Preliminary Hemispheric Analysis of fNIRS Signal Features during Motor Imagery Tasks

Group 17

***Abstract*— This study presents a preliminary analysis of inter-hemispheric differences in fNIRS during left-hand and right-hand motor imagery tasks. Using an open-access dataset, temporal features (mean, slope, and peak amplitude) were extracted from preprocessed oxygenated hemoglobin variations signals within selected cortical regions of interest. Inter-hemispheric comparisons were performed using paired parametric and non-parametric statistical tests based on data normality. Results revealed a statistically significant difference only for peak amplitude during the right-hand motor imagery task, while no significant differences were observed for the remaining features. Overall, these findings indicate the limited discriminative power of simple temporal features and support the need for more robust or multimodal approaches in fNIRS-based motor imagery analysis.**

***Index Terms*— fNIRS, Motor Imagery, Peak Amplitude, Brain Hemispheres, Features**

I. Introduction

FUNCTIONAL near-infrared spectroscopy (fNIRS) is a non-invasive optical neuroimaging technique that enables the measurement of task-related variations in the concentration of oxygenated hemoglobin (O_2Hb) and deoxygenated hemoglobin (HHb) in cortical brain tissue. Due to its characteristics, fNIRS offers several advantages, as it is safe, portable, relatively low-cost, and easy to use. However, these are associated with limitations, including limited penetration depth into cerebral tissue and lack of standardization in data acquisition and analysis procedures [1], [2].

In recent years, fNIRS has been increasingly employed in clinical and research contexts, particularly within the field of brain–computer interfaces (BCIs). In this framework, fNIRS-based BCIs have been investigated as potential tools to support rehabilitation [3], [4]. Despite these advances, several methodological challenges remain, especially concerning signal quality and the identification of reliable features related to task-induced brain activation.

The aim of this study is to perform a preliminary analysis to assess whether statistically significant differences in extracted fNIRS signal features exist between left and right hemispheres during right-hand and left-hand motor imagery (MI) tasks, considering each task condition separately. To this end, basic temporal features (mean value, slope and maximum peak amplitude) are extracted from the O_2Hb concentration changes over selected motor-related channels.

Specifically, the extracted temporal descriptors are compared to identify those with the greatest potential to support a future classification process, with the goal of assessing the feasibility of an fNIRS-based approach using the available dataset.

II. Materials and Methods

All data and experimental details were obtained from the provided open-access EEG+fNIRS dataset [5]; however, only fNIRS signals were considered in the present analysis.

A. Participants

A total of twenty-nine participants were recruited. Among them, twenty-eight were right-handed and one was left-handed. The sample consisted of sixteen females and thirteen males, with a mean age of 28.4 ± 3.8 years (mean ± Standard Deviation). All participants were healthy and naive to MI.

B. Data Acquisition

Recordings were performed in a well-lit room and signals were recorded with a sampling frequency of 10 Hz using a NIRScout system (NIRx GmbH, Berlin, Germany). The system included fourteen light sources and sixteen detectors, whose pairings formed a total of thirty-six measurement channels. Optodes were positioned according to the international 10–5 system, maintaining a source–detector distance of 30 *mm*. Nine channels covered the frontal region (around Fp1, Fp2, and Fpz), twelve channels targeted the motor cortex (around C3 and C4), and three channels acquired signals from the visual area (around Oz). Dual-wavelength light at approximately 760 *nm* and 850 *nm* was emitted to measure changes in O2H and HHb concentrations.

C. Experimental procedure

All participants completed sixty trials of MI, randomly alternating left-hand (LMI) and right-hand (RMI) tasks, organized into three sessions of twenty trials each. A 1-minute resting period was recorded at the beginning of each session. Each trial started with a 2-second visual cue, represented by a black arrow displayed on a screen pointing either left or right, followed by a 10-second imagery period and a rest interval randomly lasting 15–17 *s*. A short beep sound marked the beginning and the end of each imagery task. A fixation cross was displayed during both the imagery and rest periods to minimize eye movements.

