Regularized Gaussian Functional Connectivity Network with Post-Hoc Interpretation for Improved EEG-based Motor Imagery-BCI Classification

Daniel Guillermo García Murillo
dggarciam@unal.edu.co

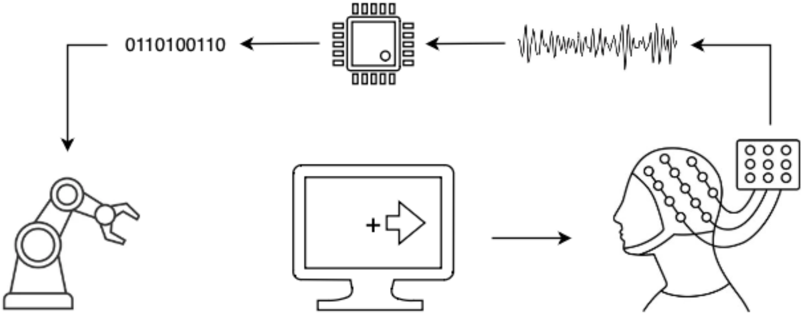
Universidad Nacional de Colombia
Signal Processing and Recognition Group - SPRG
Advisor: Andrés Marino Álvarez-Meza, Ph.D.
Co-advisor: César Germán Castellanos-Domíguez, Ph.D.

September 23, 2024

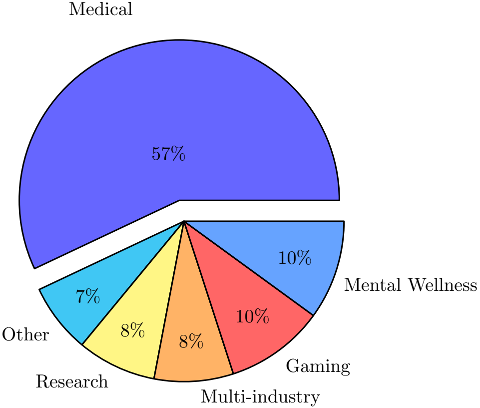
Outline I

Brain Computer Interface (BCI)

BCI provides people with external world communication by translating brain signals. [Khan et al., 2020]



Socioeconomic Facts

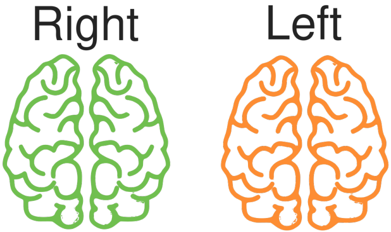


- More than 87 BCI startups across the U.S. as of 2022 [2].
- BCI market valued at 1800 million in 2022, projected to reach 6100 million by 2030 [3].

[1] **Image:** Adapted from The World Economic Forum 2024
[2] Brain Computer Interface Market Size, Report 2024-2033
[3] Brain-Computer Interface Market 2022

Motor Imagery (MI)

MI is a widely studied BCI paradigm that allows external motor communication [Cattan et al., 2018].

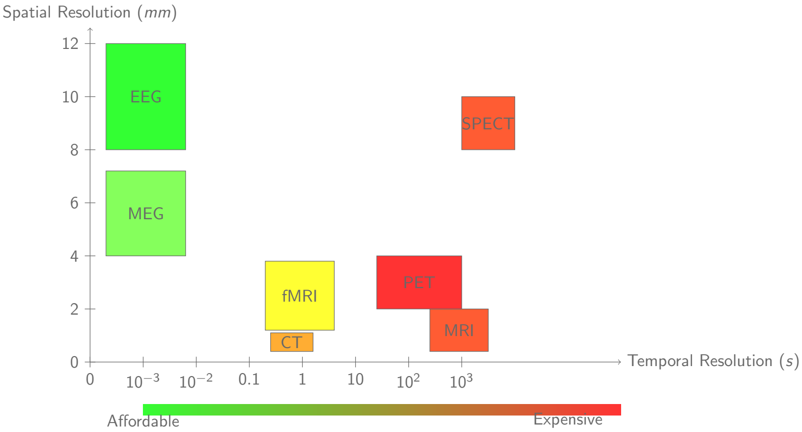


Applications:

- Recovery of motor functionality [Bonci et al., 2021].
- Motor rehabilitation [Sitaram et al., 2017].
- Virtual reality [Cattan et al., 2018].
- Gaming [Ahn et al., 2014].
- Skill acquisition [Casimo et al., 2017].

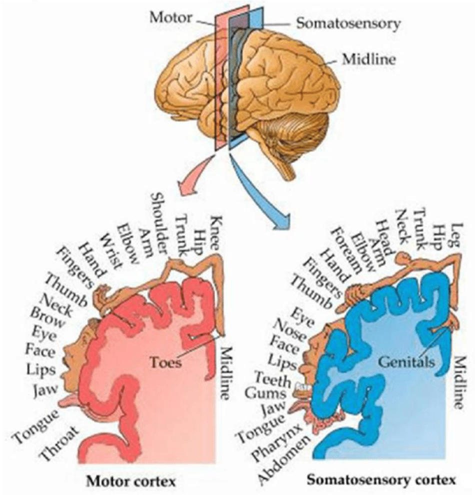
Neuroimaging Techniques

- MI involves fast-evolving cognitive processes [Värbu et al., 2022]
- EEG and MEG have remarkable temporal precision [Alsharif et al., 2020]
- EEG is portabile and cost-effective [Janapati et al., 2023, Hosseini et al., 2020]



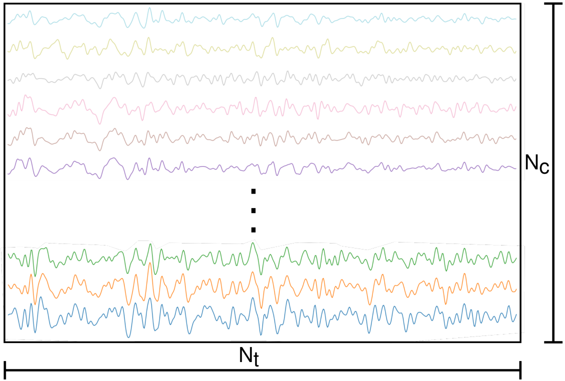
Sensorimotor Rhythms (SMRs)

- EEG signals contains multiple electrical variations (rhythms) [Barios et al., 2019].
- Sensorimotor Rhythms (SMRs) occur in the sensorimotor cortex [Altaheri et al., 2023].
- SMRs contain spectral-spatio-temporal patterns of MI tasks [Li et al., 2019].



[1] **Image:** Adapted from [Purves, 2001]

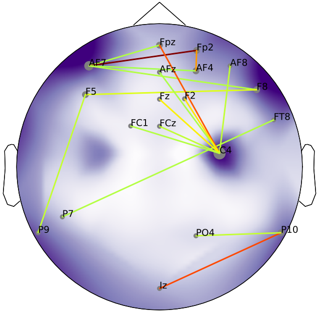
MI-EEG Feature Extraction



- High number of channels and sampling rate [Chevallier et al., 2024].
- Huge number of data points [Singh et al., 2021].
- Feature extraction strategies are required to reduce dimentionality [Ai et al., 2019].

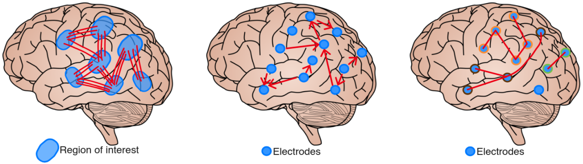
Single Channel Feature Extraction

- Capture rhythms on specific EEG channels [Samuel et al., 2017].
- Time domain: statistical [Hamedi et al., 2014], Hjorth [Yilmaz et al., 2018], etc.
- Spetral domain: Power spectral density [Oikonomou et al., 2017], Welch's periodigram [Roy et al., 2022], spectral entropy [Sarraf, 2017], etc.



Executing or imagining motor tasks activates multiple brain areas, patterns that single-channel features fail to capture [Chiarion et al., 2023].

Multi Channel Feature Extraction



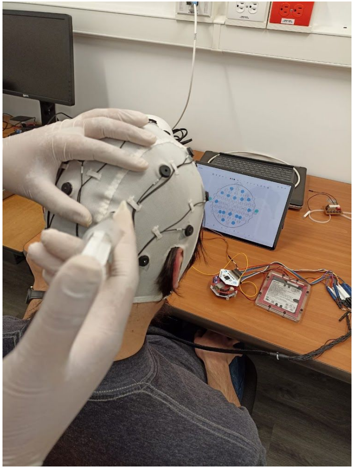
(a) Structural connectivity (b) Effective connectivity (c) Functional connectivity

- Structural connectivity (SC) focuses on physical connections, fails to capture short-living events [Thiebaut de Schotten et al., 2020]
- Effective connectivity (EC) describes direct connections and requires deep cognitive process understanding to select the best causal model [Chiarion et al., 2023].
- Functional connectivity (FC) can describe directed or non-directed connectives usually via statistical correlation [Cao et al., 2022a]

FC's simplicity, low computational demands, and lack of rigid assumptions make it ideal for MI-BCI applications [He et al., 2019].

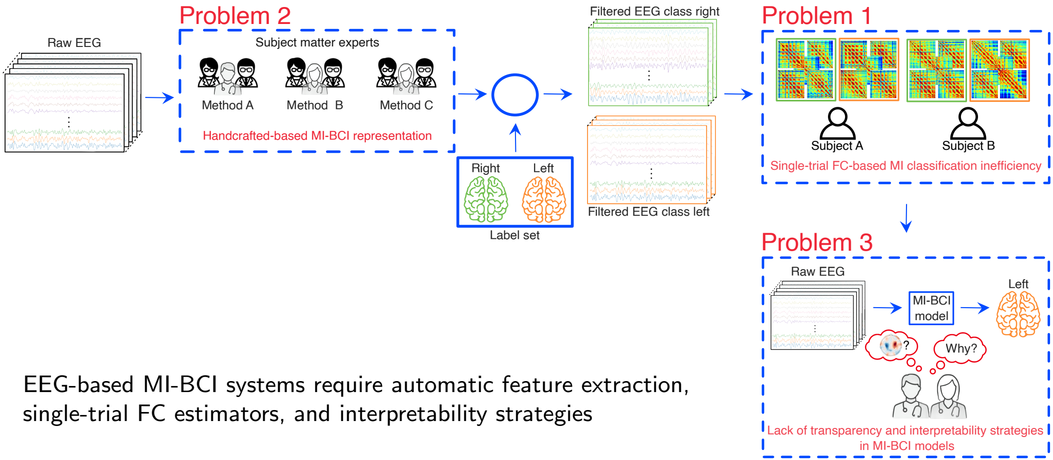
Signal Processing and Recognition Group - SPRG

The SPRG designs Machine Learning (ML) and Deep Learning (DL) models to improve the performance and explainability of EEG-based MI-BCIs [Collazos-Huertas et al., 2023].



