



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

Daniel Guillermo García Murillo

Universidad Nacional de Colombia  
Facultad de ingeniería eléctrica, electrónica y computación  
Departamento de ingeniería y arquitectura  
Manizales, Colombia  
2024

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

Daniel Guillermo García Murillo

Tesis presentada como requisito parcial para optar por el título de:  
**Doctor en Ingeniería - Ingeniería Automática**

Director(a):  
Prof. Dr. Andrés Marino Álvarez Meza  
Codirector(a):  
Prof. Dr. César Germán Castellanos Domínguez

Línea de investigación:  
Inteligencia Artificial  
Grupo de investigación:  
Grupo de control y procesamiento digital de señales (GCPDS)

Universidad Nacional de Colombia  
Facultad de ingeniería eléctrica, electrónica y computación  
Departamento de ingeniería y arquitectura  
2024

"Logic will get you from A to B. Imagination will take you everywhere."

*Albert Einstein*

"Believe you can, and you are halfway there."

*Theodore Roosevelt*

# Declaración

Me permito afirmar que he realizado ésta tesis de manera autónoma y con la única ayuda de los medios permitidos y no diferentes a los mencionados el presente texto. Todos los pasajes que se han tomado de manera textual o figurativa de textos publicados y no publicados, los he reconocido en el presente trabajo. Ninguna parte del presente trabajo se ha empleado en ningún otro tipo de tesis.

Manizales., 01-29-2024

---

Daniel Guillermo García Murillo

# Acknowledgements

Firstly, I extend my gratitude to my supervisor, Andrés Marino Álvarez, and my co-supervisor, Germán Castellanos Domínguez. Their invaluable guidance and support have been indispensable in successfully completing this research. Their contributions have influenced this study and left an indelible impact on my academic journey.

My deepest acknowledgement is extended to my beloved wife who has consistently been a source of strength, providing unwavering support throughout every challenge. Additionally, I wish to express my heartfelt thanks to my dear parents and sister who have persistently stood by me. Their words of wisdom during challenging times and unconditional love have been my pillars of support and stability. Lastly, the most special acknowledgment is reserved for my son, Emanuel, who has been a beacon of joy and hope.

I am particularly grateful to all members of the Signal Processing and Recognition Group (SPRG) at the National University of Colombia in Manizales. A special note of thanks goes to Julian Caicedo and Mateo Tobón whose camaraderie and academic discussions have significantly enriched my experience over these years.

Finally, 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.

# Resumen

Condiciones como la esclerosis lateral amiotrófica, los derrames cerebrales y las lesiones de la médula espinal pueden causar interrupciones significativas en las vías de comunicación del cerebro, afectando el control muscular voluntario y la interacción con el entorno. Las interfaces cerebro-computadora (BCIs) son sistemas basados en computadoras que se utilizan para abordar estos desafíos; interpretan las señales cerebrales y han sido usadas en aplicaciones médicas, de rehabilitación y de entretenimiento. Particularmente en el campo de BCI, la Imaginación Motora (MI), un paradigma ampliamente investigado, posee un enorme potencial para la recuperación de la funcionalidad motora. La integración de MI-BCIs con terapeutas especializados ha dado resultados positivos en los procesos de rehabilitación sensorial y motora, demostrando ser beneficiosos para las personas con trastornos neurológicos. Sin embargo, los sistemas MI-BCI no se limitan a aplicaciones clínicas; también son de gran relevancia en entornos no clínicos como la realidad virtual, los juegos y la adquisición de habilidades motoras.

Usualmente, las señales cerebrales en los sistemas MI-BCI **son** recolectada usando electroencefalografía (EEG) debido a su alta resolución temporal, asequibilidad y portabilidad. Estas características lo convierten en un método viable para capturar eventos de corta duración y patrones de actividad neuronal temporal. No obstante, la decodificación de las señales de EEG es compleja debido a su alta tasa de muestreo y al elevado número de electrodos involucrados, lo que resulta en una cantidad abrumadora de datos. Varias metodologías prometedoras, aprovechando el concepto de sincronización relacionada de eventos (ERS) y desincronización relacionada de eventos (ERD), se han desarrollado para facilitar la extracción de características pertinentes y así decodificar EEG en patrones del área sensoriomotora (SMR). Aunque son importantes, las características de un solo canal tienden a simplificar demasiado los fenómenos generales ya que no consideran las interacciones de las distintas regiones cerebrales. Teniendo en cuenta las activaciones y comunicaciones complejas dentro de varias áreas cerebrales durante la ejecución o la imaginación de tareas motoras sencillas, los enfoques de extracción de características de múltiples canales son más precisos. Por ejemplo, el método de Patrones Espaciales Comunes (CSP) ha demostrado su importancia en MI-BCI, promoviendo la extracción de características altamente discriminativas al maximizar la separabilidad de las características SMR. La conectividad cerebral (BC), un factor clave en estos enfoques, describe las interacciones dentro y entre las regiones cerebrales y se correlaciona con los mecanismos de sincronización dentro de las modulaciones oscilatorias. BC se puede dividir en conectividad estructural/anatómica (SC), conectividad efectiva (EC) y conectividad funcional (FC), cada una con características distintas. SC se centra en conexiones físicas y es independiente del contexto, mientras que EC y FC, adecuados para sistemas MI-BCI basados en EEG, decodifican estados cognitivos de evolución rápida donde FC es particularmente adecuado para aplicaciones de MI-BCI debido a su simplicidad, baja demanda computacional y falta de supuestos previos rígidos.

Esta tesis aborda tres desafíos predominantes en los sistemas MI-BCI basados en FC EEG. En primer lugar, los déficits de rendimiento en las tareas de clasificación de MI son a menudo producto de los altos niveles de ruido presentes en las señales de EEG. En segundo lugar, depender del conocimiento experto para la extracción de características a menudo presenta una barrera. Por último, existe una necesidad de mayor transparencia de dichos modelos. El estudio aporta contribuciones hacia **las FC single-trial** en los sistemas MI-BCI en respuesta a estos desafíos. Esta mejora se logra utilizando un enfoque novedoso basado en la aproximación espectral de kernel cruzado proporcionada por la extensión del teorema de Bochner. Este nuevo estimador de FC de kernel cruzado espectral (KCS) aborda problemas de ruido, conexiones espurias y variabilidad

inter-sujeto, mejorando significativamente la precisión. Además, se ha conceptualizado un enfoque end-to-end, KCS-FCnet, para mejorar la representación automática del EEG para MI-BCI. Este enfoque es fundamental para reducir los artefactos e identificar con precisión las conexiones relevantes. Además, ayuda en la extracción de características espectrales, temporales y espaciales del EEG. Reconociendo la necesidad de transparencia del modelo, este estudio propuso una estrategia para mejorar la transparencia e interpretabilidad del modelo, denominada KCS-FCnet regularizada e interpretable (IRKCS-FCnet). Capitalizando el potencial de la regularización, la interpretabilidad del modelo se fortalece minimizando las conexiones no relevantes. Esto se realiza maximizando el potencial de información cruzada utilizando la entropía de Renyi medida en  $\alpha = 2$ , junto con una función de entropía cruzada. Para complementar estas mejoras, se han empleado las estrategias de pesos en la última capa y Layer-CAM para proporcionar interpretabilidad intrínseca y post-hoc, respectivamente. El concepto de caída promedio respalda aún más esta estrategia, mejorando la transparencia y la comprensión del sistema.

Los resultados muestran que el KCS-FC propuesto logra un incremento en el rendimiento que los estimadores clásicos de FC y tiene una precisión competitiva contra estrategias más complejas basadas en FC, alcanzando el segundo lugar en DBI<sup>1</sup> con una precisión del 81.92% y el primer lugar con 74.12% en DBIII<sup>2</sup>, ambas tareas de clasificación binaria de MI. Además, el KCS-FCnet muestra un rendimiento 76.4% en DBIII manteniendo la arquitectura con 4245 parámetros entrenables, lo que lo convierte en el mejor en comparación con los modelos DL comparados. Por último, el IRKCS-FCnet muestra un mejor rendimiento, aumentando 2 puntos porcentuales para individuos con bajo rendimiento, y muestra una mejor localización e interpretaciones, demostrando que al eliminar el 5% de las características más importantes, la puntuación promedio de caída de precisión cae 22.46, 5 puntos porcentuales más alta que la versión sin regularización. Además, se puede ver el patrón contralateral en la interpretación de topoplot para ambas manos, la derecha y la izquierda. Este estudio abre el camino para usar más estimadores de FC dentro de los marcos de DL, y los trabajos futuros pueden extender y mejorar algunos de los problemas existentes, por ejemplo, incluyendo estrategias de medición multivariante y no dependiendo solo de medidas por pares, así como implementar un método de tiempo de retraso para reducir el impacto de los problemas de conducción de volumen. Finalmente, medir las matrices de FC con distancias más apropiadas que tengan en cuenta los enlaces internos entre diferentes canales, como en la geometría de Riemann, podría ayudar a decodificar mejor las señales de MI escondidas en las intrincadas grabaciones de EEG.

**Palabras clave:** Conectividad funcional, Aprendizaje profundo, Métodos de kernel, Entropía de Renyi, Ineficiencia de BCI, Densidad espectral cruzada, Teorema de Bochner.

---

<sup>1</sup>BCI2a <http://www.bbci.de/competition/iv/index.html>

<sup>2</sup>Giga Motor Imagery <http://gigadb.org/dataset/100295>

# Abstract

Conditions such as amyotrophic lateral sclerosis, brain strokes, and spinal injuries can cause significant disruptions in the brain's communication pathways, affecting voluntary muscle control and interactions with one's environment. Brain-computer interfaces (BCIs) are computer-based systems utilized to address these challenges; they interpret brain signals and have seen extensive applications in medical technology, rehabilitation, and entertainment industries. Particularly in BCI studies, motor imagery (MI), a widely researched paradigm, holds the potential for reviving motor functionality. Integrating MI-BCIs with skilled therapists has yielded positive outcomes in sensory and motor rehabilitation processes, proving to be beneficial for individuals with neurological disorders. Additionally, MI-BCI systems are not limited to clinical applications; they also find relevance in non-clinical settings like virtual reality, gaming, and skill acquisition. Brain signals in MI-BCI systems are often analyzed using electroencephalography (EEG), which is advantageous due to its high temporal resolution, affordability, and portability. These features make it a viable method for capturing short-evolving events and temporal neuronal activity patterns. Nonetheless, decoding EEG signals is complex due to their high sampling rate and the high number of electrodes involved, which results in an overwhelming amount of data points. Several promising methodologies, leveraging the concept of event-related synchronization (ERS) and event-related desynchronization (ERD), have been developed to ease the extraction of relevant features for decoding EEG into sensorimotor rhythm (SMR) patterns. While noteworthy, single-channel features tend to oversimplify overall phenomena since they need to consider the interactions of different brain regions. Considering the complex activations and communications within multiple brain areas during the execution or imagination of simple motor tasks, multi-channel feature extraction approaches are more accurate. For instance, the Common Spatial Patterns (CSP) method has demonstrated its significance in MI-BCI, promoting the extraction of highly discriminative features by maximizing the separability of SMR features. Brain connectivity (BC), a key determinant in this approach, describes interactions within and between brain regions and correlates to synchronization mechanisms within oscillatory modulations. BC can be divided into structural/anatomical connectivity (SC), effective connectivity (EC), and functional connectivity (FC), each with distinct characteristics. SC focuses on physical connections and is context-independent, while EC and FC, suitable for EEG-based MI-BCI systems, decode fast-evolving cognitive states where FC is particularly suited to MI-BCI applications due to its simplicity, low computational demands, and lack of rigid prior assumptions.

This thesis addresses three prevailing challenges experienced with FC EEG-based MI-BCI systems. Firstly, performance deficits in MI classification tasks are often a result of high noise levels present in EEG signals. Secondly, relying on expert knowledge for feature extraction often presents a barrier. Lastly, there is a necessity for increased transparency within these models. The study delivers key contributions toward a single-trial FC in MI-BCI systems in response to these challenges. This improvement is achieved by utilizing a novel approach grounded on the kernel cross-spectral approximation afforded by extending Bochner's theorem. This new single-trial kernel cross-spectral (KCS) FC estimator tackles noise issues, spurious connections, and inter-subject variability, significantly improving accuracy. Moreover, an end-to-end approach, KCS-FCnet, has been conceptualized to enhance automatic EEG representation for MI-BCI. This approach is pivotal in reducing artifacts and accurately identifying relevant connections. In addition, it assists in the extraction of critical spectral, temporal, and spatial EEG representations. Acknowledging the necessity for model transparency, this study proposed a strategy for improving model transparency and interpretability, named interpretable regularized KCS-FCnet (IRKCS-FCnet). Capitalizing on the potential of regularization,

the model's interpretability is strengthened by minimizing nonrelevant connections. This is further made possible by maximizing the cross-information potential using Renyi's entropy measured at  $\alpha = 2$ , alongside a cross-entropy function. To supplement these enhancements, Layer-CAM and last layer weight strategies have been employed to provide both post-hoc and intrinsic interpretability, respectively. The average drop concept further supports these strategies, enhancing the system's transparency and understanding.

