sigma_{t,L}\sigma_{t,R}},$$

where L and R indicate the left and the right hand emphasising the bivariate nature of our bimanual task. Specifically, two bivariate Gaussians ($M=2$) were used with the aim of facilitating the capture of correlated and anti-correlated activities in the task. Finally, the force decoded by *cnnatt-mdn* and *cnnatt-mdn-raw* can be computed as the expected value of the decoded force distribution $\hat{\mathbf{f}}_t = \mathbb{E}[p(\mathbf{f}_t | \mathbf{X}_{t-800\text{ms}}, \dots, \mathbf{X}_t)]$.

All models followed the same per subject training and testing strategy. Note we normalise each subject generated force to its maximum strength. However, the relationship between the brain signal and the hand contraction of subjects with different strengths might not share a similar mapping.

To introduce enough variance in our decoding task, we train and test in all task conditions presenting dynamic force generation (2–5 figure 2(A)). For each subject, we follow a 5-fold training approach fixing the random seed. We used a 90% of the non-left-out data for training and a 10% for validation for all neural networks approaches. Note, the *linear* model used 100% of the data in the fold for training. We

ensured that the representation of each task condition was balanced across data splits. We use the validation data to stop the training when the validation loss did not continue decreasing.

The models using MDN used a maximum likelihood approach for training. All the rest used the mean squared error (MSE) loss between the predicted and the actual force. Once the training finished, we tested the models and gathered the results across testing folds for statistical comparison purposes. Unless otherwise stated, we used Tukey's honest significance test for multiple comparisons of the means and Student's *t-test* for one-to-one mean comparisons.

3. Results

3.1. Comparison of EEG, fNIRS and multimodal (EEG & fNIRS) performance in force trajectory reconstruction

We first analyse the use of multimodal (fNIRS & EEG) signals to decode bimanual isometric force trajectories. We use the *linear* method in [14] which they used to decode continuous unimanual force from EEG. Here, we expand their approach to fNIRS and our multimodal system to decode the bimanual force task.

We quantify the decoding performance using two metrics: the Spearman's correlation coefficient (SCC) [30] (figure 3(B)), and fraction of variance accounted for (FVAF) (figure 3(A)) between the reconstructed force trajectories and the real ones. FVAF is a more stringent measure than correlation and it has been used before to evaluate force trajectory reconstruction performance [8, 31, 32]. FVAF has a value between $(-\infty, 1]$ with 1 representing a 100% reconstruction and 0 representing a 0% reconstruction that is as bad as using the average of the real trajectory as the prediction, negative values being even worse. FVAF can be computed as a percentage using, $\text{FVAF}[\%] = 100 \cdot \left(1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}\right)\%$. In the equation, y_i is the i^{th} sample of the actual signal, \hat{y}_i is the corresponding sample of the predicted signal, \bar{y} is the average of the actual signal, and N is the number of samples in the signal. We can observe that FVAF also has a more direct interpretation in terms of decoding error. Namely, we can interpret the quotient as an MSE value normalised to the actual trajectory variance. In contrast to MSE, since the quotient is subtracted to 1, the highest the FVAF the better the decoding.

The results of this analysis are shown in figures 3(A) and (B) and tables A1 and A2. At the population level, the increase of reconstruction performance of fNIRS (SCC = 0.610[L]/0.648[R]; FVAF[%] = 38.2[L]/42.3[R]%) and EEG (SCC = 0.606[L]/0.603[R]; FVAF[%] = 34.2[L]/33.9[R]%) is significantly lower than that of the multimodal system (SCC = 0.677[L]/0.699[R]; FVAF[%] = 46.9[L]/49.9[R]%; Tukey's honest significance test,

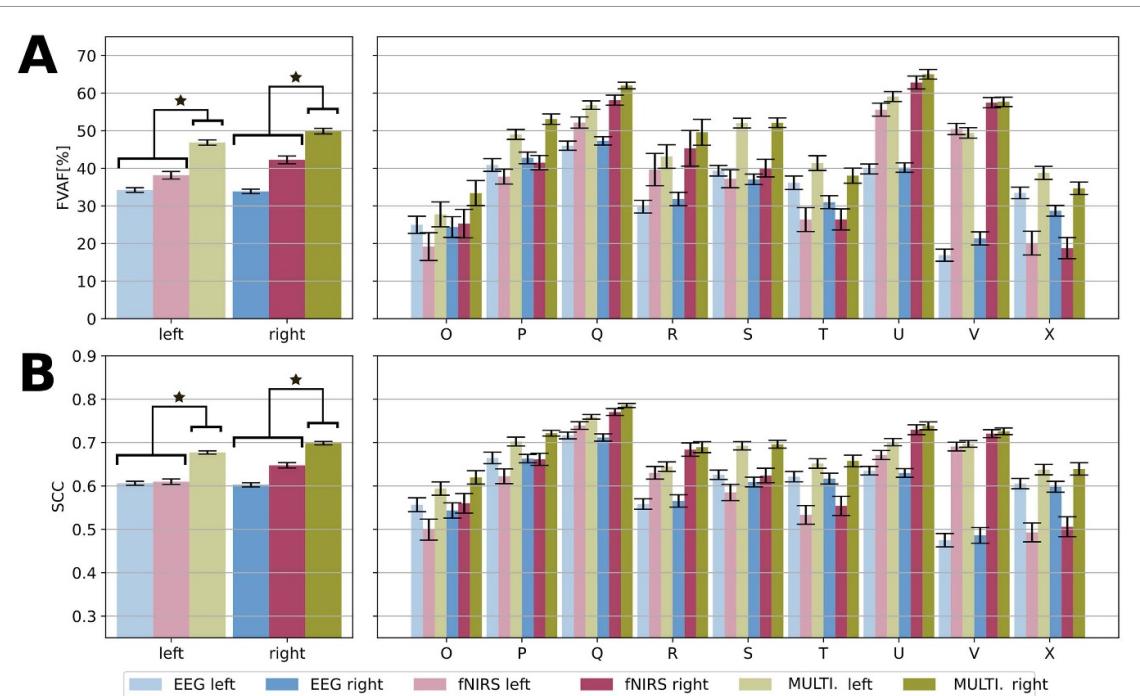


Figure 3. Force reconstruction performance metrics. Error bars represent the standard error of the mean (1080 trials). Tukey's honest significance test ($\blacktriangledown\alpha = 0.05$ and $\star\alpha = 0.01$). Performances computed between the predicted hand and the corresponding real hand (plain bars, $\hat{L}L$ and $\hat{R}R$) or the opposite real hand (shaded bars, $\hat{L}R$ and $\hat{R}L$). (A) Fraction of variance accounted for (FVAF[%]) of the *linear* baseline model for EEG, fNIRS and multimodal measures. (Left) The multimodal linear system achieves higher FVAF[%] than EEG or fNIRS alone (see table A2). (B) Ibidem for the Spearman's correlation (SCC, table A1).

$\alpha = 0.01$, figure 3(A)). In a per-subject basis, the multimodal system significantly outperforms both the EEG and the fNIRS signals in 3/18 cases for SCC (table A1) and 6/18 for FVAF[%] (table A2).

Finally, using a similar *linear* model, we trained a force decoding model using only the skin's haemodynamics. The force profile predicted with scalp hemodynamics was significantly less predictive of the real force ($SCC = -0.07/FVAF[\%] = -1\%$) than the force predicted with fNIRS ($SCC = 0.5/FVAF[\%] = 32.5\%$; t -test, $pval < 1e - 6$). Namely, force reconstructed from skin haemodynamics show insignificant reconstruction.

3.2. Comparison of linear and Deep Learning methods for bimanual force trajectories reconstruction

Linear models have shown a modest continuous force reconstruction capability from EEG [14] and fNIRS [16]. While invasive studies have shown the advantage of adding non-linearities in force decoding, [8] approaches like DL remain to be explored in continuous force decoding from non-invasive signals. In particular, these approaches might offer additional advantages in fusing EEG and fNIRS with different physical and physiological origins. To evaluate the decoding capabilities of DL in bimanual force continuous decoding, we focus on three aspects of force generation. We first evaluate the continuous force reconstruction capabilities (force trajectory

reconstruction). We then identify the aspects of this continuous reconstruction that are better decoded. In particular, we analyse how well models can distinguish a high-level force generation characteristic. Namely, how well they can distinguish between a relaxed and a contracted state (FDAcc). Then, we introduce a hand decoding specific measure that measures the models' ability to detect what hand is more contracted at each time step (hand disentanglement accuracy, HDAcc).

3.2.1. Force trajectory reconstruction, SCC and FVAF

We here examine the capabilities of DL to reconstruct bimanual force trajectories, particularly how deep-fusion methods compare to using linear approaches.