Outline I

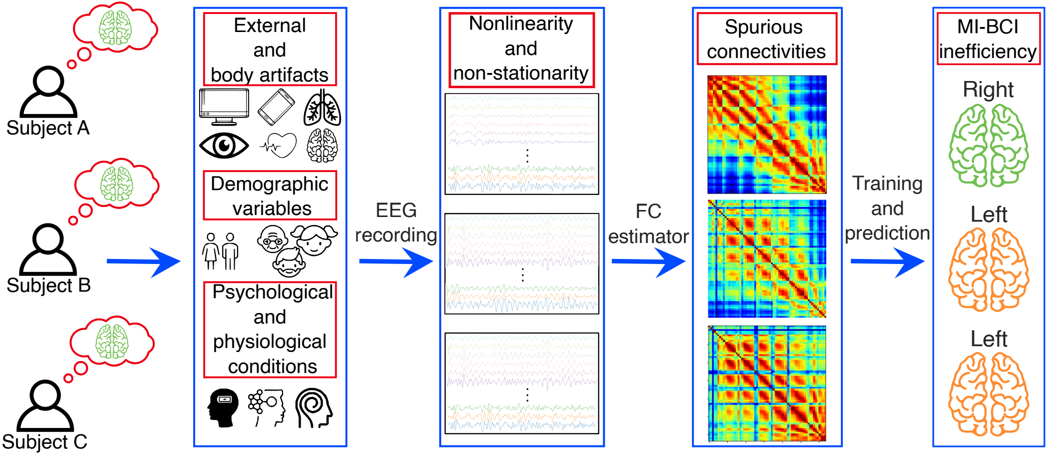
Problem Statement



EEG-based MI-BCI systems require automatic feature extraction, single-trial FC estimators, and interpretability strategies

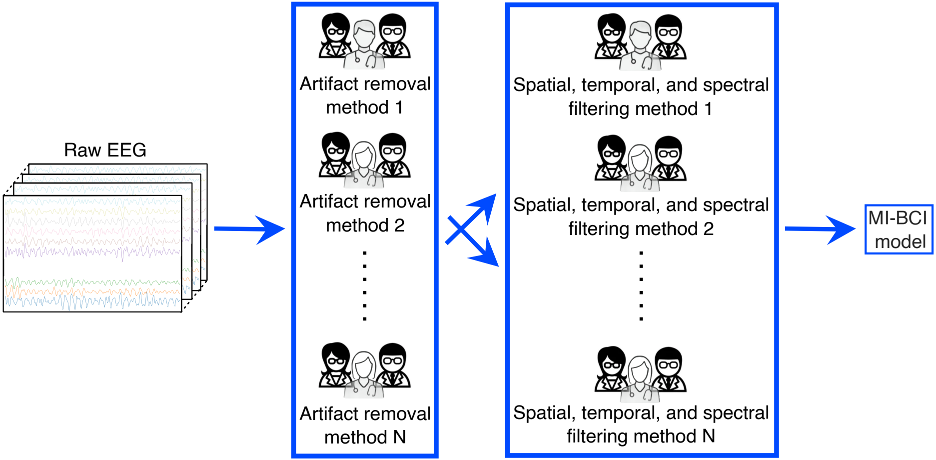
[1] [Chiarion et al., 2023, Cao et al., 2022a, Altaheri et al., 2023, Chamola et al., 2020, Xiao et al., 2018, Fan et al., 2021]

Single-Trial FC MI Classification Inefficiency



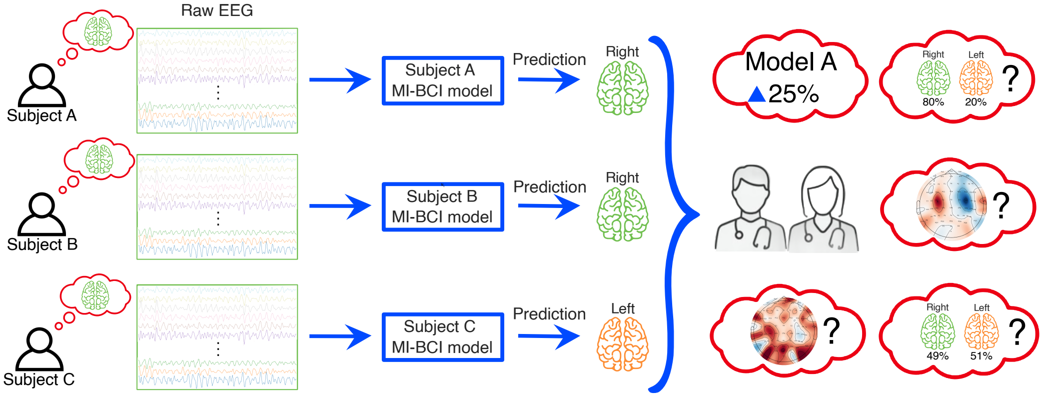
[1] [Chiarion et al., 2023, Cao et al., 2022a]

Handcrafted-based Subject-Specific EEG-based MI-BCI Representation



1 [Altaheri et al., 2023, Chamola et al., 2020]

Lack of Transparency and Interpretability Strategies in MI-BCI



[1] [Xiao et al., 2018, Fan et al., 2021]

Research Question

How can a **single-trail FC** be developed to manage non-stationary EEG subject-specific representations, handle spurious connectivities, and encode non-linear spatial, temporal, and spectral **discriminative** and **interpretable** MI patterns?

Outline I

Outline I

Functional Connectivity Estimators

	Time Domain	Frequency domain	
Indirect	Corr [Fagerholm et al., 2020]	IPC[Cao et al., 2022a] MSC [Cattai et al., 2021] PC [Gonzalez-Astudillo et al., 2020]	Linear
	MI [Gu et al., 2023] SL [Gonzalez-Astudillo et al., 2021]	PLI [Siviero et al., 2023] PLV [Cattai et al., 2021] WPLI [Gonzalez-Astudillo et al., 2020]	Nonlinear
Directed	Cross-corr [Roy et al., 2022] GC [Rezaei & Shalbaf, 2023]	DTF [Rezaei & Shalbaf, 2023] PDC[Gaxiola-Tirado et al., 2017]	Linear
	TE [Rezaei & Shalbaf, 2023]		Nonlinear

High sensitive   Sensitive   Less sensitive   Robust

- Orange shades indicate sensitivity to volume conduction (VC).
- Linear estimators are simple but may miss complex interactions, nonlinear ones capture them but are noise-sensitive [Gonzalez-Astudillo et al., 2020].
- Direct and Indirect connectivity achieve similar performance in MI, being indirect connectivity less sensitive to the VC [Cao et al., 2022b]

Feature Extraction from FC I

Common spatial patterns

- Constant regularization [Park et al., 2017]
  $\hat{\mathbf{C}} = \{\mathbf{C} + \alpha\mathbf{I}\}$
- Lasso - L_1-norm [Zhang et al., 2018a]
  $\arg\max \mathbf{W}\mathbf{C}\mathbf{W}^\top - \sum_c \|\mathbf{w_c}\|_1$
- Tikhonov L_2-norm [Fauzi et al., 2019]
  $\arg\max \mathbf{W}\mathbf{C}\mathbf{W}^\top - \sum_c \|\mathbf{w_c}\|_2$
- Elastic-net - L_1, L_2-norm [Gu et al., 2021]
  $\arg\max \mathbf{W}\mathbf{C}\mathbf{W}^\top - \sum_c \|\mathbf{w_c}\|_1 - \sum_c \|\mathbf{w_c}\|_2$
- Weighted regularization [Deng et al., 2020]
  $\hat{\mathbf{C}} = \{\mathbf{C} + diag(\mathbf{A})\}$
- $L_{q/p}$ regularization [Cai et al., 2021]
  $\arg\max \mathbf{W}\mathbf{C}\mathbf{W}^\top - \sum_c \|\mathbf{w_c}\|_{p,q}$

Easy interpretability [Akuthota et al., 2023]

Easy implementation [Khademi et al., 2023]

Usually $\mathbf{C}$ is calculated as the sample covariance matrix [Guo et al., 2020]

Ignores nonlinear interations [Ghanbar et al., 2021]

Not design for muti-class scenarios [Alizadeh et al., 2023]

Feature Extraction from FC II

Vector representation [Meng et al., 2023, Georgiadis et al., 2018]
- CSP-like when followed by linear classification [Reuderink et al., 2011]
- Frequently hard to extract features [Xu et al., 2020]
- Easy integration with deep learning [Zoumpourlis & Patras, 2022]

Graph measures [Gu et al., 2023]
- Intuitive representation [Rodrigues et al., 2022]
- Oversimplified the cognitives process [Gonzalez-Astudillo et al., 2020]

Riemannian geometry [Ding et al., 2023]
- Take into account $\mathbf{C}$ iner-links [Liu et al., 2024]
- Hard to find best $\mathbf{C}_{ref}$ [Miah et al., 2020]

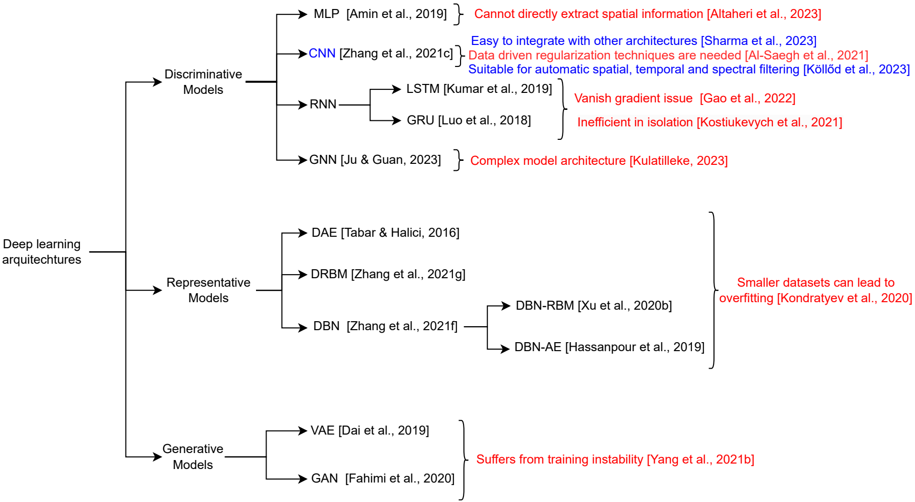
Topological representation [Altaheri et al., 2023, Collazos-Huertas et al., 2021]
- Potentially loss information [Altaheri et al., 2023]
- Performance relies on sufficient channels [Zhao et al., 2019a]
- Seamless integration with convolutional neural networks [Xu et al., 2020b]