Participants were instructed to perform kinesthetic motor imagery, specifically, imagining hand gripping (opening and closing the hand) at a pace of 1 *Hz*. Since motor imagery could not be directly monitored during the recordings, a real hand gripping movement was repeatedly demonstrated before the experiment to help participants maintain a constant rhythm.

D. Data Processing

All data processing was performed using MATLAB R2025a (MathWorks, Natick, MA, USA).

The participant 'S13', being the only left-handed individual in the dataset, was excluded to avoid an unbalanced and non-representative subgroup and to obtain a more homogeneous and statistically robust sample.

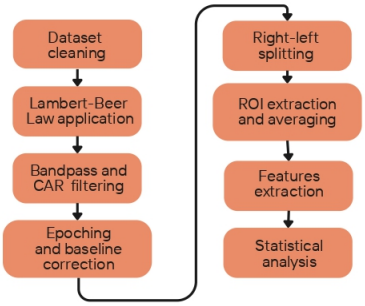


Fig. 1: Block diagram of the fNIRS signals preprocessing and analysis pipeline

As reported in Fig. 1 in the initial phase, the provided light intensity data were converted into variations of oxygenated hemoglobin (ΔO_2Hb) and deoxygenated hemoglobin (ΔHHb) concentrations by applying the modified Beer–Lambert law. The following system (1) was solved and values were expressed in *mM·mm*. It was formulated using the molar extinction coefficients [6] and the differential pathlength factor (DPF) values for brain tissue [7] at 760 *nm* and 850 *nm*, while d represents the optical path length between emitter and detector.

$$\begin{cases} \Delta A_{\lambda_1} = d \cdot B_1 \left(\alpha_{\lambda_1}^{O_2Hb} \Delta O_2Hb + \alpha_{\lambda_1}^{HHb} \Delta HHb \right) \\ \Delta A_{\lambda_2} = d \cdot B_2 \left(\alpha_{\lambda_2}^{O_2Hb} \Delta O_2Hb + \alpha_{\lambda_2}^{HHb} \Delta HHb \right) \end{cases} \quad (1)$$

Since the dataset provided only reflected light intensity values, the variation in absorbance (ΔA) was estimated by computing the optical density (OD) as the base-10 logarithm of the ratio between the mean intensity over the 60 *s* resting period preceding the first task and the intensity measured during task execution.

Subsequently, the raw signals were band-pass filtered using two cascaded Chebyshev IIR filters, applied in a double-pass configuration to compensate phase distortion. A high-pass filter with cutoff frequency 0.01 *Hz* was applied to remove baseline drifts, and a low-pass filter with cutoff frequency 0.09 *Hz* was applied to attenuate higher-frequency components mainly related to physiological and instrumental artifacts [2]. To further enhance signal quality, a Common Average Referencing (CAR) spatial filter was adopted by subtracting (separately for ΔO_2Hb and ΔHHb) the mean across all channels at each time sample, thus reducing global components shared among channels and removing physiology-based interference caused for example by blood pressure changes and increasing temperature.

Signals were then segmented into epochs corresponding to individual trials, spanning from -5 *s* to + 25 *s* relative to task onset and including: pre–task (3 *s*), instruction (2 *s*), tasks (10 *s*), and post-task (15 *s*) periods. The pre–task baseline segment was taken from the end of the preceding epoch. A baseline correction was applied by subtracting, for each epoch, the mean signal value in the interval from -5 s to -2 *s* relative to task onset. This ensured alignment to a common reference level at task onset, improving comparability across trials and subjects.

Signals were then grouped into the two task conditions (RMI and LMI).

For each subject, trials were averaged to obtain one mean signal of each channel per condition. A Grand Average (GA) across subjects was computed to identify motor cortex channels showing the strongest activation during RMI and LMI. Two regions of interest (ROIs) were defined, one for the left hemisphere (LH) and one for the right hemisphere (RH), each including three channels. From this point onward, the analysis focused on ΔO_2Hb signals, since they typically show a more pronounced and interpretable task-related response compared to ΔHHb ones.