As figure 4(A) shows, the models' reconstruction performance is in the 0.65–0.72 range for SCC and 47%–60% for FVAF (see table A4 for values). In particular, deep models outperform the *linear* model in terms of FVAF. However, the effect is less prominent in terms of SCC (figure 4(B)) where most of the differences are non-significant (Tukey's honest significance test, $\alpha = 0.05$, see table 1). Additionally, when the same models extract features from the raw signals the performance decreases significantly compared to using the 'Hilbert features' (i.e. *cnnatt* and *cnnatt-mdn* respectively compared to *cnnatt-raw* and *cnnatt-mdn-raw*, see also table 1). In terms of SCC and FVAF, the right hand is also significantly better decoded than the left hand by all models (t -test, $\alpha = 0.05$, not

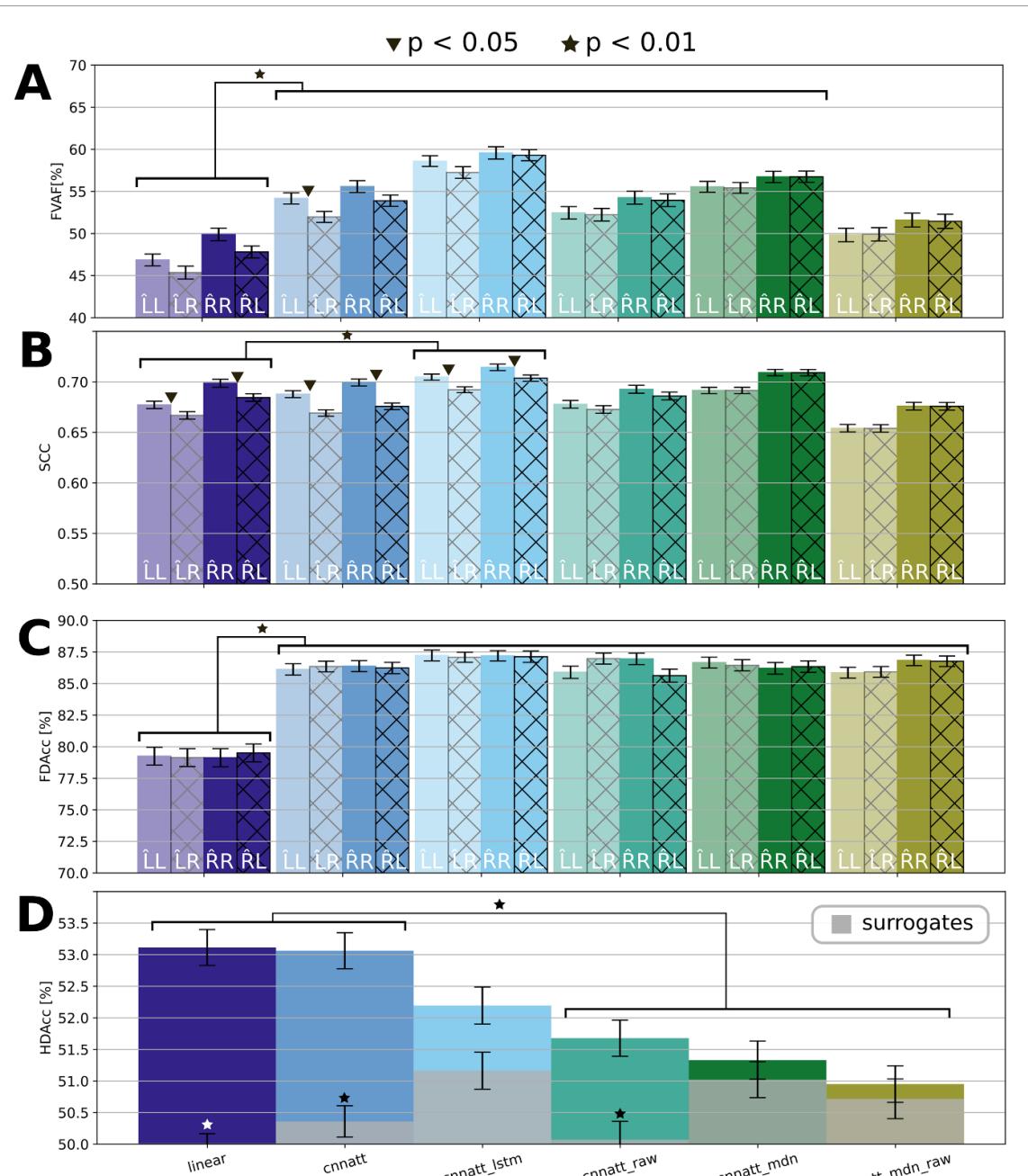


Figure 4. Force reconstruction performance metrics for *linear* vs *deep learning* methods comparison. Error bars represent the standard error of the mean (1080 trials). Tukey's honest significance test ($\blacktriangledown \alpha = 0.05$ and $\star \alpha = 0.01$). Performances computed between the predicted hand and the corresponding real hand (plain bars, LL and RR) or the opposite real hand (shaded bars, LR and RL). (A) FVAF for each method. DL generally outperforms the *linear* method (see tables 1 and A4). (B) SCC for each method. The right hand is overall better decoded than the left (see tables 1 and A4). (C) Force detection accuracy (FDAcc) per model. *deep learning* (DL) outperforms the *linear* method. (D) Hand disentanglement accuracy (HDAcc) per model (see table 1).

shown). Furthermore, in terms of SCC, all models preserve a significant sensitivity to the hand except for those using raw measures or the MDN layer (*t*-test, $\alpha = 0.05$, table 1). In terms of FVAF, the only model achieving hand specificity is the *cnnatt* for the left hand. Overall, we consider *cnnatt* (left/right hand: SCC = 0.688/0.699; FVAF[%] = 54.2/55.6%; FDAcc = 86.1/86.4%; HDAcc = 53.1%) and *cnnatt-lstm* (left/right hand: SCC = 0.705/0.714; FVAF[%] = 58.6/59.6%; FDAcc = 87.0/87.1%; HDAcc = 52.3%) the best decoders due to higher SCC, FVAF, force decoding accuracy (FDAcc) and

comparable HDAcc to the *linear* model (left/right hand: SCC = 0.677/0.699; FVAF[%] = 46.9/49.9%; FDAcc = 78.7/79.0%; HDAcc = 53.1%); tables A4 and 1). However *cnn-lstm* seems to saturate to flat contraction values (figure 6(A)) hence the lower HDAcc for the dynamic part of the force generation.

Although SCC and FVAF are valuable measures to rank decoding performance and test for hand sensitivity, it is hard to picture what is achievable in terms of decoding. Figure 6(A) presents representative decoding examples for all methods. As we can see in figure 6(A) for *cnnatt-lstm*, on the one

Table 1. Summary of bimanual force decoding methods comparison. Green indicates significant better performance of the column method compared to the row method. Red indicates significant worse performance. Grey indicates non-significant differences (Tukey's honest significance test, $\alpha = 0.05$).

		cnnatt		cnnatt lstm		cnnatt raw		cnnatt mdn		cnnatt mdn raw		
		L	R	L	R	L	R	L	R	L	R	
cnnatt-lstm>cnnattn>linear	linear	SCC	0.255	0.900	0.001	0.025	0.900	0.829	0.043	0.310	0.001	0.001
		FVAF[%]	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.033	0.576
		FDAcc	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
		HDAcc		0.900		0.219		0.006		0.001		0.001
cnnatt	cnnatt	SCC			0.008	0.037	0.315	0.751	0.900	0.390	0.001	0.001
		FVAF[%]			0.001	0.002	0.507	0.785	0.706	0.879	0.001	0.002
		FDAcc			0.664	0.010	0.900	0.900	0.900	0.900	0.900	0.846
		HDAcc			0.279		0.010		0.001		0.001	
cnnatt lstm	cnnatt lstm	SCC					0.001	0.001	0.078	0.900	0.001	0.001
		FVAF[%]					0.001	0.001	0.025	0.072	0.001	0.001
		FDAcc					0.636	0.030	0.900	0.014	0.384	0.001
		HDAcc					0.782		0.284		0.030	
cnnatt raw	cnnatt raw	SCC						0.060	0.016	0.001	0.013	
		FVAF[%]						0.022	0.176	0.086	0.114	
		FDAcc						0.900	0.900	0.900	0.658	
		HDAcc						0.900		0.486		
cnnatt mdn	cnnatt mdn	SCC								0.001	0.001	
		FVAF[%]								0.001	0.001	
		FDAcc								0.765	0.801	
		HDAcc									0.900	

hand, the prediction of force onset is very accurate, while the dynamic part of the bimanual contraction is reconstructed to a limited extent. While the reconstruction somewhat captures some of the bimanual dynamics, it struggles to reach the required amplitudes. We analyse these force detection and hand disentanglement decoding aspects in the following sections.

3.2.2. Force detection accuracy, FDAcc

Force detection accuracy (FDAcc) measures how well models detect that a hand is producing force or is relaxed. For each time step, we consider that a hand is contracting if the signal is above the threshold separating the two amplitude populations of forces (low and high level), with the low level representing relaxation (values close to zero). We also consider the hand-specificity of this measure. We do this by computing FDAcc between the predicted trajectory for the hand the model learned to decode and the opposite hand's actual trajectory.

Figure 4(C) shows the FDAcc for each hand and model. As we can see, deep models outperform the *linear* model in detecting when the hands are producing force (Tukey's honest significance test, $\alpha = 0.05$; tables A4 and 1; figure 4(C)). However, when we used the opposite decoded hand to detect the force of the actual hand, the levels of accuracy (figure 4(C) shaded bars) are similar to those using the corresponding hand (plain bars, see also A4). The lack of hand specificity at this level is an expected consequence of the task's design, as both hands produce force or are relaxed simultaneously.

3.2.3. Hand disentanglement accuracy, HDAcc

We showed that it is possible to reconstruct around a FVAF $\approx 55\%$ (*cnnatt*) of the force trajectories from multimodal signals (EEG & fNIRS). While the detection of the onset and end of the force is very accurate (FDAcc $> 85\%$), we observed that the dynamic part of the force was less well reconstructed, mainly in terms of amplitude (figure 6(A)).

To have a better idea of the models' ability to capture hand specific information during the dynamic part, we use the HDAcc. In particular, HDAcc measures how accurately the models can decide which hand is producing a higher force than the other at each time step. Figure 4(D) shows the average HDAcc achieved by each method. The *linear* and the *cnnatt* models have significantly higher HDAcc compared to the rest of the models except *cnnatt-lstm* (Tukey's honest significance test, $\alpha = 0.05$; tables A4 and 1). Nonetheless, *cnnatt* has a higher HDAcc than *cnnatt-lstm*. We empirically computed the random level of HDAcc by shuffling the fNIRS and EEG inputs in the temporal dimension ('surrogates' in figure 4(D)). The empirical random level is within the 50% – 51.5% HDAcc range depending on the model and only the *linear*, *cnnatt* and *cnnatt-raw* models had significantly higher HDAcc than the random empirical level corresponding to their models (Tukey's honest significance test, $\alpha = 0.01$).