The results show that the proposed KCS-FC achieves higher performance accuracy than classical FC estimators and has competitive accuracy against more complex FC-based strategies, reaching the second place in DBI<sup>3</sup> with an accuracy of 81.92% and first place with 74.12% in DBIII<sup>4</sup>, both binary motor imagery classification tasks. Moreover, the KCS-FCnet shows 76.4% on DBIII while keeping the architecture with 4245 trainable parameters, making it the highest against the compared DL models. Lastly, the IRKCS-FCnet shows better performance, increasing 2 percentage points for low-performing individuals, and showed better localization and interpretations, demonstrating that when removing 5% of the most important features, the average accuracy drop score goes to 22.46, 5 percentage points higher than the version without the regularization. Furthermore, the contralateral pattern can be seen in the topoplot interpretation for both the right and left hand. This study opens the path for using more FC estimators within DL frameworks, and future work can extend and improve some of the existing issues, for instance, including multivariate measuring strategies and not relying only on pairwise measures, as well as implementing a time-lag method to reduce the impact of volume conduction problems. Finally, measuring FC matrices with more appropriate distances that account for the inner links between different channels, like in the Riemann geometry, could help to better decode MI signals hindered in the intricate EEG recordings.

**Keywords:** Functional connectivity, Deep learning, Kernel methods, Renyi's entropy, BCI inefficiency, cross-spectral density, Bochner's theorem.

---

<sup>3</sup>BCI2a <http://www.bbci.de/competition/iv/index.html>

<sup>4</sup>Giga Motor Imager <http://gigadb.org/dataset/100295>

# List of Figures

<b>1-1</b>	Comparison of neuroimaging techniques: spatial and temporal resolutions, and relative costs	2
<b>1-2</b>	Experimental setup for MI-based BCIs. <b>Source :</b> Adapted from [Grigorev <i>et al.</i> , 2021]	3
<b>1-3</b>	Three types of brain connectivities: structural connectivity, effective connectivity, and functional connectivity. Blue circles and blobs represent EEG electrodes and regions of interest, respectively. Circles with different colors represent electrodes with statistical dependencies	5
<b>1-4</b>	Illustration of the three key challenges in FC EEG-based MI-BCI: noise and variability in single trial FC estimators, handcrafted-based subject-specific EEG-based MI-BCI Representation, and the need for model transparency	7
<b>1-5</b>	Single-trial FC-based MI classification inefficiency graphical scheme	8
<b>1-6</b>	Illustration of challenges in EEG signal representation, including strategies for artifact removal and implementing spatial, temporal, and frequency filters	9
<b>1-7</b>	Transparency and interpretability challenges in MI-BCI models graphical scheme	10
<b>1-8</b>	Functional connectivity estimators classified based on direct or indirect, time or frequency domain and linear or nonlinear. The varying shades of orange represent different sensitivities to volume conduction, with lighter shades indicating a higher degree of robustness and solid ones indicating a high level of sensitivity. Abbreviations: Corr, correlation [Fagerholm <i>et al.</i> , 2020]; MSC, magnitude square coherence [Cattai <i>et al.</i> , 2021]; IPC, imaginary part of the coherence [Cao <i>et al.</i> , 2022a]; PC, partial coherence [Gonzalez-Astudillo <i>et al.</i> , 2020]; PLI, phase lag index [Siviero <i>et al.</i> , 2023]; WPLI, weighted phase lag index [Gonzalez-Astudillo <i>et al.</i> , 2020]; PLV, phase locking value [Cattai <i>et al.</i> , 2021]; MI, mutual information [Gu <i>et al.</i> , 2023]; SL, synchronization likelihood [Gonzalez-Astudillo <i>et al.</i> , 2021]; Cross-corr, cross-correlation [Roy <i>et al.</i> , 2022]; GC, Granger causality [Rezaei & Shalbaf, 2023]; TE, transfer entropy [Rezaei & Shalbaf, 2023]; DTF, directed transfer function [Rezaei & Shalbaf, 2023]; PDC, partial directed coherence[Gaxiola-Tirado <i>et al.</i> , 2017]	13
<b>1-9</b>	Four feature extraction approaches for decoding FC in EEG-based MI-BCI: common spatial patterns, FC vectorization, topological representations, and Riemannian geometry	14
<b>1-10</b>	Deep Learning techniques grouped by input formulation. Featuring handcraft-based feature representation and raw EEG signals, this categorization further divides the input according to dimensions such as vector, matrix, and tensor. The weaknesses and strengths of each form are also highlighted	17
<b>1-11</b>	Deep learning model categories sorted into discriminative, representative, and generative groups, showcasing their weaknesses and strengths	18
<b>1-12</b>	Explainability strategies for machine learning and deep learning approaches	19
<b>3-1</b>	Schematic display of the main thesis contributions, including the single-trial FC estimator and feature extraction, an end-to-end approach from EEG representations to MI classification, and quantitative and qualitative strategies for model interpretation	22

<b>4-1</b>	Schema representing a one-day session of EEG data collection for the DBI MI . . . . .	24
<b>4-2</b>	Illustration of the seven-second DBI MI protocol used in the study. The protocol commences with a short acoustic warning and a cross shown on a screen for two seconds, followed by an arrow indicating the type of MI task to be performed. The black screen marks a brief break interval between tasks. . . . .	25
<b>4-3</b>	Illustration of the electrode montage including the left image adhering to the international 10-20 system used for acquiring brain activity and the right image showing the montage of the three monopolar EOG channels employed for subsequent artifact removal. <b>Source:</b> [Brunner <i>et al.</i> , 2008]. . . . .	25
<b>4-4</b>	Schema representing a one-run session of EEG data collection for the DBII ME . . . . .	26
<b>4-5</b>	Positioning of the 44 sensors on the EEG cap, covering the motor cortex used for data collection in the motor execution experiment. Source: [Borra <i>et al.</i> , 2019] . . . . .	26
<b>4-6</b>	Diagram illustrating the experiment protocol involving motor imagery tasks. The figure shows the sequence of prompting, performing, and break periods in a single run. Source: [Cho <i>et al.</i> , 2017] . . . . .	27
<b>4-7</b>	Sensor positions in a 10–10 placement electrode system, containing 64 channels. Besides, it highlights in color the main parts of the brain ( Frontal left, Frontal, Frontal right, Central right, Posterior right, Posterior, Posterior left, Central left) . . . . .	28
<b>5-1</b>	Illustration of the proposed Kernel-based Cross-Spectral FC for enhancing the MI-BCI model's accuracy. Key features include robust FC estimators using a Gaussian kernel for non-linearity, short sliding windows for addressing non-stationarity, sub-band filtering for managing inter-subject variability, vectorization representation for FC extraction and elastic-net regularization to curb spurious connectivities . . . . .	54
<b>5-2</b>	Subject accuracy at each window length $\tau$ performed by the evaluated FC measures: PLV, CCF, and KCS-FC. Subjects are displayed on the horizontal axis in decreasing order of each FC accuracy at $\tau = 2$ s. Notation CSP stands for the accuracy estimated after the spatial CSP filtering. . . . .	59
<b>5-3</b>	Accuracy performed at each $\tau$ by the evaluated measures of FC: PLV ( ), CCF( ), and KCS-FC ( ). The full-filled colored bars stand for concatenate representation, while the bars with horizontal lines present the results after CSP filtering. . . . .	60
<b>5-4</b>	Individual classifier performance at different window lengths using FC: PLV (—), CCF (—), and KCS-FC (—). The accuracy assessments performed without CSP are outlined with a continuous line, after CSP filtering—with a dashed line. The subjects are displayed on the horizontal axis in decreasing order of their CCF-based accuracy. . . . .	61
<b>5-5</b>	Clustering variability of individuals that belong to Group I (cells in yellow), Group II (Turquoise), and Group III (purple), depending on the extracting window length $\tau$ . . . . .	62
<b>5-6</b>	Spatial relevance mapped to topoplots estimated for each group of individuals from the FC measures (regarding the relevance weights $v$ and the classification accuracy). The normalized relevance weights values (between 0 to 1) are computed from Equation (5-6). The background stands for the accumulated spatial relevance mapped to the EEG channel positions. The colored links represent the normalized FC relevance weights that hold a value higher than the 99 percentile of vector $v$ . Above each plot, the average accuracy is shown over the corresponding group. . . . .	63
<b>6-1</b>	Schematic diagram illustrating the proposed Kernel Cross-Spectral FC Network (KCS-FCnet), including the use of DL strategies and a CNN layer for artifact removal and spectral and temporal feature extraction. The diagram further outlines the process of spatial extraction via a custom layer incorporating the kernel cross-spectral function . . . . .	66

<b>6-2</b> Subject specific results. EEGnet and KCS-FCnet average accuracies are depicted, with subjects being sorted based on their performance using EEGnet. The blue bars represent an improvement in performance using the KCS-FCnet, while the red bars indicate a decrease in performance. The background codes the group membership (best—G I, medium—GII, and worst—GIII performance clusters). . . . .	70
<b>6-3</b> KCS-FCnet subject group enhancement regarding the EEGnet performance. Note that green, yellow and red represent best, medium, and worst performance regarding the average accuracy along subjects. First row: Subjects organized based on the EEGnet classification. Second row: Subject membership changes based on the KCS-FCnet results. . . . .	71
<b>6-4</b> Group comparison between EEGnet (blue boxplots) and KCS-FCnet (green boxplots) concerning the accuracy's standard deviation. The background codes the group membership (best—GI, medium—GII, and worst—GIII performance clusters). . . . .	72
<b>6-5</b> <i>t</i> -SNE 2D projection of pruned functional connectivity matrices based on KSC-FCnet and two-sample Kolmogorov–Smirnov test. The color bar depicts the <i>p</i> -value of every connection for each subject matrix, where deep blue means more class separability. Therefore, the bluer the matrix, the better the discriminability. Outer boxes indicate subject group belongingness: green G I, yellow G II, and red G III. <i>p</i> -values below 5% are taken as significant. . . . .	73
<b>6-6</b> Rényi's entropy-based retained information within the estimated functional connectivity matrices using KSC-FCnet. The background codes the group membership (best, medium, and worst performance clusters). Boxplot representation is used to present the retained information within each group. . . . .	73
<b>6-7</b> KSC-FCnet functional connectivity results (Circos plots and topoplots). The first row illustrates the 99th percentile of the most significant functional connections across the centroid subjects concerning the studied best, medium, and worst performance clusters (see Figure 4-7), with the opacity representing the strength of the connectivity. The second and third rows display topoplots of the two-sampled Kolmogorov–Smirnov test and the weights of the classification layer on the KCS-FCnet, respectively. The purer the purple color, the more important the brain area is. The left to right column represents each group from G I to G III. . . . .	75
<b>6-8</b> Method comparison results: number of trainable parameters vs. average motor imagery classification accuracy for the Giga database. . . . .	76
<b>7-1</b> Illustration of the proposed Interpretable Regularized Kernel Cross-Spectral Functional Connectivity Network (IRKCS-FCnet) model, with both qualitative and quantitative strategies for improved MI-BCI model explainability . . . . .	77
<b>7-2</b> Performance comparison across EEGnet (—), KCS-FCnet (—), and RKCS-FCnet (—) models at different time windows. . . . .	81
<b>7-3</b> Comparison of subject-specific average accuracy across EEGnet, KCS-FCnet, and RKCS-FCnet models. Subjects are arranged in descending order based on the accuracy of the EEGnet model. The color-coded bars illustrate performance shifts, with dark green(■) indicating an increase in accuracy on the RKCS-FCnet over KCS-FCnet, light green(■) indicating an increase in accuracy on KCS-FCnet over EEGnet, and while red ■ indicates an accuracy reduction. . . . .	82
<b>7-4</b> Group transitions for subjects across EEGnet, KCS-FCnet, and RKCS-FCnet models. The color representation, green, yellow, and red, correspondingly signifies the best, intermediate, and worst average accuracy performance. First row: Initial group allocation of subjects based on EEGnet model results. Second row: Subject group transitions resulting from implementing the KCS-FCnet model. Third row: Subject group shifts according to the outcomes of the RKCS-FCnet model. . . . .	83
<b>7-5</b> Box and whisker plot comparing Renyi's entropy across groups for KCS-FCnet (blue boxplots) and RKCS-FCnet (green boxplots) models. The background codes the group membership (best, intermediate, and worst performance clusters) . . . . .	84

- 7-6** Visualization of the top ten channels in each group. The channels for Group I (left) are primarily concentrated in the sensorimotor cortex, central sulcus, and motor cortex. In Group II (middle), the majority of the channels are located in the motor-related brain area and parietal lobe, with visual stimuli interpretation inferred from the number of channels in the parietal-occipital area. For Group III (right), the top channels are predominantly located in non-motor-related areas, with a particular influence from visual information in the frontal and occipital lobes and possible high ocular artifacts suggested by the presence of Fp2 channels. . . . . 86
- 7-7** Illustration of the average PSD for each group plotted across the ten most important channels, categorized by importance from top to bottom. Relatively stable spectrum is visible within Group I (top), with predominant spectral power between 10 and **14Hz**. Group II (middle) exhibits a similar pattern, with a slightly left-skewed power distribution, yet still within the range of Mu oscillations. On the contrary, Group III (bottom) displays an increase in spectral power at the pre-frontal Fp2 channel, significantly influenced by  $\alpha$  and  $\theta$  waves, indicating the introduction of substantial artifacts. . . . . 87
- 7-8** Spatial relevance for each group derived from the final layer of the IRKCS-FCnet model. The background represents the accumulated spatial relevance mapped to EEG channel positions. The first row represents the left class, while the second represents the right class. The columns are dedicated to each group: the left column for Group I, the middle column for Group II, and the right column for Group III. The figure visibly presents activation in sensorimotor regions across groups, potential artifact interferences, and alignment with the low accuracy level achieved in G III. . . . . 88
- 7-9** Illustration of the PSD across subjects 4, 42, and 51 from groups I, II, and III, respectively. While Groups I and II display localized and specific frequency band importance for motor tasks, Group III suggests more spread frequencies. Zero cross-label importance across all subjects indicates that the model effectively prevents misclassification but suggests potential overfitting. . . . . 89
- 8-1** Comparison of state-of-the-art end-to-end EEG-based DL architectures (in blue) with the proposed KCS-FCnet and IRKCS-FCnet architectures (in green) in terms of accuracy and number of parameters. . . . . 92