For both LMI and RMI tasks, trials were first averaged at the channel level for each subject, obtain one signal per selected channel. These signals were then spatially averaged to derive one representative signal per ROI. From these ROI signals, three specific features were extracted: mean value, slope, and maximum peak amplitude [4].

Considering that the hemodynamic response typically peaks between +5 *s* and +15 *s* after task onset [8], and that motor imagery tasks are known to elicit delayed and more gradual responses compared to cognitive or systemic tasks, a conservative activation window ranging from +5 *s* to +22 *s* after task onset was selected. This choice is consistent with qualitative inspection of the signals and was adopted to ensure that the main task-related activation was fully captured. All features were computed within this interval. Subjects for which feature extraction was not possible due to the absence of identifiable feature values were excluded from further analysis, ensuring a clean dataset for subsequent statistical analyses.

E. Statistical test

The statistical analysis was designed to identify significant inter–hemispheric differences in the extracted features during both MI tasks.

Normality of the feature distributions was first assessed to determine the appropriate statistical test for comparisons. The Shapiro–Wilk test was applied to each feature (mean value, slope and peak amplitude) separately for each hemisphere's distribution (left and right) and for each MI condition, resulting in a total of 12 normality tests.

Comparisons (LH vs RH) were then performed using paired-sample tests, separately for the LMI and RMI conditions, as the two measurements in each comparison were obtained from the same subjects. A paired Student's *t*–test was used when the normality assumption was satisfied for both distributions in the pair. When at least one distribution deviated from normality, the non–parametric Wilcoxon signed-rank test was employed. A significance level of $\alpha = 0.05$ was adopted.

Feature distributions were visualized using boxplots to support the interpretation of the statistical comparisons.

The normality assessment was performed using a custom implementation of the Shapiro-Wilk test (function *SWTEST.p*, provided by Prof. Andrea Cereatti in a previous course). Statistical comparisons were subsequently conducted using the paired Student's *t*–test and the Wilcoxon signed-rank test, implemented via the standard MATLAB functions *ttest* and *signrank*, respectively, available in the "Statistics and Machine Learning" Toolbox.

III. Results

A. Grand average signals and ROI definition

The GA of ΔO_2Hb and ΔHHb was computed across subjects to highlight spatio–temporal activation patterns across channels. Visual inspection of the GA waveforms focused on channels over sensorimotor areas, while prefrontal and occipital channels, which are not primarily involved in MI tasks, were not considered for ROI selection.

Based on the qualitative assessment of the signal evolution relative to task onset, two groups of channels were identified. These channels exhibited the most prominent task-related hemodynamic pattern, characterized by a clear increase in ΔO_2Hb and a concurrent stability of ΔHHb, consistent with a typical fNIRS activation response [2]. The qualitative assessment considered the entire hemodynamic profile, with particular focus on the activation window defined in Section II-D.

To enable a balanced inter-hemispheric comparison, two symmetric ROIs were defined based on the GA signals. For each task, the selected ROIs consisted of three spatially adjacent channels showing the most robust and consistent task-related responses, preserving left–right topographical symmetry. As highlighted in Fig. 2a, a triplet located in the contralateral LH was selected from RMI task GA distribution, consisting of channels 20 (CP3-C3), 21 (CP3-CP1), and 24 (C1-CP1), while as shown in Fig. 2b, the RH ROI, chosen from the LMI task GA distribution, included channels 33 (CP4–C4), 32 (CP4–CP2), and 26 (C2–CP2). For both hemispheres nearby channels exhibiting weaker GA responses were not included to maintain compact and comparable ROIs.