3.2.4. Summary of comparison

We present a summary of the comparison in table 1 for all the methods and all the scoring measures (Tukey's honest significance test, $\alpha = 0.05$). The

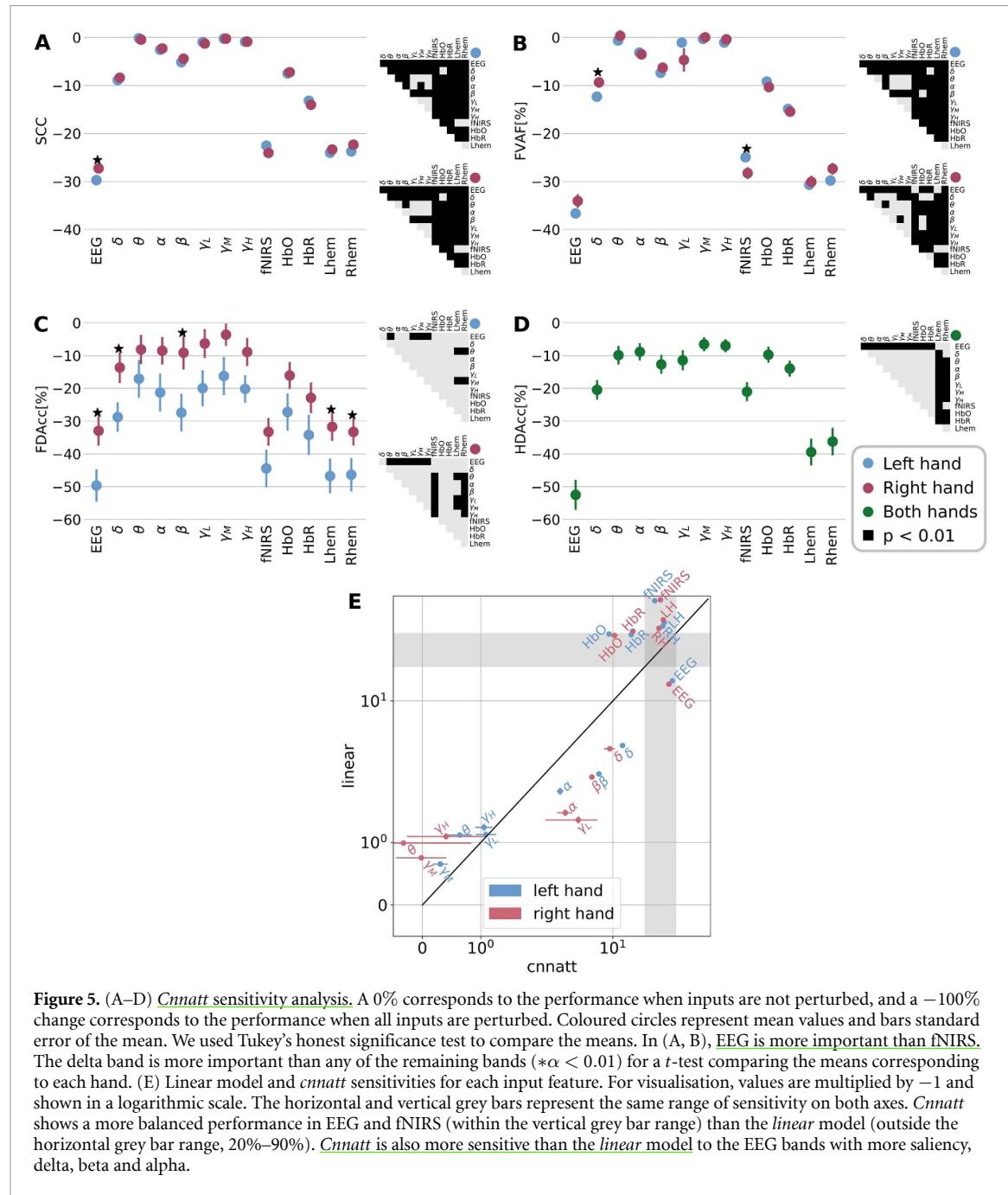


table shows how *cnnatt* and *cnnatt-lstm* significantly outperform the linear model in at least two scoring measures but perform similarly in terms of HDAcc. Consequently, we consider that *cnnatt* and *cnnatt-lstm* are among the best performing methods. Furthermore, although there is not enough statistical evidence to say that the HDAcc value of *cnnatt* is higher than that of *cnnatt-lstm*, the lower HDAcc of the latter is qualitatively captured in the corresponding column of figure 6(A), showing a flatter dynamic force decoding. In addition, *cnnatt* is the only method that preserves hand-specificity for both hands in terms of SCC and one hand in terms of FVAF (table A4). Altogether, we consider *cnnatt* as the most adequate method to decode bimanual force

trajectories from non-invasive multimodal (fNIRS and EEG) measures.

3.2.5. Model interpretation

The model's interpretation can provide insight into how the brain simultaneously controls bimanual forces and what features from the non-invasive signals offer more information. We next analyse *cnnatt* with a double-sided approach.

Figure 5 shows the sensitivity results per feature, hand and performance measure. The computation of sensitivity consists of randomly shuffling an input feature in the time dimension and then measuring its effect on the decoding. We measure the effect in decoding performance with FVAF[%], SCC, FDAcc

and HDAcc. In particular, this test evaluates how important the causal time course of the features is in the decoding of forces since we only perturb the signal's temporal structure but not the amplitude distribution. To standardise the measures, we compute the per cent change in performance with a 0% change representing the performance of the unperturbed signals and -100% the maximum performance reduction when all features are perturbed. As we can see in figure 5, the perturbation of any of the features leads to a reduction in performance (negative changes) in all measures, providing evidence that the causal order of the signal is essential in the decoding (Tukey's honest significance test, $\alpha = 0.01$).

Trajectory decoding performance metrics, SCC and FVAF, provide similar results to each other (figures 5(A) and (B)). As we can see, the perturbation of all EEG features leads to the most significant performance reduction for both measures except for the right hand for FVAF, followed by the fNIRS's perturbation. Within the EEG bands, the delta band's perturbation leads to a higher performance decrease followed by the beta and alpha bands. The remaining bands, theta and gamma, play a less significant role in force decoding. In terms of fNIRS, HbR has a more definitive role in the decoding of force compared to HbO, and the perturbation of fNIRS leads to a lower performance decrease in decoding compared to EEG. However, this decrease is more notable when compared to each of the EEG bands. The perturbation of all the signals belonging to the left or right hemisphere leads to a similar decline in decoding, with no hemisphere seeming to be more important than the other in providing information for force decoding.

In terms of FDAcc, the perturbation of features loses most of the differences found for FVAF and SCC (figure 5(C)). However, it reveals differences across hand that, to some extent, could already be observed in FVAF or SCC. Namely, the perturbation of EEG and, in particular, the delta and beta band disturbs more the left-hand decoding (t -test, $*p < 0.01$). The same occurs when each hemisphere is perturbed separately. The fact that the left-hand decoding is more sensitive to perturbations in the inputs suggests that the left hand's information is somewhat contingent on the right hand's information. Overall, the latter dominates in the detection of force generation from brain signals.

In terms of hand decoding accuracy (HDAcc, figure 5(D)), the perturbation of all the EEG bands together lead to the most significant reduction of accuracy followed by the perturbation of the left or the right hemisphere, but there are no significant differences between the different bands of EEG and fNIRS (Tukey's honest significance test, $\alpha = 0.01$).

The analysis of the sensitivity to the perturbation of the inputs' temporal order provides evidence that the decoding is causal and sheds some light on the features that are more important for decoding forces.

However, it does not reveal much of the brain signals' temporal and spatial dynamics during force generation. To reveal these temporal and spatial brain dynamics, we analyse the saliency of the inputs [33]. The saliency provides a measure of how necessary each part of *cnnatt* inputs is for force decoding. Specifically, for a given loss, the saliency specifies the direction (gradient) in the input dimension in which a change would lead to a more significant loss decrease. We compute each hand's saliency separately in two different ways related to the force and hand detection. Figures 6(B) and (C) show the saliency computed for the left and the right hand, whatever the force output during the whole trial. As we can see, the most prominent features, pointed by peaks of average saliency, are the beginning and end of dynamic force generation. Since we are averaging the saliency across all trials, these peaks might be a consequence of the beginning and end of force generation being the most temporally aligned feature across trials. In contrast, during the dynamic part of force generation, peaks and valleys of force are not aligned per trial. To make the saliency analysis more specific to the dynamic part, we computed saliency when the left hand applied more force than the right hand and vice versa (figures 6(D) and (E)).

As we can see in figures 6(B) and (C), HbO, HbR, and the beta, delta and alpha EEG bands have, in descending order, the highest saliency peaks, similar to what we observed in the sensitivity analysis. These are followed by low-gamma, high-gamma and finally theta power. Interestingly, this order changes for the phases of the features (figure 6(C)) with the phase of theta being more salient than that of beta, inverting the order of their power saliency. This saliency order is almost intact when we compute the saliency when each hand is producing a higher force than the other during the task's dynamic part (Tukey's honest significance test, $\alpha = 0.001$; figure 6(C)). This figure also shows the percentage of total saliency for each feature. By itself, fNIRS accounts for 25% of the saliency, which is more than any of the individual EEG Hilbert features. However, all EEG features account for 75% of the saliency, similar to what we observed in the perturbation analysis.

In contrast to what happens during the whole task, in the dynamic part, HbR has higher saliency than HbO. Finally, figure 6(D) highlights the order of relevance of the features during the dynamic part. As we mentioned, 25% of the saliency corresponds to fNIRS. A 50% would include delta, beta and alpha power, and a 75% would also add the phases of theta, delta and alpha, but not beta, and the power of low-gamma. Following this order of feature saliency, we will next analyse the temporal and spatial patterns of saliency.