Outline I

Input Formulation in Deep Learning

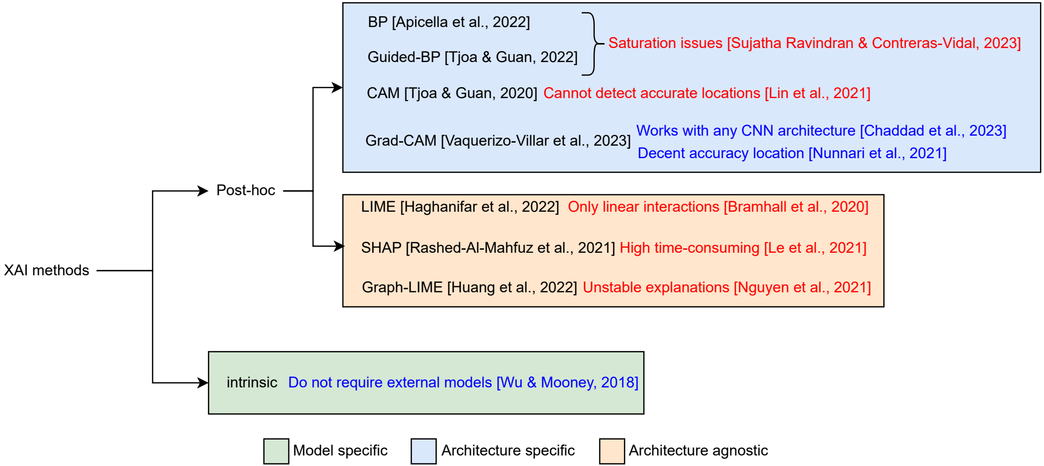
- Deep learning inputs
  - Raw EEG Matrix $[N_c \times N_t]$ [Zhang et al., 2021b]
    - Usually light architecture [Dose et al., 2018]
    - Little or no preprocessing [Amin et al., 2019]
  - Handcraft-based
    - Statistical measures
    - Graph measures
      - Vector $[N_q]$ [Gu et al., 2023]
        - Oversimplified the cognitives process [Gonzalez-Astudillo et al., 2020]
    - Spectral images (E.g., EMD, PSD, FFT, WT, STFT, CWT, FC, CSP)
      - Matrix
        - $[N_c \times N_f]$ [Ma et al., 2020]
        - $[(N_f + N_c) \times N_t]$ [Zhang et al., 2020b]
        - $[(N_t + N_c) \times N_f]$ [Kant et al., 2020]
        - $[(N_t + N_c) \times (N_f + N_c)]$ [Alwasiti et al., 2020]
        - $[N_c \times N_c]$ [Kwon et al., 2021]
      - Tensor
        - $[N_c \times N_f \times N_t]$ [Miao et al., 2020]
        - $[N_c \times N_c \times N_t]$ [Wang et al., 2019]
        - $[N_c \times N_c \times (N_f + N_t)]$ [Collazos-Huertas et al., 2023]
      - Potentially lose information [Altaheri et al., 2023]
    - Topological representation
      - Matrix
        - $[x \times y]$ [Xu et al., 2020b]
      - Tensor
        - $[N_f \times x \times y]$ [Li et al., 2019b]
        - $[N_t \times x \times y]$ [Liu & Yang, 2021]
      - Performance relies on sufficient channels [Zhao et al., 2019a]

Deep Learning Architectures



Outline I

Interpretability Strategies in MI-BCI



Outline I

General Objective

To develop a **single-trial indirect functional connectivity** framework, accompanied by **regularized deep learning** approaches, to extract pertinent subject-specific non-linear **spatio-temporal-frequency patterns** from non-stationary EEG data, improving the MI-BCI system's accuracy and **interpretability**.

Specific Objectives

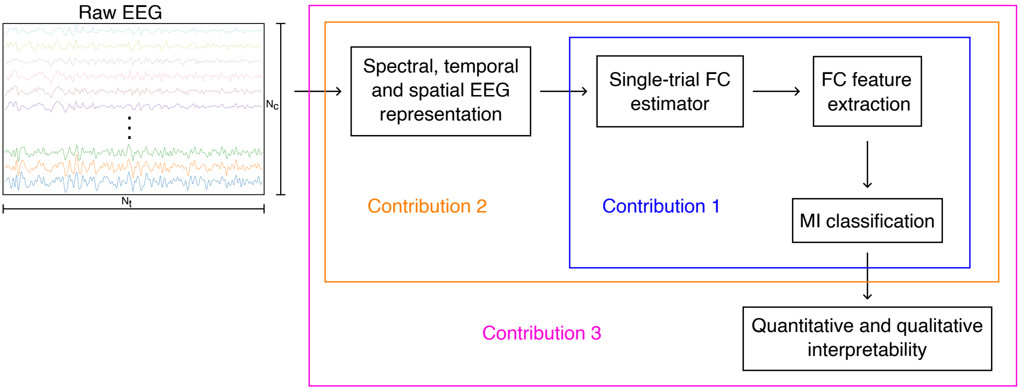
1. To develop a **single-trial indirect FC** for enhanced nonlinear feature extraction, preserving the spatio-temporal-frequency interpretability while favoring the **classification performance** in MI-BCI and avoiding spurious connectivities.

2. To extend the proposed single-trial FC within a **deep learning scheme** that handles artifacts and EEG representations, requiring **minimal preprocessing** efforts from raw signals.

3. To develop a **transparency and interpretability** strategy dedicated to MI-BCI classification that emphasizes spatial-temporal-spectral pattern domains, incorporating a **qualitative and quantitative** relevance analysis assessment.

Outline I

Contributions

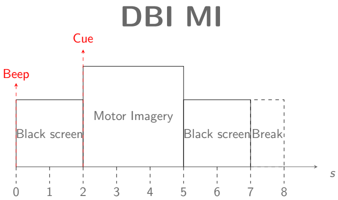


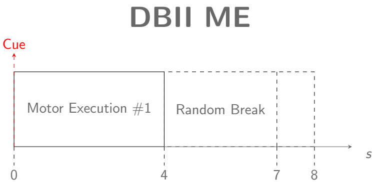
Regularized DL-based **single-trial indirect FC** framework for extracting subject-specific **spatio-temporal-spectral patterns** to improve MI-BCI system's **accuracy** and **interpretability**.

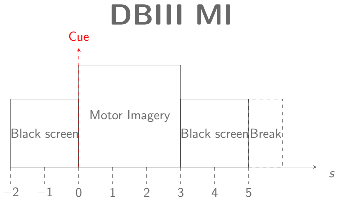
Outline I

Datasets

Dataset	Subjects	Trials	Paradigm	Classes	Channels	Sampling rate
BCI Competition IV Dataset IIa (DBI MI) [1]	9	288	Motor imagery	2	22	250Hz
Gamma Motor Execution Database (DBII ME) [2]	14	260	Motor execution	2	44	500Hz
MI BCI EEG Giga Science Database (DBIII MI) [3]	50	200	Motor imagery	2	64	500Hz







[1] http://www.bbci.de/competition/iv
[2] https://gin.g-node.org/robintibor/high-gamma-dataset
[3] http://gigadb.org/dataset/100295

Outline II

Outline II

Single-Trial Kernel-based Functional Connectivity I

1. **Wiener-Khinchin's theorem** [Cohen, 1998]:

$$R_r^c(\tau) = \int_{\varpi \in \Omega} \exp(j2\pi\tau\varpi)\, dP_r^c(\varpi),$$

where $P_r^c(\varpi) \in \mathbb{R}[0,1]$ is the spectral distribution function.

2. **Bochner's theorem** [Bochner, 2020]:

$$\kappa_r^{cc'}(\Delta_x) = \int_{\boldsymbol{\varpi} \in \Omega} \exp(j2\pi\Delta_x^\top\varpi) S_r^{cc'}(\boldsymbol{\varpi})\, d\boldsymbol{\varpi},$$

where $\Delta_x = \mathbf{x}_r^c - \mathbf{x}_r^{c'}$ is the vector delay, $\boldsymbol{\varpi} \subseteq \mathit{\Omega}$ is the frequency domain that contains the bandwidth set of analysis $\mathit{\Omega}$, and $S_r^{cc'}(\boldsymbol{\varpi})$ is the cross-spectral density.

Single-Trial Kernel-based Functional Connectivity II

3. **Cross-spectral distribution:**

$$P_r^{cc'}(\varpi) = 2\int_{\varpi\in\Omega} \mathscr{F}\left\{\kappa(\mathbf{x}_r^c, \mathbf{x}_r^{c'})\right\} d\varpi,$$

where the notation $\mathscr{F}\{\cdot\}$ stands for the Fourier transform.

4. **Kernel-based spectral distribution estimation:**

$$\hat{P}_r^{cc'}(\mathbf{u}^{cc'}, \kappa_X(\cdot;\sigma)) = \sum_{n=1}^{N_f}\sum_{w_t=1}^{N_t} u_{nw_t}^{cc'}\kappa_X\left(\mathbf{x}_{rnw_t}^c, \mathbf{x}_{rnw_t}^{c'};\sigma\right),$$

where $\mathbf{u}^{cc'} \in \mathbb{R}^{N_f N_t}$ is the spatio-temporal-frequency relevance vector.

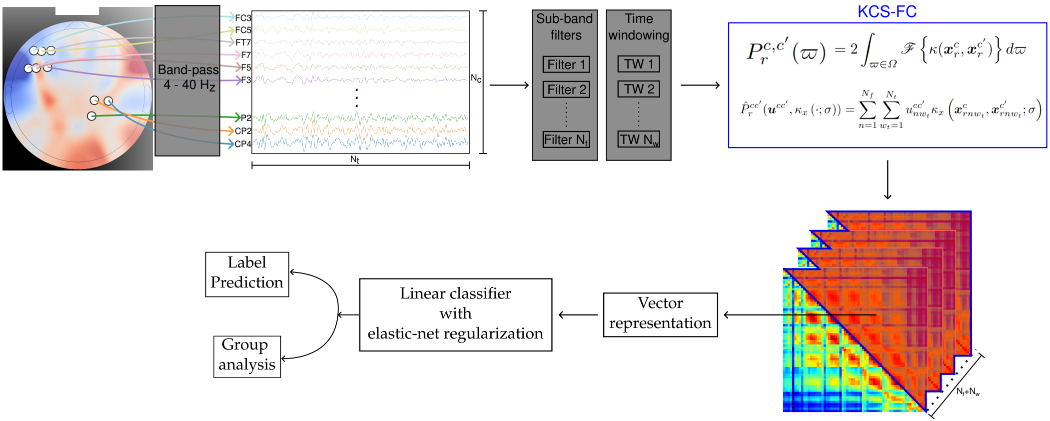
Single-Trial Kernel-based Functional Connectivity III

5 **Supervised kernel-based spectral distribution estimation:**

$$\mathbf{u}^{*} = \arg\min_{\mathbf{u}} \sum_{r=1}^{R} \left\| \sum_{c,c'=1}^{N_c} \hat{P}_r^{cc'}(\mathbf{u}^{cc'}, \kappa_x(\cdot;\sigma)) - y_r \right\|_2^2 + \alpha \sum_{c,c'=1}^{N_c} \|\mathbf{u}^{cc'}\|_1 + \frac{1-\alpha}{2} \sum_{c,c'=1}^{N_c} \|\mathbf{u}^{cc'}\|_2 \quad : \forall c < c',$$

where $\alpha \in \mathbb{R}^{+}$ is the regularization hyperparameter, y_r is the label corresponding to the r-th trial, and $\|\cdot\|_q$ is the ℓ_q-norm.