B. Feature extractions outcomes

Within the selected ROIs, three time-domain features were analyzed from the ΔO_2Hb signals to characterize the task-related hemodynamic response. The mean value was computed as the average signal value within the 5–22 s interval, while the peak amplitude was defined as the maximum one within the same window. The slope was estimated as the angular coefficient of a first-order polynomial fitted between the relative minimum preceding the peak and the peak itself.

In cases where the signal morphology did not exhibit a clearly identifiable peak or a preceding relative minimum within the selected window, feature extraction resulted in undefined values (NaN).

Since inter-hemispheric statistical comparisons are based on paired measurements, incomplete feature values prevent the application of paired statistical tests and would invalidate the corresponding hemispheric comparison. Moreover, missing values were not imputed, as this procedure could alter the feature distributions and introduce bias in the subsequent statistical tests. For this reason subjects presenting missing feature values in either hemisphere for a given MI task were excluded from both of them.

This deletion resulted in excluding 2 subjects out of 28 for RMI and 4 subjects out of 28 for LMI, affecting respectively 7% and 14% of the dataset. While the exclusion of 14% of the dataset may appear non-negligible, this decision aimed to preserve data integrity and consistency, avoiding data imputation procedures that are not suitable for highly variable signals with unpredictable temporal patterns. The excluded subjects and the corresponding missing data sources are reported in Table I.

TABLE I: Subjects excluded from the statistical analysis due to missing feature values.

Task	Subject ID	Missing Data Source (ROI)
LMI	S5	Left ROI features
LMI	S9	Both ROIs (LH and RH)
LMI	S19	Right ROI features
LMI	S28	Left ROI features
RMI	S19	Left ROI features
RMI	S27	Right ROI features

C. Statistical test

As summarized in Table II, Shapiro–Wilk normality tests indicated that both hemispheric distributions were normal only for the mean feature in the LMI task. Therefore, this pair was analyzed using a paired Student's t-test, while for the remaining feature pairs the Wilcoxon signed-rank test was applied.

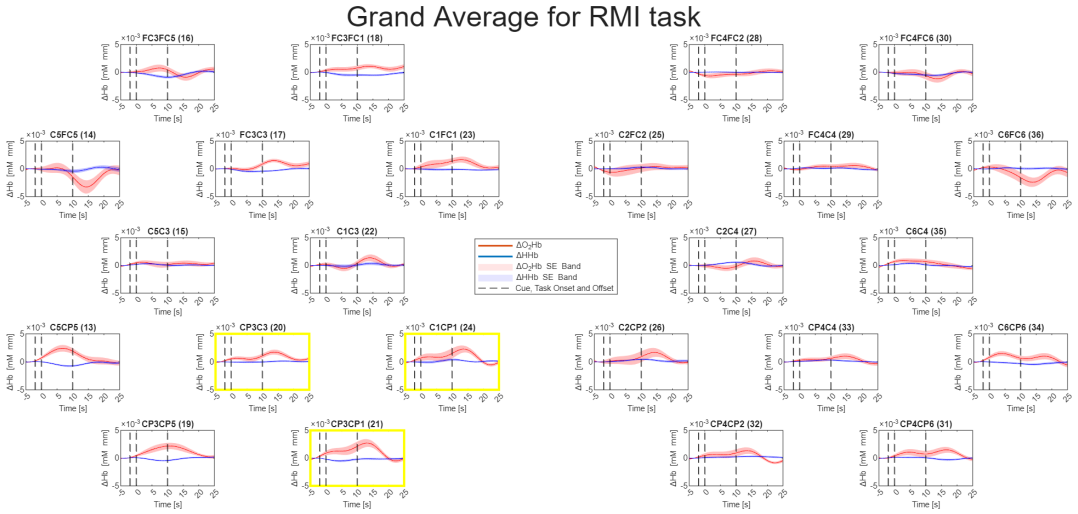
Feature distributions were visualized using paired boxplots (Fig. 3) to provide a descriptive overview of inter-hemispheric differences and distributional properties. A statistically significant inter-hemispheric difference was observed only for the peak amplitude feature during the RMI task ($p < 0.05$), whereas no significant differences were found for the remaining features.