Figure 6(E) shows the topographical distribution of saliency associated with the left and the right hand when each produces a higher force than the

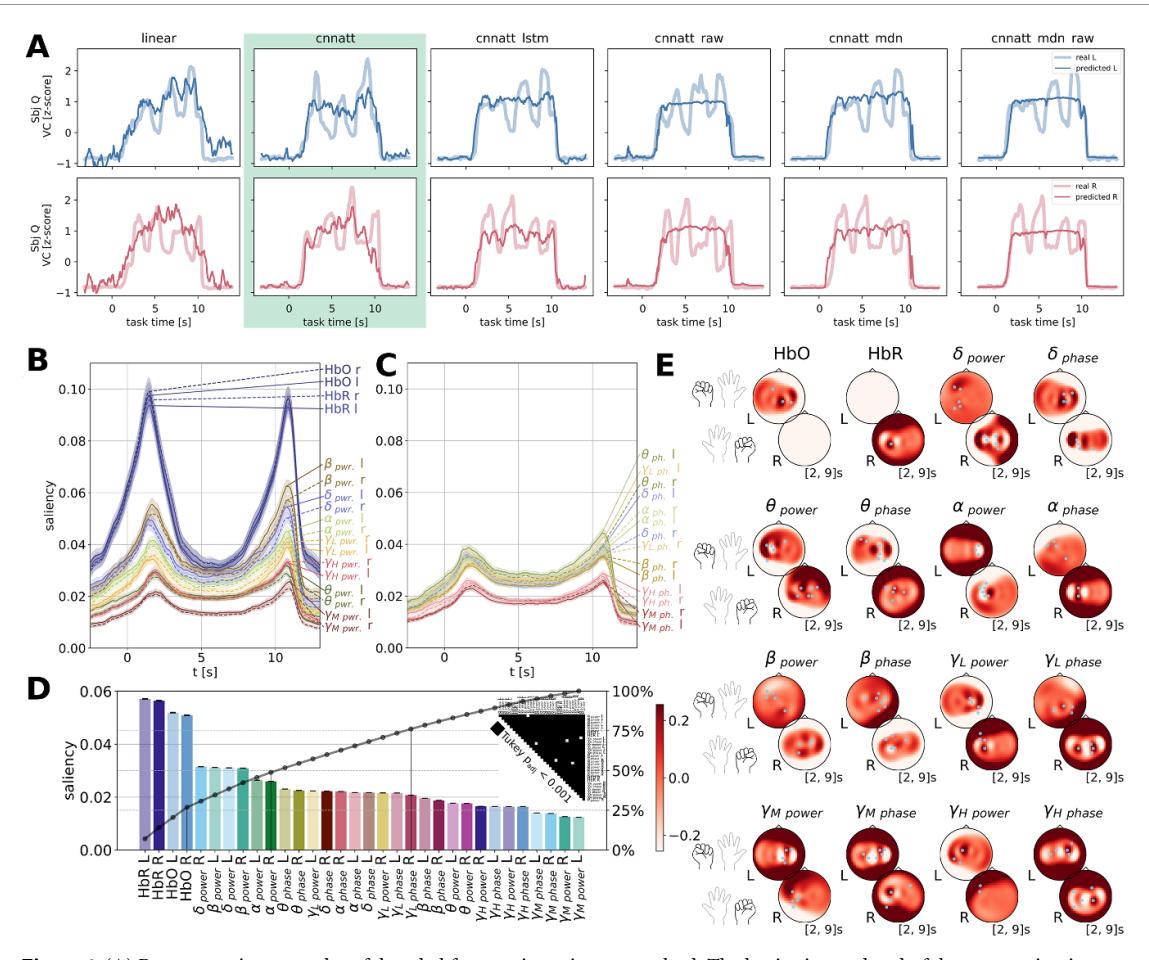


Figure 6. (A) Representative examples of decoded force trajectories per method. The beginning and end of the contraction is particularly well decoded by all DL methods. Dynamic changes of forces better captured by the linear and *cnnatt* methods. *Cnnatt-raw*, *cnnatt-mdn* and *cnnatt-mdn-raw* notably struggle at disentangling the hands (see figure 4(D)). (B, C) The mean and standard error of each hand's saliency associated with each input (phases in C). The saliency prominently peaks at the beginning and end of the dynamic force generation. (D) Saliency (mean and standard error) associated with each hand per input modality during the dynamic part of force generation (left y-axis). The right y-axis represents the percentage of accumulated saliency. A 25% of the total saliency corresponds to fNIRS. (E) Topoplot of saliency associated with each hand when producing greater force than the other during the dynamic part of force generation. Channels with significant higher saliency for one hand than the other represented in blue, and their saliency included in the topoplot (Tukey's honest significance test, $\alpha = 0.001$).

other. The heat-map represents the saliency value while the blue dots represent channels that have significantly (t -test, $p < 0.001$) higher saliency for a hand compared to the other. Remarkably, despite the low accuracy in HDAcc, the use of each hand leads to disentangled patterns of hand-specific activity at the cortical level during dynamic force generation. Some of the patterns are purely contralateral, others are ipsilateral, and others show specific bilateral activations. HbR shows a specific significant contralateral saliency for the right hand while HbO shows the same for the left hand and somewhat more spread. Delta power shows ipsilateral activity for the left hand and bilateral activity for the right hand, with delta phases showing stronger contralateralisation. Beta power shows a bilateral pattern with ipsilateral frontal projections and contralateral dorsal projections of saliency. Finally, alpha power shows pure contralateralisation, similar to low-gamma.

Some of these patterns survive when the saliency is computed for each hand regardless of the force applied. Figure 7, shows the time course of the saliency for each feature associated with each hand during the whole task. If we focus on the dynamic part of the task, [2, 10] s, we can see how some of the patterns identified for each hand in the dynamic part (figure 6(E)) survive when we use the forces generated in both hands to obtain the each hand's saliency. Namely, the contralateralisation of the alpha power is preserved, delta shows a similar bilateral pattern, and beta maintains a bilateral activation with ipsilateral frontal projections.

We further inspect the representation of force in non-invasive brain signals in figure 8. We compute the lateralisation index of the saliency per pair of symmetric channels around C3 and C4. We compute the saliency during aligned dynamic changes of force (leftmost column, figure 8) across conditions and force trajectories showing opposite trends per hand.

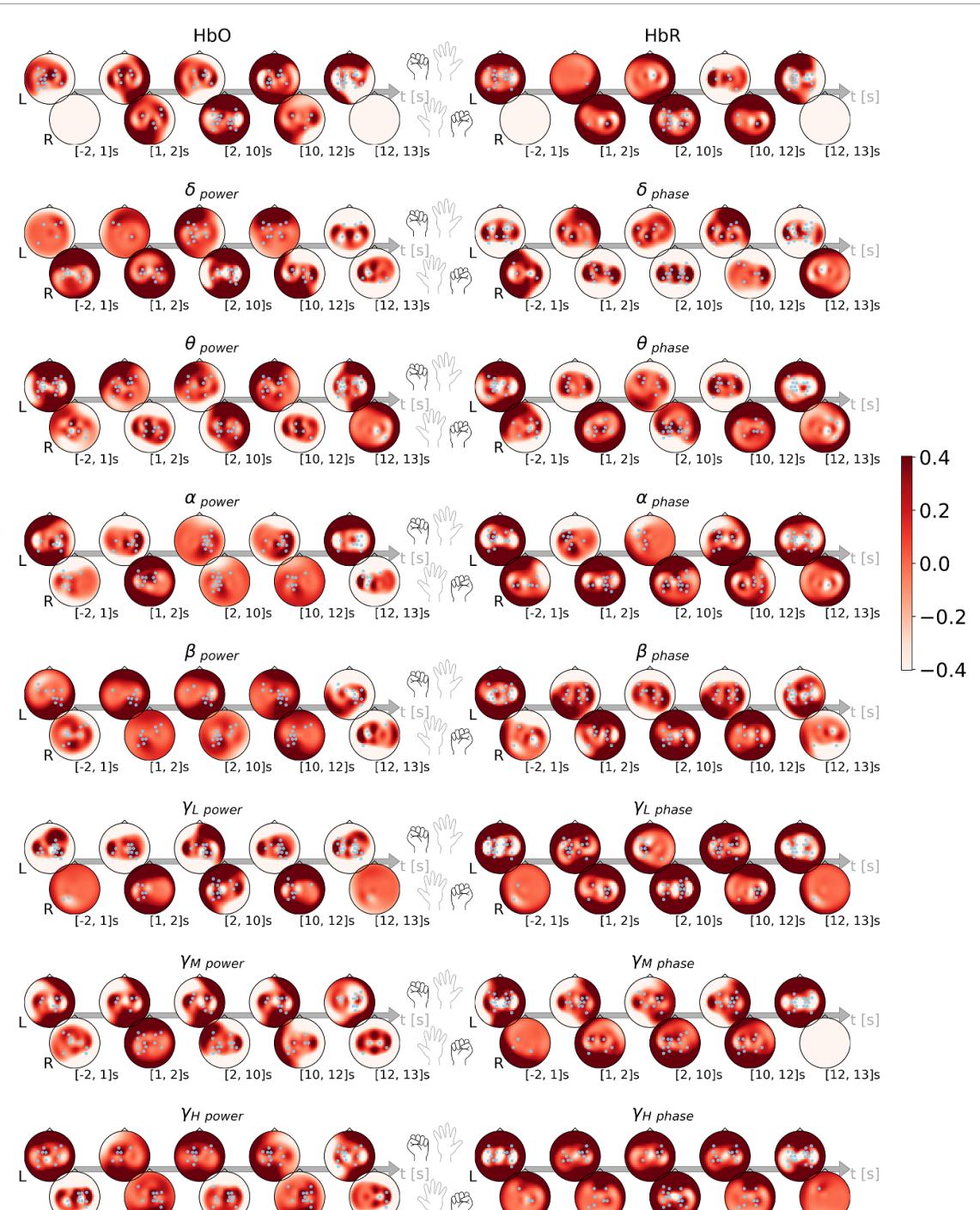


Figure 7. Topoplots of saliency associated with each generated force. We compute the saliency using the error associated with the decoded force; differences across hands emerge from the different ways in which each hand controls its contraction. In contrast to figure 6(E), the saliency is computed during the complete trial in the specified periods in the figure. Some patterns observed in figure 6(E), in particular, delta, alpha and beta, can also be observed here.