Single-Trial Kernel-based Functional Connectivity Proposal



Experimental Set-up

1. Sliding window of length $\tau = [0.5, 1.0, 1.5, 2.0]$ s, and overlap of 75%.
2. Frequency bands from 4 Hz to 40 Hz, window bandwidth of 4 Hz, and overlap of 50%.
3. Data split using 5-fold 80-20 scheme.
4. Gaussian kernel

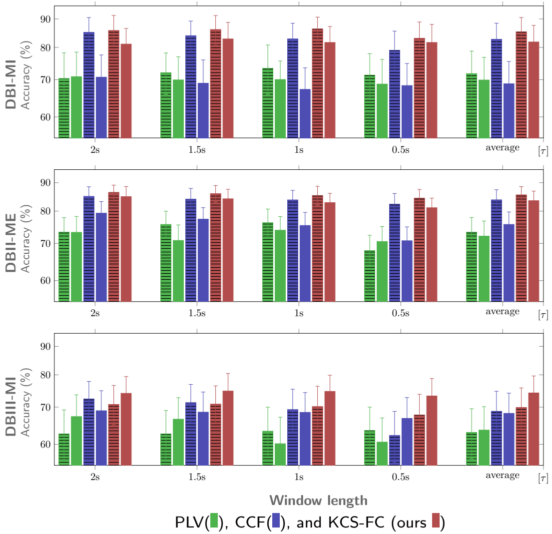
$$\kappa_X\left(\mathbf{x}^c_{rnw_t}, \mathbf{x}^{c'}_{rnw_t}; \sigma\right) = \exp\left(-\|\mathbf{x}^c_{rnw_t} - \mathbf{x}^{c'}_{rnw_t}\|_2^2/2\sigma^2\right),$$

5. We compare our proposal with Cross-Correlation Coefficient (CCF) and Phase Lag Value (PLV)

$$\rho(\mathbf{x}^c_{rnw_t}, \mathbf{x}^{c'}_{rnw_t}) = \left\langle \mathbf{x}^c_{rnw_t}, \mathbf{x}^{c'}_{rnw_t} \right\rangle$$

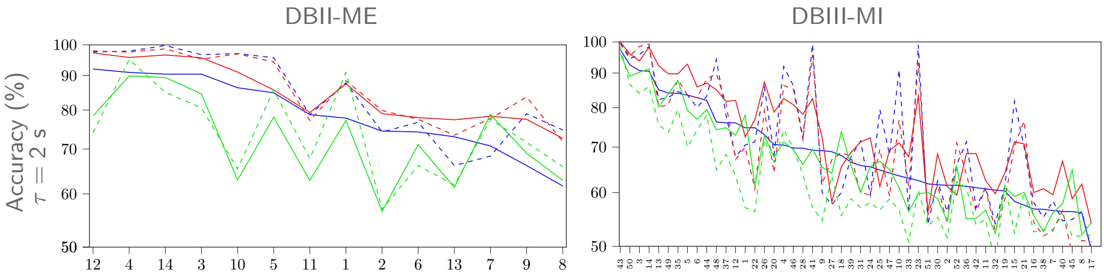
$$\Delta\phi(\mathbf{x}^c_{rnw_t}, \mathbf{x}^{c'}_{rnw_t}) = |\exp(j(\phi^c_{rnw_t} - \phi^{c'}_{rnw_t}))|$$

Influence of Sliding Windows



- Solid bars represent concatenated data; striped bars show CSP-filtered.
- DBIII-MI shows the highest subject variability, with all FC measures, especially PLV, dropping in accuracy.
- CSP filtering effectiveness depends more on the selected sliding window.
- PLV and CCF accuracy are sensitive to sliding window size, while KCS-FC (ours) remains consistent across τ.

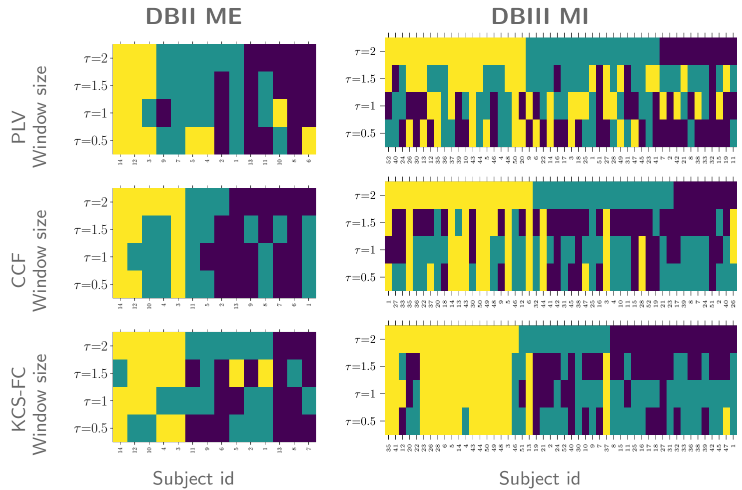
Classifier Accuracy of Individuals



PLV (—), CSP-PLV (- -), CCF (—), CSP (- -), KCS-FC (—), and CSP-KCS-FC (- -)

- CSP algorithm reduces the effectiveness of KCS-FC (ours) in handling subject variability.
- Both versions of the PLV algorithm achieve the lowest accuracy.
- KCS-FC (ours) ensures several subjects exceed the BCI-inefficiency threshold (below 70%).

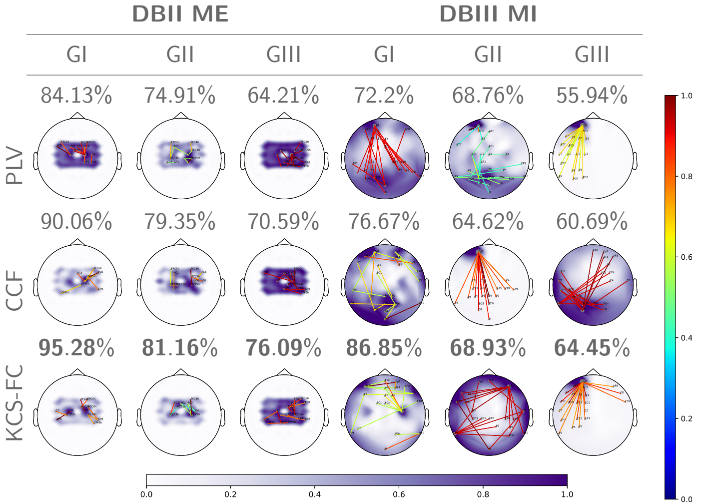
Interpretation of Subject Clusters



- Groups are ranked in decreasing order of accuracy: Group I ( ), Group II ( ), Group III ( ).
- Different window values cause frequent changes between groups for PLV and CCF measures.
- KCS-FC (ours) is less affected by different window values.

Functional Connectivity Analysis

- PLV shows high background activity, CCF has fewer amplitudes, and KCS-FC shows localized activity over SMR hemispheres.
- GII and GIII generate abnormally connections over the frontal and occipital lobe.



Classifier Accuracy Comparison of FC Approaches

Data	Time Window	Filter Band	Interpretation	Feature Extraction	Accuracy (%)
DBI-MI	✓	✓	✓	TSGSP [Zhang et al., 2018]	**82.50 ± 12.2**
	-	-	✓	STR connectivity [Rodrigues et al., 2019]	69.56±15.02
	✓	-	✓	Renyi's α-entropy [De La Pava Panche et al., 2019]	72.40 ± 6.50
	✓	✓	✓	Proposed KCS-FC	81.92 ± 9.44
DBIII-MI	-	✓	✓	CSP [Cho et al., 2017]	67.60 ± 13.17
	✓	✓	-	OPTICAL [Kumar et al., 2019]	68.19 ± 9.36
	-	-	✓	STR connectivity [Rodrigues et al., 2019]	62.00 ± 13.00
	✓	✓	✓	Proposed KCS-FC	**74.12 ± 12.13**

KCS-FC (ours) achieves competitive classifier performance compared to state-of-the-art FC strategies.

Outline II

Kernel Cross-Spectral Functional Connectivity Network I

1. **1-D convolutional feature extraction and Gaussian pairwise similarity:**

$$\hat{\boldsymbol{P}}_r(\boldsymbol{w}_f) = \tilde{K}(\cdot;\sigma) \circ \varphi(\boldsymbol{X}_r;\boldsymbol{w}_f),$$

where $\tilde{K}(\tilde{\boldsymbol{X}}_r;\sigma) = [\boldsymbol{K}_{r1}, \boldsymbol{K}_{r2}, \cdots, \boldsymbol{K}_{rf}, \cdots, \boldsymbol{K}_{rN_f}]$ and $\boldsymbol{K}_{rf}$ is defined as:

$$\boldsymbol{K}_{rf} = \begin{bmatrix} \kappa_X(\boldsymbol{x}_{rf}^1, \boldsymbol{x}_{rf}^1;\sigma) & \kappa_X(\boldsymbol{x}_{rf}^1, \boldsymbol{x}_{rf}^2;\sigma) & \cdots & \kappa_X(\boldsymbol{x}_{rf}^1, \boldsymbol{x}_{rf}^{N_c};\sigma) \\ \kappa_X(\boldsymbol{x}_{rf}^2, \boldsymbol{x}_{rf}^1;\sigma) & \kappa_X(\boldsymbol{x}_{rf}^2, \boldsymbol{x}_{rf}^2;\sigma) & \cdots & \kappa_X(\boldsymbol{x}_{rf}^2, \boldsymbol{x}_{rf}^{N_c};\sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa_X(\boldsymbol{x}_{rf}^{N_c}, \boldsymbol{x}_{rf}^1;\sigma) & \kappa_X(\boldsymbol{x}_{rf}^{N_c}, \boldsymbol{x}_{rf}^2;\sigma) & \cdots & \kappa_X(\boldsymbol{x}_{rf}^{N_c}, \boldsymbol{x}_{rf}^{N_c};\sigma). \end{bmatrix}$$

2. **Average functional connectivity measure:**

$$\tilde{\boldsymbol{P}}_r = \text{AvgPooling}_f\left(\hat{\boldsymbol{P}}_r(\boldsymbol{w}_f)\right),$$

Kernel Cross-Spectral Functional Connectivity Network II

3. **Vectorized version of $\tilde{P}_r$:**

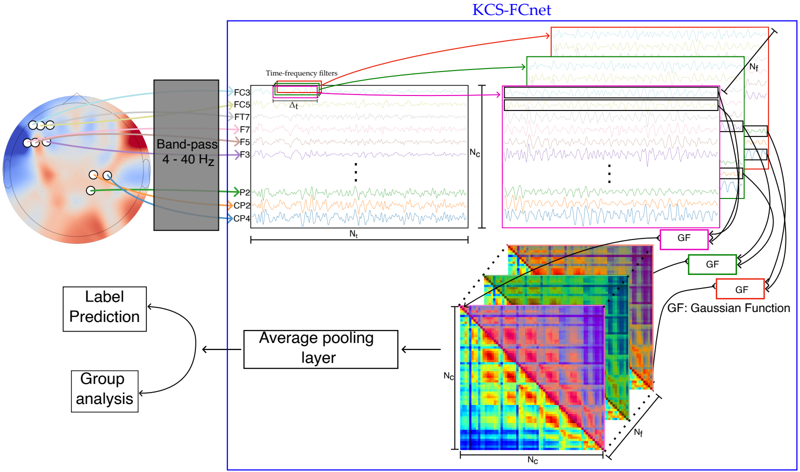
$$\overline{\boldsymbol{p}}_r = \left[\tilde{p}_r^{12}, \tilde{p}_r^{13}, \cdots, \tilde{p}_r^{cc'}, \cdots, \tilde{p}_r^{(N_c-1)N_c}\right]; \forall c < c',$$

4. **Optimization problem:**

$$\Theta^* = \arg\min_{\Theta} \quad \mathbb{E}_r\{\mathcal{L}(\boldsymbol{y}_r, \hat{\boldsymbol{y}}_r|\Theta); \forall r \in \{1, 2, \ldots, R\}\},$$

where $\mathcal{L}\{\cdot\}$ is a given loss function (i.e., cross-entropy) and $\hat{\boldsymbol{y}}_r = \mathrm{softmax}\left(\boldsymbol{V}\overline{\boldsymbol{p}}_r + \boldsymbol{b}\right)$

