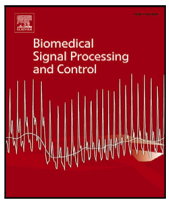
Biomedical Signal Processing and Control 110 (2025) 108256

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Characterizing motion artifacts in functional near-infrared spectroscopy signals using ground-truth movement information and computer vision

Andrea Bizzego [a,*], Alessandro Carollo [a], Seraphina Fong [a], Cesare Furlanello [b,c], Gianluca Esposito [a]

[a] Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy
[b] HK3 Lab, Rovereto, Italy
[c] LIGHT Center, Brescia, Italy

ARTICLE INFO

Dataset link: https://gitlab.com/a.bizzego/computer-vision-fnirs

Keywords:
Functional near-infrared spectroscopy
Motion artifacts correction
Computer vision
Neuroimaging
Motion artifacts

ABSTRACT

Functional near-infrared spectroscopy (fNIRS) is a preferred neuroimaging technique for studies requiring high ecological validity, allowing participants greater freedom of movement. Despite its relative robustness against motion artifacts (MAs) compared to traditional neuroimaging methods, fNIRS still faces challenges in managing and correcting these artifacts. Many existing MA correction algorithms notably lack validation on real data with ground-truth movement information. In this work, we combine computer vision, ground-truth movement data, and fNIRS signals to preliminarily characterize the association between specific head movements and MAs. Fifteen participants (age = 22.27 ± 2.62 years) took part in a whole-head fNIRS study, performing controlled head movements along three main rotational axes. Movements were categorized by axis (vertical, frontal, sagittal), speed (fast, slow), and type (half, full, repeated rotation). Experimental sessions were video recorded and analyzed frame-by-frame using the SynergyNet deep neural network to compute head orientation angles. Maximal movement amplitude and speed were extracted from head orientation data, while spikes and baseline shifts were identified in the fNIRS signals. Results showed that head orientation and movement metrics extracted via computer vision closely aligned with participant instructions. Additionally, repeated as well as upward and downward movements tended to compromise fNIRS signal quality. The occipital and pre-occipital regions were particularly susceptible to MAs following upwards or downwards movements, whereas the temporal regions were most affected by bend left, bend right, left, and right movements. These findings underscore the importance of cap adherence and fit in the relationship between movements and MAs. Overall, the present work lays the foundation for an automated approach to developing and validating fNIRS MA correction algorithms.

1. Introduction

Functional near-infrared spectroscopy (fNIRS) is an optical method for studying brain activity by measuring changes in blood oxygenation levels. fNIRS leverages the near-infrared (NIR) optical window to transmit NIR light from light sources to detectors through various biological tissues [1,2].

In recent years, there has been a surge in fNIRS usage compared to traditional neuroimaging methods, such as electroencephalography and functional magnetic resonance imaging, due to its distinctive advantages [3,4]. Unlike other techniques, fNIRS allows participants to remain mobile rather than confined to a scanner, and it is less sensitive to motion artifacts (MAs). Additionally, its portability and cost-effectiveness make it particularly suitable for conducting ecologically valid and participant-friendly experiments (e.g., [5,6]).

Although fNIRS is less sensitive to movement than traditional neuroimaging methods, effectively managing MAs in fNIRS signals is crucial for drawing reliable scientific conclusions [7]. This is particularly pertinent in naturalistic experimental paradigms, which often involve significant movement components. Two primary approaches have emerged in addressing this challenge: one advocates avoiding corrections to the collected signals, while the other proposes correcting them using specific criteria (e.g., changes in the amplitude and frequency of the signal) [8]. Opting not to correct the collected signals ensures the use of the original data without modifications. However, this conservative approach demands a large number of experimental trials and a minimum occurrence of MAs. Conversely, correcting signals with MAs allows for retaining all the available neuroimaging data, which are typically challenging and time-consuming to collect [8].

* Corresponding author.
E-mail address: andrea.bizzego@unitn.it (A. Bizzego).

https://doi.org/10.1016/j.bspc.2025.108256
Received 16 December 2024; Received in revised form 11 April 2025; Accepted 23 June 2025
Available online 10 July 2025

However, this approach strongly depends on the reliability of the methods used to objectively identify and correct MAs.

While identification and correction methods for MAs are pivotal in fNIRS-based neuroimaging experiments, achieving a consensus on the optimal approach to use remains challenging [9,10]. Currently, two primary methods are widely regarded as the gold standard for correcting MAs: spline interpolation [11] and wavelet filtering [12]. Spline interpolation is particularly effective in addressing shifts in fNIRS signals, whereas wavelet filtering excels in mitigating spikes within the signals.

The application of specific corrections for fNIRS signals is often guided by intrinsic signal characteristics or external parameters, such as accelerometer data [13]. However, uncertainty remains about whether these methods effectively target MAs or unintentionally alter genuine brain activity data. Additionally, even when the suggested criteria are applied, some MAs may go undetected and remain uncorrected. This challenge arises from the lack of a definitive ground truth on how movements manifest in fNIRS-recorded brain signals. Instead, most criteria and correction methods rely on assumptions about the hemodynamic brain response or expected experimental outcomes for their development and validation. Consequently, the evaluation of MAs identification and correction methods largely relies on simulated or theoretically expected brain data (e.g., [14,15]), with limited assessment on real neural signals.

To address the lack of ground truth, some studies have explored the association between movements and artifacts in recorded signals. However, these investigations are often constrained by a scarcity of controlled movements and ground-truth information. In this regard, the use of computer vision based on artificial intelligence holds promise in annotating movements from videos and obtaining ground-truth data to elucidate their impact on brain signals collected (e.g. [16,17]). For example, [18] demonstrated the feasibility of real-time electroencephalography artifact annotation by using computer vision to detect blinks and head movements, reinforcing the value of visual data for identifying motion-related artifacts.

Another line of research adopts a theoretical approach to characterize the statistical properties of MAs and, based on this characterization, to define robust statistical models for estimating brain responses [19–22]. The main advantage of this approach is that it minimizes the need for extensive pre-processing of the fNIRS signal. In fact, the processing can be limited to the conversion of raw data into changes in oxy- and deoxyhemoglobin concentrations, as the statistical model is designed to be resilient to MAs and other sources of noise.

To summarize, existing studies have primarily focused on characterizing MAs, evaluating how MAs affect signal quality, often without accounting for the specific characteristics of the underlying movements, or developing algorithms for detecting and correcting MAs, frequently without the support of ground-truth data regarding their presence and magnitude. In the present study, we examine the relationship between specific types of head movements and the resulting signal quality, aiming to understand how different movement patterns impact signal integrity.

The findings of this study can provide valuable insights for the development of more effective MA detection and correction algorithms by identifying when head movements are most likely to compromise signal quality and which regions of the head are most sensitive to such artifacts.

In a previous feasibility study [23], computer vision techniques were shown to effectively annotate movements from head orientation signals extracted from videos. The current work aims to extend those findings by integrating neuroimaging data. Specifically, the present study utilizes controlled movement in conjunction with artificial intelligence-derived movement metrics to investigate their manifestation in brain signals recorded via fNIRS. By integrating neuroimaging data with computer vision, this study aims to further establish an artificial intelligence-guided approach to identify and correct MAs.

2. Methods

2.1. Study design

The study merges neuroimaging data with cutting-edge computer vision algorithms to investigate the impact of controlled movements on the fNIRS signal. Participants took part in an experimental session where they were instructed to perform specific head movements, recorded via a webcam while their brain activity was simultaneously measured using fNIRS. Movements were annotated from the video recording using computer vision, and objective measures of the magnitude and speed of the movements were computed. The impact of MA was quantified in terms of spike amplitude and baseline shift on the portions of fNIRS signals affected by the movements. The relation between types of movements, movement characteristics, and MA in brain data was then investigated.

Data collection was approved by the University of Trento (2023-054). The conduction of the experiment followed the guidelines provided by the Declaration of Helsinki. Informed consent was obtained from all participants.

2.2. Participants

A total of 15 participants ($n = 8$ females, *Mean age* = 22.27, *Standard Deviation age* = 2.62) were recruited for the study. The sample size was chosen to be comparable to that of other preliminary studies on MAs in fNIRS signals (e.g., [19,20]). Participants were recruited among university students via convenience and snowball sampling facilitated through social media sites and a dedicated recruiting platform. All participants reported no known or diagnosed health or neurological conditions.

2.3. Experimental procedure

During the experiment, participants were asked to mimic a series of head movements displayed in videos on a monitor in front of them. While they performed these movements, brain activity (NIRSport2, NIRx; 10.2 Hz), physiological signals (AIM Physiological Monitor, CGX; 500 Hz), and a video recording (C310 Webcam, Logitech; 30 Hz) of the participants were acquired. The data streams from the different devices were synchronized using the Lab Streaming Layer (LSL). Physiological signals are not considered in this study.

The experimental video began with an instructional segment, prompting participants to "Please, imitate the movements represented in the video". An avatar then demonstrated the movements to be replicated, accompanied by text descriptions of the actions at the bottom of the screen (e.g., "Move your head backward slowly"). Each movement example was followed by a 3-second countdown, after which participants were given 7 s to perform the required movement.

In the experimental video, the avatar demonstrated two categories of movements: head movements and facial expressions, but only the former are considered in this study. Head movements involved head rotations along the three main axes: vertical, frontal, and sagittal. To facilitate the association between the axes and the types of head movements, the following names will be used instead of the anatomical names of the axes: "Left/Right" axis instead of "vertical" axis (yaw rotations); "Up/Down" axis instead of "frontal" axis (pitch rotations); and "bendLeft/bendRight" axis instead of "sagittal" axis (roll rotations).