# List of Tables

<b>1-1</b>	Common neuroimaging techniques principle and working mechanism description . . . . .	2
<b>1-2</b>	Brain signals classification by frequency range with description . . . . .	4
<b>4-1</b>	EEGnet arquitecture . . . . .	48
<b>4-2</b>	Shallowconvnet architecture . . . . .	48
<b>4-3</b>	Deepconvnet architecture . . . . .	49
<b>4-4</b>	TCNet-Fusion architecture . . . . .	50
<b>5-1</b>	Classifier accuracy comparison of approaches using functional connectivity features recently reported against the Gaussian FC performance in discriminating MI tasks. Notation TSGSP is temporally constrained sparse group spatial pattern, STR is space-time recurrence, and OPTICAL is Optimized CSP with long short term memory (LSTM). The best value performed for each database is marked in bold. . . . .	63
<b>6-1</b>	Detailed KCS-FCnet architecture for MI classification. . . . .	68
<b>6-2</b>	Group-based accuracy results for EEGnet and KCS-FCnet. The average accuracy for the best, medium, and worst-grouped subjects is depicted. The KCS-FCnet average increase for each cluster is also reported. . . . .	71
<b>6-3</b>	Method Comparison results regarding the average MI classification for the Giga database. . .	74
<b>7-1</b>	Comparison of accuracy performance across groups and models . . . . .	83
<b>7-2</b>	p-values comparing accuracy drop of KCS-FCnet and RKCS-FCnet models vs Random (* indicates p-value < 0.01) . . . . .	84
<b>7-3</b>	Average accuracy drop across feature removal percentages for KCS-FCnet and RKCS-FCnet, with higher difference drops indicated by (■) for GI, (□) for GII, and (■) for GIII. . . . .	85

# Content

<b>Acknowledgements</b>	<b>ii</b>
<b>Symbols and Abbreviations</b>	<b>iii</b>
Nomenclature . . . . .	iii
List of Abbreviations . . . . .	iii
<b>Resumen</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xi</b>
<b>Content</b>	<b>xiv</b>
<b>1 Preliminaries</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Problem Statement . . . . .	7
1.2.1 Single-Trial FC MI Classification Inefficiency . . . . .	7
1.2.2 Handcrafted-based Subject-Specific EEG-based MI-BCI Representation . . . . .	9
1.2.3 Lack of Transparency and Interpretability Strategies in MI-BCI . . . . .	10
1.2.4 Research Question . . . . .	11
1.3 State of the Art . . . . .	12
1.3.1 Single-Trial FC in MI-BCI . . . . .	12
1.3.2 Subject-Specific EEG Representation for MI-BCI . . . . .	16
1.3.3 Interpretability Strategies in MI-BCI . . . . .	19
1.3.4 Summary . . . . .	20
<b>2 Aims</b>	<b>21</b>
2.1 General Objective . . . . .	21
2.2 Specific Objectives . . . . .	21
<b>3 Outline and Contributions</b>	<b>22</b>
3.1 Single-Trial FC for Enhancing MI-BCI Classification . . . . .	22
3.2 Automatic Subject-Specific EEG Representation for MI-BCI . . . . .	23

3.3 Qualitative and Quantitative Post-Hoc and Intrinsic Explainability . . . . .	23
<b>4 Materials and Methods</b>	<b>24</b>
4.1 EEG-based Motor Imagery Datasets . . . . .	24
4.1.1 BCI Competition IV Dataset IIa - DBI MI . . . . .	24
4.1.2 Gamma Motor Execution Database—DBII ME . . . . .	25
4.1.3 MI BCI EEG Giga Science Database - DBIII MI . . . . .	27
4.2 Theoretical Background . . . . .	28
4.2.1 Motor Imagery Classification from EEG Signals . . . . .	28
4.2.2 Stochastic Processes . . . . .	30
4.2.3 EEG Classical Feature Extraction Techniques . . . . .	32
4.2.4 Functional Connectivity Estimators . . . . .	36
4.2.5 Reproducing Kernel Hilbert Spaces for Machine Learning . . . . .	42
4.2.6 Information Theoretic Learning from Kernel Matrices . . . . .	43
4.2.7 EEG Deep Learning Techniques . . . . .	44
4.2.8 Explainable AI in MI-BCI . . . . .	49
4.3 Computing Resources . . . . .	52
<b>5 Kernel-based Functional Connectivity</b>	<b>53</b>
5.1 Kernel-based Covariance Function . . . . .	53
5.2 Gaussian Functional Connectivity from Kernel-based Spectral Distribution . . . . .	54
5.3 Experimental Set-Up . . . . .	55
5.4 Results and Discussion . . . . .	57
5.4.1 Influencing FC Parameters on Accuracy Estimation . . . . .	57
5.4.2 Estimated Classifier Accuracy of Individuals . . . . .	58
5.4.3 Interpretation of Subject Clusters Using Functional Connectivity Patterns . . . . .	58
5.4.4 Selecting the Sliding Time Window . . . . .	62
5.4.5 Prior Spatial Filtering Versus to Feed the Concatenation of FC Feature Sets . . . . .	62
5.5 Summary . . . . .	64
<b>6 Kernel Cross-Spectral Functional Connectivity Network</b>	<b>65</b>
6.1 Kernel Cross-Spectral Functional Connectivity Network . . . . .	65
6.2 Experimental Set-Up . . . . .	67
6.2.1 KCS-FCnet Implementation Details . . . . .	67
6.2.2 Functional Connectivity Pruning and Visualization . . . . .	68
6.2.3 Method Comparison . . . . .	69
6.3 Results and Discussion . . . . .	69
6.3.1 Subject Dependent and Group Analysis Results . . . . .	69
6.3.2 Estimated Functional Connectivity Results . . . . .	72
6.4 Summary . . . . .	76
<b>7 Interpretable Regularized Kernel Cross-Spectral FC Network</b>	<b>77</b>
7.1 Cross-Information Potential Regularization . . . . .	77
7.2 Experimental Set-Up . . . . .	78
7.3 Results and Discussion . . . . .	81

---

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

---

7.3.1	Model Level Analysis . . . . .	81
7.3.2	Group Level Analysis . . . . .	82
7.3.3	Individual Subject Analysis . . . . .	87
7.4	Summary . . . . .	89
<b>8</b>	<b>Final Remarks</b>	<b>91</b>
8.1	Conclusions . . . . .	91
8.2	Future Work . . . . .	92
8.3	Academic Products . . . . .	93
<b>References</b>		<b>94</b>

# 1 Preliminaries

## 1.1 Motivation

Neurological and cognitive conditions such as amyotrophic lateral sclerosis, brain strokes, and spinal injuries can potentially disrupt the communication pathways within the human brain, affecting voluntary muscle control used to communicate with the external environment, such as locked-in syndrome [Sisti *et al.*, 2022, McFarland & Wolpaw, 2011]. To address this, Brain-Computer Interfaces (BCIs) have surged in interest as they are computer-based systems that capture and analyze brain signals, translating them into commands [Khan *et al.*, 2020, Ramadan *et al.*, 2015].

Motor Imagery (MI), an extensively studied paradigm within BCI research, particularly holds promise in the recovery of motor functionality [Bonci *et al.*, 2021]. Clinical integrations of MI-BCIs with skilled therapists have proven beneficial in aiding the sensory or motor rehabilitation process, facilitating better environmental interaction for people with neurological disorders [Sitaram *et al.*, 2017, Hochberg *et al.*, 2012, Lebedev & Nicolelis, 2006]. Beyond clinical applications, MI-BCI systems have seen increased use in non-clinical settings such as virtual reality, gaming, and skill acquisition among healthy subjects [Cattan *et al.*, 2018]. Iconic video games such as Pong and Pacman can be played using MI, and studies on healthy subjects show promise in facilitating skills acquisition across various disciplines [Ahn *et al.*, 2014, Ganin *et al.*, 2013, Johnson *et al.*, 2018, Perdikis *et al.*, 2018].

Nowadays, interest in BCI systems has grown to such a point that before the COVID-19 pandemic, there were more than 150 BCI-based companies distributed in sectors such as automotive, aerospace, medical technology, rehabilitation, robotics, entertainment, marketing, and technology. Besides, the European Commission leads a program called 'The Future of Brain/Neural Computer Interaction: Horizon 2020' to establish a worldwide coordination effort by funding and supporting actions as well as defining a roadmap for future research, taking into account market needs until 2025 [Brunner *et al.*, 2015]. For instance, the BCI technology market, valued at USD 228.1 million in 2022, is expected to witness substantial growth, with an estimated value of USD 460.1 million by 2029, according to a recent market report <sup>1</sup>. From a demographic perspective, the value of BCI technology in global socioeconomic development is acknowledged. However, the method for studying this technology varies depending on the specific context of each country, with the USA and China leading in the number of scientific publications, followed by Germany, Japan, and India [Maiseli *et al.*, 2023]. In this sense, BCI technology development must adhere to the fundamental ideals of humanity, including convenience, ease-of-use, privacy, security, and safety [Willett *et al.*, 2021, Matemba *et al.*, 2020].

Neuroimaging techniques are crucial for acquiring brain signals in MI-BCI systems, each playing distinct roles in understanding brain functionality. For instance, Electroencephalography (EEG) and Magnetoencephalography (MEG) capture electrical activity and magnetic fields in the brain, respectively [Altaheri *et al.*, 2023, Philip *et al.*, 2022]. Functional Magnetic Resonance Imaging (fMRI) detects changes

---

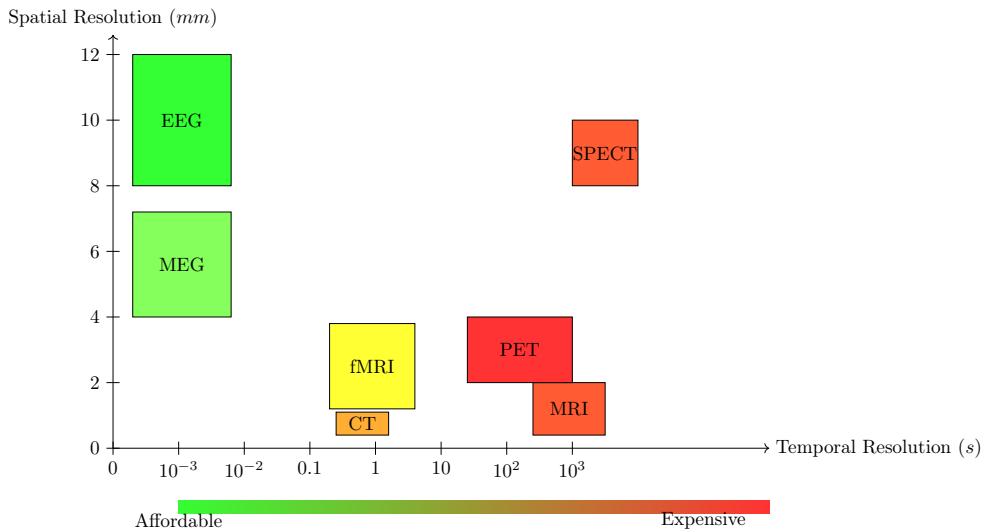
<sup>1</sup><https://www.marketreportsworld.com/enquiry/request-sample/23789843>

in blood flow, indicating brain activity [Sorger & Goebel, 2020]. Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) provide detailed structural images of the brain, useful in pathology identification [Singh et al., 2021]. Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT) map functional processes in the brain using radioactive substances, highlighting areas with high metabolic activity [Veena & Anitha, 2020]. These neuroimaging techniques are summarized in Table 1-1.

**Table 1-1:** Common neuroimaging techniques principle and working mechanism description

Technique	Principle	Working Mechanism
EEG	Records electrical activity	Electrodes on scalp capture voltage fluctuations
MEG	Measures magnetic fields	Superconducting sensors detect magnetic fields
fMRI	Measures blood oxygenation	Detects blood flow changes associated with neural activity
CT	Uses X-rays	X-rays create cross-sectional images
MRI	Uses strong magnetic fields	Protons emit signals for detailed structural images
PET	Measures emitted gamma rays	Radioactive tracers indicate areas with high metabolic activity
SPECT	Measures emitted gamma rays	Tracers indicate areas with high blood flow

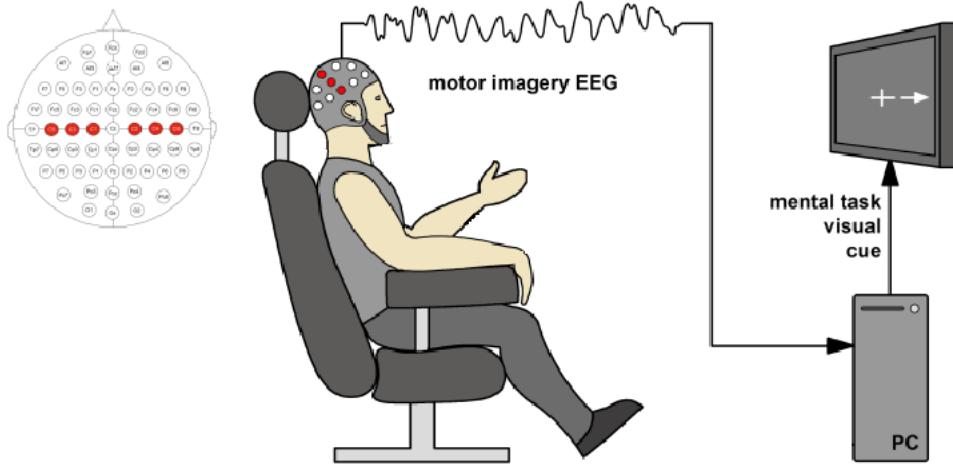
Figure 1-1 shows the spatial and temporal resolutions and relative cost comparison among the previously mentioned techniques. The color gradient in the figure represents the relative cost of each neuroimaging technique, with green indicating more affordable options and red signifying higher expenses. EEG and MEG are notable for their high temporal resolution, capturing rapid changes in brain activity with millisecond precision [Shao et al., 2020]. In contrast, fMRI provides excellent spatial resolution, offering detailed insights into brain structures and activity [Yuan et al., 2021]. However, fMRI has a slower temporal resolution than EEG and MEG, making it less suitable for capturing rapid neural dynamics [Li et al., 2021a]. CT scans, which use X-rays to create cross-sectional images, and MRIs, which use powerful magnetic fields and radio waves, thus avoiding ionizing radiation, both provide detailed anatomical information about the brain's morphology [Zhuang et al., 2020]. PET and SPECT focus on functional imaging by detecting radioactive tracers and can reveal metabolic and neurotransmitter activity, but their temporal resolution is relatively lower than EEG and MEG [Ramadan & Vasilakos, 2017]. While spatial and temporal resolutions are critical factors in selecting the appropriate neuroimaging method, the cost aspect is equally important for decision-making in various research and clinical contexts [Rashid et al., 2020]."