TABLE II: Summary of normality assessment, selected statistical tests, and significance outcomes for inter-hemispheric comparisons.

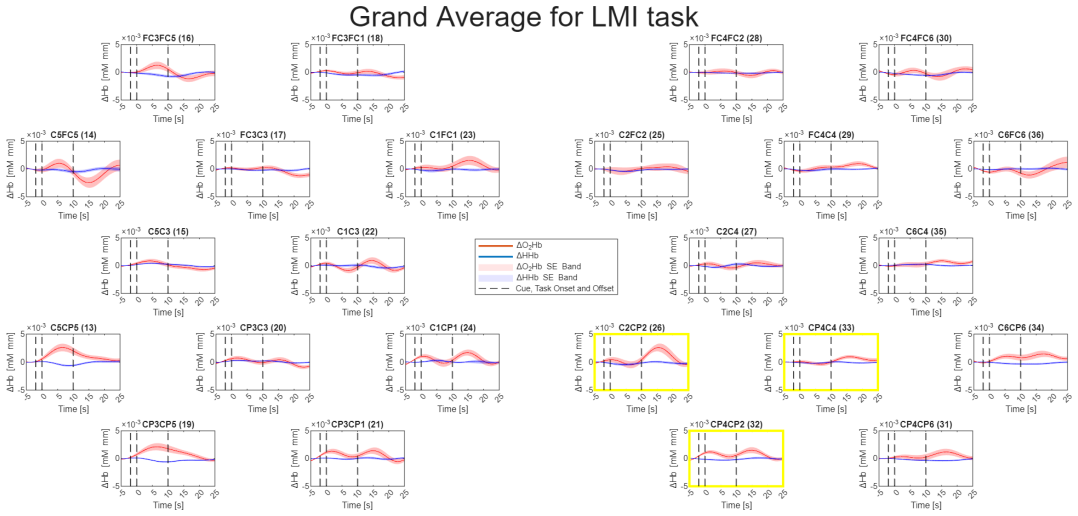
Task	Feature	Normality (LH)	Normality (RH)	Statistical Test Used	p<0.05
LMI	Mean	Yes	Yes	Paired t-test	No
LMI	Slope	No	No	Wilcoxon signed-rank	No
LMI	Peak amplitude	No	Yes	Wilcoxon signed-rank	No
RMI	Mean	No	No	Wilcoxon signed-rank	No
RMI	Slope	No	No	Wilcoxon signed-rank	No
RMI	Peak amplitude	No	No	Wilcoxon signed-rank	Yes

IV. Discussion

The findings of this study highlight the challenges associated with detecting robust inter-hemispheric differences in fNIRS signals during MI tasks using simple temporal features. Although a statistically significant effect emerged only for the



(a) RMI task. Channels selected for the left-hemisphere ROI are highlighted.



(b) LMI task. Channels selected for the right-hemisphere ROI are highlighted.

Fig. 2: Grand Average of ΔO_2Hb and ΔHHb signals across all channels during motor imagery tasks: (a) RMI and (b) LMI. Highlighted channels indicate the ROIs selected for subsequent analyses.