The figure shows that some channel pairs and features have a significant modulation that, to some extent, follows the bimanual force. For example, for channel 9, the HbO saliency has a similar modulation as the force generated with an initial right hemisphere dominance for the left hand (blue), generating more force at the beginning. A higher left hemisphere dominance for the left hand follows when the right-hand force is larger. We can observe similar patterns for other features and channels, especially for beta power.

3.3. Linear vs cnnatt comparison

We finally want to understand how each method exploits the information in the inputs to decode force trajectories. To do so, we compare the linear and *cnnatt* models (figure 5(E)). We follow the same previous time-perturbation analysis and compare the FVAF[%] performance decrease for each of them. For readability, we multiply the decrease in performance values by -1 to get positive values. These values represent how much the decoding of a model depends

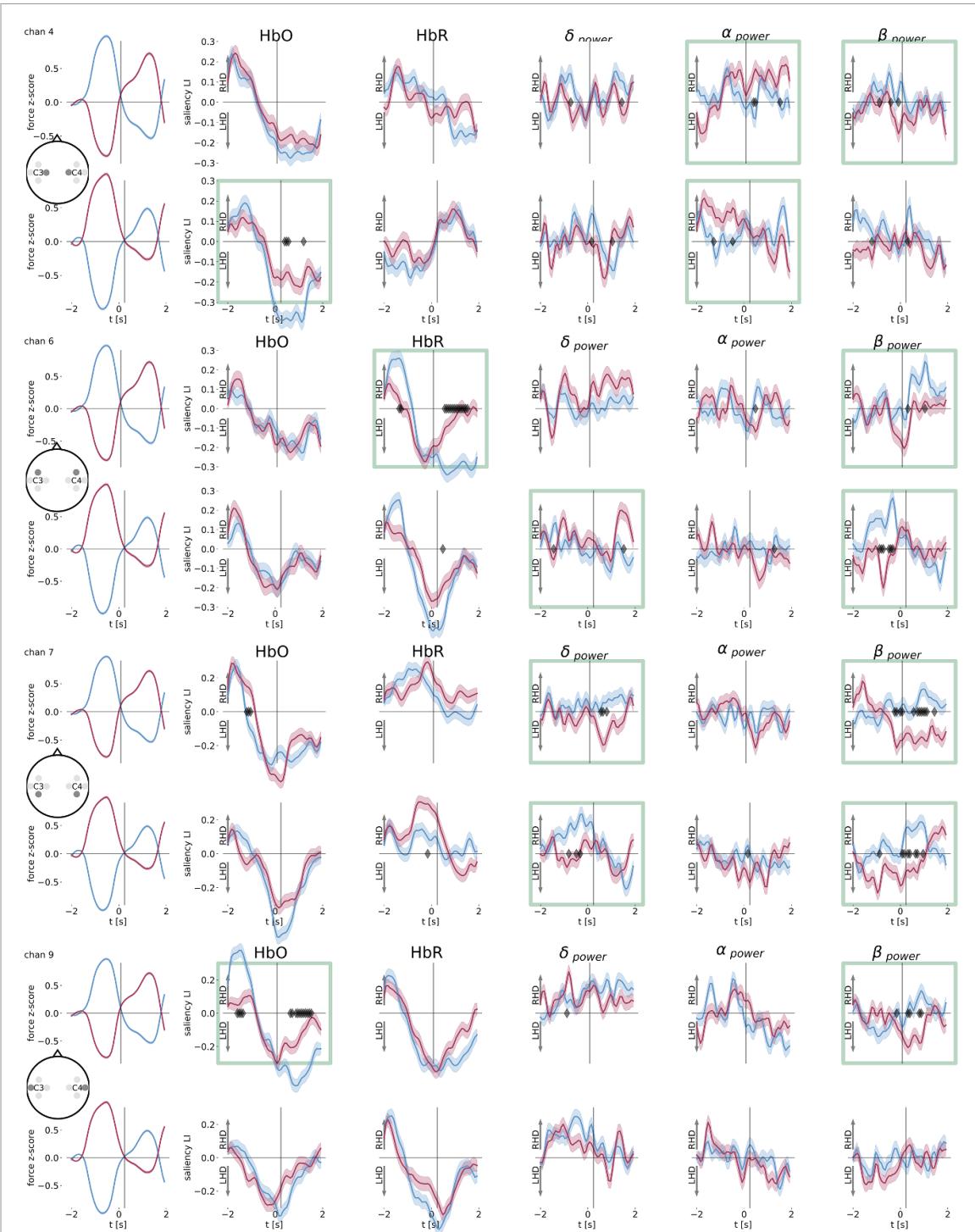


Figure 8. Saliency average time course during the dynamic generation of bimanual force. The leftmost column shows the average and standard deviation of the z -score contraction aligned so that peaks and valleys coincide. We use the time-course of saliency associated with each shown input (HbO , HbR and delta, alpha and beta power) to compute the lateralisation. Except for the leftmost column, in the remaining of the columns, positive y -axis values indicate right hemisphere dominance (RHD, upwards arrow) and vice versa for the left hemisphere dominance (LHD, downwards arrow). Diamonds represent significant differences between the lateralisation of the saliency associated with each hand (t -test, $\blacklozenge p < 0.05$). Blue represents the left hand and red the right hand. The examples in green boxes mark cases with clear differences across hands.

on a given feature for a given model. Namely, a 0% indicates that a model does not depend on that feature to decode force, while 100% means that it depends entirely on it.

The vertical and horizontal grey bars in the figure are visual guides marking the same interval of dependence in both axes (20% – 90%). For

cnnatt, the perturbation of EEG or fNIRS leads to a decrease in force reconstruction within the range indicated by the vertical grey bar. We can observe the opposite for the linear model, where perturbations in EEG lead to a lower reconstruction decrease than for fNIRS (both outside the horizontal grey bar interval). Indeed, the linear model

presented a worse decoding with only EEG ($\text{SCC} = 0.606[L]/0.603[R]$; $\text{FVAF}[\%] = 34.2[L]/33.9[R]\%$) than with only fNIRS ($\text{SCC} = 0.610[L]/0.648[R]$; $\text{FVAF}[\%] = 38.2[L]/42.3[R]\%$, Tukey's honest significance test, $\alpha = 0.01$, figure 3(A)). This indicates that the *cnnatt* strikes a more balanced dependence between EEG and fNIRS than the *linear* model which mainly depends on fNIRS.

4. Discussion

We set out to explore EEG and fNIRS deep-fusion to reconstruct bimanual force trajectories on a time-step basis. To ensure we exploit brain activities, we rejected ocular sources of activity using independent component analysis (ICA) and confirmed the negligible decoding power of skin haemodynamic activity.

4.1. Multimodal decoding outperforms each the decoding of each separate modality

We first found an increase of 0.1 SCC and 17% FVAF[%] when combining fNIRS and EEG (table A3), showing that the *linear* fusion of multimodal brain signals outperforms each single modality at the population level. For every subject, the multimodal system led in performance which was significant in 17% (SCC) and 33% (FVAF[%]) of the cases, corresponding to subjects with balanced EEG-fNIRS performance (tables A1 and A3). In the remaining cases, the multimodal system performed as the best subsystem but doubled the chances of acquiring valuable signals. Our results expand previous multimodal work, showing that combining EEG and fNIRS improves the classification of force trials [11, 13, 34]. In particular, our multimodal *linear* bimanual force reconstruction (0.70–0.79 SCC, 30–65 FVAF[%]) improved previous unimanual reconstructions using EEG (0.42–0.51 SCC, [14], mean 60%FVAF[%] [7]) and was similar to that of a non-causal fNIRS decoder (32–68 FVAF[%]) [16].

4.2. Deep learning improves the decoding of bimanual forces

The reconstruction (FVAF[%]: *cnnatt* 54.9%, *linear* 48.5%) and force detection (FDAcc: *cnnatt* 86.3%, *linear* 78.9%) were further improved by our *cnnatt* method, performing similarly in correlation (SCC *cnnatt* 0.694, *linear* 0.688) and hand disentanglement (*cnnatt* and *linear* 53.1% HDAcc, tables 1, A4). *Cnnatt-lstm* reached higher reconstruction (60%FVAF[%]) and force detection (87%FDAcc), but led to lower hand disentanglement (52.3% HDAcc, figure 4(D), table 1, A4). Overall, our EEG-fNIRS deep-fusion methods outperformed the multimodal *linear* fusion. Nonetheless, hand disentanglement was much lower (53.1% HDAcc,

50%–51% empirical baseline) than force detection (*cnnatt-lstm* 87% FDAcc, *linear* 78% FDAcc, figure 4(C)) and most saliency corresponded to the initiation and completion of forces (figure 6(B)), suggesting that the better ability to detect and model sharp changes of force provided the highest reconstruction gain. Our deep-fusion approach (*cnnatt-lstm*, 87% FDAcc) favourably compares to attempts at decoding unimanual force using linear decoders that achieved 74.8% accuracy using EEG [14], or 81% accuracy on unilateral foot force detection [11].

Interestingly, although by task design the force detection is not hand-specific (see table A4), when the delta or beta bands, or the left or right hemispheres, were perturbed, force detection decreased more for the left than for the right hand (*t*-test, $\alpha = 0.01$, figure 5(C)). The higher robustness of the right-hand suggests that force detection is hand-specific and that the right-hand cortical representation is more widespread across hemispheres. However, while many real-world manipulations require both hands contracting simultaneously [35, 36], this aspect is rarely tested since most studies focus on unilateral tasks [19, 34, 37–39]. In contrast, our EEG-fNIRS deep-fusion method offers more robust real-time grasp decoding, reflecting in real-world tasks which need decoders immune to hand cross-talk.