**Figure 1-1:** Comparison of neuroimaging techniques: spatial and temporal resolutions, and relative costs

In particular, MI involves fast-evolving cognitive processes that require superior temporal resolution [Värbu et al., 2022]. Therefore, EEG and MEG offer valuable options for capturing short-living events

and temporal patterns of neural activity, allowing researchers to observe and analyze brain activity with remarkable temporal precision [Alsharif *et al.*, 2020]. In addition, due to portability and cost-effectiveness, EEG is the neuroimaging technique most extensively studied by researchers and budget-constrained institutions focusing on MI-BCI applications [Janapati *et al.*, 2023, Hosseini *et al.*, 2020b, Shao *et al.*, 2020, Khan *et al.*, 2020, Lopez *et al.*, 2020, Ramadan & Vasilakos, 2017, Fogelson *et al.*, 2013]. A typical EEG-based MI-BCI system uses a processor unit and an EEG headset equipped with  $N_c$  electrodes (channels), varying from 1 to 256 [Grigorev *et al.*, 2021, Rashid *et al.*, 2020]. These electrodes, attached with an elastic cap, make direct contact with the scalp, recording electrical variations that encode brain activities [Värbu *et al.*, 2022]. In addition, some setups may include a monitor for instructing the individual to execute a specific MI task [Gu *et al.*, 2021b]. A visual representation of the setup can be found in Figure 1-2.



**Figure 1-2:** Experimental setup for MI-based BCIs. **Source :** Adapted from [Grigorev *et al.*, 2021]

Electrical variations recorded by EEG channels are the combination of multiple rhythms that can be used to describe spatially and temporally dynamic patterns of cognitive states within an individual [Li *et al.*, 2019a]. The specific neural rhythms that occur in the brain's sensorimotor cortex, an area of the brain involved in planning, controlling, and executing motor functions [Leeuwis *et al.*, 2021], are known as Sensorimotor Rhythms (SMRs) [Altaheri *et al.*, 2023]. SMRs depend upon individuals and manifest as frequency, shape, and amplitude changes in electrical activity [Barrios *et al.*, 2019]. In general, most of the cognitive processes implicate six frequency bands: delta [0.5-4]Hz, theta [4-8]Hz, alpha [8-13]Hz, beta [13-30]Hz, lower gamma [30-80]Hz, and upper gamma [80-150]Hz [Khademi *et al.*, 2023, Rashid *et al.*, 2020]. Specifically, motor execution and motor imagery actions implicate a frequency range between approximately [4-40] Hz [Cattai *et al.*, 2021]. Table 1-2 describe the most common band types of brain signals according to their frequency ranges.

One of the most important phenomena associated with SMRs is the increase and decrease of oscillatory activity in a particular frequency band during MI tasks, referred to as Event-Related Synchronization (ERS) and Event-Related Desynchronization (ERD) [Belwafi *et al.*, 2020]. ERS and ERD serve as a unique feature of the specific motor action an individual is imagining or executing [Zapała *et al.*, 2020]. This is why several methodologies leverage this critical concept for decoding EEG signals into SMR patterns, aiding in discrimination between MI tasks in EEG-based BCI systems [Brusini *et al.*, 2021].

In general, decoding EEG signals is intricate due to the high sampling rate and number of electrodes, leading to a huge number of data points [Singh *et al.*, 2021]. Therefore, to achieve decent discriminatory capabilities, it is mandatory to perform feature extraction strategies, ensuring a small number of relevant features [Ai *et al.*, 2019]. Feature extraction for MI-BCI systems can be classified into single- or multi-channel approaches.

Single channel perspective relies on the concept of SMRs, and thus, its ultimate goal is to capture oscillatory

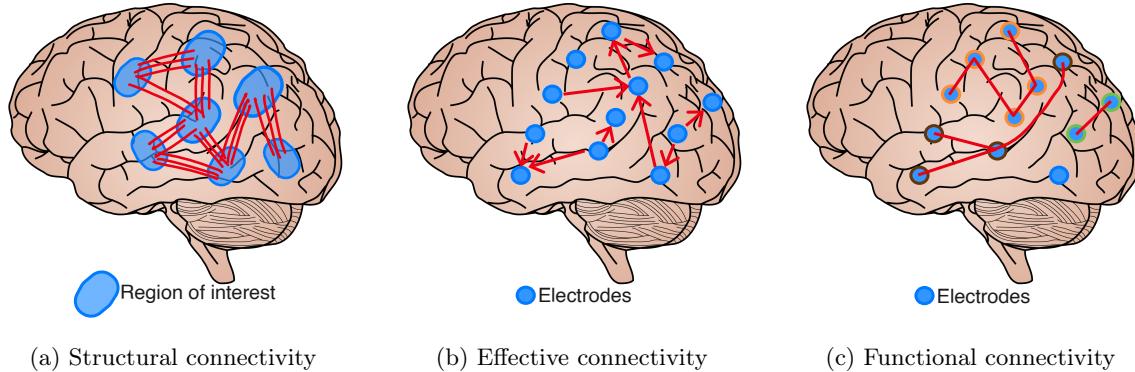
**Table 1-2:** Brain signals classification by frequency range with description

Brain signal	Frequency range	Description
Delta ( $\delta$ )	0.5 to 4 Hz	Slowest brain wave with the highest amplitude, dominant during deep sleep
Theta ( $\theta$ )	4 to 8 Hz	Dominant during deep relaxation and dreaming in light sleep. While normal in children, can be abnormal in awake adults
Alpha ( $\alpha$ )	8 to 13 Hz	Dominant in wakeful but relaxed states with closed eyes. <b>Seen in all ages and can indicate white matter health.</b>
Beta ( $\beta$ )	13 to 30 Hz	Dominant when alert, concentrated, attentive, and calculating. Typically associated with problem-solving, task engagement, and decision-making
Lower Gamma ( $\gamma_l$ )	30 to 80 Hz	Active during high-level cognitive tasks
Upper Gamma ( $\gamma_u$ )	80 to 150 Hz	Associated with perception, learning, and language processing

modulations on specific EEG channels. Time domain strategies collect different temporal information related to a specific MI task, constructing the feature set for a single trial. For instance, statistical features like mean, root mean square, power of the signal, standard deviation variance, skewness, and kurtosis were widely used for MI classification [Samuel *et al.*, 2017, Hamedi *et al.*, 2014]. Other strategies include Hjorth features like activity, mobility, and complexity that indicate the signal power, mean frequency, and change in frequency, respectively [Yilmaz *et al.*, 2018]. Simpler strategies like Peak-valley representations, which extract local maximum and local minimum points within a defined timestamp, can also be used to predict MI tasks [Batres-Mendoza *et al.*, 2016]. On the frequency domain side, strategies try to condensate cognitive states within specific frequency bands like in band-power features, that calculates the spectral power in delta, theta, alpha, and beta bands [Luo *et al.*, 2020]; fast Fourier transform, which decomposes EEG signals into its frequency components [Rashkov *et al.*, 2019]; Power Spectral Density (PSD), measuring how the signal power is distributed over frequency [Oikonomou *et al.*, 2017]. Welch's periodogram, Lomb-Scargle periodogram, and spectral entropy are methods for extracting or quantifying the PSD of a signal [Roy *et al.*, 2022, Sarraf, 2017, Li *et al.*, 2015].

Even though single-channel features can provide useful insights for distinguishing between trials from different MI tasks, it has been shown that oscillatory modulations are not necessarily confined to the sensorimotor cortex [Singh *et al.*, 2021]. Furthermore, the execution or imagination of simple motor tasks activates multiple brain areas that communicate with each other, demonstrating functional integration [Ladda *et al.*, 2021], as opposed to isolated activation of specialized brain regions, known as functional segregation [Chiariion *et al.*, 2023]. In brief, by focusing only on single channels, these features may oversimplify the overall phenomenon as they overlook interactions with other regions. Therefore, it is advised to leverage multi-channel feature extraction approaches to provide a more accurate decoding of MI tasks from EEG [Leeuwis *et al.*, 2021]. Among these approaches, the Common Spatial Patterns (CSP) method has been significant in MI-BCI [Yang *et al.*, 2021a]. CSP allows the extraction of highly discriminative features by maximizing the separability of SMR features, identifying spatial filters that increase the variance of relevant patterns within one MI process while diminishing the variance of unrelated ones [Gaur *et al.*, 2021b].

In addition to CSP strategies, Brain Connectivities (BC), which have been used to describe interactions within and between brain regions [Ismail & Karwowski, 2020], can be linked to SMR synchronization mechanisms within EEG recordings [Maksimenko *et al.*, 2017, Chiariion *et al.*, 2023]. Key factors of BC include going beyond isolated EEG signals and giving a comprehensive graph visualization that enables researchers to understand brain dynamics between different cognitive states [Collazos-Huertas *et al.*, 2023, Graña & Morais-Quilez, 2023, Tafreshi *et al.*, 2019, Van Mierlo *et al.*, 2014, Sakkalis, 2011]. These factors assist in diagnosing conditions, monitoring progress, and tailoring treatment strategies by tracking how patterns evolve during recovery, therapy, or disease progression [Yen *et al.*, 2023, Lim *et al.*, 2021].



**Figure 1-3:** Three types of brain connectivities: structural connectivity, effective connectivity, and functional connectivity. Blue circles and blobs represent EEG electrodes and regions of interest, respectively. Circles with different colors represent electrodes with statistical dependencies

In general, BC estimates non-directed or directed connections between brain regions and can be divided into three types as depicted in Figure 1-3: Structural/anatomical Connectivity (SC), Effective Connectivity (EC), and Functional Connectivity (FC) [Leeuwis *et al.*, 2021, Friston, 2002]. SC, as shown in Figure 1-3a focuses on physical connections between different anatomical brain regions, cannot be directly estimated using merely EEG signals and fails to capture short-living events during MI tasks because it is context-independent [Thiebaut de Schotten *et al.*, 2020, Yeh *et al.*, 2021]. EC and FC, commonly estimated directly from EEG measures, are suitable for EEG-based MI-BCI systems because they can decode fast-evolving cognitive states [Cao *et al.*, 2022a]. According to [Friston *et al.*, 2013], EC is always directed and aims to determine how one EEG signal influences another by using parametric models of causal influences as shown in Figure 1-3b, while FC can describe directed or non-directed connectives usually via statistical correlation or covariance between EEG signals as shown in Figure 1-3c. On one hand, EC requires a deep understanding of the cognitive process to select the most suitable causal model and also demands high computational resources [Chiarion *et al.*, 2023, Lee *et al.*, 2020]. On the other hand, the simplicity of FC, its low computational demands, and the absence of a need for rigid prior assumptions make it particularly well-suited for MI-BCI applications [He *et al.*, 2019, Hamed *et al.*, 2016, Sakkalis, 2011, Friston, 2011].

