Joint-Level Motor Imagery Decoding in fNIRS via a Unified Pairwise Machine Learning Framework

Tayyaba Nosheen[1 [0000−0002−8612−0534]], Jacek Jurkojć[1 [0000−0003−0827−5335]], Piotr Wodarski[1 [0000−0001−6701−1546]], Soyiba Jawed[3 [0000−0002−9984−3889]], and Aamir Saeed Malik[2 [0000−0003−1085−3157]]

[1] Department of Biomechatronics, Faculty of Biomedical Engineering, Silesian University of Technology, Gliwice, Poland
tayyaba.nosheen@polsl.pl
[2] Department of Computer and Software Engineering, National University of Sciences and Technology (NUST), Islamabad, Pakistan
[3] Department of Computer Systems, Faculty of Information Technology, Brno University of Technology, Brno, Czechia
malik@fit.vut.cz

Abstract. Precise decoding of intended joint movements from brain activity is essential for next-generation neurorehabilitation interfaces, yet most motor imagery studies focus on single joints or coarse movement categories rather than systematically evaluating all upper-limb joints. We present a joint-aware decoding framework that systematically quantifies joint-level separability across eight upper-limb motor imagery tasks through binary pairwise classification in a unified machine learning framework. All preprocessing, feature selection, dimensionality reduction, and hyperparameter optimization were performed within subject-wise nested cross-validation to eliminate any risk of data leakage. Evaluation on a public dataset of eighteen participants yielded consistent performance across all joint pairs with mean accuracies ranging from 83% to 88% (SD = 0.04). Logistic regression (LR) and linear discriminant analysis (LDA) achieved the strongest and most stable performance. Statistical testing using the Wilcoxon signed rank test confirmed that linear discriminant analysis significantly outperformed the baseline model with $p < 0.001$. These findings establish a systematic and reproducible framework for assessing fine-grained upper-limb joint separability in fNIRS-based motor imagery decoding, enabling interpretable and clinically meaningful evaluation.

Keywords: Neurorehabilitation · fNIRS-based BCI · Motor imagery · Upper-limb motor decoding.

1 Introduction

Stroke remains one of the leading causes of long-term disability worldwide, imposing a substantial burden on individuals, families, and the healthcare system.

According to the World Health Organization, approximately 15 million people experience stroke annually, of whom nearly one-third die, and another one-third are permanently disabled [6]. Motor imagery (MI) based brain–computer interfaces (BCIs) provide a promising non-invasive approach for motor rehabilitation by decoding imagined movements without physical execution [12]. Electroencephalography (EEG) captures rapid neural dynamics with a high temporal resolution, whereas functional near-infrared spectroscopy (fNIRS) measures task-related cortical hemodynamic responses [7] and offers improved spatial specificity and localized cortical sensitivity. However, decoding fine-grained MI remains challenging because imagined movements typically produce weaker and temporally delayed hemodynamic responses compared to actual motor execution [10]. The existing MI decoding studies can be broadly categorized into three groups. First, many studies have focused on coarse paradigms such as MI versus rest or left versus right hand imagery [3], which yield robust but limited representations of motor control. Second, several studies have investigated a restricted subset of joint-related tasks, including wrist and finger movements [2], hand open–close or wrist flexion–extension, and hand clenching [8], often within prosthetic control contexts [1, 9, 14]. Third, hybrid EEG–fNIRS approaches aim to exploit complementary temporal and spatial information [11], yet typically rely on task-specific pipelines and heterogeneous preprocessing strategies. While these studies provide valuable insights, they do not systematically characterize joint-level separability under a unified analytical framework. In particular, there remains a lack of structured evaluation of intrinsic discriminability among different upper-limb joints in MI decoding. Although recent studies have advanced neural signal modeling and connectivity-based analysis using machine learning [15], systematic evaluation of intrinsic task separability under unified analytical conditions remains limited, particularly in fine-grained motor imagery decoding.