Kernel Cross-Spectral Functional Connectivity Network Proposal



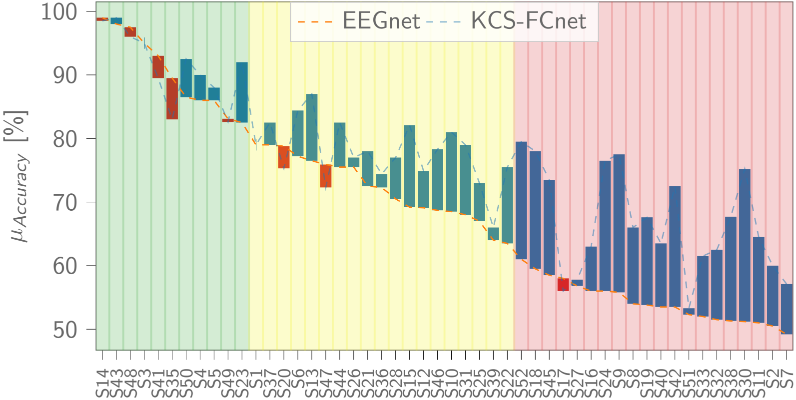
Kernel Cross-Spectral Functional Connectivity Network Architecture

Layer	Output Dimension	Params.
Input	$N_c \times N_t \times 1$	·
Conv2D	$N_c \times (N_t - \Delta_t + 1) \times N_f$	max norm $= 2.0$, kernel size $= (1, \Delta_t)$ Stride size $= (1, 1)$, Bias $=$ False
BatchNormalization	$N_c \times (N_t - \Delta_t + 1) \times N_f$	·
ELU activation		
KCS-FCblock	$N_f \times (N_c \cdot (N_c - 1)/2) \times 1$	·
AveragePooling2D	$1 \times (N_c \cdot (N_c - 1)/2) \times 1$	·
BatchNormalization	$1 \times (N_c \cdot (N_c - 1)/2) \times 1$	·
ELU activation		
Flatten	$N_c \cdot (N_c - 1)/2$	·
Dropout	$N_c \cdot (N_c - 1)/2$	Dropout rate $= 0.5$
Dense	N_y	max norm $= 0.5$
Softmax		

Experimental Set-up

1. **Raw EEG Preprocessing:**
   - Database used DBIII MI
   - Donwsampling from 512 Hz to 128 Hz.
   - Filtering from 4 Hz to 40 Hz.
   - Records clipped from 0.5 s to 2.5 s post cue.
2. **KCS-FCnet Training** [Lawhern et al., 2018, Schirrmeister et al., 2017]**:**
   - Data split using 5-fold 80-20 scheme.
   - 1-D convolutional kernel length set to 20
   - Number of filters were searched within the set $\{2, 3, 4\}$
   - Gaussian kernel $\sigma = \varsigma Median(\boldsymbol{D})$ with $\boldsymbol{D}$ being the distance matrix.
3. **Group-Level Analysis:**
   - Scaled scoring matrix with subjects and accuracy, Cohen's kappa, AUC.
   - Cluster subjects in three groups based on base line EEGnet.
   - PCA was used to reduce the dimensions to two, enabling us to plot it.

Subject Dependent and Group Analysis Results I



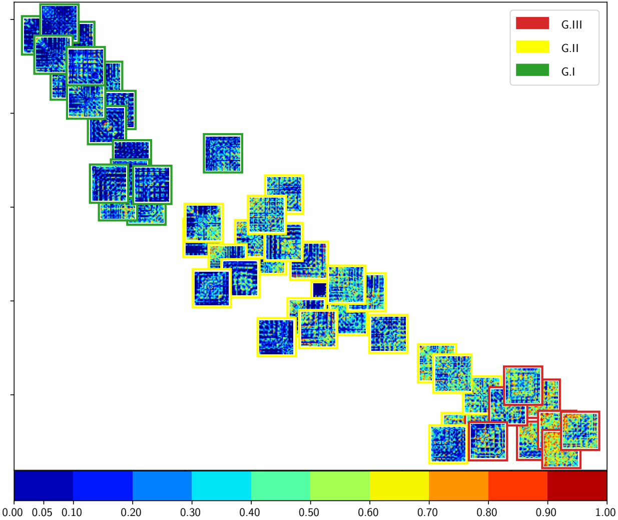
- KCS-FCnet (ours) shows strong impact on subjects in the third group.
- Seven subjects experienced a decrease in accuracy.
- 47 subjects experience an increase of more than five points in accuracy.

Subject Dependent and Group Analysis Results II

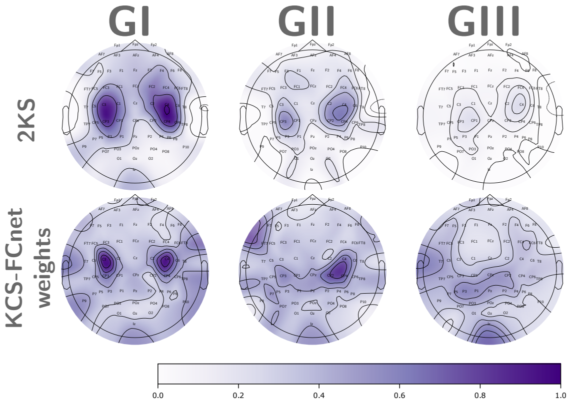
- Eleven subjects significantly improved their performance, transitioning to the GII cluster, while nine remained in GIII and six advanced to GI.
- Subjects initially in the best group retained their status.
- Our proposal boosts GII accuracy by 5.6% and GIII by 12.4%.

Functional Connectivity Analysis I

- Two-sample Kolmogorov–Smirnov (2KS) test for MI separability.
- Lower p-values indicate higher class separability and more informative results.
- For GIII, most class PDFs are indistinguishable.



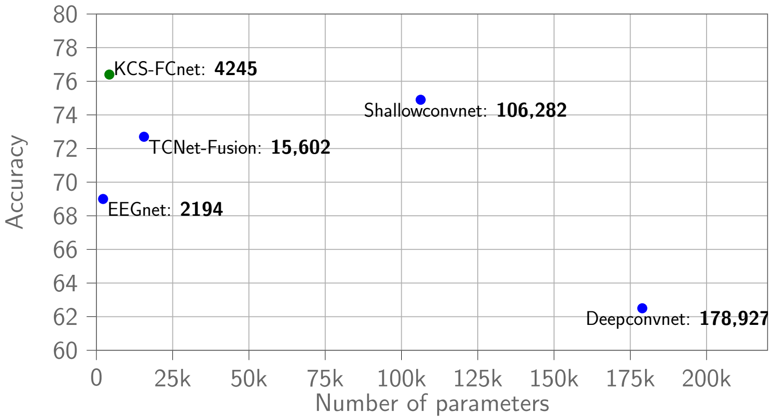
Functional Connectivity Analysis II



- GI shows similar results for both approaches.
- In GII, weighting approach highlights relevance around C4, while 2KS highlights C3 and C4.
- For GIII, the 2KS test reveals some importance around C3 and C4, whereas the weight-based approach lacks a clear pattern.

Classifier Accuracy Comparison of DL Approaches

- A higher number of trainable parameters does not guarantee better classification accuracy.
- KCS-FCNet outperforms ShallowConvNet while using 25 times fewer parameters.



Outline II

Interpretable Regularized Kernel Cross-Spectral FC Network I

1. **Parzen density estimation:**

$$\hat{\zeta}(\chi) = \frac{1}{N_c} \sum_{c=1}^{N_c} \kappa(\chi, x^c; \sigma),$$

2. **Renyi's entropy $\alpha = 2$:**

$$\begin{aligned} H_2(\chi) &= -\log \left(\int_{\chi} \hat{\zeta}(\chi)^2 \, d\zeta \right) \\ &= -\log \left(\frac{1}{N_c^2} \sum_{c,c'=1}^{N_c} \kappa(\mathbf{x}^c, \mathbf{x}^{c'}; \sigma) \right). \end{aligned}$$

Interpretable Regularized Kernel Cross-Spectral FC Network II

3. **Concept extended to function composition:**

$$\hat{\zeta}(\chi) = \frac{1}{N_c}\sum_{c=1}^{N_c} \kappa_X(\chi, \cdot; \sigma) \circ \varphi(\boldsymbol{x}^c; \boldsymbol{w}_f),$$

where $\varphi(\cdot; \boldsymbol{w}_f)$ is a 1-D convolutional layer.

4. **Optimization problem with Renyi's-based regularization:**

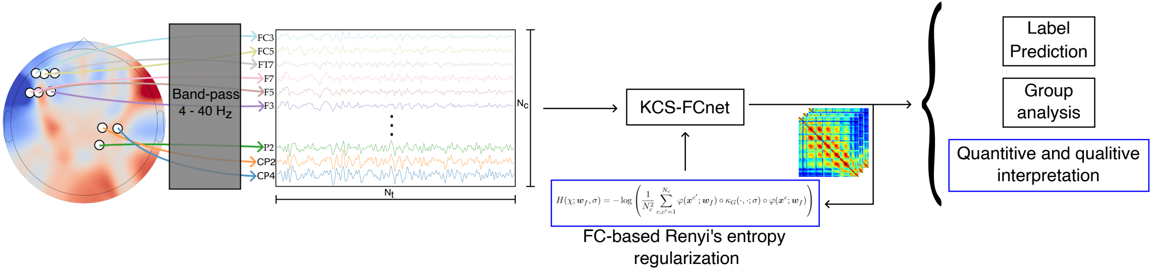
$$\Theta^* = \arg\min_{\Theta} \quad \mathbb{E}_r\left\{\mathcal{L}(\boldsymbol{y}_r, \hat{\boldsymbol{y}}_r|\Theta) - \rho H(\tilde{\boldsymbol{P}}_r; \boldsymbol{w}_f, \sigma)\right\}; \forall r \in \{1, 2, \dots, R\},$$

where

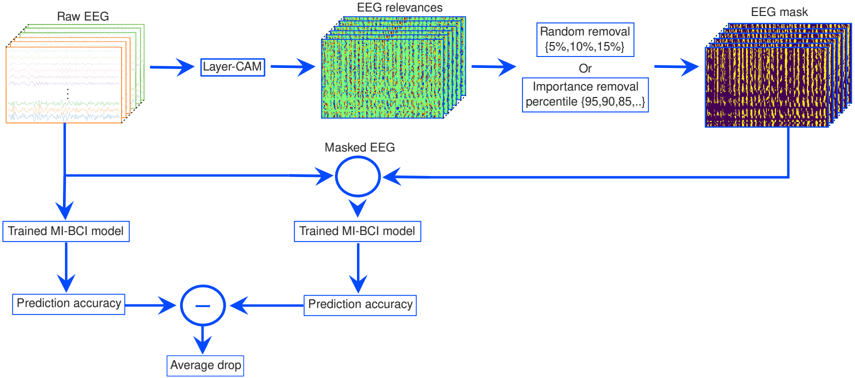
$$H(\tilde{\boldsymbol{P}}_r; \boldsymbol{w}_f, \sigma) = -\log\left(\frac{1}{N_c^2}\sum_{c,c'=1}^{N_c} \kappa_X(\cdot, \cdot; \sigma) \circ \left(\tilde{p}_r^{c'}(\boldsymbol{w}_f), \tilde{p}_r^{c}(\boldsymbol{w}_f)\right)\right).$$

[1] [Yu et al., 2019, Giraldo et al., 2014]