The Digital Signal Processing and Control Group (DSP&CG) of Universidad Nacional de Colombia has been working on bio-signal data analysis to propose and develop machine learning methodologies related to automatic systems for the assisted diagnosis of mental conditions [Cárdenas-Peña *et al.*, 2017], automated analysis of human activity recognition [Pulgarin-Giraldo *et al.*, 2017], video analysis based on Multi-Kernel representation [Molina-Giraldo *et al.*, 2015], and biomedical data analysis [Hurtado-Rincón *et al.*, 2016]. More recently, DSP&CG works have been focused on extracting relevant patterns from EEG to recognize learning tasks, mainly under the MI paradigm. The used methodologies address dynamic modeling for estimating the spatial relevance of everyday neural activity across subjects [Velasquez-Martinez *et al.*, 2020c], estimation of ERD/S using information theory learning measures [Velasquez-Martinez *et al.*, 2020a], enhancement of feature representation based on kernel functional connectivity [García-Murillo *et al.*, 2021], improving both MI classification and interpretation using convolutional neural networks [Collazos-Huertas *et al.*, 2020]. Lastly, BCI performance predictors from deep learning architectures [Caicedo-Acosta *et al.*, 2021]. Additionally, within the group a variety of research projects are developed, supported by COLCIENCIAS/MINCIENCIAS, Dirección Nacional de Investigaciones de Manizales (DIMA), and Vicerrectoría de Investigaciones de la Universidad Nacional de Colombia, such as:

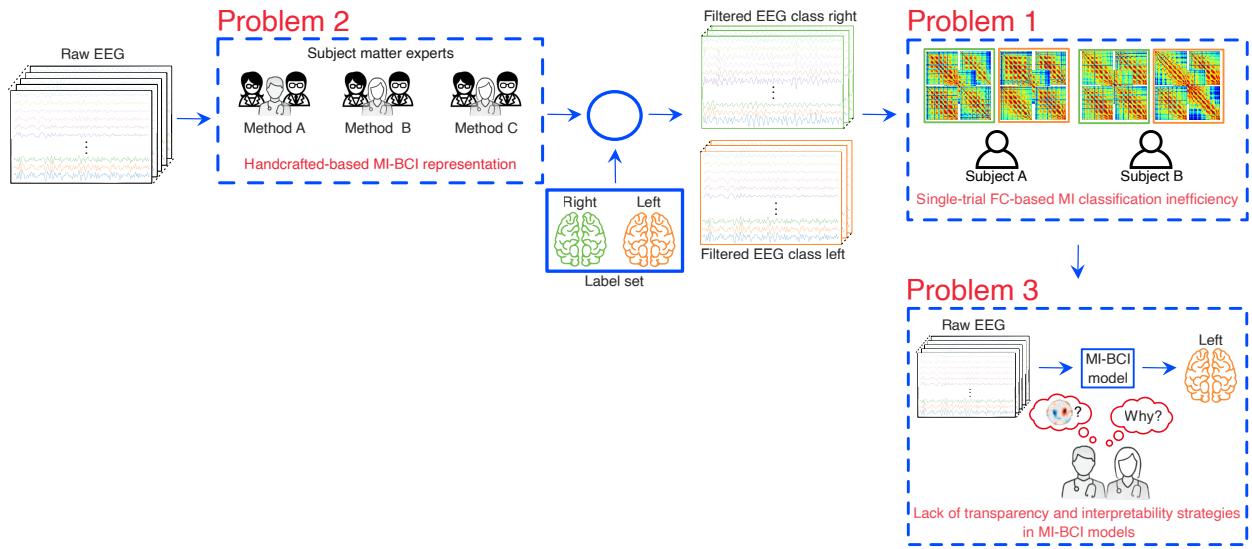
- Prototipo de interfaz cerebro-computador de bajo costo para la detección de patrones relevantes de actividad eléctrica cerebral relacionados con TDAH.

- Prototipo de interfaz cerebro-computador multimodal para la detección de patrones relevantes relacionados con trastornos de impulsividad.

From local and general perspectives, integrating MI-BCIs into automatic and semi-automatic rehabilitation systems can provide crucial support for individuals with neurological disorders, aiding in diagnosing conditions, monitoring progress, and tailoring treatment plans. In addition, MI-BCI systems have expanded beyond clinical settings to various applications for healthy subjects, including virtual reality environments, video gaming, and skill acquisition paradigms. EEG's high temporal resolution and affordability make it the most used in MI-BCI clinical and research investigations. Among different feature extraction approaches, multi-channel techniques offer better insights into the actual cognitive state. Specifically, brain connectivity is extensively used in EEG-based MI-BCI systems because it can distinguish between MI tasks; particularly, functional connectivity presents the most suitable strategy for such applications due to their interpretability, low computational demands, and the absence of a need for rigid prior assumptions. The most common issues associated with functional connectivity in EEG-based MI-BCI systems are described in detail in Section 1.2.

## 1.2 Problem Statement

As shown in Figure 1-4 this thesis will focus its attention on three critical challenges for FC EEG-based MI-BCI. First, despite the valuable insights derived from standard functional connectivity, it faces a performance deficit in MI classification tasks due to the inherently high noise in the EEG signals, issues of volume conduction, and notable inter-subject variability. Second, the heavy reliance on subject matter experts in FC-based feature extraction affects the system's efficiency. This dependence is amplified in scenarios where the same MI cognitive task generates various neural responses in different trials from the same subject, leading to poor and inconsistent predictions. Third, the importance of model transparency to achieve trust among users. With most of the complex models functioning as 'black boxes', predictions are made without clear justifications that can explain why certain decisions are chosen over others, which is crucial for understanding the cognitive processes occurring underneath.



**Figure 1-4:** Illustration of the three key challenges in FC EEG-based MI-BCI: noise and variability in single trial FC estimators, handcrafted-based subject-specific EEG-based MI-BCI Representation, and the need for model transparency

### 1.2.1 Single-Trial FC MI Classification Inefficiency

While standard FC, extracted from various EEG recordings of the same cognitive process (multiple trials), provides invaluable insights for interpreting brain activity patterns, it often struggles to achieve high accuracy in MI classification tasks because single-trial FC estimation only gets a fraction of the data (one trial) [Chiarion *et al.*, 2023, Rodrigues *et al.*, 2020, Billinger *et al.*, 2013]. Since FC quantifies the statistical dependencies between EEG channels, its accuracy estimation is heavily influenced by challenges such as a low signal-to-noise ratio, volume conduction problems, and inter-subject variability [Abiri *et al.*, 2019, Bastos & Schoffelen, 2016]. Therefore, single-trial estimation is an intricate task that usually yields MI classification inefficiency systems [Yang *et al.*, 2021c].

The inherent noise in signals recorded on the scalp significantly hinders the underlying neural activity, challenging the single-trial estimation from EEG data [Si *et al.*, 2020, Hata *et al.*, 2016]. Specifically, external or environmental artifacts such as electromagnetic interferences by lighting, AC power lines, and electronic devices; and noise arising from body activities like other cognitive processes, skin resistance fluctuations, electrocardiographic activity (heart functioning), electrooculographic activity (eye movements),

electromyographic activity (any muscle movements), and respiration diminish the fidelity of EEG signals [Nentwich *et al.*, 2020, Somers *et al.*, 2018]. In addition, functional activities recorded using EEG show volume conduction problems even before they can be measured on the scalp's surface, causing several EEG channels to be correlated [Qiao *et al.*, 2023, Varone *et al.*, 2021]. As a result, many spurious connectivities are induced due to the low signal-to-noise ratio present in EEG signals, impeding competitive performance [Bakhshayesh *et al.*, 2019].

Aside from these issues, individual brain activity's intrinsic diversity and complexity also lead to inter-subject variability [Wriessnegger *et al.*, 2020]. The inter-subject variability is not considered a noise per se but an expression of the subjective motor process [Saha & Baumert, 2020]. For instance, single trial FC-based MI-BCI performance can be negatively affected by the fact that the motor learning process varies over subjects due to intrinsic human kinematic parameters, motivational components, and factors ranging from demographic variables like gender and age to psychological and physiological conditions such as fatigue, relaxation, concentration, and living habits [Antonakakis *et al.*, 2020, Huang *et al.*, 2023]. This is especially evident when different individuals imagine performing the same motor task, leading to varied SMR synchronization patterns that affect FC features [Wriessnegger *et al.*, 2020, Xie & Oniga, 2020].

Most of the existing single-trial FC estimation methods are based on covariance matrices calculated from the original EEG signal domain and do not account for brain activity's inherent nonlinearity and non-stationarity, inevitably compromising their accuracy [Cao *et al.*, 2022a, EPMoghaddam *et al.*, 2022, Miladinović *et al.*, 2021]. Moreover, the high variance in brain activity during task execution adds further difficulties in identifying relevant feature patterns. While improvements in covariance matrix estimation might be achieved through mapping and regularization techniques, selecting an appropriate distance metric to encapsulate variability remains challenging [Wang *et al.*, 2020b, Congedo *et al.*, 2017b].

In summary, single-trial FC-based MI-BCI systems must enhance the model's accuracy without compromising its interpretability using innovative strategies to effectively manage the complex issues of nonlinearity, non-stationarity, and inter-subject variability. Figure 1-5 shows a representation of these challenges.

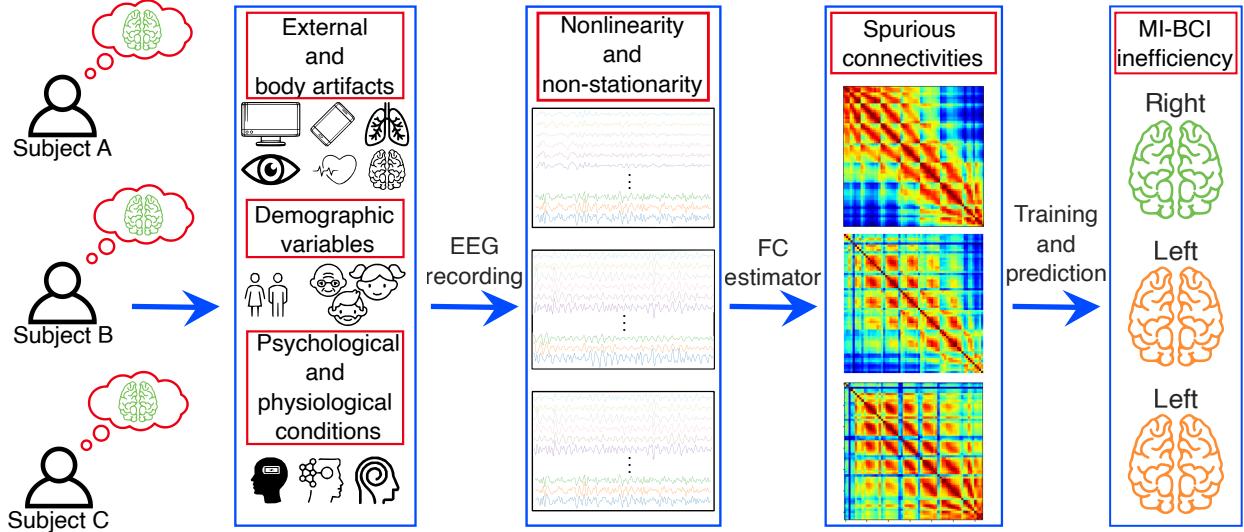


Figure 1-5: Single-trial FC-based MI classification inefficiency graphical scheme

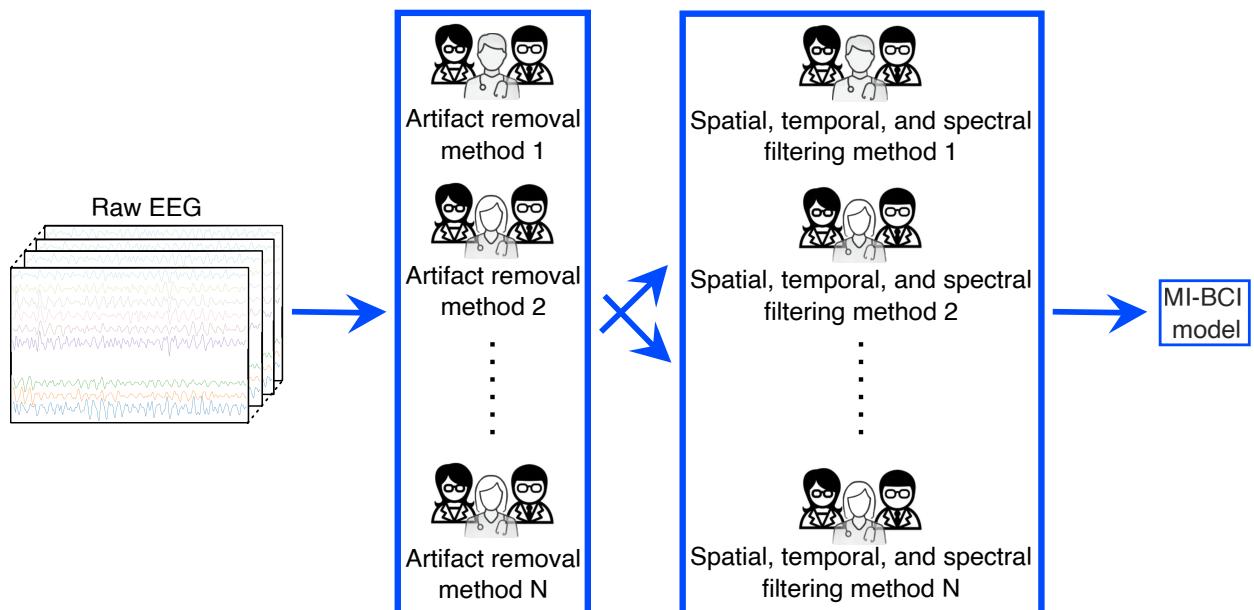
### 1.2.2 Handcrafted-based Subject-Specific EEG-based MI-BCI Representation

Despite the substantial progress made in the EEG signals representation in the MI-BCI field, functional connectivity analysis suffers from a high rate of spurious connections due to noise, affecting mainly the accuracy of low-performing individuals [Ismail & Karwowski, 2020, Wang et al., 2020b, Congedo et al., 2017b]. This noise within EEG signals significantly contributes to high variability, hindering the underlying neural activity and potentially leading to inaccurate interpretation [Caicedo-Acosta et al., 2021]. Therefore, EEG signal representation still presents a significant challenge [Rashid et al., 2020].

These challenges arise predominantly from the complexity, degree of sensitivity, and necessity of human experts in EEG representation, such as artifact removal strategies and spatial, temporal, and frequency filtering [Chamola et al., 2020, Padfield et al., 2019] as graphically illustrated in Figure 1-6. For instance, the handcrafted nature of EEG representation leads to high information loss and limits the capacity to detect discriminative correlations among EEG channels, particularly in low-performing individuals [Altaheri et al., 2023].

While multiple strategies exist for addressing EEG representation issues that can be beneficial in facing MI classification inefficiency, they are often carried out by exhaustive search, limiting the range of artifact removal strategies and spatial, temporal, and spectral filtering selection that can be accomplished [Saha & Baumert, 2020, Ismail & Karwowski, 2020, Meers et al., 2020, Talukdar et al., 2019]. In addition, optimizing these steps proves particularly challenging when handling low-performing subjects, even with the insights from subject matter experts [Vidaurre et al., 2020].