To address these limitations, we propose a time, spectral, and correlation index-based machine-learning framework for joint-level MI decoding (TSCIML-MI). The proposed framework reformulates the problem into structured pairwise binary decoding across eight upper-limb joints, resulting in 28 joint-pair classification tasks evaluated within a single unified pipeline. Six (linear and nonlinear) classifiers were systematically benchmarked to assess the decoding performance. Compared with a conventional baseline, the proposed framework achieved an average accuracy improvement of up to 34%, with statistically validated gains across subjects. This joint-aware evaluation paradigm establishes a principled and reproducible foundation for fine-grained MI decoding and provides practical implications for personalized neurorehabilitation.

2 Methodology

We propose a TSCIML-MI binary classification pipeline designed to evaluate all pairwise combinations of eight motor imagery joints and compare classification performance across multiple classifiers. The architecture of this framework is shown in Fig 1. Let $S = \{1, \ldots, 18\}$ denote the set of subjects and $C =$

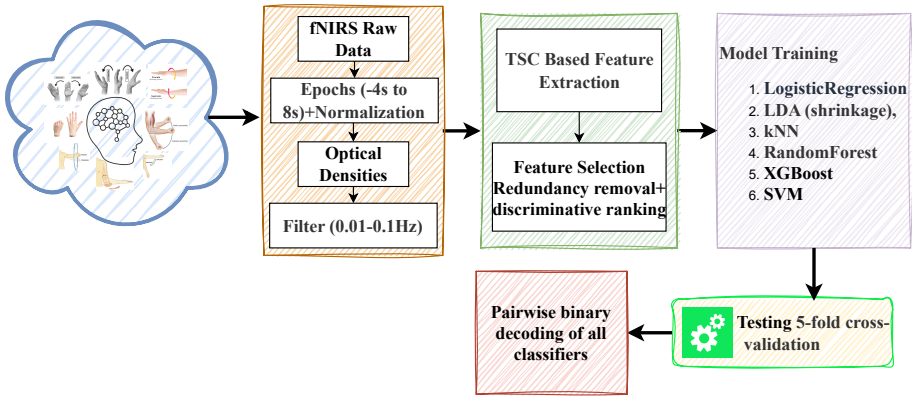

Fig. 1. Overview of TSCIML-MI decoding framework.

$\{c_1, \ldots, c_8\}$ denote the eight upper-limb motor imagery classes. For each subject s, the preprocessed fNIRS time-series signal is represented as $X^{(s)} \in \mathbb{R}^{T \times L}$ where T denotes the number of temporal samples, and L denotes the number of fNIRS measurement channels. Motor imagery events were extracted from stimulus annotations, yielding eight joint-specific class labels. Continuous signals were segmented into epochs spanning $[-4, 8]$ s relative to the task onset, After epoch segmentation, each subject's dataset is defined as

$$\mathcal{D}^{(s)} = \{(X_i^{(s)}, y_i^{(s)})\}_{i=1}^{N_s}, \tag{1}$$

where $X_i^{(s)} \in \mathbb{R}^{T_e \times L}$ represents the i-th trial and $y_i^{(s)} \in C$ represents its corresponding joint label and $N_s = 40$ for each joint. To systematically evaluate joint-level separability, we constructed all possible pairwise class combinations as follows:

$$\mathcal{P} = \{(c_i, c_j) \mid c_i < c_j,\ c_i, c_j \in C\}, \tag{2}$$

resulting in

$$|\mathcal{P}| = \binom{8}{2} = 28 \tag{3}$$

For each joint pair $p = (c_i, c_j)$, we restrict the dataset to trials belonging to these two classes and define binary labels as

$$y = \begin{cases} 0, & \text{if } y_i = c_i \\ 1, & \text{if } y_i = c_j \end{cases}$$