Four types of rotations were considered: half rotation (both directions, i.e., left, right), full rotation, and repeated rotation (three times). Participants were instructed to end each movement with their head facing forward in a neutral position (with no residual rotations). Additionally, head movements along the Left/Right and Up/Down axes were performed at two speeds: fast and slow. Consequently, the entire experiment encompassed a total of 20 head movements: 4 types of rotations and two speeds for the Left/Right and Up/Down axes (4

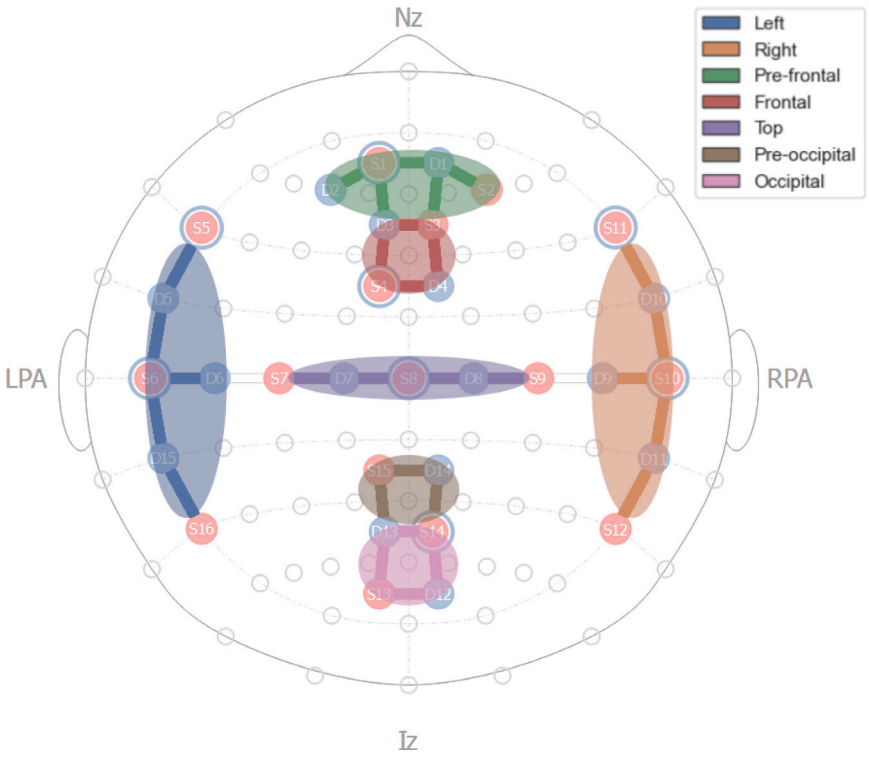


Fig. 1. Functional near-infrared spectroscopy (fNIRS) cap setup used in the study. Light sources are indicated in red, while detectors are shown in blue. The blue circles around the light sources represent short-distance channels. Channels are grouped by color to represent the regions of interest used in the analysis of the study.

types ×2 speeds ×2 axes = 16 movements), and 4 types of rotations (only slow) for the bendLeft/bendRight axis. The sequence of head movements was the same for all participants. The administration of the videos depicting the movements to mimic was controlled with *PsychoPy*.

2.4. Data acquisition and pre-processing

Brain activity was monitored using fNIRS. Each fNIRS cap was equipped with 16 LED sources emitting NIR light at wavelengths of 760 nm and 850 nm, along with 16 light detectors. Among these, one detector collected signals from 8 short-distance channels. The placement of optodes followed the standard international 10–20 electroencephalography layout, aiming to achieve comprehensive coverage of the participant's brain activity. This arrangement resulted in a total of 32 channels (see Fig. 1).

Optode stabilizers were employed to maintain a consistent distance between sources and detectors, ensuring that it never exceeded 3 cm. This setup was crucial for achieving a robust signal-to-noise ratio [24]. Participants' head circumference was measured prior to the experiment, and the appropriate fNIRS cap was selected accordingly. The fNIRS data were acquired using a NIRSport2 device (NIRx Medical Technologies LLC), operating at a sampling rate of 10.17 Hz.

The raw fNIRS signals were then converted to relative densities of oxygenated and deoxygenated hemoglobin, based on the modified Beer–Lambert Law [25], and resampled to 10 Hz. Then, the 32 channels were grouped into seven regions of interest: Left, Right, Pre-frontal, Frontal, Top, Pre-occipital, and Occipital (Fig. 1). For each region of interest, a unique brain activity signal was computed by averaging the individual region channels.

The video recording was processed frame-by-frame (Fig. 2A) by the SynergyNet deep neural network model to compute the head orientation angles [23,26] on the three axes: $\gamma_x, \gamma_y, \gamma_z$. SynergyNet is a convolutional neural network designed for face orientation estimation and was among the top-performing models on the benchmark dataset Annotated Facial Landmarks in the Wild [27].

The detected head orientation angles from all video frames were concatenated to obtain the head orientation signal, which was then resampled at 10 Hz. Subsequently, manual annotation of each movement onset and offset was performed based on the head orientation signal. The data processing was based on the *pyphysio* [28], and *opencv-python* [29] Python packages.

2.5. Movement and motion artifact metrics

The objective of this study was to characterize how different head movements impact the quality of the fNIRS data. To this aim, two types of quantitative metrics were computed (Fig. 2B): movement metrics to characterize the head movement, and MA metrics to describe the effect of the movement on the fNIRS signal.

The movement metrics were the maximal speed (Max Speed) of the head rotation and the maximal amplitude (Max Amplitude) of the head rotation, both computed based on the value of the norm of the head orientation vector at each frame:

$$\|\boldsymbol{\gamma}(n)\| = \sqrt{\gamma_x(n)^2 + \gamma_y(n)^2 + \gamma_z(n)^2}; \tag{1}$$

where n is the sample, and $\gamma_x(n), \gamma_y(n), \gamma_z(n)$ are the three components of the orientation vector that are computed frame-by-frame by SynergyNet.

Specifically, the Max Speed was computed as the maximal value of the norm of the orientation vector:

$$\max_{n_{on} \le n \le n_{off}} \left\| \frac{\partial \gamma(n)}{\partial n} \right\|. \tag{2}$$

The Max Amplitude was computed as the maximal value of the norm between the onset (n_{on}) and offset (n_{off}) of the movement:

$$\max_{n_{on} \le n \le n_{off}} \|\gamma(n)\|. \tag{3}$$

The MA metrics focused on two different types of MAs: spikes and baseline shifts. The Spike metric was computed as the range of the fNIRS signal, considering the portion between the onset and offset of the movement:

$$\max_{n_{on} \leq n \leq n_{off}} s(n) - \min_{n_{on} \leq n \leq n_{off}} s(n); \quad (4)$$

where $s(n)$ is the fNIRS signal.

The Shift metric was computed as the difference in the average fNIRS signal value between the 1-second portion before the movement onset and the 1-second portion after the movement offset:

$$\mu_{n_{on-1} \leq n \leq n_{on}} - \mu_{n_{off} \leq n \leq n_{off+1}}; \quad (5)$$

where $\mu_{n_i \leq n \leq n_j}$ is the average of the fNIRS signal between samples n_i and n_j.

To allow for comparing the MA metrics from different regions, which differ in terms of baseline values and range of variation of the signals, the MA metrics were normalized based on the 5-second portion before the movement onset. The Spike metric was divided by the standard deviation of the signal in that portion, and the Shift metric was divided by the average of the absolute value of the signal in that portion.

For all metrics, the average across the three-movement repetitions was finally computed.

2.6. Data analysis

The data analysis aimed to highlight the effects of different types of movements on MAs by comparing the previously described movement and MA metrics. The following steps were performed:

(a) Validation of the video-based measurement of head orientation and movement metrics;
(b) Investigation of the characteristics of MAs across different types of movements;
(c) Analysis of the relationship between movement metrics and MA metrics.

The study focused on providing a descriptive and statistical characterization of the relationship between head movement and MAs. Specifically, we addressed the following research questions:

1. Are certain types of head movements more prone to causing MAs than others?
2. Are MAs more prevalent in specific brain regions? If so, which regions are most sensitive, and which head movements have the greatest impact?
3. Does the magnitude of MAs vary based on the speed or amplitude of the movement?
4. Can practical recommendations be derived to experimentally reduce the occurrence of MAs?

Linear mixed models were employed to account for both fixed and random effects in the data [30]. In each model, the inclusion of random effects – such as participant ID or movement axis – allowed us to model variability associated with these factors, thereby accounting for participant-specific and context-specific differences.

The dataset and the analytical scripts are available at https://gitlab.com/a.bizzego/computer-vision-fnirs.

3. Results

3.1. Movement specificity

In the first step of our analysis, we implemented two linear mixed models with movement metrics (i.e., Max Speed and Max Amplitude) as the dependent variables. Both models included the target axis (i.e., target *vs.* non-target), speed (i.e., slow *vs.* fast), and movement type (i.e., half *vs.* complete *vs.* repeated) as fixed effects. Participant ID and movement axis were included as random effects. Overall, the results demonstrate that the movement metrics obtained using the SynergyNet model closely matched the ground-truth movement information provided by the experimental instructions. Specifically, we observed that participants' movements primarily impacted the main rotational axis specified in the instructions. Significantly higher Max Speed was observed along the target axis – the axis along which the movement was intended – compared to the other two axes (β = 8.25, Standard Error (SE) = 0.49, $t(653.99) = 16.72$, p <.001; Fig. 3A). Similarly, Max Amplitude was significantly greater on the target axis compared to the other axes (β = 36.51, SE = 0.94, $t(654.00) = 38.94$, p <.001; Fig. 3B).

Moreover, participants executed movements at two distinct speeds as instructed. Fig. 3C shows that increased speed along the target axis occurred only in accordance with the experimental instructions. Specifically, Max Speed on the target axis was significantly lower when participants were instructed to move slowly, compared to when they were instructed to move fast ($\beta = -4.39$, SE = 0.52, $t(655.74) = -8.46$, p <.001). However, instructing participants to perform movements at both fast and slow speeds had no effect on the amplitude of their movements on the target axis ($\beta = -1.81$, SE = 0.99, $t(655.97) = -1.83$, p = .067; Fig. 3D).