Despite the considerable advancements in the EEG signal representation within the MI-BCI field, the necessity for human expert intervention in the analysis still poses a significant challenge. In particular, formulating strategies for artifact removal and implementing spatial, temporal, and frequency filters can be exhaustively demanding, especially when dealing with low-performing subjects. The complexity and sensitivity of these strategies often result in a loss of crucial information and limitations in accurately identifying genuine connections between signals, even when guided by subject matter expert knowledge.



**Figure 1-6:** Illustration of challenges in EEG signal representation, including strategies for artifact removal and implementing spatial, temporal, and frequency filters

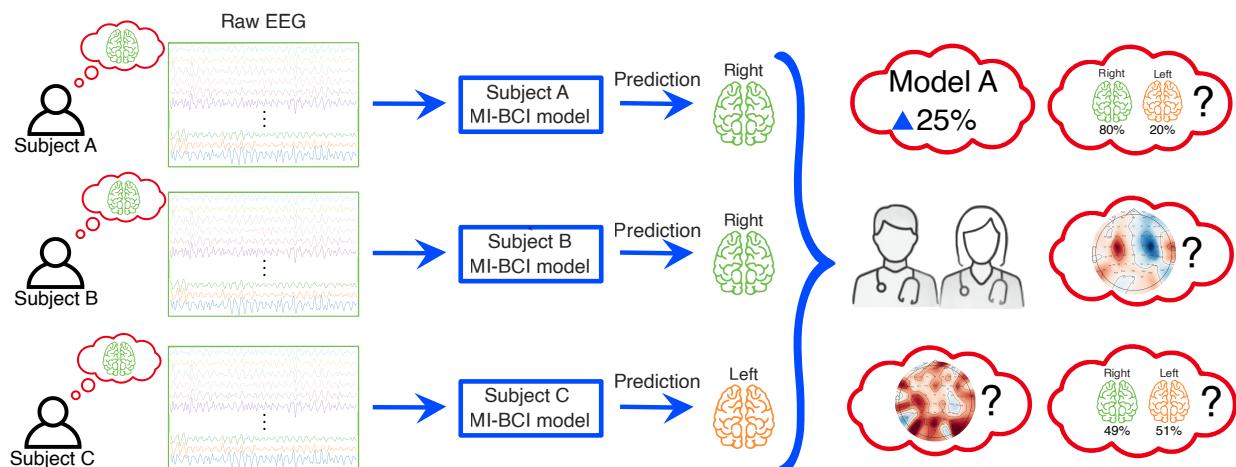
### 1.2.3 Lack of Transparency and Interpretability Strategies in MI-BCI

The third significant challenge for FC-based MI-BCI systems is model transparency and interpretability, linking back to the core aim of Explainable AI (XAI), which seeks to provide meaningful insights into the reasoning behind model predictions, increasing the confiability, causality, transferability, accuracy, accessibility, and interaction [Smuha, 2019]. Integrating various features from different domains into MI-BCI systems poses a unique level of complexity [Guillot & Debarnot, 2019]. For instance, complex models often act like ‘black boxes’, offering outcomes without explanations or justifications [Ras et al., 2018], as seen in Figure 1-7.

Understanding why models favor specific predictions over others is crucial, not only for predictive efficiency but also in healthcare and medical decision-making scenarios [Miotto et al., 2018]. The interpretation process should highlight the prioritized features during model training and illustrate how these features’ interrelationships impact such decisions [Zeiler & Fergus, 2014, Chakraborty et al., 2017]. As models become more refined, the need for transparency and interpretability amplifies, especially in clinical research where reliance on model recommendations is critical [Xiao et al., 2018].

Moreover, the ongoing debate over the definition of interpretability leads to diverse interpretation methods [Fan et al., 2021]. Specifically for FC EEG-based MI-BCI, the challenge lies in creating a quantitative and qualitative interpretability framework that integrates the distinct contributions of each connection to classification across different MI tasks [Zuk et al., 2020]. Identifying and excluding irrelevant features, especially those generated from noisy functional connectivity patterns, is crucial for model refinement, enhancing its generality, reducing overfitting risks, and expanding the use of MI-BCI in clinical applications like diagnosis, monitoring, and computer-assisted learning [Qian et al., 2018].

In brief, one of the significant challenges in XAI within MI-BCI systems is ensuring model transparency and interoperability. However, introducing several features from different domains into these systems can be complex. This complexity is exacerbated by models that function like ‘black boxes,’ providing results without explanations. Understanding why specific predictions are favored and how interrelated features influence these decisions is crucial, especially in critical fields like healthcare. In the MI-BCI context, the task lies in forming a quantitative and qualitative framework that can map the individual contributions of each connection within different domains across varying MI tasks for optimizing the model, reducing overfitting risks, and expanding the clinical applications of MI-BCI.



**Figure 1-7:** Transparency and interpretability challenges in MI-BCI models graphical scheme

### 1.2.4 Research Question

Given the challenges identified in EEG-based MI-BCI systems, including single-trial FC inefficiency, complexities in EEG representation, and issues of transparency and interpretability in complex models, we propose the following 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?

## 1.3 State of the Art

This section presents an overview of the state-of-the-art in the realm of single-trial functional connectivity within EEG-based motor imagery brain-computer interface systems. Our focus is on the cutting-edge approaches designed to tackle the problems described in Section 1.2, with a subsequent discussion addressing their respective strengths and drawbacks. Initially, we explore single-trial FC feature extraction and estimation. Following this, we discuss end-to-end models that merely depend on the raw EEG data. Lastly, we provide a comprehensive outline of the different interpretability strategies, focusing on deep learning. This overview will allow the reader to comprehend new strategies to clarify the underlying complexities of single-trial FC and its critical role in EEG-based MI-BCI systems.

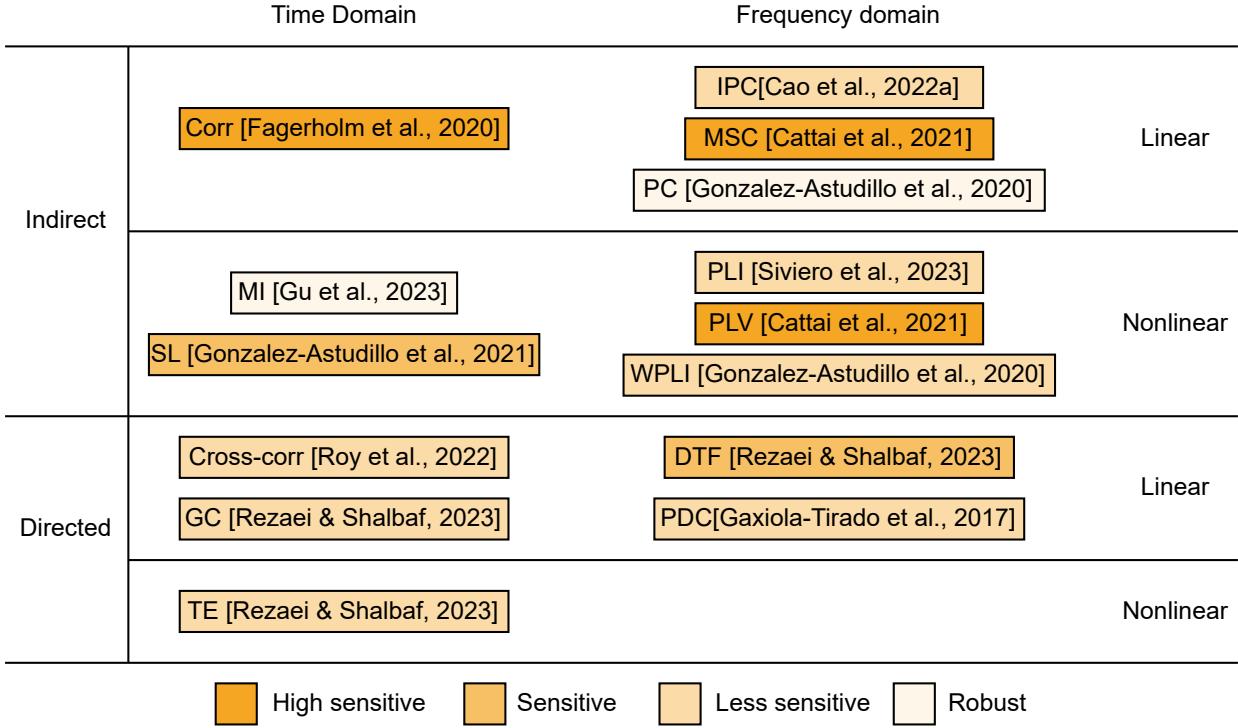
### 1.3.1 Single-Trial FC in MI-BCI

In the context of MI-BCI systems, using FC estimators becomes crucial in decoding the intricate links between different brain parts during MI tasks. These linear and non-linear estimators attempt to measure the statistical dependencies between neurophysiological recordings from different regions, providing crucial insights into brain mechanisms. In addition, FC-based feature extraction techniques play a significant role in effectively translating the nuanced connectivity information into distinguishable patterns related to different MI tasks. When properly extracted and processed, these features enable improved classification and prediction in MI-BCI systems. Over the years, numerous studies have delved into the exploration and enhancement of FC in MI-BCI, leading to significant advancements and understanding. Yet, the field presents challenges and opportunities, particularly considering the complexities and variabilities inherent to brain function and MI processes.

#### Functional Connectivity Estimators

This section offers a review of the most commonly used functional connectivity estimators in the field of EEG-based BCI systems, highlighting their utility and relevance. Given the range of available brain connectivity estimators, selecting an appropriate one is a critical initial step when working with single-trial FC in EEG-based MI-BCI systems. These estimators can generally be broadly categorized based on several factors. First, the domain of interest can be either time or frequency. Second, whether the connectivity is direct or indirect, the former referring to simple zero lag correlation and the latter about the k-lag or cross-correlation. Third, whether the interactions are grounded in a linear or nonlinear formulation. A detailed representation of these categorizations can be found in Figure 1-8.

Linear estimators, due to their straightforward computation and interpretability, have been widely investigated for predicting brain connectivity in MI-BCI systems [Gonzalez-Astudillo *et al.*, 2020, Van Diessen *et al.*, 2015]. The simplest method of estimating linear intercommunication between EEG signals is based on the widely recognized Pearson correlation (Corr) for indirect connectivity and cross-correlation (Cross-corr) for direct connectivity [Fagerholm *et al.*, 2020]. It's important to note that Cross-corr is essentially an expanded version of correlation (zero time lag) that considers a k-time lag. As a result, the correlation may not accurately represent linear relationships if there is a delay between specific patterns from two EEG channels [Roy *et al.*, 2022]. A frequency domain counterpart of correlation is coherence, which is based on the amplitude synchrony of the Power Spectrum Density (PSD) between two EEG signals [Cattai *et al.*, 2021]. Specifically, Magnitude Squared Coherence (MSC), and the Imaginary Part of Coherency (IPC) are two coherence-based FC estimators that are widely used, with the latter being less sensitive to the volume conduction problem [Cao *et al.*, 2022a]. While most of the FC estimators rely on the interactions between two signals without removing the influence of other signals, Partial Coherence (PC) and Partial Directed



**Figure 1-8:** Functional connectivity estimators classified based on direct or indirect, time or frequency domain and linear or nonlinear. The varying shades of orange represent different sensitivities to volume conduction, with lighter shades indicating a higher degree of robustness and solid ones indicating a high level of sensitivity. Abbreviations: Corr, correlation [Fagerholm et al., 2020]; MSC, magnitude square coherence [Cattai et al., 2021]; IPC, imaginary part of the coherence [Cao et al., 2022a]; PC, partial coherence [Gonzalez-Astudillo et al., 2020]; PLI, phase lag index [Siviero et al., 2023]; WPLI, weighted phase lag index [Gonzalez-Astudillo et al., 2020]; PLV, phase locking value [Cattai et al., 2021]; MI, mutual information [Gu et al., 2023]; SL, synchronization likelihood [Gonzalez-Astudillo et al., 2021]; Cross-corr, cross-correlation [Roy et al., 2022]; GC, Granger causality [Rezaei & Shalbaf, 2023]; TE, transfer entropy [Rezaei & Shalbaf, 2023]; DTF, directed transfer function [Rezaei & Shalbaf, 2023]; PDC, partial directed coherence[Gaxiola-Tirado et al., 2017]

Coherence (PDC) are designed to alleviate spurious connectivities due to common influence from other signals [Gonzalez-Astudillo et al., 2020]. Additionally, Directed Transfer Function (DTF) in the frequency domain and Granger Causality (GC) in the time domain employ a multivariate autoregressive model to express one signal as a linear combination of past samples from all the channels [Rezaei & Shalbaf, 2023]. Nonetheless, these methods are rooted in linear correlations and do not account for the non-linearity inherently present in EEG signals [Cao et al., 2022b, Mirzaei & Ghasemi, 2021].