The extracted epochs consist of raw light-intensity signals. These raw light intensity signals were converted to optical density (OD) using the modified Beer–Lambert law [4, 5]:

$$\mathrm{OD}_i^{(s)}(t, \ell) = -\log\left(\frac{I_i^{(s)}(t, \ell)}{I_{0,\ell}}\right), \tag{4}$$

where $I_i^{(s)}(t,\ell)$ denotes the detected light intensity and $I_{0,\ell}$ represents the baseline light intensity.

$$\Delta\mathbf{C}_i^{(s)}(t,\ell) = \frac{1}{d_\ell \,\mathrm{DPF}} \varepsilon^{-1} \Delta\mathbf{OD}_i^{(s)}(t,\ell), \tag{5}$$

Thus, each epoch $X_i^{(s)}$ is transformed into hemodynamic concentration signals $\Delta\mathbf{C}_i^{(s)} \in \mathbb{R}^{T\times L\times 2}$, corresponding to HbO and HbR changes. $\Delta\mathbf{C}_i^{(s)}(t,\ell)$ denotes the change in chromophore concentration (e.g., HbO or HbR), $\Delta\mathbf{OD}_i^{(s)}(t,\ell)$ is the change in optical density, ε is the molar extinction coefficient of the chromophore at a given wavelength, d_ℓ represents the source–detector separation distance, DPF is the differential pathlength factor, accounting for photon scattering and the increased effective optical pathlength in biological tissue. To suppress physiological noise and slow drifts, the signals were band-pass filtered between 0.01 and 0.1 Hz. This preprocessing ensures physiological consistency and inter-subject comparability.

2.1 TSCIML-MI model

The processed signals are then passed to the TSCIML-MI model to classify MI based on fNIRS. The model follows three steps: TSCI feature extraction, fNIRS feature selection, and classification. From each normalized trial, the feature mapping function is as follows:

$$\Phi : \mathbb{R}^{T_e\times L} \to \mathbb{R}^d \tag{6}$$

was used to extract complementary information from each trial. Given a normalized epoch signal $x(t), t = 1,\ldots,T$, the following feature sets are extracted:

1. Temporal and Statistical Features: Mean, Variance, standard deviation, Range, RMS, Time to peak (HbO), Skewness, Kurtosis, Slope, Activity, Mobility, Complexity.
2. Spectral Features: Magnitude spectra of FFT bins (excluding the DC component).
3. Coupling Features: Pearson correlation, Area under the curve ratio of HbO and HbR, Mean concentration difference.

All extracted features were concatenated into a single feature vector per trial:

$$F_i = \Phi(\tilde{X}_i). \tag{7}$$

This multi-domain representation enables the model to capture both localized hemodynamic magnitude and distributed cortical interactions. Highly correlated features were removed to reduce multicollinearity. $|\mathrm{corr}(f_i, f_j)| > \tau, \quad \tau = 0.85$ For each pairwise task, ANOVA ranking $d_{\max} = 100$ was performed within the training folds only, and the top 100 features were selected. For each classifier $m \in \mathcal{M}$, the predictions are computed $\hat{y} = f_m(F)$. The performance was evaluated using 5-fold cross-validation.

$$Acc = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}(\hat{y}_i = y_i), A_{m,p}^{(s)} = \frac{1}{K}\sum_{k=1}^{K} Acc_k, \bar{A}_{m,p} = \frac{1}{|S|}\sum_{s\in S} A_{m,p}^{(s)}$$

where Acc, $A_{m,p}^{(s)}$ and $A_{m,p}^{(s)}$ are per fold, subject-wise decoding and mean accuracy across subjects, respectively. The final results are reported as subject-wise accuracies, along with the mean and standard deviation across subjects for all classifiers.