4.3. Deep learning improves the fusion of EEG and fNIRS

Several reasons explain the better performance of our deep-fusion approach. First, the *linear* model performs worse with EEG than with fNIRS (table A3), reflecting on its poor capacity to combine the different EEG frequency bands and its greater dependence on the slowest fNIRS signals (figure 5(E)). In contrast, *cnnatt* is more dependent on EEG (figures 5(A) and (B), 75% of total saliency, figure 6(D)) and strikes a more balanced cross-modality dependence than the *linear* model (figure 5(E)). The different physiological origins of EEG and fNIRS lead to very different signal velocities whose coupling require extensive feature engineering [34, 38]. Our deep-fusion method provides better decoding by automatically finding the best temporal features across modalities.

Second, to further exploit these time-scale differences, we added an LSTM layer to *cnnatt*, which led to better force detection but worse hand disentanglement. Additionally, explicitly modelling the bimanual covariance (*cnnatt-mdn*) led to a reduction of hand disentanglement (table 1, figure 6(A)). We think that these architectures' internal mechanisms might be biasing the models to detect the force edges. These edges are shared across hands and might be driving the gradient through all time-steps, masking more subtle hand-specific features.

Third, compared to previous DL classification approaches using raw fNIRS [19, 20], EEG [18, 40] or both [39], we early realised the benefit of including residual [41] and attention modules [42] to reconstruct force in a time-step basis. Without them, we could not outperform the *linear* method.

Finally, using raw EEG reduced the hand disentanglement performance compared to using Hilbert features (table 1). Raw EEG required a sampling frequency that increased 20 times the temporal input dimensionality. Since the output is redundantly over-sampled, this raw EEG approach may require approximately 20 times more examples to start learning temporal features outperforming the Hilbert transform [19, 34, 37–39].

4.4. Cortical encoding of hand-specific activities during simultaneous hand use

Remarkably, where previous studies were limited to unimanual [14] or a left/right sequential [16] tasks, our deep-fusion method can dissociate hand-specific cortical activities during simultaneous hand use despite the limited hand-specificity effect in the decoding (53% HDAcc).

First, we find a contralateral dominant activation for the left hand in HbO and a similar contralateral activation for the right hand in HbR ([2, 9] s in figure 7), suggesting that the network learns hand-related contralateral activations across haemoglobin types similar to those previously found during unimanual tasks [20, 43].

In the EEG delta band, there is a clear frontoparietal ipsilateral pattern for the left hand that becomes more bilateral for the right. This wider cortical representation of the right hand in delta (figure 5(E)) might explain why the dominant hand is generally better decoded than the non-dominant left hand (table A3, 47/50% L/R FVAF[%] and 0.676/0.698 L/R SCC). This kind of low frequency frontoparietal bilateral activations were previously found in unilateral grasping [14, 44, 45]. In particular, the ipsilateral component is more prevalent during fine motor control [45], reflecting the activity of a network in charge of modulating efferent motor commands with sensory afferent information [46–50]. Our bimanual task shows that this kind of patterns survive during bimanual control, supporting evidence of bilateral hemispheric interactions during bimanual control [51] and a dominant cortical representation of the dominant hand [52].

Alpha shows a clear contralateralisation (figures 6(D) and 7), suggesting that during bimanual control, forces mostly rely on the contralateral hemisphere as in the unimanual case [53, 54].

Beta reveals interesting fronto-parietal patterns across hemispheres (figures 6(D) and 7). Namely, there is a frontal projection in the ipsilateral

hemisphere and a dorsal projection in the contralateral. While previous unimanual studies revealed bilateral frontoparietal patterns as separated entities for the dominant hand [14, 45], our bimanual task suggests that in beta, these fine control networks span across hemispheres. Activity in alpha and beta has been previously associated with top-down control [55–57]. Our results suggest that beta may also serve a hand coordination function across hemispheres [51, 58].

Note that our deep-fusion method primarily relies on the delta, beta and alpha bands and fNIRS for the decoding (figures 5(A) and (B)), assigning similar amounts of saliency to each modality (25% for delta, beta and alpha and 25% for HbO and HbR, figure 6(D)).

Despite lower saliences in the gamma band, contralateral gamma activities are also revealed but do not survive when we compute saliency for both hands (figure 7). In general, higher frequencies have lower saliences (figure 6(D)) in line with previous invasive research showing that high-frequency bands are more attenuated in EEG [59], and more relevant for sensorimotor feedback [32, 60, 61].

Interestingly, the theta band was the only low-frequency band with more saliency in its phase than in its power (figure 6(D)). Theta's phase has been related to behaviour monitoring, and errors [62, 63] in the medial and lateral prefrontal cortex [64]. Here, the left-hand saliency is more frontal than the right-hand saliency ([2, 10] s figure 7), suggesting that the left hand is more subject to monitoring that result in higher frontal engagement.

Overall, our results suggest that some unimanual cortical activities are also present in the bimanual case and can be detected non-invasively with DL. However, differences in cortical resource allocation for each hand can reflect differences in hand decoding error.

Moreover, *cnnatt*'s showed hand-specific lateralised saliences that followed the time-course of the force produced by the corresponding hand (figure 8). These modulation traces suggest that some of the minor hand disentanglement effects we observed (53% HDAcc, 50%–51% random level) is captured by the non-invasive techniques and enhanced at the population level. Previous studies discarded the linear modulation of non-invasive brain signals with force [14, 15]. Moreover, invasive approaches provide evidence that graded forces modulate neuronal activity non-linearly [8, 65]. Our results suggest that deep-fusion techniques can reveal non-linear modulations non-invasively with lower signal-to-noise ratio (SNR). Next-generation non-invasive sensors, combining photonics, electrical fields and acoustics [66–69] could increase the SNR and benefit from the deep-fusion methods that we advance here, to unlock

real-time, continuously graded and bimanual control in non-invasive BMI.

4.5. Limitations of the study

We took measures to reduce EOG artefacts by ICA components with strong EOG content. We also recorded breathing and pulse signals on the forehead recorded with a NONIN 8000 R (previously characterised on the neck [26]) and confirmed their negligible decoding power. While these measures do not guarantee a total rejection of artefacts, our results show hand-specific cortical encoding for all signals used. In particular, since both hands are used at the same time in our study, if the decoding was dominated by artefacts we would have expected similar cortical topographic distribution of activities and, more importantly, a lack of hand-specificity in the encoding. However, these physiological artefacts may have acted as a confound in the decoding, resulting in a decrease in decoding power. In this scenario, better artefact rejection using setups with more EEG channels [70] and short fNIRS channels [25, 71] could improve the decoding results at the cost of more expensive equipment.

Another limitation in our study is the complexity of our bimanual task. We asked participants to track independent force profiles with each hand. In pilot trials, most subjects required some anticipation to prepare for the go signal, which led us to introduce a 'get ready' signal 2 s before the start of the task. This preparation signal could have triggered an advanced hemodynamic response that favoured the decoding of the start of force activity. We expect that the lack of the preparation signal would lead to worse or delayed fNIRS decoding of force profiles [16].

Finally, another limitation comes with the selection of the window length at 800 ms affecting all

models except those with LSTM modules that can learn different history depths of the time-series they exploit [72]. For the remaining models, 800 ms might not be enough time to resolve the dynamics of the hemodynamic response, classically considered to have a time constant of 0.65 s^{-1} [73, 74]. Nonetheless, at 800 ms windows while not able to resolve dynamics, we can resolve different states in the HbO and HbR amplitudes that can define high and low states of haemoglobin concentrations used by the decoders to define when force is being applied and when not. An interesting venue in this regard is the development of an LSTM variant that can learn different memory depths for fast and slow modalities at the same time, which we leave for future work.

Competing interests

The authors declare there are no competing financial interests.

Data availability statement

The data that support the findings of this study will be openly available following an embargo at the following URL/DOI: [10.6084/m9.figshare.13643243](https://doi.org/10.6084/m9.figshare.13643243). Data will be available from 30 August 2020.

Acknowledgment

We acknowledge the financial support of EPSRC: PO was supported by EPSRC CDT HiPEDS (EP/L016796/1) to PO and a UKRI Turing AI Fellowship (EP/V025449/1) to AAF. The fNIRS system was available through a capital equipment support grant by the EPSRC.

Appendix. Supplementary tables

Table A1. Comparison of SCC force reconstruction mean values (1080 measures) per hand and between EEG, fNIRS and multimodal (fNIRS & EEG) systems (Tukey's honest significance test $\alpha = 0.05$). In bold, the highest SCC, which always correspond to the Hybrid system. The asterisk (*) indicates where the multimodal system is significantly ($p < 0.05$) better than the EEG and fNIRS systems used separately. At the population (POP) level the differences are significant ($p < 0.05$).

Subject	Hand	EEG	fNIRS	Multim.	Tukey adjusted p value		
					EEG/fNIRS	EEG/Multim.	fNIRS/Multim.
O	Left	0.560	0.495	0.597	0.045	0.351	0.001
	Right	0.554	0.557	0.628*	0.900	0.014	0.019
P	Left	0.664	0.621	0.705	0.073	0.088	0.001
	Right	0.660	0.658	0.718*	0.900	0.001	0.001
Q	Left	0.709	0.737	0.767	0.026	0.001	0.094
	Right	0.698	0.770	0.786	0.001	0.001	0.254
R	Left	0.550	0.630	0.647	0.001	0.001	0.585
	Right	0.555	0.680	0.689	0.001	0.001	0.895
S	Left	0.606	0.580	0.686*	0.440	0.001	0.001
	Right	0.589	0.621	0.690	0.032	0.101	0.069
T	Left	0.613	0.523	0.648	0.001	0.270	0.001
	Right	0.611	0.534	0.652	0.005	0.200	0.001
U	Left	0.619	0.667	0.697	0.001	0.001	0.068
	Right	0.612	0.729	0.738	0.001	0.001	0.784
V	Left	0.456	0.689	0.697	0.001	0.001	0.862
	Right	0.468	0.718	0.726	0.001	0.001	0.882
X	Left	0.601	0.488	0.635	0.001	0.312	0.001
	Right	0.602	0.498	0.640	0.001	0.280	0.001
POP	Left	0.597	0.606	0.676*	0.369	0.001	0.001
	Right	0.594	0.642	0.698*	0.001	0.001	0.001

Table A2. Comparison of FVAF force reconstruction mean values (1080 measures) per hand and between EEG, fNIRS and multimodal systems (Tukey's honest significance test, $\alpha = 0.05$). In bold, the highest SCC means that always correspond for the means. The asterisk (*) indicates where the multimodal system is significantly ($p < 0.05$) better than both the EEG and the fNIRS system alone. At the population (POP) level the differences are significant ($p < 0.05$).