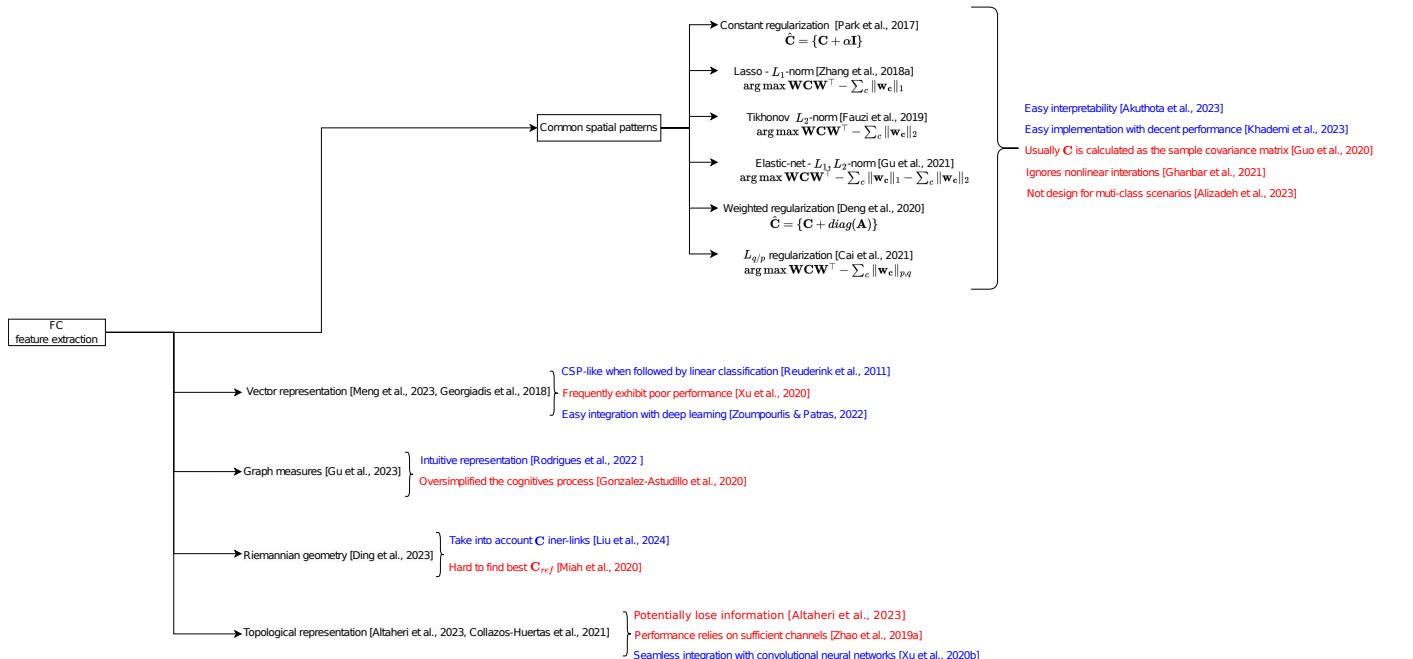
To better understand the nonlinear interactions between EEG channels phase synchronization, usually less affected by volume conduction problems and having direct connection with time delays, becomes a relevant concept [Bastos & Schoffelen, 2016]. Techniques such as the Phase Locking Value (PLV) and Phase Lag Index (PLI) often rely on the phase coupling of oscillatory patterns and are frequently used to measure phase synchronization strength [Siviero et al., 2023]. The Weighted Phase Lag Index (WPLI), created to remove contributions from zero lag, helps to limit misleading connections, but it may miss instantaneous signal interactions [Gonzalez-Astudillo et al., 2020]. Information theory techniques have also been effective in revealing nonlinear or complex interactions among EEG signals [Jin et al., 2021]. For example, Mutual Information (MI) and Synchronization Likelihood (SL) can help estimate the likelihood of two EEG channels being correlated [Gonzalez-Astudillo et al., 2021]. However, it's important to note that SL can be sensitive

to volume conduction problems [Chiarion *et al.*, 2023]. On the other hand, Transfer Entropy (TE), used to discover direct connections, stands up well against issues with volume conduction [Cao *et al.*, 2022a]. Although nonlinear methods can incorporate more complex FC estimators, they all are prone to problems with spurious connectivities since most noise induced in the EEG leads to nonlinear interactions recorded by different signals that may not reflect genuine relationships [Chiarion *et al.*, 2023].

In brief, linear and nonlinear methods have been developed to extract functional connectivity data from EEG signals. Linear estimators are simple to implement and computationally efficient, but they might not accurately capture complex interactions between EEG channels. Nonlinear methods can address some of these limitations, but they are typically more complex to implement and can be more sensitive to noise in the EEG. While direct connectivity provides information about the direction of information flow between two channels, indirect connectivity only measures the strength of the relationship [Gonzalez-Astudillo *et al.*, 2020]. Both direct and indirect connectivity have been shown to achieve similar performance in EEG-based MI-BCI systems, with indirect connectivity being less sensitive to the volume conduction problem [Cao *et al.*, 2022b].

### Feature Extraction from FC

In conjunction with the selection of single-trial FC estimators, feature extraction techniques are vital for decoding connectivity information, ensuring decent discriminatory capacities between MI tasks [Ai *et al.*, 2019]. Broadly, these techniques can be divided into four primary approaches as shown in Figure 1-9: CSP which finds linear spatial filters that enhance class discriminability; vector representations that leverage the symmetry of indirect FC to extract the upper triangular values from the adjacency matrix; topological representations that depict FC as brain activation topoplots; and approaches using Riemannian geometry that leverage the manifold of symmetric positive definite matrices.



**Figure 1-9:** Four feature extraction approaches for decoding FC in EEG-based MI-BCI: common spatial patterns, FC vectorization, topological representations, and Riemannian geometry

Undoubtedly, CSP is the most extensive feature extraction technique that typically utilizes the simple correlation adjacency matrix to decode connectivities into distinct features, distinguishing between MI classes [Miljevic *et al.*, 2022, Guo *et al.*, 2020]. While it allows direct interpretability using linear spatial filters, its

effectiveness is limited as it relies on sample covariance [Akuthota *et al.*, 2023, Darvish Ghanbar *et al.*, 2021]. CSP requires sufficient sample data and is susceptible to noise, leading to spurious connectivities and overfitting [Khademi *et al.*, 2023]. Consequently, two main approaches to regularization techniques have been explored. The first involves regularizing the covariance matrix through either constant regularization [Park *et al.*, 2017] or weighted regularization, which impacts each channel differently [Deng *et al.*, 2020]. The second approach regularizes the projection matrix  $\tilde{W}$ , specifically, lasso using the  $L_1$ -norm [Zhang *et al.*, 2018a], Tikhonov with the  $L_2$ -norm [Fauzi *et al.*, 2019], Elastic-net combining  $L_1$  and  $L_2$  norms [Gu *et al.*, 2021a], and the  $L_{q/p}$  regularization, which applies the  $q/p$  norm [Cai *et al.*, 2021]. Despite the numerous attempts to mitigate spurious connectivities and overfitting, CSP-based techniques often overlook the nonlinearity interactions between EEG signals [Ghanbar *et al.*, 2021]. Furthermore, they do not offer a direct multi-class scenario [Alizadeh *et al.*, 2023].

Other methodologies, such as vectorizing the adjacency matrix, leverage the symmetry of indirect FC to extract the upper triangular values, which contain insights about the mental state [Meng *et al.*, 2023, Georgiadis *et al.*, 2018]. As exhibited in [Reuderink *et al.*, 2011], it has been demonstrated that when vectorizing is paired with linear classification, it mimics the CSP approach. This strategy allows us to classify directly from the values of the FC matrix. Moreover, vectorizing is easily adaptable to new deep learning approaches, ensuring a seamless integration [Zoumpourlis & Patras, 2022]. However, this approach is prone to overfitting and often exhibits poor performance [Xu *et al.*, 2020a]. Since FC can be represented as graphs, various studies have employed graph measures as a feature to differentiate between MI classes [Gu *et al.*, 2023]. In particular, global measures, such as Characteristic path, global efficiency, and clustering coefficient, alongside node measures like local efficiency, nodal centrality, and degree centrality, have been used as features [Desbois *et al.*, 2021, Rodrigues *et al.*, 2022]. Nonetheless, graph measures offer oversimplified brain functioning features, often resulting in weak performance, especially in subjects with noisy EEG signals [Gonzalez-Astudillo *et al.*, 2020].

So far, we have primarily discussed FC representations in Euclidean spaces, which do not account for the internal link interactions between signals in the adjacency matrix. Therefore, Riemannian geometry, which leverages the links between the coefficients of symmetric and positive definite (SPD) matrices, has been proposed [Freer *et al.*, 2019]. This particular branch of differential geometry allows each matrix to be represented by a point in the manifold  $\mathcal{M}$ , where scalar products can be defined in an Euclidean tangent space [Mishra *et al.*, 2018]. Therefore, the distance between two FC matrices can be approximately measured in this tangent space and be used as a feature to distinguish between two MI tasks [Ding *et al.*, 2023]. However, the main drawback of Riemannian geometry is the necessity for a reference matrix, typically obtained from the average of all the adjacency matrices, where minor changes can significantly affect the distance [Miah *et al.*, 2020].

On the other hand, imaging techniques have gained popularity due to the recent success of convolutional neural networks (CNNs) [Xu *et al.*, 2020c]. That is why some studies have used a topological representation of the FC connectivity to benefit from the easy integration with CNNs [Collazos-Huertas *et al.*, 2021]. While often achieving competitive performance, this approach heavily depends on the number of channels and suffers from potential information loss when projecting the adjacency matrix into a 2D topoplot [Zhao *et al.*, 2019a, Altaheri *et al.*, 2023].

In summary, techniques such as CSP and vectorizing give different methodologies for feature extraction, each with advantages and disadvantages. CSP is a widely recognized extraction strategy that uses a correlation adjacency matrix to find linear spatial filters. However, it heavily relies on sample covariance and is susceptible to noise. Conversely, vectorizing the adjacency matrix lets us classify directly from the FC matrix values and is easily integrated with deep learning approaches. Nevertheless, both methods are susceptible to overfitting, resulting in inferior performance. Riemannian geometry provides a possible solution by allowing each matrix to represent a point on the manifold  $\mathcal{M}$ . However, it necessitates a reference matrix, making it sensitive to minor changes. The topological representation of FC connectivity integrates well with convolutional neural networks, often demonstrating competitive performance. However, it heavily relies on the number of channels and can potentially lose information when projecting the adjacency matrix into a 2D topoplot.

### 1.3.2 Subject-Specific EEG Representation for MI-BCI

In recent years, there has been an increase in the number of methodologies introduced to handle the problem of artifacts in EEG-based MI-BCI systems [Kotte & Dabbakuti, 2020]. Techniques such as Canonical Correlation Analysis (CCA), Principal Component Analysis (PCA), and Independent Component Analysis (ICA) have become common tools to remove contaminated components from EEG signals [Stergiadis et al., 2022]. Notably, CCA has been identified as particularly effective in removing various artifacts [Rashid et al., 2020, Lahane et al., 2019]. In addition, spatial methodologies like Common Average Reference (CAR) and Surface Laplacian (SL) have also been used to reduce noise [Uribe et al., 2019]. However, despite the effectiveness of these techniques, they might introduce complications by adding noise into other sensors [Mridha et al., 2021]. This problem occurs because the implemented filter removes the average electrical activity from neighboring sensors, leading to alterations in the original signal information [Xu et al., 2018]. **On the other hand, Deep Learning (DL) strategies have shown no significant difference in testing results between models with and without classical artifact removing strategies. This suggests that merely using DL models may be sufficient to handle artifact removal [Hassanpour et al., 2019, Altaheri et al., 2023].** Generally, DL methods can be subcategorized based on the nature of their inputs, which may be raw EEG data or handcrafted features such as statistical measures, graph measures, spectral images, and topological representations. Furthermore, the input dimension, which could be a vector, matrix, or tensor representation, also contributes to the categorization. The type of architecture employed also serves as a basis for differentiating DL, thus forming groups such as discriminative, representative, and generative models. The following section reviews the strengths and weaknesses of each model class.

## Input Formulation in Deep Learning

Regarding input formulation, DL techniques can be categorized into models that depend on handcraft-based feature representation or raw EEG signals, as depicted in Figure 1-10. Furthermore, the input can be classified according to its dimensionality, represented as either a vector, matrix, or tensor.

The simplest handcraft method employs either graph measurements or signal statistics, with the input represented as a vector  $[N_q]$ , where  $N_q$  denotes the total number of features [Gu et al., 2023]. Graph measurements are metrics that capture significant properties of connectivity architectures, such as node degree and cluster coefficient. On the other hand, signal statistics represent fundamental characteristics of the signal, like mean, variance, skewness, and kurtosis. Despite the simple nature of these methods, they tend to oversimplify the complex processes within EEG signals, leading to less accurate performance [Gonzalez-Astudillo et al., 2020].