3 Experiment

3.1 Dataset and Baseline

All experiments were conducted on a publicly available fNIRS motor imagery dataset from 18 healthy participants. All right limb MI tasks were recorded in 4s randomized trials using 24-channel fNIRS recordings. A total of 320 trials were carried out per subject in 8 experimental blocks and stored in the NIRx format. [13]. To ensure a principled comparison, the baseline protocol consists of an SVM classifier operating on two conventional features [13] compared with the proposed framework, which integrates multi-domain features applied uniformly across six classifiers, ensuring that performance differences arise from model characteristics rather than methodological inconsistencies.

3.2 Results

Table 1 and Fig 2 summarize the performance of all classifiers across 28 joint-pair decoding tasks. LDA achieved the most consistent and highest performance, with accuracies ranging from 83% to 88% across joint pairs. In particular, joint pairs with proximal shoulder movements showed slightly higher erroneous classification rates, suggesting that cortical representations of anatomically related movements overlapped. Logistic Regression showed comparable stability, while non-linear models such as Random Forest, XGBoost, kNN, and SVM demonstrated lower and more variable performance. Across all joint pairs, the proposed framework improved accuracy over the baseline by approximately 22 to 34 percentage points across joint pairs, indicating that the multi-domain feature representation and unified decoding strategy substantially enhance joint-level separability. Fig 4 presents pooled row-normalized confusion matrices for LDA, showing strong diagonal dominance with most class-wise recalls between 85% and 89%. All confusion matrices are row-normalized and averaged across subjects. The highest confusions were limited to a small subset of anatomically related pairs, such as Wrist FE vs Shoulder AA and Shoulder PS vs Shoulder AA, where misclassification reached 16–17%. Fig 3 revealed that spectral components, particularly mid-frequency FFT features derived from HbR channels, were consistently among the most discriminative features.SHAP values were computed using a linear explainer applied to the LDA classifier. Domain-level aggregation further showed that spectral and HbO–HbR coupling features contributed more strongly than classical amplitude-based statistics. These findings suggest that joint separability is driven primarily by frequency-specific deoxygenation dynamics and coordinated oxygenation–deoxygenation interactions