Subject	Hand	EEG	fNIRS	Multim.	Tukey adjusted p value		
					EEG/fNIRS	EEG/Multim.	fNIRS/Multim.
O	Left	26.1	19.5	29.9*	0.262	0.632	0.038
	Right	26.3	25.3	35.3	0.900	0.107	0.064
P	Left	41.4	37.6	49.8*	0.223	0.001	0.001
	Right	42.5	40.9	52.9*	0.759	0.001	0.001
Q	Left	45.0	51.8	56.7*	0.001	0.001	0.023
	Right	45.6	57.7	61.8*	0.001	0.001	0.035
R	Left	29.4	39.7	43.8	0.041	0.002	0.584
	Right	30.9	45.0	49.8	0.091	0.001	0.570
S	Left	37.0	36.8	51.3*	0.900	0.001	0.001
	Right	34.6	39.5	51.2	0.121	0.001	0.001
T	Left	35.6	25.4	41.3	0.007	0.214	0.001
	Right	30.4	24.8	37.5	0.168	0.058	0.001
U	Left	37.8	54.7	58.7	0.001	0.001	0.129
	Right	38.1	62.1	64.6	0.001	0.001	0.419
V	Left	16.4	49.8	49.9	0.001	0.001	0.900
	Right	20.2	56.8	57.7	0.001	0.001	0.880
X	Left	33.2	19.2	38.6	0.001	0.236	0.001
	Right	29.2	18.0	34.7	0.001	0.160	0.001
POP	Left	33.5	37.7	47.1*	0.001	0.001	0.001
	Right	33.0	41.6	49.8*	0.001	0.001	0.001

Table A3. Cross-subject average of force trajectory reconstruction measures, left/right hand SCC and FVAF, for the linear model using EEG, fNIRS or both (multimodal). At the population level and for both hands and measures the multimodal approach achieves significant higher performance than either signal, EEG or fNIRS, alone (see table A2).

	EEG	fNIRS	Multim.	Max. increase
SCC	0.597/0.594	0.606/0.642	0.676/0.698	0.1
FVAF	34/33	38/42	47/50	17

Table A4. Method comparison. The asterisk (*) indicates hand-specific significant differences (t-test, $\alpha = 0.05$). Metrics are computed between decoded and real left hand ($\hat{L}L$), decoded left and real right ($\hat{L}R$), decoded and real right hand ($\hat{R}R$) and decoded right and real left hand ($\hat{R}\hat{L}$). The computations with the opposite real hand are used to test for hand specificity. Cnnatt is selected due to better FDAcc and SCC and FVAF and similar HDAcc than the linear model while still preserving hand specificity in more cases compared to other DL models.

Method	Case	FDAcc	p-adj	SCC	p-adj	FVAF	p-adj	HDAcc
linear	$\hat{L}L/\hat{L}R$	78.7/78.3	0.704	0.677/0.667*	0.046	46.9/45.4	0.148	53.1
	$\hat{R}R/\hat{R}\hat{L}$	79.0/79.0	0.967	0.699/0.684*	0.009	49.9/47.8	0.105	
cnnatt	$\hat{L}L/\hat{L}R$	86.1/86.3	0.713	0.688/0.669*	0.001	54.2/52.0*	0.018	53.1
	$\hat{R}R/\hat{R}\hat{L}$	86.4/86.2	0.816	0.699/0.676*	0.001	55.6/54.0	0.084	
cnnatt	$\hat{L}L/\hat{L}R$	87.0/86.9	0.864	0.705/0.692*	0.003	58.6/57.2	0.149	52.3
lstm	$\hat{R}R/\hat{R}\hat{L}$	87.1/87.2	0.807	0.714/0.704*	0.017	59.6/59.3	0.778	
cnnatt	$\hat{L}L/\hat{L}R$	85.9/87.0	0.099	0.678/0.673	0.342	52.5/52.2	0.834	51.7
raw	$\hat{R}R/\hat{R}\hat{L}$	87.0/85.6	0.054	0.693/0.686	0.226	54.3/54.0	0.773	
cnnatt	$\hat{L}L/\hat{L}R$	86.7/86.4	0.732	0.691/0.691	0.997	55.5/55.4	0.885	51.3
mdn	$\hat{R}R/\hat{R}\hat{L}$	86.2/86.3	0.838	0.709/0.709	0.990	56.7/56.7	0.980	
cnnatt	$\hat{L}L/\hat{L}R$	85.9/85.9	0.912	0.654/0.654	0.974	49.8/49.9	0.946	51.0
mdn raw	$\hat{R}R/\hat{R}\hat{L}$	86.8/86.8	0.905	0.676/0.676	0.989	51.6/51.5	0.907	

ORCID iDs

Pablo Ortega  <https://orcid.org/0000-0003-4532-209X>

A Aldo Faisal  <https://orcid.org/0000-0003-0813-7207>

References

- [1] Bouton C E *et al* 2016 Restoring cortical control of functional movement in a human with quadriplegia *Nature* **533** 247–50
- [2] Ganzer P D *et al* 2020 Restoring the sense of touch using a sensorimotor demultiplexing neural interface *Cell* **181**(4) 763–73
- [3] Carmena J M, Lebedev M A, Crist R E, O'Doherty J E, Santucci D M, Dimitrov D F, Patil P G, Henriquez C S and Nicolelis M A 2003 Learning to control a brain-machine interface for reaching and grasping by primates *PLoS Biol.* **1** 193–208
- [4] Cherian A, Krucoff M O and Miller L E 2011 Motor cortical prediction of EMG: evidence that a kinetic brain-machine interface may be robust across altered movement dynamics *J. Neurophysiol.* **106** 564–75
- [5] Chhatbar P Y and Francis J T 2013 Towards a naturalistic brain-machine interface: hybrid torque and position control allows generalization to novel dynamics *PLoS One* **8** e52286
- [6] Bansal A K, Truccolo W, Vargas-Irwin C E and Donoghue J P 2012 Decoding 3D reach and grasp from hybrid signals in motor and premotor cortices: spikes, multiunit activity and local field potentials *J. Neurophysiol.* **107** 1337–55
- [7] Flint R D *et al* 2014 Extracting kinetic information from human motor cortical signals *Neuroimage* **101** 695–703
- [8] Naufel S, Glaser J I, Kording K P, Perreault E J and Miller L E 2019 A muscle-activity-dependent gain between motor cortex and EMG *J. Neurophysiol.* **121** 61–73
- [9] Makin T R, de Vignemont F and Faisal A A 2017 Neurocognitive barriers to the embodiment of technology *Nat. Biomed. Eng.* **1** 1–3
- [10] Nicolas-Alonso L F and Gomez-Gil J 2012 Brain computer interfaces, a review *Sensors* **12** 1211–79
- [11] Hansen M H, Kassebaum E, Plocharska M A, Jochumsen M and Kamavuako E N 2014 Improved detection and force decoding through combined near-infrared spectroscopy and electroencephalographic measurements *Replace, Repair, Restore, Relieve—Bridging Clinical and Engineering Solutions in Neurorehabilitation* (Berlin: Springer) pp 411–419
- [12] Yin X, Xu B, Jiang C, Fu Y, Wang Z, Li H and Shi G 2015 A hybrid BCI based on EEG and fNIRS signals improves the performance of decoding motor imagery of both force and speed of hand clenching *J. Neural Eng.* **12** 036004
- [13] Fu Y, Xiong X, Jiang C, Xu B, Li Y and Li H 2016 Imagined hand clenching force and speed modulate brain activity and are classified by fNIRS combined with EEG *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** 1641–52
- [14] Paek A Y, Gailey A, Parikh P J, Santello M and Contreras-Vidal J L 2019 Regression-based reconstruction of human grip force trajectories with noninvasive scalp electroencephalography *J. Neural Eng.* **16** 066030
- [15] Hayashi M, Tsuchimoto S, Mizuguchi N, Miyatake M, Kasuga S and Ushiba J 2019 Two-stage regression of high-density scalp electroencephalograms visualizes force regulation signaling during muscle contraction *J. Neural Eng.* **16** 056020
- [16] Nambu I, Osu R, Aki Sato M, Ando S, Kawato M and Naito E 2009 Single-trial reconstruction of finger-pinch forces from human motor-cortical activation measured by near-infrared spectroscopy (NIRS) *NeuroImage* **47** 628–37
- [17] Ahn S and Jun S C 2017 Multi-modal integration of EEG-fNIRS for brain-computer interfaces—current limitations and future directions *Front. Hum. Neurosci.* **11** 503
- [18] Schirrmeister R T, Springenberg J T, Fiederer L D J, Glasstetter M, Eggensperger K, Tangermann M, Hutter F,