Interpretable Regularized KCS-FCnet Proposal I



Interpretable Regularized KCS-FCnet Proposal II

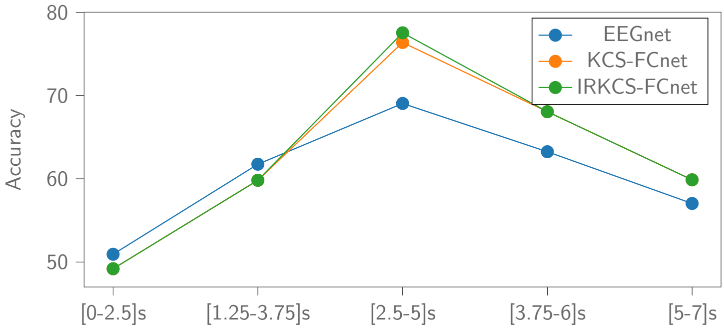


Quantitative average drop interpretation comparison

Experimental set-up

1. **Raw EEG Preprocessing:**
   - Database used DBIII MI
   - Donwsampling from 512 Hz to 128 Hz.
   - Filtering from 4 Hz to 40 Hz.
2. **IRKCS-FCnet Training [Lawhern et al., 2018, Schirrmeister et al., 2017]:**
   - Data split using 5-fold 80-20 scheme.
   - 1-D convolutional kernel length set to 20.
   - Number of filters were taken from the best results within the original model.
   - Gaussian kernel $\sigma = \varsigma Median(\boldsymbol{D})$ with $\boldsymbol{D}$ being the distance matrix and ς a trainable parameter.
   - Regularization hyperparamter ρ is searched within the set $\{0, 0.2, 0.5, 0.7, 1, 5, 10\}$.
3. **Group-Level Analysis:**
   - Scaled scoring matrix with subjects and accuracy, Cohen's kappa, AUC.
   - Cluster subjects in three groups absed on base line EEGnet.
   - PCA was used to reduce the dimensions to two, enabling us to plot it.

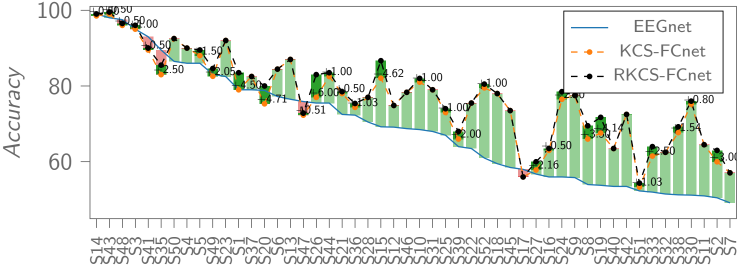
Model Level Analysis I



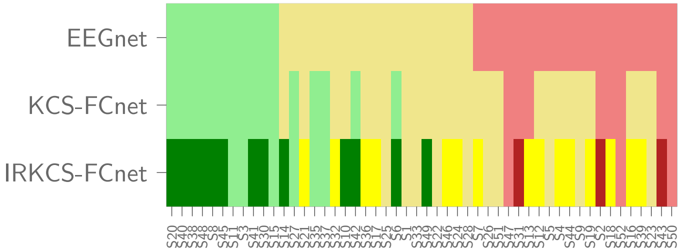
- Average performance comparison at different time windows.
- All three models show reduced accuracy with no or insufficient MI information.
- All models achieve peak accuracy in the [2.5 - 5]s window.

Model Level Analysis II

- Regularized version consistently performs at least as well as, or better than, the plain version.
- Accuracy boost of more than four percentage points observed for subjects 1, 20, 26, and 15.
- Increases in accuracy were not significant in subjects already achieving more than 80% accuracy.



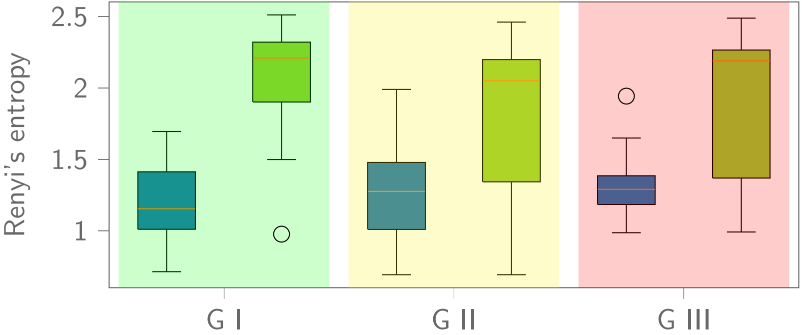
Group Level Analysis I



- Light-colored symbols mark subjects who remain with same accuracy.
- Subjects 14, 10, and 49 are promoted from the intermediate to the best group.
- Subjects 13 and 18 moved up from the worst to the intermediate Group
- Accuracy improvement of GII from 72.15% in EEGnet to nearly 80% in IRKCS-FCnet (ours).

Group Level Analysis II

- IRKCS-FCnet (ours) consistently shows higher entropy values across all groups.
- Except for one, all samples exceed 1.5 in GI for IRKCS-FCnet (ours).
- Best-performing subjects may rely on fewer connections to achieve high accuracy.



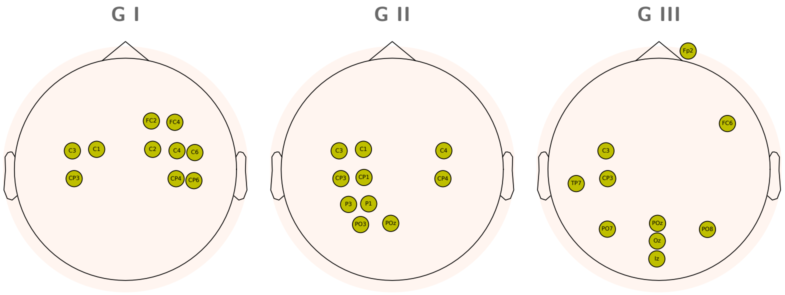
Post-Hoc Interpretability

Average accuracy drop across feature removal percentage

Strategy	Group	25%	20%	15%	10%	5%
KCS-FCnet	G I	40.51	37.93	35.97	30.25	17.12
	G II	40.10	39.00	36.46	28.60	17.72
	G III	39.05	36.98	31.99	25.28	15.02
IRKCS-FCnet	G I	44.80	39.21	37.79	30.55	22.46
	G II	41.08	40.01	36.77	27.28	18.84
	G III	39.65	37.30	33.12	27.10	16.80

- The highest difference per group is indicated by ( ) for GI, ( ) for GII, and ( ) for GIII.
- Plain model drops 17.12 accuracy points, while the regularized model drops 22.46 points for 5% case.
- As feature removal increases, the accuracy drop discrepancy diminishes, plateauing around 40% from 15% to 25% removal.

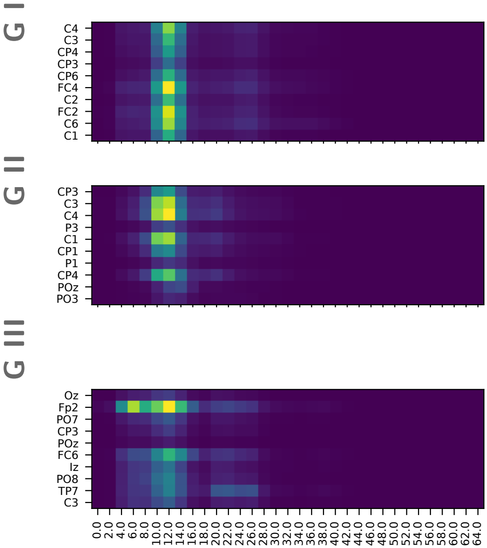
Intrinsic Interpretability I



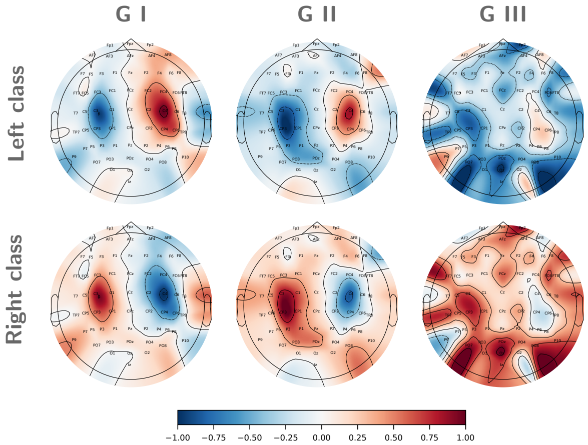
- Ten most relevant channels for each group based on the last layer weights of IRKCS-FCnet.
- For GI, the best-performing group, most channels are located in the sensorimotor cortex (CP3, CP4, CP6), central sulcus (C1, C2, C3, C4, C6), and motor cortex (FC2, FC4).
- In GII, 60% of the most relevant channels are located in the motor-related brain area, while in GIII, this value drops to only 30%.

Intrinsic Interpretability II

- In GI, the spectrum is stable, with power concentrated between 10 and 14 Hz, consistent with studies showing MI present on Mu rhythm (8–13 Hz) [Al-Saegh et al., 2021, Hobson and Bishop, 2017, Llanos et al., 2013].
- GII shows a similar pattern, but the spectral power is slightly skewed to the right.
- In GIII, the Fp2 channel dominates spectral power, indicating significant artifact introduction [Han et al., 2023].



Intrinsic Interpretability III



- For GI and GII, positive values are seen in the contralateral region while negative for the ipsilateral region [Van der Lubbe et al., 2021].
- GIII exhibits intriguing behavior, showing only faint significance in the contralateral region.
- GIII also displays numerous activations outside motor-related areas, likely due to artifact interference.

Outline II

Conclusions I

- The proposed **single-trial KCS-FC** achieves **competitive classifier performance** while maintaining FC **interpretability** (Objective 1).
- **KCS-FCnet** automatically extracts **spatio-temporal-spectral feature** maps, outperforming other end-to-end architectures. Slightly more complex than EEGnet, it offers **greater accuracy**, especially compared to models with 25 times more parameters (Objective 2).
- The **IRKCS-FCnet** is an **intrinsically interpretable model**, providing valuable insights into spatial connections for MI tasks, particularly emphasizing the **contralateral nature of motor-related** neurological states within the brain network (Objective 3).

Outline II

Future work I

- Extracting more information from time lags, rather than relying solely on instantaneous relationships, could help address volume conduction issues [Uribe et al., 2019, Bakhshali et al., 2020, De La Pava Panche et al., 2019].

- The current method's reliance on pairwise relationships may overlook the influence of a third channel, suggesting the potential benefits of more complex multivariate approaches that account for information from other electrodes [Vidaurre et al., 2023].

Future work II

- Incorporate Riemannian geometry to capture the internal relationships between different channels in the FC matrix's feature extraction process [Carrara and Papadopoulo, 2023].

- To compare with state-of-the-art methods we employ a 5-fold cross-validation (80-20 split) for our proposals, which may lead to hyperparameter overfitting. Future work could explore subject-independent or cross-dataset approaches to mitigate this risk [Wei et al., 2023]..