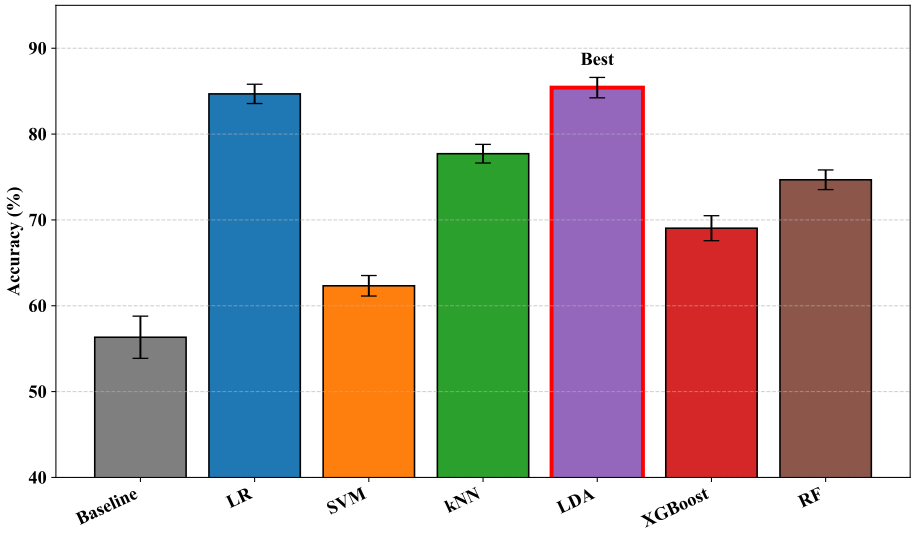

Fig. 2. Comparison of average joint accuracy

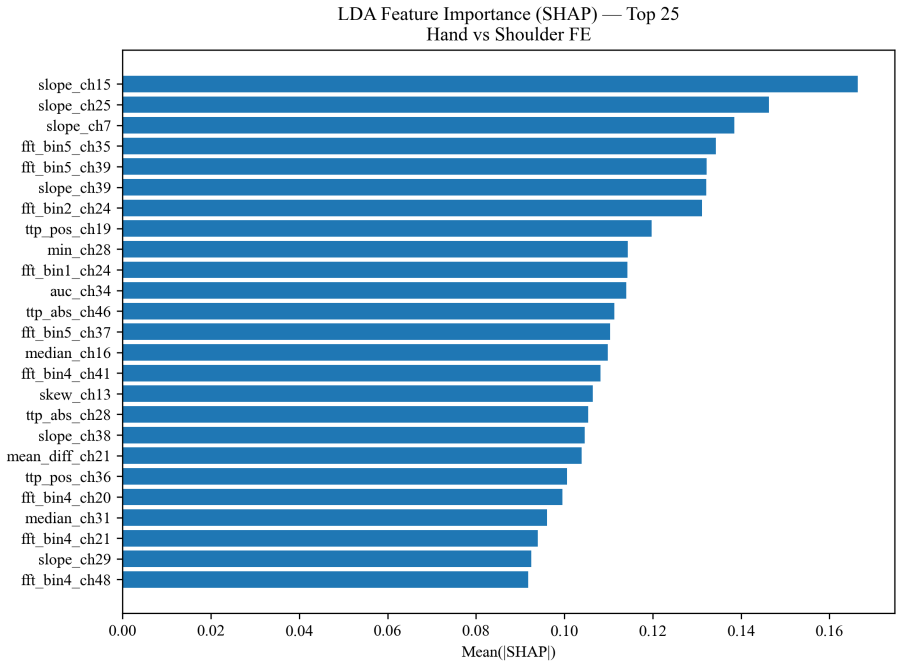

Fig. 3. LDA Feature Importance (SHAP)

Table 1. Subject-wise mean accuracies (%) across 28 joint-pair combinations where (HOC: Hands Open/close), (WFE and WAA: Wrist Flexion/Extension, Adduction/Abduction), (EFE and EPS: Elbow Flexion/Extension, Pronation/Supination), (SFE, SAA, SPS: Shoulder Flexion/Extension, Adduction/Abduction, Pronation/Supination).

Joint Combination	Baseline	LR	SVM	kNN	LDA	RF	XGB
HOC/WFE	57	84	63	78	**85**	74	70
HOC/WAA	57	85	62	78	**85**	74	70
HOC/EPS	52	84	62	76	**86**	74	71
HOC/EFE	57	86	62	78	**85**	76	69
HOC/SPS	60	86	63	79	**86**	76	70
HOC/SAA	62	86	63	80	**87**	77	70
HOC/SFE	61	85	64	79	**88**	74	70
WFE/WAA	56	85	60	77	**87**	73	66
WFE/EPS	55	84	60	79	**85**	74	69
WFE/EFE	53	**86**	64	78	**86**	75	70
WFE/SPS	56	86	62	78	**87**	74	70
WFE/SAA	60	82	63	78	**83**	74	66
WFE/SFE	59	**85**	63	77	**85**	75	70
WAA/EPS	54	85	60	78	**86**	76	70
WAA/EFE	58	84	63	77	**85**	73	68
WAA/SPS	57	**84**	62	78	**84**	75	70
WAA/SAA	57	**85**	61	77	**85**	76	69
WAA/SFE	58	84	63	77	**85**	76	70
EPS/EFE	53	85	62	77	**85**	74	68
EPS/SPS	55	86	64	78	**87**	75	70
EPS/SAA	56	83	62	78	**84**	73	67
EPS/SFE	57	83	62	76	**84**	75	68
EFE/SPS	54	**86**	62	77	**86**	74	68
EFE/SAA	56	84	64	77	**85**	73	67
EFE/SFE	56	84	63	77	**85**	75	69
SPS/SAA	56	**84**	62	75	83	73	68
SPS/SFE	53	**87**	62	80	**87**	77	73
SAA/SFE	54	86	62	79	**87**	74.31	70