The type of movement (half, complete, or repeated) also influenced movement speed (Fig. 3E). Specifically, Max Speed was significantly lower during half movements compared to complete movements ($\beta = -1.96$, SE = 0.57, $t(653.99) = -3.45$, p <.001). Conversely, Max Speed was significantly higher during repeated movements compared to complete movements (β = 2.14, SE = 0.57, $t(653.99) = 3.76$, p <.001). Therefore, faster movements were observed when participants performed repeated movements, followed by complete movements, and then half movements. The type of movement also influenced movement amplitude (Fig. 3F), although the effects were less pronounced than those observed for speed. Specifically, Max Amplitude was significantly lower during half movements compared to complete movements ($\beta = -3.07$, SE = 1.08, $t(654.00) = -2.83$, p = .005). No significant difference in Max Amplitude was observed between repeated and complete movements (β = 1.47, SE = 1.08, $t(654.00) = 1.36$, p = .174). Therefore, reduced movement amplitude was primarily associated with half movements, while repeated and complete movements showed comparable amplitude values.

Finally, a positive association between movement amplitude and speed along the target axis is observed, for both movement speeds (Fig. 4). To assess the overall relationship between Max Speed and Max Amplitude, we performed a linear mixed model with Max Amplitude as the dependent variable, and Max Speed, target speed, and their interaction as fixed effects. Participant ID, movement axis, and movement type were included as random effects. The analysis revealed a statistically significant positive relationship between Max Amplitude and Max Speed (β = 1.03, SE = 0.10, $t(186.66) = 9.84$, p <.001). Although no significant main effect of target speed on Max Amplitude was found, a significant interaction emerged: the relationship between Max Speed and Max Amplitude was more strongly positive during slow movements compared to fast movements (β = 0.39, SE = 0.14, $t(207.67) = 2.79$, p = .006).

Overall, patterns observed in the movement metrics and obtained through the SynergyNet model are consistent with the type of controlled movements that participants were asked to perform during the experiment. This supports the idea that the video-based detection of the head orientation is a promising approach that can help identify MAs during experiments with fNIRS.

3.2. Shifts and spikes in fNIRS signals

To investigate the distribution MAs across movements, we implemented two linear mixed models with MA metrics (i.e., Shift and Spike)

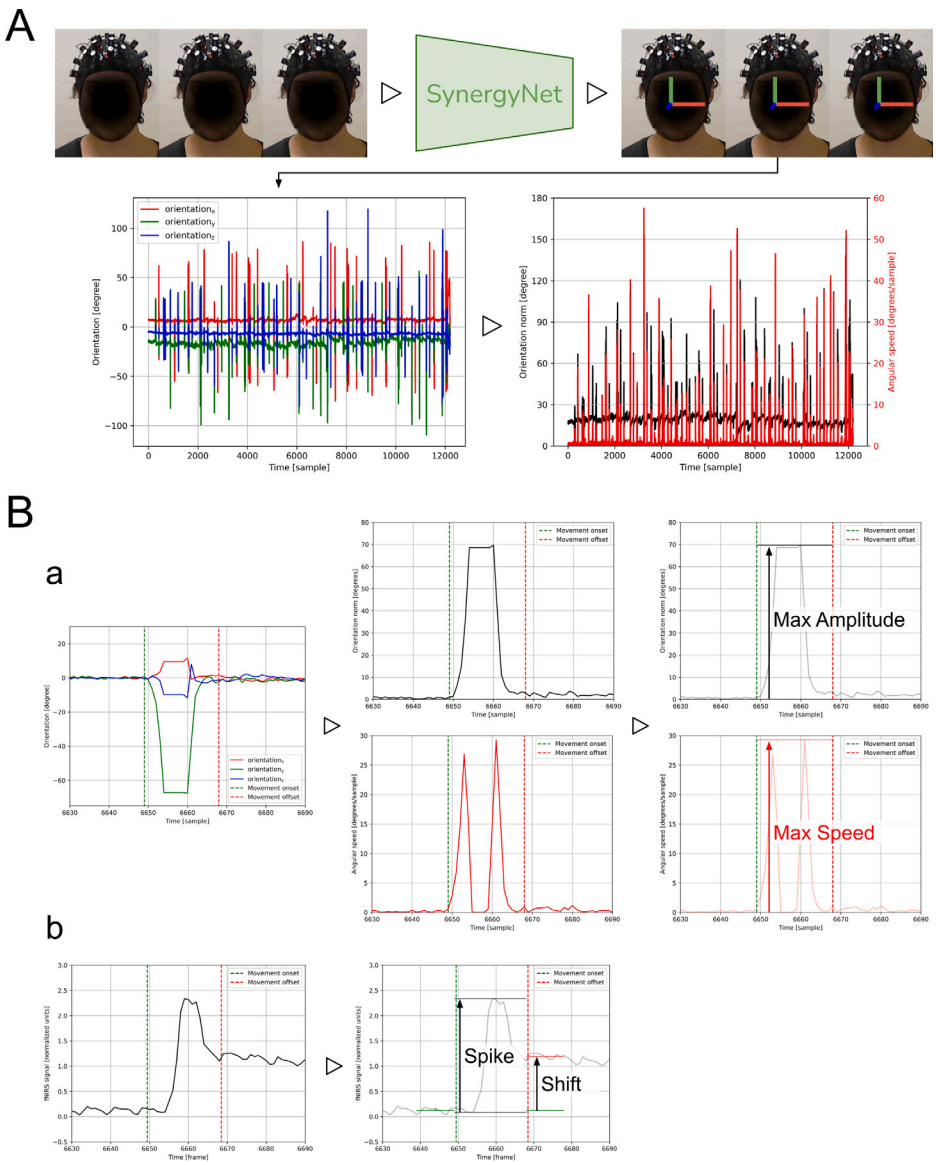


Fig. 2. Representation of the main signal processing procedures applied. **(A)** Extraction of movement signals from the video. The video frames were processed by the SynergyNet network to obtain the three components of the head orientation: γ_x (red), γ_y (green), and γ_z (blue). The components obtained from each frame were concatenated to obtain the three components of the orientation signal, from which the norm of the orientation vector and the angular speed (norm of the gradient) were computed. **(B)** Computation of the metrics to characterize the movement characteristics from the movement signals (orientation and speed; **a**) and the MAs from the fNIRS signal (**b**).

as the dependent variables. Both models included the rotational axis (i.e., Left/Right *vs.* Up/Down *vs.* bendLeft/bendRight), speed (i.e., slow *vs.* fast), and movement type (i.e., half *vs.* complete *vs.* repeated) as fixed effects. Participant ID and region of interest were included as random effects. Overall, significantly higher Shift values were observed during Up-Down movements compared to movements along the other two rotational axes (β = 0.09, SE = 0.01, $t(1549.00) = 6.21$, p <.001; Fig. 5A). Similarly, Spike values were significantly higher during Up-Down movements compared to the other axes (β = 1.02, SE = 0.07, $t(1549.00) = 14.17$, p <.001; Fig. 5B).

Moreover, no statistically significant differences in Shift or Spike values were observed across different movement speeds. In particular, slow movements did not generate a significantly different amount of Shift (β = −0.01, SE = 0.01, $t(1549.00) = -0.68$, p = .498; Fig. 5C) or Spike (β = −0.08, SE = 0.07, $t(1549.00) = -1.17$, p = .244; Fig. 5D) compared to fast movements.

Generally, the type of movement (i.e., half, complete, or repeated) had a slight impact on Shift and Spike values detected in the signals. Specifically, repeated movements generated significantly higher Shift values as compared to other movement type (β = 0.04, SE = 0.02, $t(1549.00) = 2.23$, p = .026; Fig. 5E), while half movements produced significantly lower Spike values (β = −0.34, SE = 0.08, $t(1549.00) = -4.24$, p <.001; Fig. 5F).

In summary, movements along the Up/Down axis appear to have the greatest impact in terms of shifts and spikes, especially for repeated movements. Interestingly, movement speed does not seem to influence the MA metrics.

To investigate the relationship between movements and MAs in individual regions of interest, we performed eight linear mixed models. Two models were run with one MA metric (i.e., Shift or Spike) as the dependent variable and region of interest as the only fixed effect. In these models, movement axis, participant ID, target speed, and movement type were included as random effects. The remaining six models were computed on data from individual regions of interest. Three of these models, one for each movement axis, used Shift values as the dependent variable, while the other three models used Spike values as the dependent variable. All of these models included the