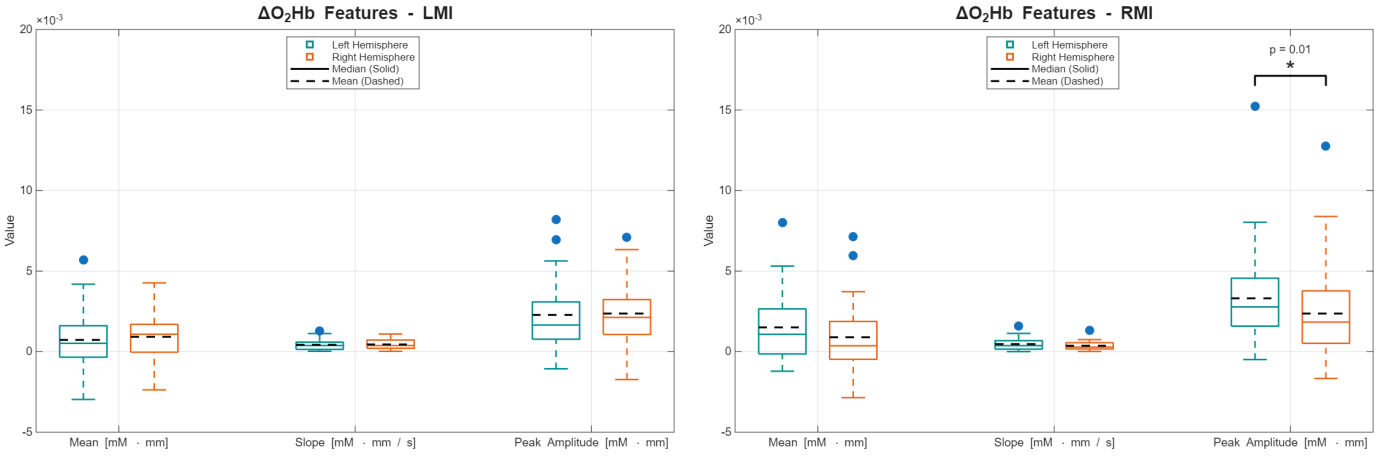
peak amplitude during the RMI task, the overall pattern of results provides useful insights into the characteristics and limitations of MI-based fNIRS analyses.

The lack of statistical significance may be partly explained by the fact that the recorded waveforms did not systematically exhibit the fNIRS activation pattern, typically characterized by a well-defined and temporally consistent increase in ΔO_2Hb, which may have reduced the sensitivity of the extracted features to detect inter-hemispheric differences. As briefly discussed in the following, this limitation is largely intrinsic to MI–based fNIRS applications.

Qualitatively, RMI tasks showed better results than LMI tasks. In particular, in the contralateral hemisphere, peak amplitudes mean during RMI was higher than the one observed during LMI. This finding is consistent with previous studies reporting stronger and more reliable cortical activation during imagery of movements performed with the dominant hand [9]. This interpretation is further supported by the observation that p-values associated with LMI features were consistently higher and farther from the significance threshold ($\alpha = 0.05$) compared to those obtained for RMI, suggesting weaker and more variable task-related responses.

The statistical analysis of the mean value feature led to less prominent results compared to the peak amplitude. This can be attributed to the low-frequency nature of the hemodynamic response. While the mean value feature is highly informative for signals characterized by rapid, high-frequency variations, it tends to underestimate the activation intensity in slow-evolving signals like ΔO_2Hb.

The interpretation of slope-related features was less straightforward. Due to the adopted representation, slope distributions appeared as more compact and overlapped between hemispheres, making potential differences difficult to appreciate visually. Moreover, the intrinsic variability of the hemodynamic response shape across subjects suggests that slope, as defined in Section III–B, may be less sensitive to minimal inter-



(a) LMI (b) RMI

Fig. 3: Inter-hemispheric comparison of ROI-extracted ΔO_2Hb features (mean, slope, peak amplitude) for (a) LMI and (b) RMI tasks, displayed as paired boxplots. For each feature, left- and right-hemisphere distributions are shown. Solid lines indicate the median, while dashed lines represent the mean, included to support the interpretation of parametric tests based on mean values (t–test). The joint visualization of median and mean also provides qualitative insight into distribution normality. Since the extracted features have different physical units (mean and peak in mM·mm, slope in mM·mm/s), they are reported along the horizontal axis rather than sharing a common vertical scale. Statistically significant differences are indicated by asterisks (* if $0.01 \leq p < 0.05$), with the corresponding p–value reported above the paired boxplots.

hemispheric differences in MI tasks compared to amplitude-based descriptors.