rather than static amplitude differences alone. Wilcoxon signed-rank test revealed that the proposed framework using LDA significantly outperformed the baseline ($W = 0, z = -3.72, p < 0.001$), with a maximal rank-biserial correlation of -1.00, indicating consistent improvement across all subjects. No significant difference was observed between LDA and LR ($p = 0.151$), suggesting comparable performance stability between the linear models. Further pairwise comparisons demonstrated that LDA significantly outperformed Random Forest, XGBoost, kNN, and SVM (all $p < 0.001$). These findings confirm that linear

classifiers provide a superior and more stable decoding performance under the proposed feature representation.

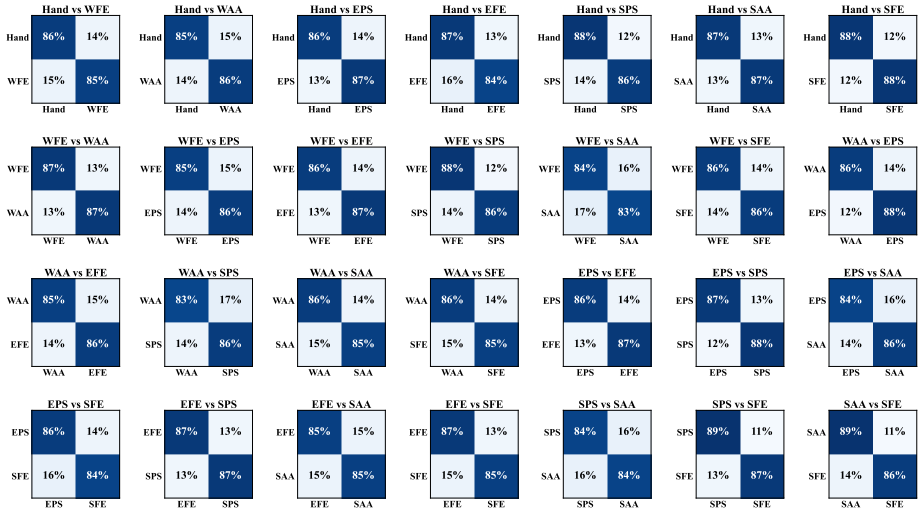

Fig. 4. Confusion matrix LDA of all 28 Pairs

4 Conclusion

Decoding fine-grained motor imagery at the joint level is a clinically relevant yet challenging problem. MI hemodynamic responses are weaker in amplitude, more variable across trials, and delayed due to neurovascular coupling, which collectively reduces the effective signal-to-noise ratio (SNR). The proposed TSCIML-MI framework explicitly targets this challenge through a sequence of stabilizing steps. First, band-pass filtering (0.01–0.1 Hz) and epoch-wise baseline correction mitigate slow drifts and physiological contamination while preserving task-related hemodynamic. Second, within-epoch Z-score normalization reduces inter-trial scale variability. Third, instead of relying on a single descriptor, TSCIML-MI employs a multi-domain feature bank (time, spectral, and HbO–HbR coupling) that captures complementary signatures of MI, increasing robustness when any one feature family is weak. The proposed framework enables reproducible and interpretable comparison across movements. Most importantly, the proposed approach reveals relative decoding difficulty among joints while achieving stable performance gains over conventional baselines, laying the groundwork for future hybrid EEG–fNIRS extensions and clinically adaptive neurorehabilitation systems. Although decoding performance was relatively consistent across joint pairs, several anatomically related movements exhibited lower separability, suggesting