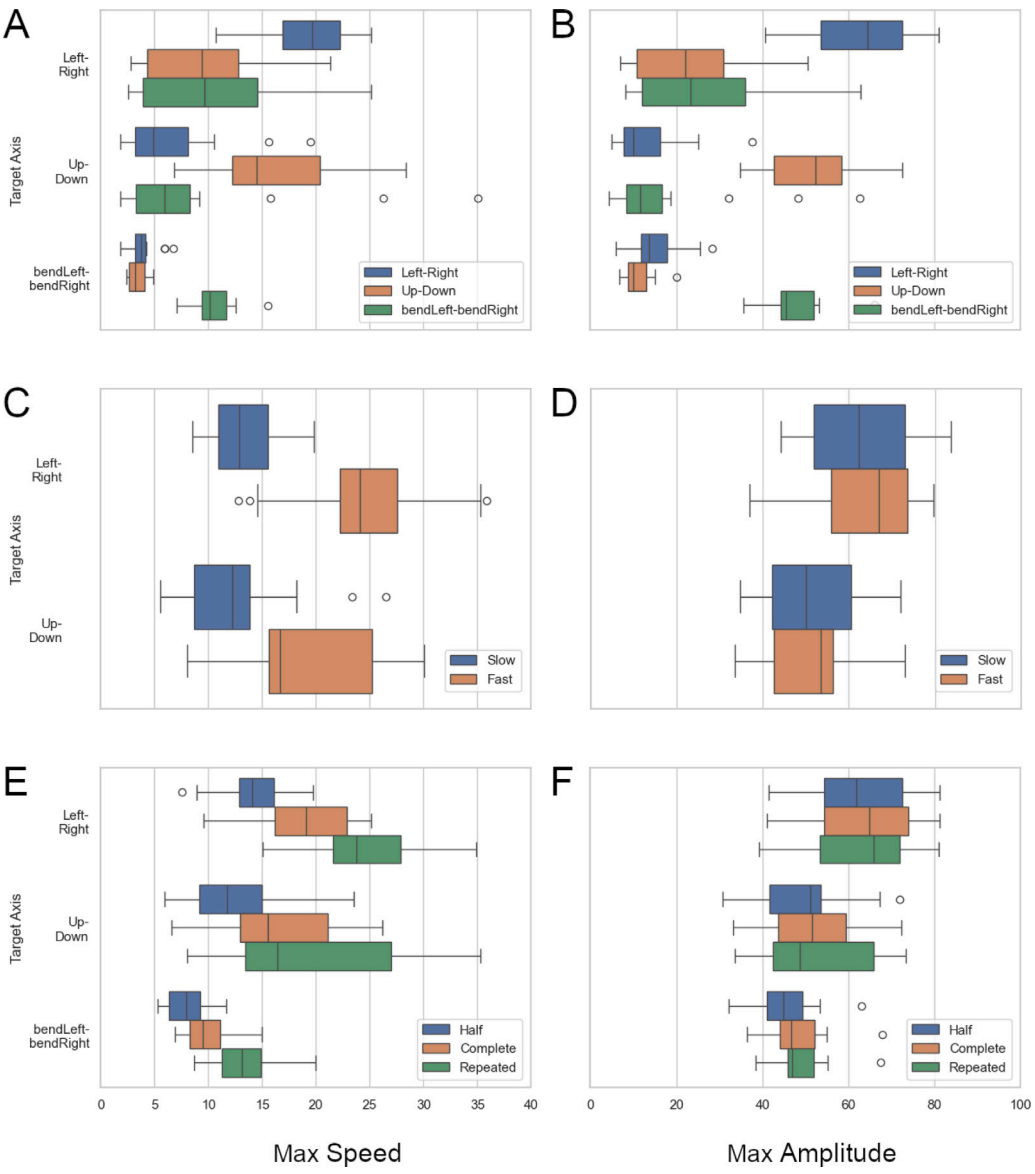


Fig. 3. Validation of movement metrics extracted from the videos using the SynergyNet deep neural network model [23,26]. The *x*-axis represents the values of Max Speed and Max Amplitude. The *y*-axis includes the movement axes on which the participant was instructed to move their head. **(A)** Max Speed of movements on the three axes. **(B)** Max Amplitude of movements on the three axes. **(C)** Max Speed of movement on the target axis across fast and slow movements. **(D)** Max Amplitude of movement on the target axis across fast and slow movements. **(E)** Max Speed of movement on the target axis across types of movement (i.e., half, complete, repeated). **(F)** Max Amplitude of movement on the target axis across types of movement (i.e., half, complete, repeated).

region of interest as the fixed effect, and participant ID, target speed, and movement type as random effects.

Overall, the results indicated that, across movement axes, significantly higher Shift values were observed in the right (β = 0.05, SE = 0.02, t(1550.00) = 2.28, p = .023), top (β = 0.05, SE = 0.02, t(1550.00) = 2.07, p = .039), pre-occipital (β = 0.37, SE = 0.02, t(1550.00) = 15.22, p <.001), and occipital brain regions (β = 0.20, SE = 0.02, t(1550.00) = 8.22, p <.001). Similarly, significantly higher Spike values were found in the top (β = 0.61, SE = 0.12, t(1548.91) = 5.04, p <.001), pre-occipital (β = 2.04, SE = 0.12, t(1548.91) = 16.91, p <.001), and occipital regions (β = 1.39, SE = 0.12, t(1548.91) = 11.51, p <.001).

The models conducted on individual regions of interest revealed distinct patterns of MAs depending on the movement performed. Specifically, the pre-occipital and occipital brain regions showed significantly higher Shift values compared to other regions during Left/Right (pre-occipital: β = 0.24, SE = 0.04, t(292.00) = 5.87, p <.001; occipital: β = 0.15, SE = 0.04, t(292.00) = 3.65, p <.001), Up/Down (pre-occipital: β = 0.50, SE = 0.04, t(606.00) = 11.86, p <.001; occipital: β = 0.31, SE = 0.04, t(606.00) = 7.32, p <.001), and bendLeft/bendRight movements (pre-occipital: β = 0.24, SE = 0.04, t(292.00) = 5.87, p <.001; occipital: β = 0.15, SE = 0.04, t(292.00) = 3.65, p <.001). Significantly higher Shift values were observed in brain regions at the top of the brain during Up/Down movements (β = 0.09, SE = 0.04, t(606.00) = 2.16, p = .032).

With regard to Spike values, the right (β = 0.20, SE = 0.09, t(606.00) = 2.38, p = .018), pre-occipital (β = 0.89, SE = 0.09, t(606.00) = 10.41, p <.001), and occipital brain regions (β = 0.36, SE = 0.09, t(606.00) = 4.23, p <.001) were the most significantly affected regions of interest during Left/Right movements. Similarly, top (β = 1.17, SE = 0.24, t(606.00) = 4.98, p <.001), pre-occipital (β = 3.73, SE = 0.24, t(606.00) = 15.84, p <.001), and occipital brain regions (β = 2.81, SE = 0.24, t(606.00) = 11.95, p <.001) were the most significantly affected regions of interest during Up/Down movements. Finally, during bendLeft/bendRight movements, we observed significantly higher Spike values in left (β = 0.38, SE = 0.14,

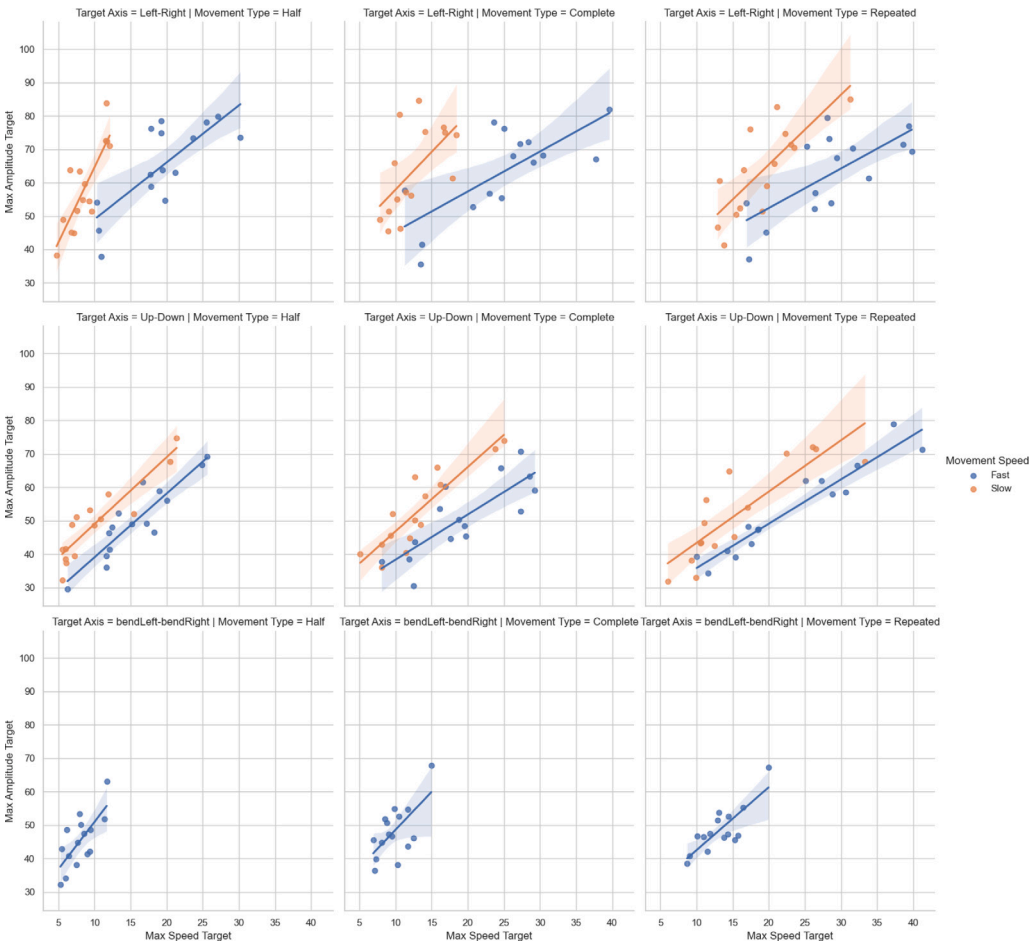


Fig. 4. Relationship between Max Amplitude and Max Speed of movements on the target axis across all movements performed in the experiment.

$t(292.00) = 2.75$, $p = .006$), right ($\beta = 0.51$, SE = 0.14, $t(292.00) = 3.70$, $p < .001$), top ($\beta = 0.39$, SE = 0.14, $t(292.00) = 2.80$, $p = .006$), pre-occipital ($\beta = 0.99$, SE = 0.14, $t(292.00) = 7.12$, $p < .001$), and occipital ($\beta = 0.61$, SE = 0.14, $t(292.00) = 4.40$, $p < .001$) brain regions of interest.

Overall, channels situated in the occipital and pre-occipital regions of the cap exhibited greater sensitivity to MAs, both in terms of Shift and Spike values, particularly during Up/Down movements (Figs. 6A and 6B). Additionally, after bendLeft/bendRight and Left/Right movements, we noted a difference in Shift and Spike values between the left and right regions of interest (Figs. 6A and 6B).

These patterns suggest that changes in the shape and adherence of the fNIRS cap to the participant's head might play a significant role in generating MAs, rather than just the movement itself. In fact, during Up/Down movements, the occipital and pre-occipital parts of the cap are the most affected and are compressed between the head and neck. Similarly, during bendLeft/bendRight movements, the lateral parts of the cap, on the left and right sides, respectively, experience the most stress. These findings will be discussed in more detail in the Discussion section.