- Burgard W and Ball T 2017 Deep learning with convolutional neural networks for EEG decoding and visualization *Human Brain Mapp.* **38** 5391–420
- [19] Trakoolwilaiwan T, Behboodi B, Lee J, Kim K and Choi J-W 2017 Convolutional neural network for high-accuracy functional near-infrared spectroscopy in a brain–computer interface: three-class classification of rest, right- and left-hand motor execution *Neurophotonics* **5** 011008
- [20] Ortega P and Faisal A 2021 HemCNN: deep learning enables decoding of fNIRS cortical signals in hand grip motor tasks *IEEE Neural Eng. (NER)* **10** 718–21
- [21] Fatemi M and Daliri M R 2020 Nonlinear sparse partial least squares: an investigation of the effect of nonlinearity and sparsity on the decoding of intracranial data *J. Neural Eng.* **17** 016055
- [22] Oldfield R C 1971 The assessment and analysis of handedness: the Edinburgh inventory *Neuropsychologia* **9** 97–113
- [23] Bhatt M, Ayyasomayajula K R and Yalavarthy P K 2016 Generalized Beer–Lambert model for near-infrared light propagation in thick biological tissues *J. Biomed. Opt.* **21** 1–11
- [24] Oppenheim A V and Schafer R W *Discrete-Time Signal Processing* 3rd edn (London: Pearson) pp 987–1023
- [25] Gagnon L, Cooper R J, Yücel M A, Perdue K L, Greve D N and Boas D A 2012 Short separation channel location impacts the performance of short channel regression in NIRS *Neuroimage* **59** 2518–28
- [26] García-López I and Rodriguez-Villegas E 2020 Characterization of artifact signals in neck photoplethysmography *IEEE Trans. Biomed. Eng.* **67** 2849–61
- [27] Ortega P, Zhao T and Faisal A 2021 hygrip2 (<https://doi.org/10.6084/m9.figshare.13643243>)
- [28] Sussillo D, Nuyujukian P, Fan J M, Kao J C, Stavisky S D, Ryu S and Shenoy K 2012 A recurrent neural network for closed-loop intracortical brain–machine interface decoders *J. Neural Eng.* **9** 026027
- [29] Graves A, Mohamed A-r and Hinton G 2013 Speech recognition with deep recurrent neural networks *IEEE Int. Conf. on Acoustics, Speech and Signal Processing* (IEEE) pp 6645–9
- [30] Zwillinger D and Kokoska S 1999 *CRC Standard Probability and Statistics Tables and Formulae* (Boca Raton, FL: CRC Press)
- [31] Fagg A H, Ojakangas G W, Miller L E and Hatsopoulos N G 2009 Kinetic trajectory decoding using motor cortical ensembles *IEEE Trans. Neural Syst. Rehabil. Eng.* **17** 487–96
- [32] Flint R D, Rosenow J M, Tate M C and Slutsky M W 2016 Continuous decoding of human grasp kinematics using epidural and subdural signals *J. Neural Eng.* **14** 016005
- [33] Simonyan K, Vedaldi A and Zisserman A 2013 Deep inside convolutional networks: visualising image classification models and saliency maps (arXiv:[1312.6034](https://arxiv.org/abs/1312.6034))
- [34] Li R, Potter T, Huang W and Zhang Y 2017 Enhancing performance of a hybrid EEG-fnirs system using channel selection and early temporal features *Front. Hum. Neurosci.* **11** 462
- [35] Lima I R, Haar S, Di Grassi L and Faisal A A 2020 Neurobehavioural signatures in race car driving: a case study *Sci. Rep.* **10** 1–9
- [36] Haar S and Faisal A A 2020 Brain activity reveals multiple motor-learning mechanisms in a real-world task *Front. Hum. Neurosci.* **14** 354
- [37] Chiarelli A M, Croce P, Merla A and Zappasodi F 2018 Deep learning for hybrid EEG-fNIRS brain–computer interface: application to motor imagery classification *J. Neural Eng.* **15** 036028
- [38] Khan M U and Hasan M A 2020 Hybrid EEG-fNIRS BCI fusion using multi-resolution singular value decomposition (MSVD) *Front. Hum. Neurosci.* **14** 536
- [39] Ghonchi H, Fateh M, Abolghasemi V, Ferdowsi S and Rezvani M 2020 Deep recurrent-convolutional neural network for classification of simultaneous EEG–fnirs signals *IET Signal Process.* **14** 142–53
- [40] Ortega P, Colas C and Faisal A A 2018 Compact convolutional neural networks for multi-class, personalised, closed-loop EEG-BCI *7th IEEE Int. Conf. on Biomedical Robotics and Biomechatronics (Biorob)* (IEEE) pp 136–41
- [41] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition (*Proc. Conf. Computer Vision and Pattern Recognition*) pp 770
- [42] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L and Polosukhin I 2017 Attention is all you need *Advances in Neural Information Processing Systems* pp 5998–6008
- [43] Shibuya K, Kuboyama N and Yamada S 2016 Complementary activation of the ipsilateral primary motor cortex during a sustained handgrip task *Eur. J. Appl. Physiol.* **116** 171–8
- [44] Di Bono M G, Begliomini C, Castiello U and Zorzi M 2015 Probing the reaching–grasping network in humans through multivoxel pattern decoding *Brain Behav.* **5** e00412
- [45] Iturrate I, Chavarriaga R, Pereira M, Zhang H, Corbet T, Leeb R and Millán J d R 2018 Human EEG reveals distinct neural correlates of power and precision grasping types *NeuroImage* **181** 635–44
- [46] Olivier E, Davare M, Andres M and Fadiga L 2007 Precision grasping in humans: from motor control to cognition *Curr. Opinion Neurobiol.* **17** 644–8
- [47] Castiello U and Begliomini C 2008 The cortical control of visually guided grasping *Neuroscientist* **14** 157–70
- [48] Mulliken G H, Musallam S and Andersen R A 2008 Forward estimation of movement state in posterior parietal cortex *Proc. Natl Acad. Sci.* **105** 8170–7
- [49] Andersen R A and Cui H 2009 Intention, action planning and decision making in parietal-frontal circuits *Neuron* **63** 568–83
- [50] Davare M, Kraskov A, Rothwell J C and Lemon R N 2011 Interactions between areas of the cortical grasping network *Curr. Opinion Neurobiol.* **21** 565–70
- [51] Serrien D J, Ivry R B and Swinnen S P 2006 Dynamics of hemispheric specialization and integration in the context of motor control *Nat. Rev. Neurosci.* **7** 160–6
- [52] Volkmann J, Schnitzler A, Witte O and Freund H-J 1998 Handedness and asymmetry of hand representation in human motor cortex *J. Neurophysiol.* **79** 2149–54
- [53] Penfield W and Boldrey E 1937 Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation *Brain* **60** 389–443
- [54] Pfurtscheller G, Brunner C, Schlogl A and Da Silva F L 2006 Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks *NeuroImage* **31** 153–9
- [55] Klimesch W, Sauseng P and Hanslmayr S 2007 EEG alpha oscillations: the inhibition–timing hypothesis *Brain Res. Rev.* **53** 63–88
- [56] Fries P 2015 Rhythms for cognition: communication through coherence *Neuron* **88** 220–35
- [57] Sadaghiani S and Kleinschmidt A 2016 Brain networks and α -oscillations: structural and functional foundations of cognitive control *Trends Cogn. Sci.* **20** 805–17
- [58] Ames K C and Churchland M M 2019 Motor cortex signals corresponding to the two arms are shared across hemispheres, mixed among neurons, yet partitioned within the population response *Elife* **8** e46159
- [59] Buzsáki G, Anastassiou C A and Koch C 2012 The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes *Nat. Rev. Neurosci.* **13** 407–20
- [60] Branco M P, de Boer L M, Ramsey N F and Vansteensel M J 2019 Encoding of kinetic and kinematic movement parameters in the sensorimotor cortex: A Brain-Computer Interface perspective *Eur. J. Neurosci.* **50** 2755–72
- [61] Branco M P, Geukes S H, Aarnoutse E J, Vansteensel M J, Freudenburg Z V and Ramsey N F 2019 High-frequency

- band temporal dynamics in response to a grasp force task *J. Neural Eng.* **16** 056009
- [62] Wang C, Ulbert I, Schomer D L, Marinkovic K and Halgren E 2005 Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity and orienting *J. Neurosci.* **25** 604–13
- [63] Womelsdorf T, Johnston K, Vinck M and Everling S 2010 Theta-activity in anterior cingulate cortex predicts task rules and their adjustments following errors *Proc. Natl Acad. Sci.* **107** 5248–53
- [64] Cavanagh J F, Cohen M X and Allen J J 2009 Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring *J. Neurosci.* **29** 98–105
- [65] Ethier C, Acuna D, Solla S A and Miller L E 2016 Adaptive neuron-to-EMG decoder training for FES neuroprostheses *J. Neural Eng.* **13** 046009
- [66] Torricelli A, Contini D, Pifferi A, Caffini M, Re R, Zucchelli L and Spinelli L 2014 Time domain functional NIRS imaging for human brain mapping *Neuroimage* **85** 28–50
- [67] Phan P, Highton D, Lai J, Smith M, Elwell C and Tachtsidis I 2016 Multi-channel multi-distance broadband near-infrared spectroscopy system to measure the spatial response of cellular oxygen metabolism and tissue oxygenation *Biomed. Opt. Express* **7** 4424–40
- [68] de Roever I, Bale G, Cooper R J and Tachtsidis I 2017 Functional NIRS measurement of Cytochrome-C-Oxidase demonstrates a more brain-specific marker of frontal lobe activation compared to the haemoglobins *Oxygen Transport to Tissue XXXIX* (Berlin: Springer) pp 141–7
- [69] Tang J, Coleman J E, Dai X and Jiang H 2016 Wearable 3-D photoacoustic tomography for functional brain imaging in behaving rats *Sci. Rep.* **6** 25470
- [70] Lantz G, De Peralta R G, Spinelli L, Seeck M and Michel C 2003 Epileptic source localization with high density EEG: how many electrodes are needed? *Clin. Neurophysiol.* **114** 63–9
- [71] Gagnon L, Perdue K, Greve D N, Goldenholz D, Kaskhedikar G and Boas D A 2011 Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling *Neuroimage* **56** 1362–71
- [72] Graves A 2012 Long short-term memory *Supervised Sequence Labelling With Recurrent Neural Networks* (Berlin: Springer) pp 37–45
- [73] Buxton R B, Wong E C and Frank L R 1998 Dynamics of blood flow and oxygenation changes during brain activation: the balloon model *Magn. Reson. Med.* **39** 855–64
- [74] Buxton R B, Uludağ K, Dubowitz D J and Liu T T 2004 Modeling the hemodynamic response to brain activation *Neuroimage* **23** S220–33