These results highlight some well-known challenges associated with fNIRS-based MI analysis. Unlike tasks such as breath-holding or mental arithmetic, MI causes hemodynamic responses with relatively small amplitude compared to real movements, and no clear deterministic activation pattern can be defined a priori [7]. Furthermore, fNIRS responses during MI are characterized by strong inter-subject variability, as not all participants are equally capable of performing motor imagery and subjects may instead rely on visual imagery strategies, leading to weaker or less localized cortical activation. Since the correctness of MI execution cannot be directly verified during acquisition, unlike in cognitive tasks, the use of complementary assessment tools (such as imagery questionnaires) and preliminary training sessions could improve task compliance and signal quality, as suggested in previous works [3].

In addition, cortical activation patterns and optimal activated channels can vary substantially across individuals, further complicating the identification of consistent activation patterns and common ROIs [7]. While averaging signals across trials and subjects, as performed in this study, improves statistical robustness, it also carries the risk of attenuating subject-specific responses and masking channels with stronger localized activation. This effect is intensified by variability in hemodynamic response timing, as early or delayed peaks may arise from task execution differences, residual effects from previous trials, or systemic physiological factors [10]. Overall, the temporal features considered in this study were not sufficiently discriminative to reliably distinguish between LMI and RMI tasks based on fNIRS data alone. Future work should therefore explore more robust and informative features, potentially incorporating frequency-domain descriptors. Moreover, fNIRS is rarely employed as a standalone modality for classification purposes; multimodal approaches combining fNIRS with EEG are commonly adopted to exploit complementary information and improve decoding performance [5], [7].

References

[1] P. Pinti, I. Tachtsidis, A. Hamilton, J. Hirsch, C. Aichelburg, S. Gilbert, and P. W. Burgess, "The present and future use of functional near-infrared spectroscopy (fnirs) for cognitive neuroscience," *Annals of the New York Academy of Sciences*, vol. 1464, no. 1, pp. 5–29, 2020.

[2] F. Scholkmann, S. Kleiser, A. J. Metz, R. Zimmermann, J. Mata Pavia, U. Wolf, and M. Wolf, "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology," *NeuroImage*, vol. 85, pp. 6–27, 2014.

[3] A. E. Hramov, V. Grubov, A. Badarin, V. A. Maksimenko, and A. N. Pisarchik, "Functional near-infrared spectroscopy for the classification of motor-related brain activity on the sensor-level," *Sensors*, vol. 20, no. 8, p. 2362, 2020.

[4] N. Naseer and K.-S. Hong, "fnirs-based brain–computer interfaces: a review," *Frontiers in Human Neuroscience*, vol. 9, p. 3, 2015.

[5] J. Shin, A. von Lühmann, B. Blankertz, D.-W. Kim, J. Jeong, H.-J. Hwang, and K.-R. Müller, "Open access dataset for eeg+nirs single-trial classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1735–1745, 2017.

[6] N. Kollias and W. Gratzer, "Tabulated molar extinction coefficient for hemoglobin in water," http://omlc.org/spectra/hemoglobin/summary.html, 1999, accessed: Dec. 2025.

[7] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, and B. Blankertz, "Enhanced performance by a hybrid nirs–eeg brain–computer interface," *NeuroImage*, vol. 59, no. 1, pp. 519–529, Jan. 2012.

[8] L. Holper and M. Wolf, "Single-trial classification of motor imagery differing in task complexity: A functional near-infrared spectroscopy study," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, p. 34, 2011.

[9] N. Naseer and K.-S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface," *Neuroscience Letters*, vol. 553, pp. 84–89, 2013.

[10] W.-L. Chen, J. Wagner, N. Heugel, J. Sugar, Y.-W. Lee, L. Conant, M. Malloy, J. Heffernan, B. Quirk, A. Zinos *et al.*, "Functional near-infrared spectroscopy and its clinical application in the field of neuroscience: Advances and future directions," *Frontiers in Neuroscience*, vol. 14, p. 724, Jul. 2020.