Finally, we estimated the relationship between MA metrics using a linear mixed model. In the model, Spike values were used as the dependent variable while Shift values, target speed, and their interaction were used as fixed effects. In the model, participant ID, movement axis, movement type, and region of interest were included as random effects. Overall, a statistically significant positive correlation between shifts and spikes across movements was observed in the experiment ($\beta = 2.99$, SE = 0.13, $t(1547.41) = 23.92$, $p < .001$; Fig. 7). Conversely, no significant main effect of speed ($\beta = 0.02$, SE = 0.07, $t(1546.51) = 0.22$, $p = .828$) and no interaction effect ($\beta = -0.33$, SE = 0.18, $t(1549.58) = -1.85$, $p = .064$) were observed.

3.3. Movements and motion artifacts

The analysis of the relation between movement and MA metrics across movements reveals no consistent pattern. In particular, we computed four linear mixed models to test the association between individual movement (i.e., Max Speed and Max Amplitude) and MA (i.e., Shift and Spike) metrics. Each model included one of the two MA metrics as the dependent variable and one of the two movement metrics as the fixed effect. All models included target speed, participant ID, movement axis, movement type, and region of interest as random effects.

No statistically significant effect was observed between Max Speed and Shift values ($\beta = 0.001$, SE = 0.001, $t(483.80) = 0.681$, $p = .497$; Fig. 8). Conversely, we observed a statistically significant negative relationship between Max Speed and Spike values ($\beta = -0.015$, SE = 0.007, $t(117.63) = -2.099$, $p = .038$; Fig. 9).

Concerning the maximum amplitude of movements and MA metrics, no statistically significant pattern emerged. In particular, the relationship between Max Amplitude and Shift values (see Fig. 10) was not statistically significant across movements ($\beta = 0.001$, SE = 0.001, $t(911.30) = 1.545$, $p = .123$). Moreover, there is no discernible trend regarding the relation between Max Amplitude of movements and Spike

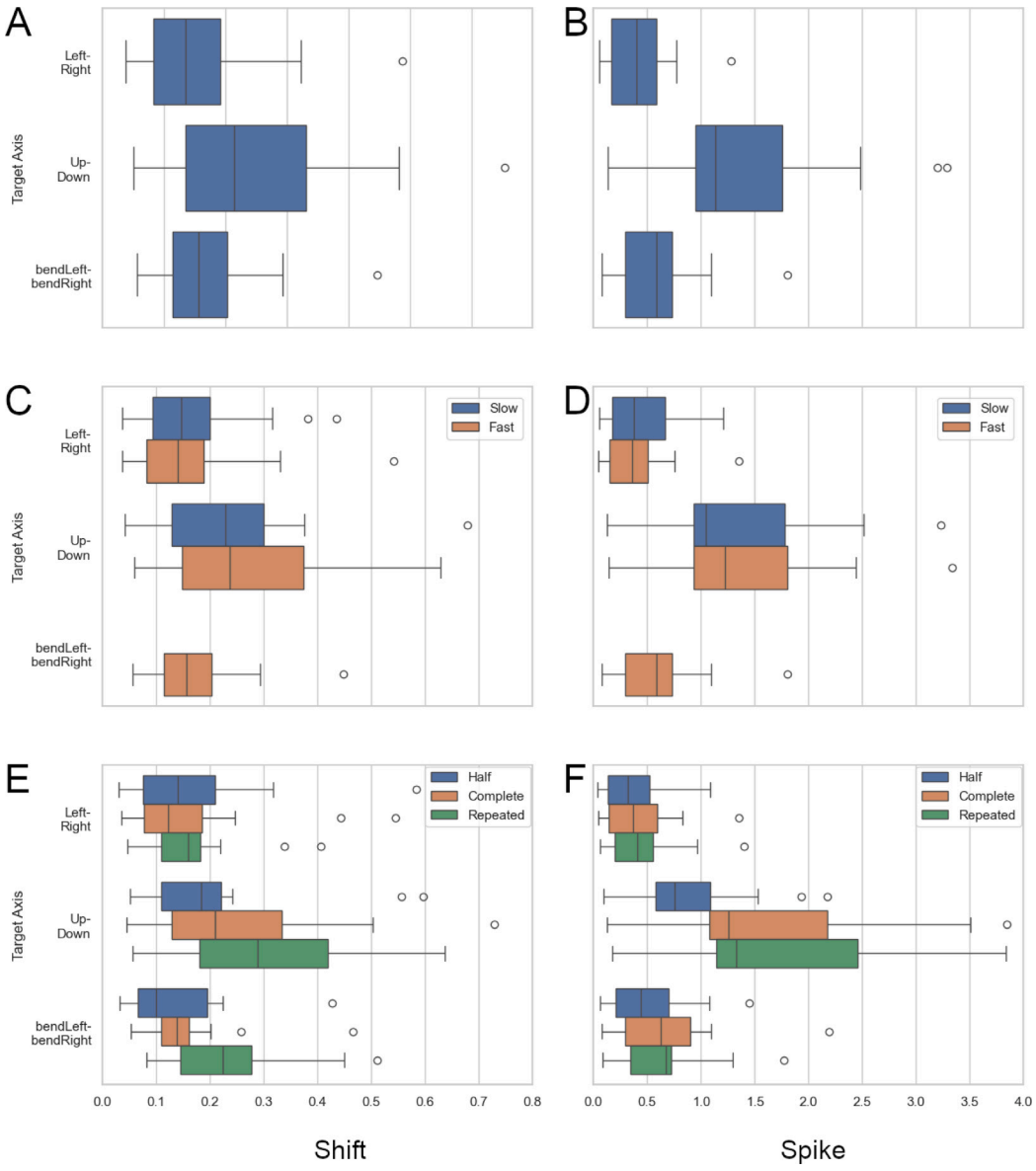


Fig. 5. Motion artifact metrics (i.e., Shift and Spike values) across movements. **(A)** Shift values after movements on the three main rotational axes. **(B)** Spike values after movements on the three main rotational axes. **(C)** Shift values after movements on the target axis across fast and slow movements. **(D)** Spike values after movements on the target axis across fast and slow movements. **(E)** Shift values after movements on the target axis across types of movement (i.e., half, complete, repeated). **(F)** Spike values after movements on the target axis across types of movement (i.e., half, complete, repeated).

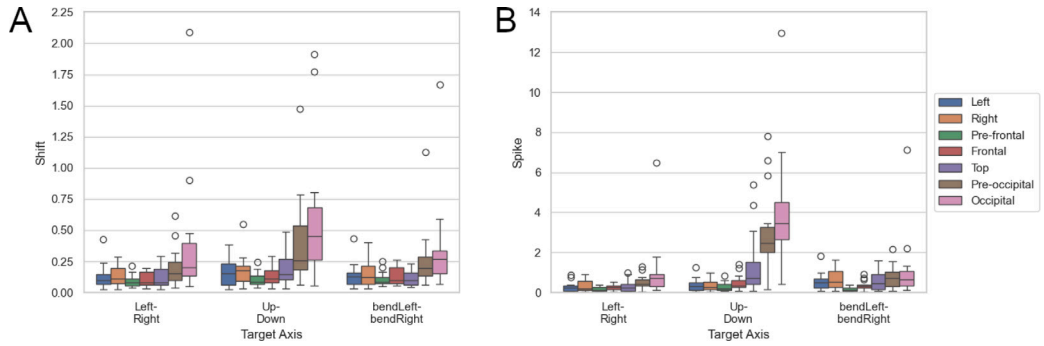


Fig. 6. Motion artifact metrics (i.e., Shift and Spike values) across movements and regions of interest. **(A)** Shift values across regions of interest depending on the target axis. **(B)** Spike values across regions of interest depending on the target axis.

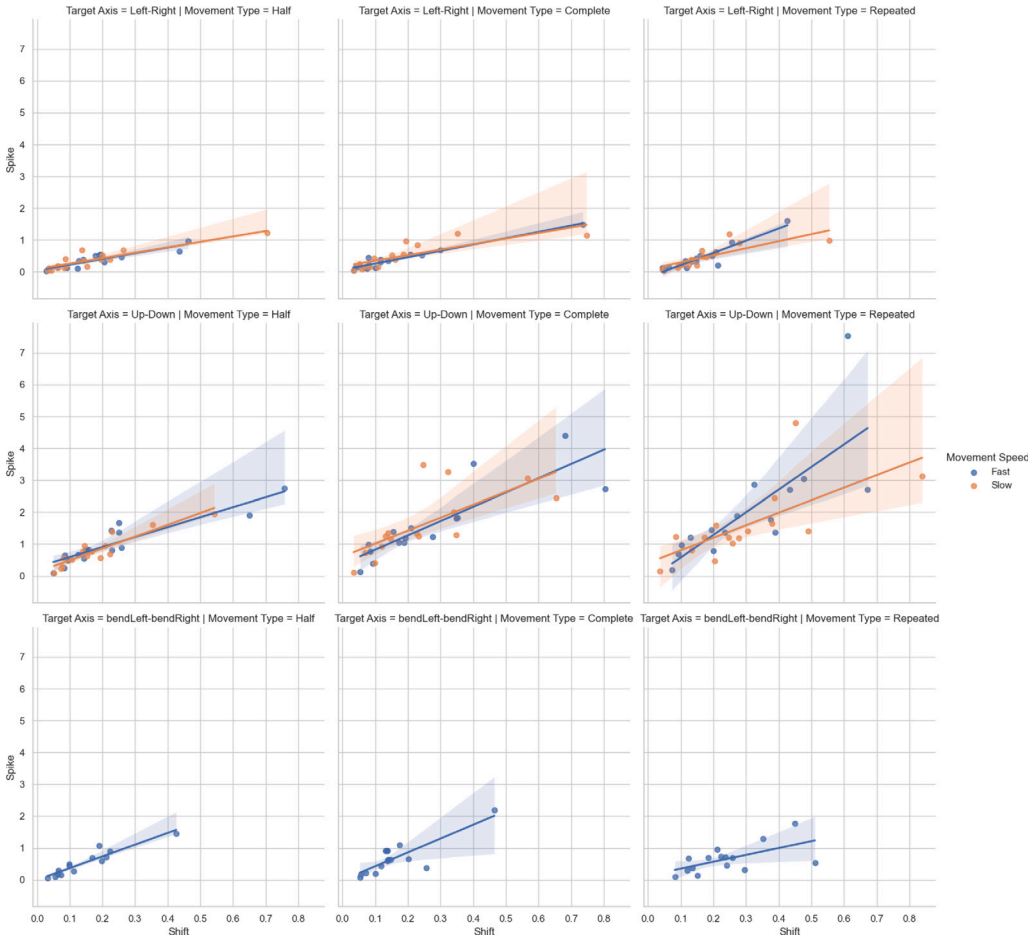


Fig. 7. Relationship between motion artifact metrics (Shift and Spike values) across all movements performed in the experiment.

values ($\beta = -0.002$, SE = 0.004, $t(1224.00) = -0.471$, $p = .638$; Fig. 11).