Outline II

Academic Products

1. García-Murillo, D.G.; Alvarez-Meza, A.; Castellanos-Dominguez, G. Single-Trial Kernel-Based Functional Connectivity for Enhanced Feature Extraction in Motor-Related Tasks. Sensors 2021, 21, 2750. https://doi.org/10.3390/s21082750 (Q1-A1)
2. García-Murillo, D.G.; Álvarez-Meza, A.M.; Castellanos-Dominguez, C.G. KCS-FCnet: Kernel Cross-Spectral Functional Connectivity Network for EEG-Based Motor Imagery Classification. Diagnostics 2023, 13, 1122. https://doi.org/10.3390/diagnostics13061122 (Q2-A2)
3. García-Murillo, D.G.; Álvarez-Meza, A.M.; Castellanos-Dominguez, C.G. IRKCS-FCnet: Interpretable Regularized Kernel Cross-Spectral Functional Connectivity Network with Qualitative and Quantitative Post-Hoc and Intrinsic Explainability (Sent)
4. Github repository containing all notebooks used in this study IRKCS-FCnet reporsitory

Acknowledgements

I recognize that this research would not have been possible without the support given by the projects: "Alianza científica con enfoque comunitario para mitigar brechas de atención y manejo de trastornos mentales y epilepsia en Colombia (ACEMATE)." (Code 111091991908, Hermes Code 56118) funded by MINCIENCIAS, and " Sistema de visión artificial para el monitoreo y seguimiento de efectos analgésicos y anestésicos administrados vía neuroaxial epidural en población obstétrica durante labores de parto para el fortalecimiento de servicios de salud materna del Hospital Universitario de Caldas - SES HUC." (Hermes Code 57661), funded by Universidad Nacional de Colombia.

Outline II

References I

Ahn, M., Lee, M., Choi, J., and Jun, S. C. (2014).
A review of brain-computer interface games and an opinion survey from researchers, developers and users.
Sensors, 14(8):14601–14633.

Ai, Q., Chen, A., Chen, K., Liu, Q., Zhou, T., Xin, S., and Ji, Z. (2019).
Feature extraction of four-class motor imagery EEG signals based on functional brain network.
Journal of Neural Engineering, 16(2):026032.

Al-Saegh, A., Dawwd, S. A., and Abdul-Jabbar, J. M. (2021).
Deep learning for motor imagery EEG-based classification: A review.
Biomedical Signal Processing and Control, 63:102172.

Alsharif, A., Salleh, N., Baharun, R., and Safaei, M. (2020).
Neuromarketing approach: An overview and future research directions.
Journal of Theoretical and Applied Information Technology, 98(7):991–1001.

Altaheri, H., Muhammad, G., Alsulaiman, M., Amin, S. U., Altuwaijri, G. A., Abdul, W., Bencherif, M. A., and Faisal, M. (2023).
Deep learning techniques for classification of electroencephalogram (eeg) motor imagery (mi) signals: A review.
Neural Computing and Applications, 35(20):14681–14722.

Bakhshali, M., Khademi, M., Ebrahimi, A., and Moghimi, S. (2020).
Eeg signal classification of imagined speech based on riemannian distance of correntropy spectral density.
Biomedical Signal Processing and Control, 59:101899.

References II

Barios, J. A., Ezquerro, S., Bertomeu-Motos, A., Nann, M., Badesa, F. J., Fernandez, E., Soekadar, S. R., and Garcia-Aracil, N. (2019).
Synchronization of slow cortical rhythms during motor imagery-based brain–machine interface control.
International journal of neural systems, 29(05):1850045.

Bochner, S. (2020).
Harmonic analysis and the theory of probability.
University of California press.

Bonci, A., Fiori, S., Higashi, H., Tanaka, T., and Verdini, F. (2021).
An introductory tutorial on brain–computer interfaces and their applications.
Electronics, 10(5):560.

Cao, J., Zhao, Y., Shan, X., Wei, H.-l., Guo, Y., Chen, L., Erkoyuncu, J. A., and Sarrigiannis, P. G. (2022a).
Brain functional and effective connectivity based on electroencephalography recordings: A review.
Human brain mapping, 43(2):860–879.

Cao, L., Wang, W., Huang, C., Xu, Z., Wang, H., Jia, J., Chen, S., Dong, Y., Fan, C., and de Albuquerque, V. H. C. (2022b).
An effective fusing approach by combining connectivity network pattern and temporal-spatial analysis for eeg-based bci rehabilitation.
IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30:2264–2274.

Carrara, I. and Papadopoulo, T. (2023).
Classification of bci-eeg based on augmented covariance matrix.
arXiv preprint arXiv:2302.04508.

References III

Casimo, K., Weaver, K. E., Wander, J., and Ojemann, J. G. (2017).
Bci use and its relation to adaptation in cortical networks.
IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25(10):1697–1704.

Cattan, G., Mendoza, C., Andreev, A., and Congedo, M. (2018).
Recommendations for integrating a p300-based brain computer interface in virtual reality environments for gaming.
Computers, 7(2):34.

Chamola, V., Vineet, A., Nayyar, A., and Hossain, E. (2020).
Brain-computer interface-based humanoid control: A review.
Sensors, 20(13):3620.

Chevallier, S., Carrara, I., Aristimunha, B., Guetschel, P., Sedlar, S., Lopes, B., Velut, S., Khazem, S., and Moreau, T. (2024).
The largest eeg-based bci reproducibility study for open science: the moabb benchmark.
arXiv preprint arXiv:2404.15319.

Chiarion, G., Sparacino, L., Antonacci, Y., Faes, L., and Mesin, L. (2023).
Connectivity analysis in eeg data: A tutorial review of the state of the art and emerging trends.
Bioengineering, 10(3):372.

Cho, H., Ahn, M., Ahn, S., Kwon, M., and Jun, S. (2017).
EEG datasets for motor imagery brain–computer interface.
GigaScience, 6(7):gix034.

References IV

Cohen, L. (1998).
The generalization of the wiener-khinchin theorem.
In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, volume 3, pages 1577–1580. IEEE.

Collazos-Huertas, D. F., Álvarez-Meza, A. M., Cárdenas-Peña, D. A., Castaño-Duque, G. A., and Castellanos-Domínguez, C. G. (2023).
Posthoc interpretability of neural responses by grouping subject motor imagery skills using cnn-based connectivity.
Sensors, 23(5):2750.

De La Pava Panche, I., Alvarez-Meza, A., and Orozco-Gutierrez, A. (2019).
A data-driven measure of effective connectivity based on renyi's α-entropy.
Frontiers in neuroscience, 13:1277.

Fan, F.-L., Xiong, J., Li, M., and Wang, G. (2021).
On interpretability of artificial neural networks: A survey.
IEEE Transactions on Radiation and Plasma Medical Sciences, 5(6):741–760.

Giraldo, L. G. S., Rao, M., and Principe, J. C. (2014).
Measures of entropy from data using infinitely divisible kernels.
IEEE Transactions on Information Theory, 61(1):535–548.

Gonzalez-Astudillo, J., Cattai, T., Bassignana, G., Corsi, M.-C., and Fallani, F. D. V. (2020).
Network-based brain computer interfaces: principles and applications.
Journal of Neural Engineering.

References V

Grigorev, N. A., Savosenkov, A. O., Lukoyanov, M. V., Udoratina, A., Shusharina, N. N., Kaplan, A. Y., Hramov, A. E., Kazantsev, V. B., and Gordleeva, S. (2021).
A bci-based vibrotactile neurofeedback training improves motor cortical excitability during motor imagery.
IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29:1583–1592.

Hamedi, M., Salleh, S.-H., Noor, A. M., and Mohammad-Rezazadeh, I. (2014).
Neural network-based three-class motor imagery classification using time-domain features for bci applications.
In *2014 IEEE region 10 symposium*, pages 204–207. IEEE.

Han, S., Zhang, C., Lei, J., Han, Q., Du, Y., Wang, A., Bai, S., and Zhang, M. (2023).
Cepstral analysis based artifact detection, recognition and removal for prefrontal eeg.
IEEE Transactions on Circuits and Systems II: Express Briefs.

He, B., Astolfi, L., Valdés-Sosa, P. A., Marinazzo, D., Palva, S. O., Bénar, C.-G., Michel, C. M., and Koenig, T. (2019).
Electrophysiological brain connectivity: theory and implementation.
IEEE Transactions on Biomedical Engineering, 66(7):2115–2137.

Hobson, H. M. and Bishop, D. V. (2017).
The interpretation of mu suppression as an index of mirror neuron activity: past, present and future.
Royal Society Open Science, 4(3):160662.

Hosseini, M.-P., Hosseini, A., and Ahi, K. (2020).
A review on machine learning for EEG signal processing in bioengineering.
IEEE reviews in biomedical engineering.

References VI

Janapati, R., Dalal, V., and Sengupta, R. (2023).
Advances in modern eeg-bci signal processing: A review.
Materials Today: Proceedings, 80:2563–2566.

Khan, M. A., Das, R., Iversen, H. K., and Puthusserypady, S. (2020).
Review on motor imagery based BCI systems for upper limb post-stroke neurorehabilitation: From designing to application.
Computers in Biology and Medicine, page 103843.

Kumar, S., Sharma, A., and Tsunoda, T. (2019).
Brain wave classification using long short-term memory network based OPTICAL predictor.
Scientific reports, 9(1):1–13.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J. (2018).
Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces.
Journal of neural engineering, 15(5):056013.

Li, S., Xie, X., Gu, Z., Yu, Z. L., and Li, Y. (2019).
Motor imagery classification based on local isometric embedding of riemannian manifold.
In *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 2368–2372. IEEE.

Llanos, C., Rodriguez, M., Rodriguez-Sabate, C., Morales, I., and Sabate, M. (2013).
Mu-rhythm changes during the planning of motor and motor imagery actions.
Neuropsychologia, 51(6):1019–1026.

References VII

Musallam, Y. K., AlFassam, N. I., Muhammad, G., Amin, S. U., Alsulaiman, M., Abdul, W., Altaheri, H., Bencherif, M. A., and Algabri, M. (2021).
Electroencephalography-based motor imagery classification using temporal convolutional network fusion.
Biomedical Signal Processing and Control, 69:102826.

Oikonomou, V. P., Georgiadis, K., Liaros, G., Nikolopoulos, S., and Kompatsiaris, I. (2017).
A comparison study on eeg signal processing techniques using motor imagery eeg data.
In *2017 IEEE 30th international symposium on computer-based medical systems (CBMS)*, pages 781–786. IEEE.

Purves, W. K. (c 2001.).
Life, the science of biology /.
Sinauer Associates,, Sunderland, MA :, 6th ed. edition.
Includes index.

Rodrigues, P., Stefano, C., Attux, R., Castellano, G., and Soriano, D. (2019).
Space-time recurrences for functional connectivity evaluation and feature extraction in motor imagery brain-computer interfaces.
Medical & biological engineering & computing, 57(8):1709–1725.

Roy, G., Bhoi, A., and Bhaumik, S. (2022).
A comparative approach for mi-based eeg signals classification using energy, power and entropy.
IRBM, 43(5):434–446.

Samuel, O. W., Geng, Y., Li, X., and Li, G. (2017).
Towards efficient decoding of multiple classes of motor imagery limb movements based on eeg spectral and time domain descriptors.
Journal of medical systems, 41:1–13.

References VIII

Sarraf, S. (2017).
Eeg-based movement imagery classification using machine learning techniques and welch's power spectral density estimation.
American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 33(1):124–145.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T. (2017).
Deep learning with convolutional neural networks for EEG decoding and visualization.
Human brain mapping, 38(11):5391–5420.