task-specific differences in cortical representations. Future work incorporating cortical activation mapping, external validation datasets, and multimodal neuroimaging will further clarify the extent to which the identified features reflect fine-grained joint representations versus broader motor imagery processes. These findings support emerging evidence that neurovascular coupling dynamics provide richer information than static amplitude features alone.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Abiri, R., Borhani, S., Zhao, X., Jiang, Y.: Real-time brain machine interaction via social robot gesture control. In: Dynamic Systems and Control Conference. vol. 58271, p. V001T37A002. American Society of Mechanical Engineers (2017)
2. Alazrai, R., Abuhijleh, M., Alwanni, H., Daoud, M.I.: A deep learning framework for decoding motor imagery tasks of the same hand using eeg signals. IEEE Access **7**, 109612–109627 (2019)
3. Baghaeifar, E., Shamekhi, S., Shalchizadeh, F.: An fnirs-bci study: Effective channels selection in imagining right and left hand movements via brain functional connectivity. Biomedical Signal Processing and Control **109**, 107915 (2025)
4. Cope, M., Delpy, D.T.: System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. Medical and Biological Engineering and Computing **26**(3), 289–294 (1988)
5. Delpy, D.T., Cope, M., van der Zee, P., Arridge, S., Wray, S., Wyatt, J.: Estimation of optical pathlength through tissue from direct time of flight measurement. Physics in Medicine & Biology **33**(12), 1433–1442 (1988)
6. Feigin, V.L., et al.: World stroke organization: Global stroke fact sheet 2025. International Journal of Stroke **20**(2), 132–144 (Feb 2025). https://doi.org/10.1177/17474930241308142
7. Finnis, R., Mehmood, A., Holle, H., Iqbal, J.: Exploring imagined movement for brain–computer interface control: An fnirs and eeg review. Brain sciences **15**(9), 1013 (2025)
8. Hanafy, M.A., Yusufjonov, S., Yusufjonov, D., Sharafianardakani, P., Rayguru, M.M., Popa, D.O.: A robust eeg brain-computer interface approach for decoding hand and foot motor imagery. In: Proceedings of the 18th ACM International Conference on PErvasive Technologies Related to Assistive Environments. pp. 92–100 (2025)
9. Khan, M.J., Zafar, A., Hong, K.S.: Comparison of brain areas for executed and imagined movements after motor training: An fnirs study. In: 2017 10th International Conference on Human System Interactions (HSI). pp. 125–130. IEEE (2017)
10. Kumarasinghe, K., Owen, M., Taylor, D., Kasabov, N., Kit, C.: Faneurobot: a framework for robot and prosthetics control using the neucube spiking neural network architecture and finite automata theory. In: 2018 IEEE international conference on robotics and automation (ICRA). pp. 4465–4472. IEEE (2018)
11. Shi, J., Chen, D., Zhao, X., Zhao, Z., Li, S., Xu, Y., Ding, T., Zhu, Z., Zhang, P., Ye, Q., et al.: Hefmi-ich: a hybrid eeg-fnirs motor imagery dataset for brain-computer interface in intracerebral hemorrhage. Scientific Data **12**(1), 1816 (2025)

12. Vavoulis, A., Figueiredo, P., Vourvopoulos, A.: A review of online classification performance in motor imagery-based brain–computer interfaces for stroke neurorehabilitation. Signals **4**(1), 73–86 (2023)
13. Yi, W., Chen, J., Wang, D., Hu, X., Xu, M., Li, F., Wu, S., Qian, J.: A multimodal dataset of electroencephalography and functional near-infrared spectroscopy recordings for motor imagery of multi-types of joints from unilateral upper limb. Scientific Data **12**(1), 953 (2025)
14. Yun, Y.D., Jeong, J.H., Cho, J.H., Kim, D.J., Lee, S.W.: Reconstructing degree of forearm rotation from imagined movements for bci-based robot hand control. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 3014–3017. IEEE (2019)
15. Zaheer, M.A., Malik, A.S.: Unveiling neural signatures: A comprehensive review of eeg biomarkers in stress, anxiety, and depression. IEEE Transactions on Affective Computing (2026)