4. Discussion

Despite fNIRS being reportedly less sensitive to participants' movements, reaching a consensus on the treatment of MAs remains pivotal [7]. One key challenge in evaluating the performance of available MAs correction strategies is the lack of ground-truth movement information. In the current study, we combined ground-truth movement data with computer vision techniques to evaluate MAs in fNIRS signals. Participants performed controlled movements while their brain activity was monitored using fNIRS. Computer vision was utilized to extract essential movement metrics from session videos [23].

Our results show a high correlation between movement metrics (i.e., Max Speed and Amplitude) extracted using computer vision and ground-truth movement data (provided by experimental instructions). This result serves a dual function: confirming participants' compliance with task instructions and validating the precision of metrics generated by artificial intelligence models in representing experimental movements. As shown in our feasibility study [23], the latter aspect is particularly significant as it illustrates how artificial intelligence models can aid in annotating movements in neuroimaging experiments and potentially leverage this information when correcting for MAs in the signals. Video-based motion detection carries several advantages over traditional manual annotations or the use of external sensors (e.g., accelerometer) [23]. Firstly, it is minimally invasive, allowing for the collection of movement data without requiring participants to wear additional devices that might interfere with their natural behavior. Secondly, runtime algorithms can be adopted to preserve the anonymity of participants, ensuring that personal identification is not compromised during the data collection process. In its current state, this approach requires validation using ground-truth data, such as data collected from inertial sensors (e.g., [31]), to ensure accuracy and reliability.

Unlike existing methods that primarily focus on the statistical properties of MAs (e.g., [19–22]), our study establishes a direct connection between the occurrence of MAs and specific, identifiable head movement patterns. By combining computer vision with neuroimaging data, we uncovered an association between MAs (such as Shift and Spike values) and specific head movements. Notably, complete and repetitive Up/Down movements resulted in higher Shift and Spike values, particularly in the occipital and pre-occipital brain regions. Additionally, Left/Right movements predominantly affected the corresponding Left and Right regions of interest, with the amplitude and speed of these movements being related to Shift values. These patterns suggest that a degradation of the cap fit, rather than the movement alone, might play a significant role in corrupting the fNIRS signals [13,32]. Our findings therefore indicate that in order to collect high-quality data, the fNIRS cap must adhere well to the participant's head. During Up/Down movements, the occipital part of the cap gets compressed between the head and the neck. Similarly, during Left/Right movements, the inertia of the participant's hair might cause the cap to move, reducing its adherence to the head. These shifts in cap positioning can result in unreliable readings, ultimately undermining the accuracy of the data. Thus, we recommend that researchers carefully assess the fit of the fNIRS cap before each experimental session to ensure that it remains

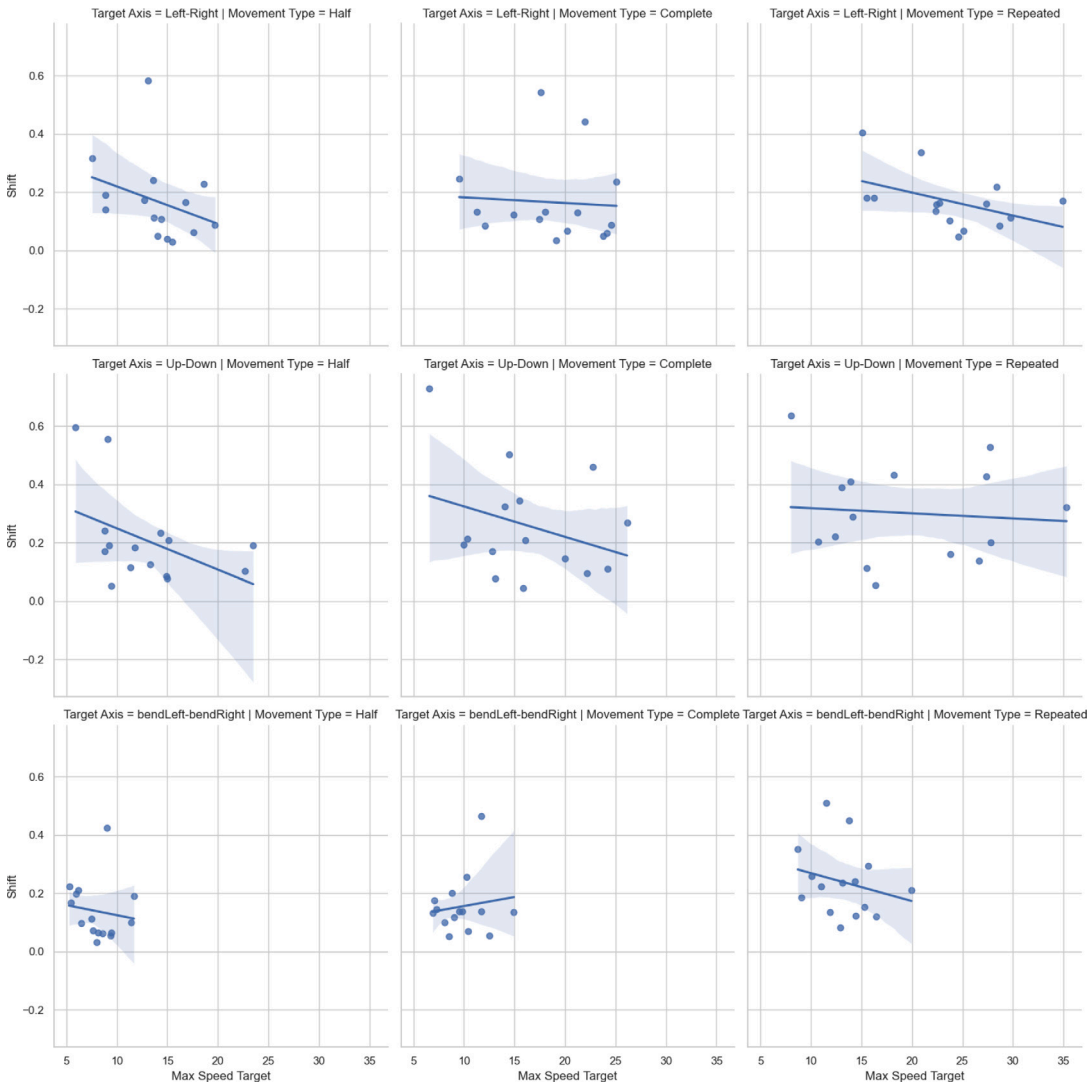


Fig. 8. Relationship between Max Speed of movements on the target axis and Shift values across all movements performed in the experiment.

secure throughout the course of an experiment. Doing so can potentially minimize the risk of artifacts caused by cap displacement, thereby improving the overall quality of the collected data.

This study introduces a quantitative computer vision approach for extracting movement patterns, in contrast to traditional methods that often rely on accelerometers (e.g., [13]). While accelerometers provide information on the magnitude of movement, they lack the detailed spatial and kinematic data that our vision-based system is able to capture. This study highlights the potential of computer vision to enhance the pre-processing of fNIRS signals for MA correction. Existing correcting methods frequently rely on simulated and expected brain responses [14,15], and often lack evaluation against ground-truth movement information regarding actual movements. By accurately capturing and analyzing head movements, computer vision techniques offer a valuable tool to identify and mitigate the impact of MAs, ultimately improving the reliability and accuracy of neuroimaging data.

A key strength of our study lies in the use of ground-truth information regarding the specific types of head movements performed by participants. This represents a major improvement over prior studies, which often depend on theory-driven assumptions or simulated data to infer how MAs affect signal quality. By leveraging empirically validated movement data, our approach enables a more robust and realistic analysis of the relationship between head motion and signal quality. This paves the way for the development of more accurate artifact identification methods and potentially more effective correction strategies.

This manuscript provides initial insights into which movements generate more significant MAs and which brain regions are most susceptible to such artifacts under specific movement conditions. These findings can guide experiment design by helping researchers identify movements that participants should avoid to optimize signal quality in the regions of interest. For example, greater movement control is essential when assessing brain activity in the occipital and pre-occipital regions. Similarly, although to a lesser extent, limiting Left/Right head movements is important when evaluating temporal Left and Right regions. To mitigate movement-related artifacts, researchers can consider incorporating visual cues or gentle reminders throughout the experimental procedure to encourage participants to maintain a stable head position. These cues can be strategically integrated into the setup to promote minimal and controlled movement, thereby reducing the likelihood of MAs. By incorporating these insights into experimental protocols, researchers can design experiments that minimize movement-induced artifacts, thereby improving the quality and reliability of fNIRS data.