Singh, A., Hussain, A. A., Lal, S., and Guesgen, H. W. (2021).
A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain-computer interface.
Sensors, 21(6):2173.

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M. L., Rana, M., Oblak, E., et al. (2017).
Closed-loop brain training: the science of neurofeedback.
Nature Reviews Neuroscience, 18(2):86–100.

Thiebaut de Schotten, M., Foulon, C., and Nachev, P. (2020).
Brain disconnections link structural connectivity with function and behaviour.
Nature communications, 11(1):5094.

Uribe, L. F. S., Stefano Filho, C. A., de Oliveira, V. A., da Silva Costa, T. B., Rodrigues, P. G., Soriano, D. C., Boccato, L., Castellano, G., and Attux, R. (2019).
A correntropy-based classifier for motor imagery brain-computer interfaces.
Biomedical Physics & Engineering Express, 5(6):065026.

References IX

Van der Lubbe, R. H., Sobierajewicz, J., Jongsma, M. L., Verwey, W. B., and Przekoracka-Krawczyk, A. (2021).
Frontal brain areas are more involved during motor imagery than during motor execution/preparation of a response sequence.
International journal of psychophysiology, 164:71–86.

Vărbu, K., Muhammad, N., and Muhammad, Y. (2022).
Past, present, and future of eeg-based bci applications.
Sensors, 22(9):3331.

Vidaurre, C., Gurunandan, K., Idaji, M. J., Nolte, G., Gómez, M., Villringer, A., Müller, K.-R., and Nikulin, V. V. (2023).
Novel multivariate methods to track frequency shifts of neural oscillations in eeg/meg recordings.
NeuroImage, 276:120178.

Wei, M., Yang, R., Huang, M., Ni, J., Wang, Z., and Liu, X. (2023).
Sub-band cascaded csp-based deep transfer learning for cross-subject lower limb motor imagery classification.
IEEE Transactions on Cognitive and Developmental Systems.

Xiao, C., Choi, E., and Sun, J. (2018).
Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review.
Journal of the American Medical Informatics Association, 25(10):1419–1428.

Yilmaz, C. M., Kose, C., and Hatipoglu, B. (2018).
A quasi-probabilistic distribution model for eeg signal classification by using 2-d signal representation.
Computer methods and programs in biomedicine, 162:187–196.

References X

Yu, S., Giraldo, L. G. S., Jenssen, R., and Principe, J. C. (2019).
Multivariate extension of matrix-based rényi's α-order entropy functional.
IEEE transactions on pattern analysis and machine intelligence, 42(11):2960–2966.

Zhang, Y., Nam, C., Zhou, G., Jin, J., Wang, X., and Cichocki, A. (2018).
Temporally constrained sparse group spatial patterns for motor imagery BCI.
IEEE transactions on cybernetics, 49(9):3322–3332.

Group	KCS-FCnet vs. Random	RKCS-FCnet vs. Random
G I	0.00077*	0.00230*
G II	0.00368*	0.00434*
G III	0.00108*	0.00074*

Strategy	Group	25%	20%	15%	10%	5%
KCS-FCnet	G I	40.51	37.93	35.97	30.25	17.12
	G II	40.10	39.00	36.46	28.60	17.72
	G III	39.05	36.98	31.99	25.28	15.02
RKCS-FCnet	G I	44.80	39.21	37.79	30.55	22.46
	G II	41.08	40.01	36.77	27.28	18.84
	G III	39.65	37.30	33.12	27.10	16.80

- The accuracy drop upon important feature removal in both models is statistically different from the random drop.
- Plain model drops 17.12 accuracy points, while the regularized model drops 22.46 points, a difference of over 5 points.
- As feature removal increases, the accuracy drop discrepancy diminishes, plateauing around 40% from 15% to 25% removal.

Group Level Analysis II

Strategy	G I	G II	G III
EEGnet	90.55 $\pm$ 5.88	72.15 $\pm$ 4.87	54.27 $\pm$ 3.21
KCS-FCnet	91.46 $\pm$ 5.31	77.85 $\pm$ 4.76	66.66 $\pm$ 7.88
IRKCS-FCnet	**92.28** $\pm$ **4.79**	**79.26** $\pm$ **4.93**	**67.77**$\pm$ **7.83**

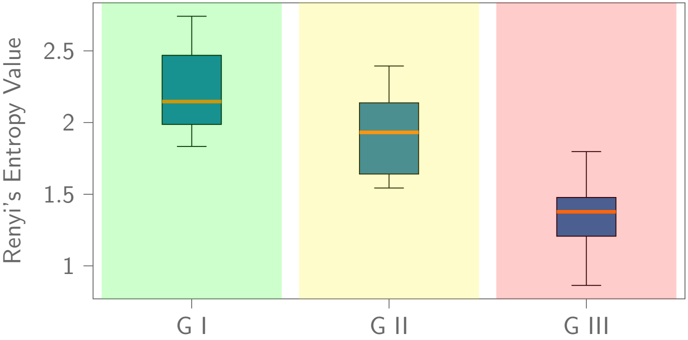
- The standard deviation for the regularized version decreases by over one percentage point, indicating enhanced performance consistency.
- Accuracy improvement of GII from 72.15% in EEGnet to nearly 80% in IRKCS-FCnet (ours).

Classifier Accuracy Comparison of DL Approaches I

Approach	Accuracy	Kappa	AUC
Deepconvnet [Schirrmeister et al., 2017]	62.5 ± 13.0	24.5 ± 25.9	68.9 ± 17.8
EEGnet [Lawhern et al., 2018]	69.0 ± 14.6	38.0 ± 29.1	75.4 ± 16.6
TCNet-Fusion [Musallam et al., 2021]	72.7 ± 14.0	45.0 ± 28.2	79.6 ± 15.9
Shallowconvnet [Schirrmeister et al., 2017]	74.9 ± 13.9	49.5 ± 27.8	79.9 ± 15.1
KCS-FCnet	76.4 ± 11.3	52.6 ± 22.7	82.2 ± 12.2

- DeepConvNet performs the worst, making it unsuitable for high intra-class variability.
- KCS-FCNet (ours) achieves the highest scores with the lowest standard deviation.

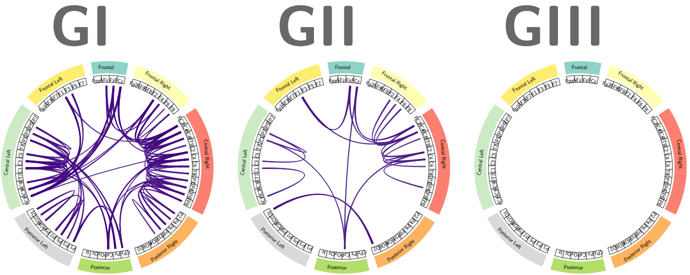
Functional Connectivity Analysis II



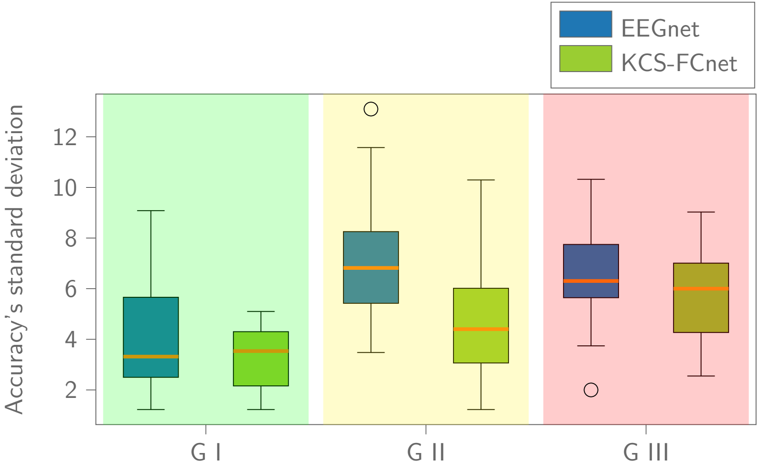
- The Rényi's entropy $\alpha = 2$ quantifies the interpretability performance from 2KS test matrices.
- Higher entropy values indicates a higher class distribution separability.
- Groups that perform better show higher entropy values.

Functional Connectivity Analysis III

- Relevant connections identified from 2KS test matrices.
- In GI, most relevant connections are found in motor-related areas.
- In GII lacks relevant connections.



Subject Dependent and Group Analysis Results IV



- For GI, KCS-FCnet (ours) reduces variability.
- For GII, KCS-FCnet (ours) lowers standard deviation, with similar variability proportion.
- For GIII, both approaches behave similarly.

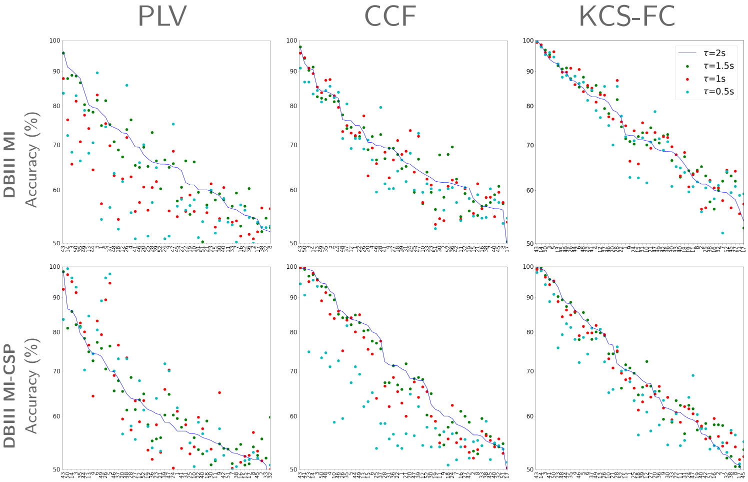
Subject Dependent and Group Analysis Results III

Approach	Group	Accuracy	KCS-FCnet Gain
EEGnet	G I	90.6 ± 4.3	·
	G II	72.2 ± 7.3	·
	G III	54.3 ± 6.6	·
KCS-FCnet	G I	91.5 ± 3.3	**0.9**
	G II	77.8 ± 4.7	**5.6**
	G III	66.7 ± 5.6	**12.4**

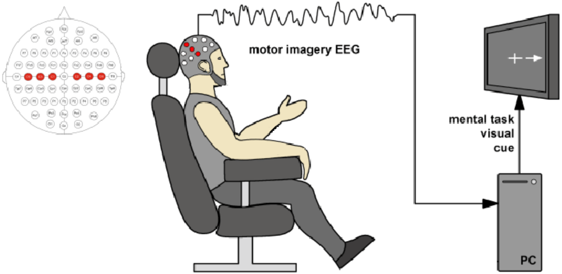
Our proposal not only outperforms EEGnet in terms of accuracy but also reduces the variability for all clusters.

Impact of Prior CSP Filtering

- PLV shows high variance and low accuracy.
- CSP accuracy drops at $\tau = 0.5$ due to its reliance on FC estimation, where less information leads to poorer estimates.
- KCS-FC (ours) outperforms other FC measures, with less variance in accuracy across subjects.



Experimental Setup for MI-based BCIs



- EEG headsets are equipped with 1 to 256 electrodes [Grigorev et al., 2021].
- Visual cues are often used to guide MI tasks during EEG recordings [Hosseini et al., 2020].

[1] **Image:** Adapted from [Grigorev et al., 2021]