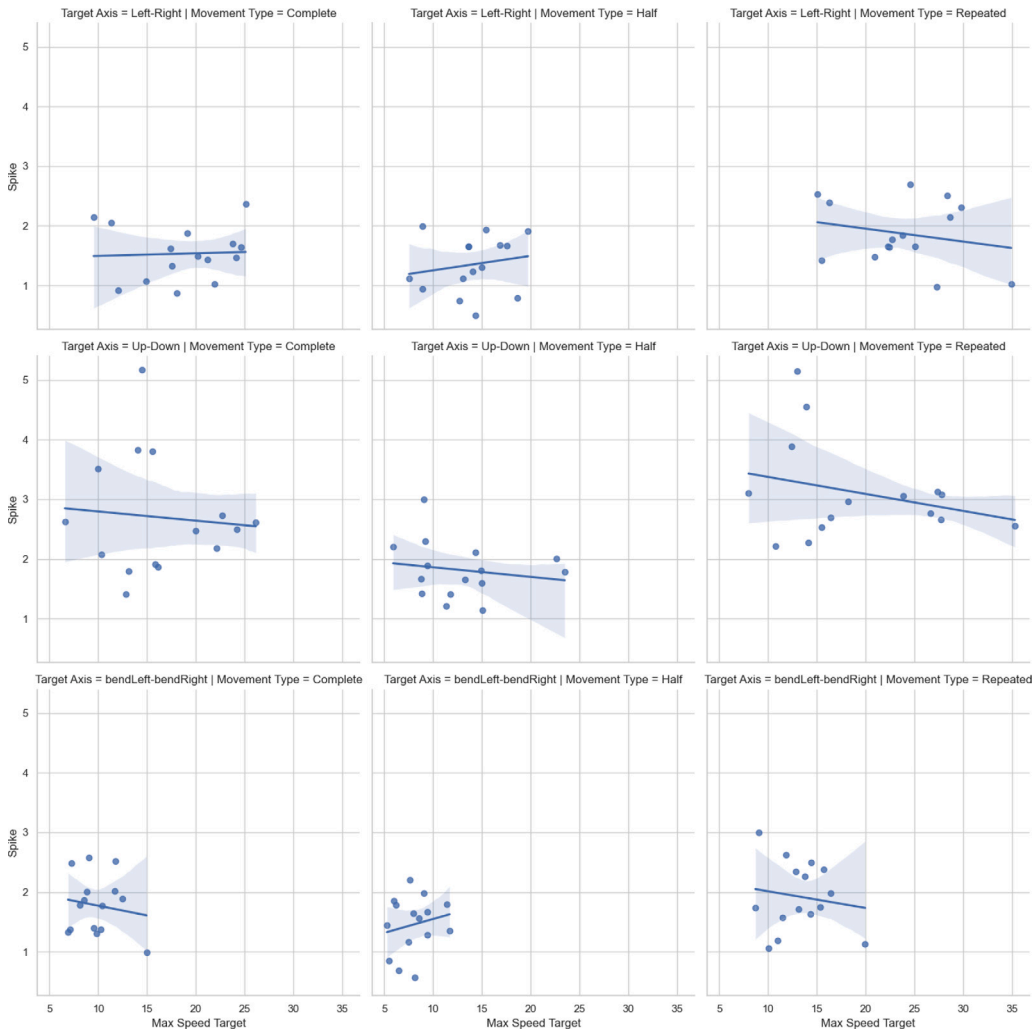


Fig. 9. Relationship between Max Speed of movements on the target axis and Spike values across all movements performed in the experiment.

4.1. limitations of the study

The limited sample size in this study may constrain the generalizability of the findings. However, it is important to note that the computer vision-based approach we adopted is highly scalable and can be readily applied to larger samples, as it only requires video recordings of participants' movements. Future studies could leverage this method to investigate movement and MAs in experimental settings involving more participants and potentially less controlled movement conditions.

Another limitation of the study is the potential for biases in the computer vision-based movement detection, which may have influenced the results. Although the model employed achieved state-of-the-art performance in head pose estimation on a benchmark dataset [27], further research is needed to assess its reliability and accuracy in other experimental research settings. That said, computer vision is a rapidly evolving field, driven by advancements in artificial intelligence, and it is likely that newer and more efficient models will soon become available for use in similar applications.

Finally, although cap sizes were selected to properly fit each participant's head, our interpretation of the differences in MAs characteristics across head regions, particularly the greater susceptibility observed in occipital areas, requires further investigation. Future studies could benefit from incorporating photogrammetry-based monitoring of optode positioning to better understand and control for these regional differences [17].

5. Conclusion

In this work, we introduced an artificial intelligence-guided approach to characterize the impact of specific head movements on fNIRS signals. Building on our previous study [23], we demonstrated that state-of-the-art computer vision algorithms can effectively extract real-time head orientation and movement information from video recordings of neuroimaging sessions. Our findings indicate that certain head movements – particularly repeated movements and Up/Down movements – significantly compromise the quality of fNIRS signals, largely independent of the movement speed. The occipital and pre-occipital brain regions are particularly susceptible to MAs following Up/Down movements, while the temporal regions are most affected by bendLeft/bendRight and Left/Right movements. These results suggest that the primary source of MAs may be related to alterations in the cap's adherence or shape, rather than the movement itself, emphasizing the importance of addressing cap fit and stability to minimize MAs in fNIRS studies [13,32]. By integrating state-of-the-art computer vision techniques with neuroimaging data, this work lays the groundwork for an artificial intelligence-driven development, evaluation, and application of MA correction algorithms for fNIRS signals. We argue that the adoption of automated methods in fNIRS pre-processing and processing pipelines can enhance the reproducibility and cross-comparability of fNIRS data and results [33].

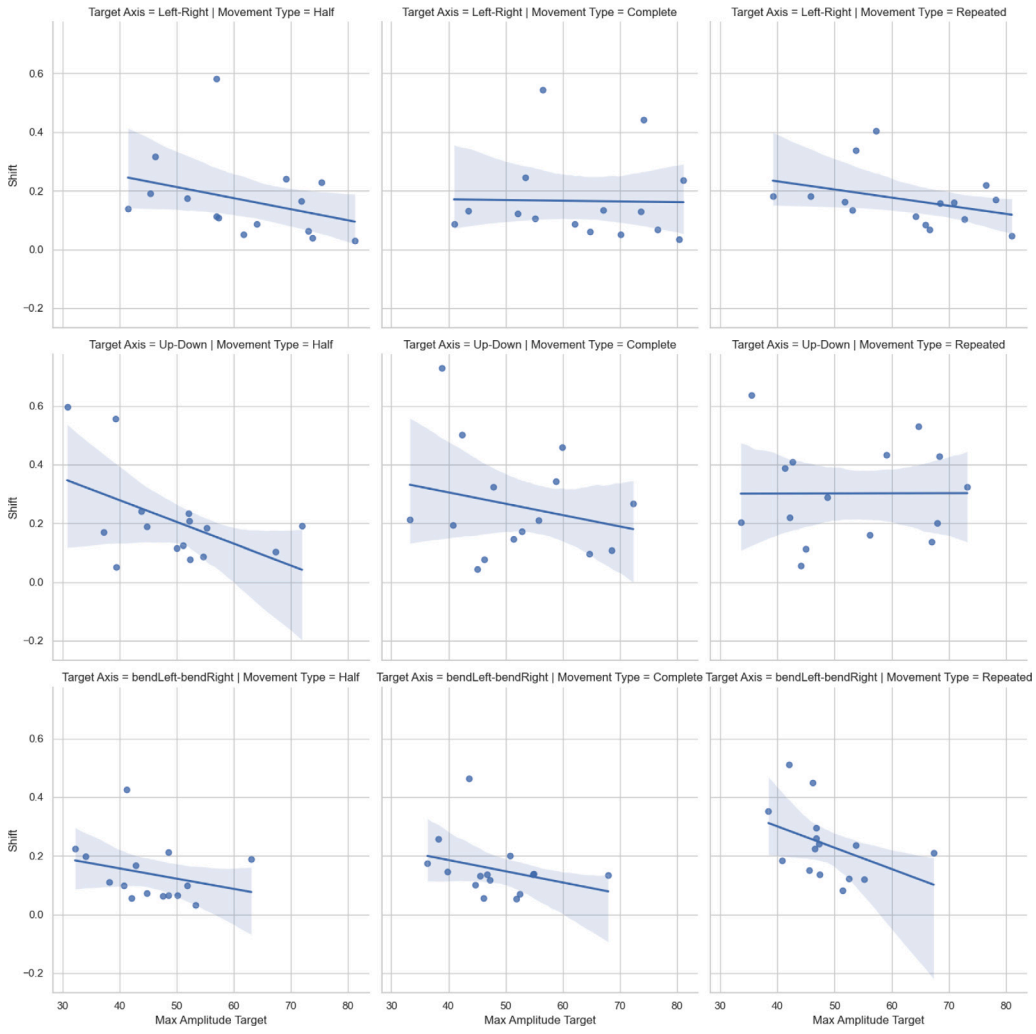


Fig. 10. Relationship between Max Amplitude of movements on the target axis and Shift values across all movements performed in the experiment.

Abbreviations

The following abbreviations are used in this manuscript:

fNIRS	functional near-infrared spectroscopy
LSL	Lab Streaming Layer
MA	motion artifact
NIR	near-infrared

CRediT authorship contribution statement

Andrea Bizzego: Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Alessandro Carollo:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Seraphina Fong:** Writing – review & editing, Methodology. **Cesare Furlanello:** Writing – review & editing. **Gianluca Esposito:** Writing – review & editing, Supervision.

Funding

A.B. was funded by the European Union - FSE-REACT-EU, PON Research and Innovation 2014–2020 DM1062/2021.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrea Bizzego reports financial support was provided by European Union. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Alessia Chistè, Nicole Dossi, Andrea Englaro, Dorina Shermadhi, Soraija Testerini, and Lucrezia Torre for their support in conducting the experiment and recruiting the participants.

Data availability

The dataset and the analytical scripts are available at https://gitlab.com/a.bizzego/computer-vision-fnirs.

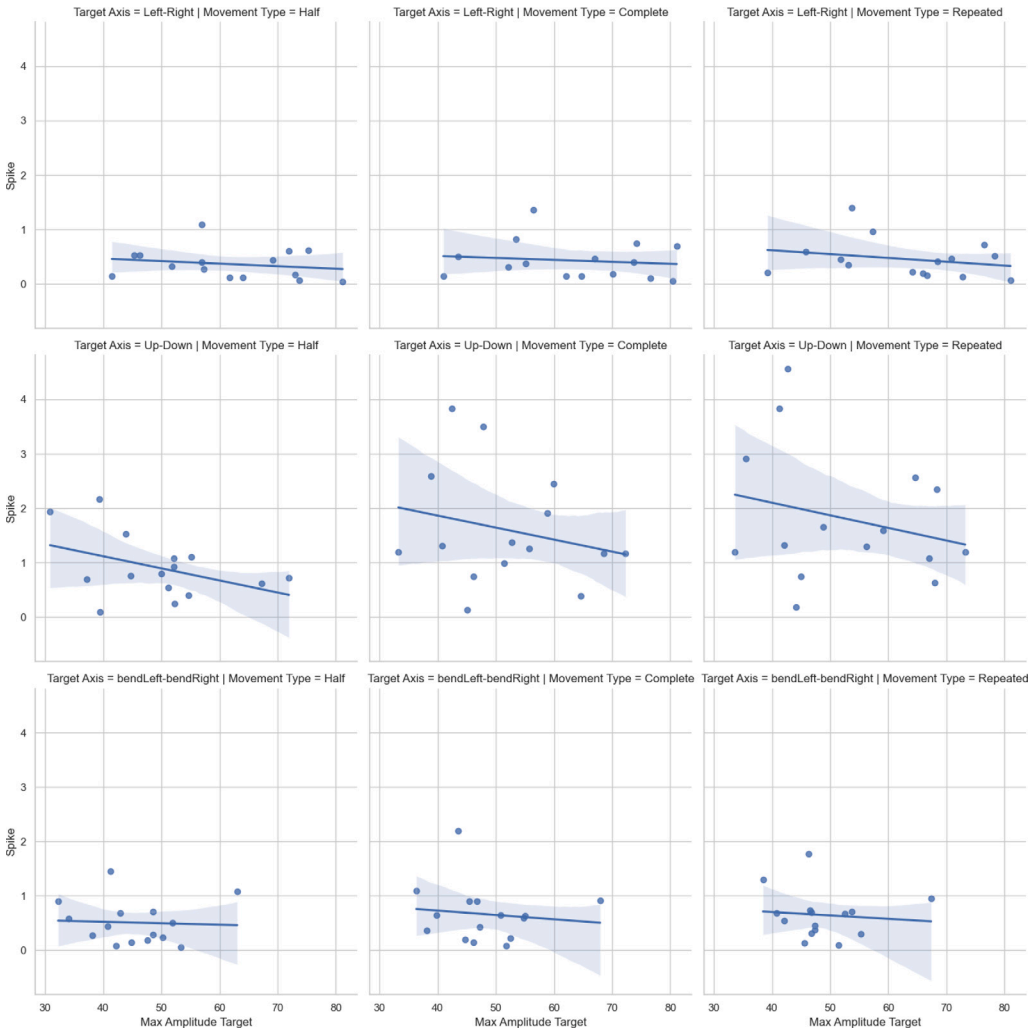


Fig. 11. Relationship between Max Amplitude of movements on the target axis and Spike values across all movements performed in the experiment.

References

[1] F.F. Jöbsis, Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters, Science 198 (4323) (1977) 1264–1267.

[2] P. Pinti, I. Tachtsidis, A. Hamilton, J. Hirsch, C. Aichelburg, S. Gilbert, P.W. Burgess, The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience, Ann. New York Acad. Sci. 1464 (1) (2020) 5–29.

[3] A. Carollo, I. Cataldo, S. Fong, O. Corazza, G. Esposito, Unfolding the real-time neural mechanisms in addiction: Functional near-infrared spectroscopy (fNIRS) as a resourceful tool for research and clinical practice, Addict. Neurosci. 4 (2022) 100048.

[4] A. Carollo, G. Esposito, Hyperscanning literature after two decades of neuroscientific research: A scientometric review, Neuroscience 551 (2024) 345–354.

[5] A. Bizzego, G. Gabrieli, A. Azhari, M. Lim, G. Esposito, Dataset of parent-child hyperscanning functional near-infrared spectroscopy recordings, Sci. Data 9 (1) (2022) 625.

[6] M. Lim, A. Carollo, A. Bizzego, A.S. Chen, G. Esposito, Culture, sex and social context influence brain-to-brain synchrony: an fNIRS hyperscanning study, BMC Psychol. 12 (1) (2024) 350.

[7] D. Perpetuini, D. Cardone, C. Filippini, A.M. Chiarelli, A. Merla, A motion artifact correction procedure for fNIRS signals based on wavelet transform and infrared thermography video tracking, Sensors 21 (15) (2021) 5117.

[8] S. Brigadoi, L. Ceccherini, S. Cutini, F. Scarpa, P. Scatturin, J. Selb, L. Gagnon, D.A. Boas, R.J. Cooper, Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data, Neuroimage 85 (2014) 181–191.

[9] A. Bizzego, J.P.M. Balagtas, G. Esposito, Commentary: Current status and issues regarding pre-processing of fNIRS neuroimaging data: An investigation of diverse signal filtering methods within a general linear model framework, Front. Hum. Neurosci. 14 (2020) 247.

[10] P. Pinti, F. Scholkmann, A. Hamilton, P. Burgess, I. Tachtsidis, Current status and issues regarding pre-processing of fNIRS neuroimaging data: an investigation of diverse signal filtering methods within a general linear model framework, Front. Hum. Neurosci. 12 (2019) 505.

[11] F. Scholkmann, S. Spichtig, T. Muehlemann, M. Wolf, How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation, Physiol. Meas. 31 (5) (2010) 649.

[12] B. Molavi, G.A. Dumont, Wavelet-based motion artifact removal for functional near-infrared spectroscopy, Physiol. Meas. 33 (2) (2012) 259.

[13] J. Virtanen, T. Noponen, K. Kotilahti, J. Virtanen, R.J. Ilmoniemi, Accelerometer-based method for correcting signal baseline changes caused by motion artifacts in medical near-infrared spectroscopy, J. Biomed. Opt. 16 (8) (2011) 087005.

[14] M. Kim, S. Lee, I. Dan, S. Tak, A deep convolutional neural network for estimating hemodynamic response function with reduction of motion artifacts in fNIRS, J. Neural Eng. 19 (1) (2022) 016017.

[15] Y. Gao, H. Chao, L. Cavuoto, P. Yan, U. Kruger, J.E. Norfleet, B.A. Makled, S. Schwaitzberg, S. De, X. Intes, Deep learning-based motion artifact removal in functional near-infrared spectroscopy, Neurophotonics 9 (4) (2022) 041406.

[16] X.-S. Hu, N. Wagley, A.T. Rioboo, A.F. DaSilva, I. Kovelman, Photogrammetry-based stereoscopic optode registration method for functional near-infrared spectroscopy, J. Biomed. Opt. 25 (9) (2020) 095001.

[17] I. Mazzonetto, M. Castellaro, R.J. Cooper, S. Brigadoi, Smartphone-based photogrammetry provides improved localization and registration of scalp-mounted neuroimaging sensors, Sci. Rep. 12 (1) (2022) 10862.

[18] X. Qian, M. Wang, X. Wang, Y. Wang, W. Dai, Intelligent method for real-time portable EEG artifact annotation in semiconstrained environment based on computer vision, Comput. Intell. Neurosci. 2022 (1) (2022) 9590411.

[19] A. Aarabi, T.J. Huppert, Characterization of the relative contributions from systemic physiological noise to whole-brain resting-state functional near-infrared spectroscopy data using single-channel independent component analysis, Neurophotonics 3 (2) (2016) 025004.

[20] J.W. Barker, A. Aarabi, T.J. Huppert, Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS, Biomed. Opt. Express 4 (8) (2013) 1366–1379.
[21] T.J. Huppert, Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy, Neurophotonics 3 (1) (2016) 010401.
[22] H. Santosa, A. Aarabi, S.B. Perlman, T.J. Huppert, Characterization and correction of the false-discovery rates in resting state connectivity using functional near-infrared spectroscopy, J. Biomed. Opt. 22 (5) (2017) 055002.
[23] A. Bizzego, A. Carollo, B. Senay, S. Fong, C. Furlanello, G. Esposito, Computer vision-driven movement annotations to advance fNIRS pre-processing algorithms, Sensors 24 (21) (2024) 6821.
[24] G.E. Strangman, Z. Li, Q. Zhang, Depth sensitivity and source-detector separations for near infrared spectroscopy based on the Colin27 brain template, PLoS One 8 (8) (2013) e66319.
[25] W.B. Baker, A.B. Parthasarathy, D.R. Busch, R.C. Mesquita, J.H. Greenberg, A. Yodh, Modified Beer-Lambert law for blood flow, Biomed. Opt. Express 5 (11) (2014) 4053–4075.
[26] C.-Y. Wu, Q. Xu, U. Neumann, Synergy between 3dmm and 3d landmarks for accurate 3d facial geometry, in: 2021 International Conference on 3D Vision, 3DV, IEEE, 2021, pp. 453–463.
[27] M. Koestinger, P. Wohlhart, P.M. Roth, H. Bischof, Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization, in: 2011 IEEE International Conference on Computer Vision Workshops, ICCV Workshops, IEEE, 2011, pp. 2144–2151.
[28] A. Bizzego, A. Battisti, G. Gabrieli, G. Esposito, C. Furlanello, pyphysio: A physiological signal processing library for data science approaches in physiology, SoftwareX 10 (2019) 100287.
[29] J. Howse, OpenCV Computer Vision with Python, Vol. 27, Packt Publishing Birmingham, UK, 2013.
[30] V.A. Brown, An introduction to linear mixed-effects modeling in R, Adv. Methods Pr. Psychol. Sci. 4 (1) (2021) 2515245920960351.
[31] O. Pearl, S. Shin, A. Godura, S. Bergbreiter, E. Halilaj, Fusion of video and inertial sensing data via dynamic optimization of a biomechanical model, J. Biomech. 155 (2023) 111617.
[32] T. Noponen, K. Kotilahti, I. Nissilä, T. Kajava, P. Meriläinen, Effects of improper source coupling in frequency-domain near-infrared spectroscopy, Phys. Med. Biol. 55 (10) (2010) 2941.
[33] A. Bizzego, A. Carollo, M. Lim, G. Esposito, Effects of individual research practices on fNIRS signal quality and latent characteristics, IEEE Trans. Neural Syst. Rehabil. Eng. (2024).