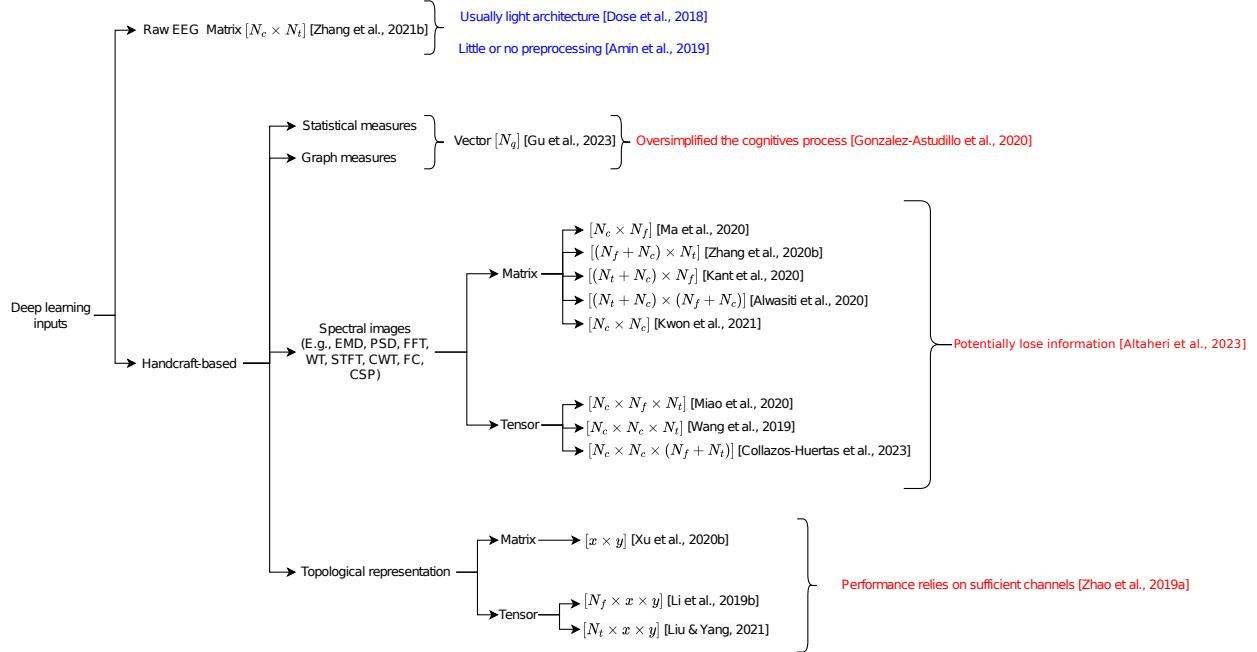
Spectral image approaches are another classification of techniques in which spectrograms, visual representations of the spectrum of frequencies of a signal as it varies with time, are used to portray the EEG signal as spatial-temporal-frequency images. Numerous methods can be employed to extract the spectrogram as a matrix. These techniques include Empirical Mode Decomposition (EMD), a method that decomposes a signal into intrinsic oscillatory components called intrinsic mode functions [Tang et al., 2020]; PSD, a measure of signal's power content versus frequency [Ma et al., 2020]; Fast Fourier Transform (FFT), a mathematical technique that transforms a function of time into a function of frequency [Hassanpour et al., 2019]; Wavelet Transform (WT), a mathematical transform used to split up data, functions or signals into different frequency components [Ortiz-Echeverri et al., 2019]; Short-Time Fourier Transform (STFT), a technique for determining the sinusoidal frequency and phase content of local sections of a signal as it changes over time [Tayeb et al., 2019]; and Continuous Wavelet Transform (CWT), a tool that allows for the analysis of non-stationary signals at many different scales or sizes [Lee & Choi, 2019].

Once the spectrogram is derived, several strategies can be executed for its organization, ranging from simple two-dimensional matrices to complex three-dimensional tensors. The possible permutations include but

## 1. Preliminaries

## 1.3. State of the Art

are not limited to the following formats:  $[N_c \times N_f]$  [Ma et al., 2020],  $[N_c \times N_c \times N_t]$  [Wang et al., 2019b],  $[N_c \times N_c \times (N_f + N_t)]$  [Collazos-Huertas et al., 2023],  $[N_c \times N_c]$  [Kwon et al., 2021],  $[(N_f + N_c) \times N_t]$  [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], and  $[N_c \times N_f \times N_t]$  [Miao et al., 2020].  $N_t$  stands for time points,  $N_c$  number of channels, and  $N_f$  number of filters. The selected arrangement largely depends on the specific requirements of the EEG signal analysis and the complexity of the information to be analyzed. Each organizational structure offers a unique perspective and could be more effective for different tasks or scenarios. Despite this flexibility, it's worth noting that there is a potential loss of information in each representation [Altaheri et al., 2023]



**Figure 1-10:** Deep Learning techniques grouped by input formulation. Featuring handcraft-based feature representation and raw EEG signals, this categorization further divides the input according to dimensions such as vector, matrix, and tensor. The weaknesses and strengths of each form are also highlighted

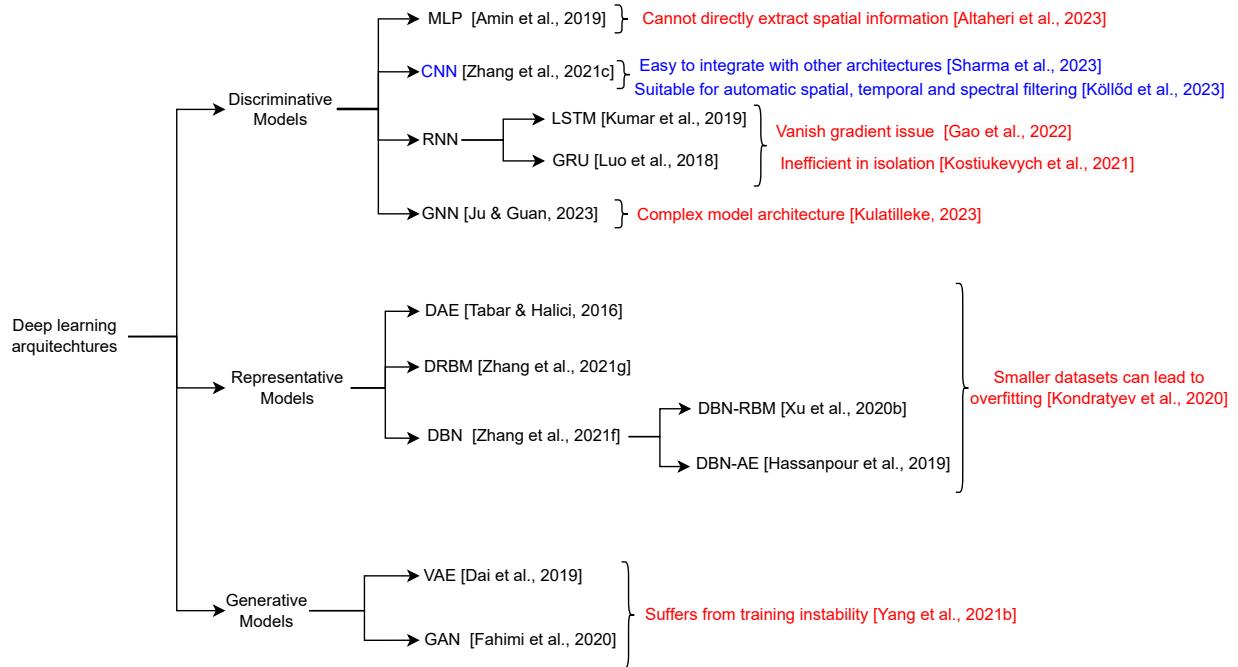
The topological representation approach maps the brain's connectivity to topoplots and then processes them through deep learning architectures. These representations can range from matrix forms to more complex tensor forms. They can appear in the form of a matrix  $[x \times y]$  [Xu et al., 2020c], where  $x$  and  $y$  represent the dimensions of the pixel array in the topoplot image. However, higher dimensional structures are favored to capture the complexities in signal variability across the frequency and time axes. For instance, tensors are often utilized to hold additional dimensions. These structures come in the forms of  $[N_f \times x \times y]$  [Li et al., 2019b], and  $[N_t \times x \times y]$  [Liu & Yang, 2021]. These formats offer a more volumetric view of the information in the EEG data. However, it is to be noted that while these approaches can maintain a greater amount of information, they also heavily rely on a sufficient number of channels being available [Zhao et al., 2019b].

DL models that directly utilize raw EEG data as input are increasingly becoming widespread owing to their unique advantages [Dose et al., 2018]. The typical input structure for these models is in the form of a matrix  $[N_c \times N_t]$ . This configuration allows the model to process the data directly without requiring complex transformations or feature extractions. One of the major strengths of this approach is the lightweight architecture associated with the model. As a result, raw EEG-based DL models are less computationally intensive and require little to no preprocessing [Amin et al., 2019].

## Deep Learning Architectures

Depending on the underlying architecture, DL models can be categorized into discriminative models, representative models, and generative models, as displayed in Figure 1-11.

Discriminative models encompass architectures such as Multilayer Perceptron (MLP) [Amin et al., 2019], a type of feedforward artificial neural network model that maps sets of input data onto appropriate outputs. It's worth noting, however, that MLP is relatively inefficient at extracting spatial information [Hossain et al., 2023, Altaheri et al., 2023]. Convolutional Neural Networks (CNN) [Zhang et al., 2021c] represent another type of discriminative model. Like the connectivity pattern between neurons in the human brain, CNNs process data using a grid-like topology. This architecture is particularly apt for automatic spatial, temporal, and spectral filtering [Köllöd et al., 2023]. Furthermore, CNNs readily integrate with other architectures [Sharma et al., 2023]. Discriminative models also include Recurrent Neural Networks (RNN), which utilize previous model states in current inputs. RNNs branch into Long Short-Term Memory (LSTM) units [Kumar et al., 2019] and Gated Recurrent Units (GRU) [Luo et al., 2018]. LSTM networks, by using gates, can selectively filter temporal information to effectively capture longer dependencies. On the other hand, GRU is a variation of LSTM with only two gates and fewer parameters. Despite these advantages, such networks can lead to inefficiencies when used independently [Kostiukeych et al., 2021] and can still face vanishing gradient issues [Gao et al., 2022]. Graph Neural Networks (GNN), tailored specifically for MI-BCI [Ju & Guan, 2023], have emerged recently. These networks extend the concept of Neural Networks (NNs) to graph-structured data, typically involving complex architectures [Kulatilleke, 2023].



**Figure 1-11:** Deep learning model categories sorted into discriminative, representative, and generative groups, showcasing their weaknesses and strengths

Concerning representative models, architectures such as Deep Autoencoders (DAE) [Tabar & Halici, 2016], Deep Restrictive Boltzmann Machines (DRBM) [Zhang et al., 2021g], and Deep Belief Networks (DBN) are included [Zhang et al., 2021f]. DAE is a specific variant of a Stacked autoencoder and one of the pioneering deep learning techniques. DBN is a generative probabilistic model with deep architecture that contains many layers of hidden units, and DRBM is a deep learning model capable of learning complex temporal behavior. DBN can further be integrated with autoencoders DBN-AE [Hassanpour et al., 2019] and restrictive Boltzmann machines DBN-RBM [Xu et al., 2020b]. However, These models are prone to

overfitting with small datasets [Kondratyev *et al.*, 2020].

Generative models such as Variational Autoencoders (VAE) [Dai *et al.*, 2019] and Generative Adversarial Networks (GAN) [Fahimi *et al.*, 2020] are also prevalent among DL architectures. VAEs are extensions of traditional autoencoders that learn to encode input data into a set of parameters from which data can be generated, while GANs consist of two networks, a generator and a discriminator, that compete against each other to become more accurate in their predictions. However, GANS and VAE suffer from training instability [Yang *et al.*, 2021b].

Finally, hybrid architectures integrated from different models have also been developed to exploit the strengths of different models and mitigate their weaknesses. In practice, hybrid methods for EEG-based MI-BCI sytems include CNN+LSTM [Zhang *et al.*, 2021d], GAN + CNN [Zhang *et al.*, 2020c], LSTM+SVM [Kumar *et al.*, 2021], CNN+SVM [Alazrai *et al.*, 2019], GNN+CNN [Tang *et al.*, 2023], and DBN + CNN [Dai *et al.*, 2019]. to note that CNN is the most relevant architecture being used for most of the hybrid approach since they can extract relevant spatial, temporal, and spectral features.

### 1.3.3 Interpretability Strategies in MI-BCI

To enhance the interpretability of models in MI-BCI systems that handle multiple domain features, the primary goal is to unmask the complexity underlying deep learning models and leverage this knowledge to improve model predictions. As shown in Figure 1-12, XAI methodologies can generally be divided into two categories: post-hoc and intrinsic.

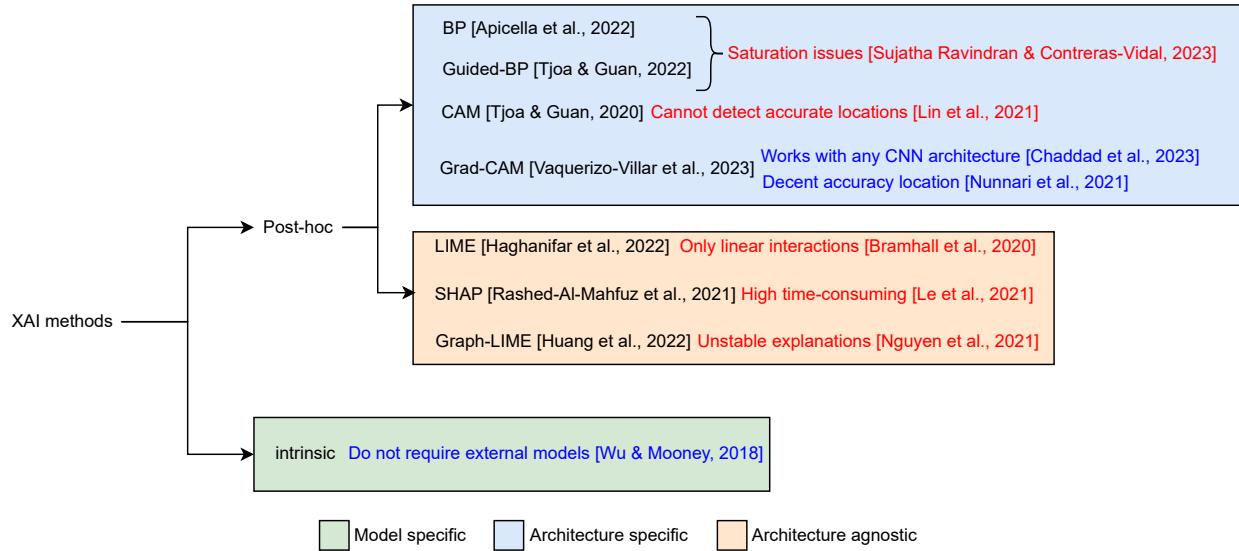


Figure 1-12: Explanability strategies for machine learning and deep learning approaches

Intrinsic models, characterized by their explicit model structures and inherent interpretability, include logistic/linear regression, Random Fourier Features (RFF), k-Nearest Neighbors (KNN), and Bayesian methods [Chaddad *et al.*, 2023]. These techniques directly offer interpretability and do not require external models, but they are model-specific [Wu & Mooney, 2018].

In contrast, post-hoc methods aim to interpret existing models' outputs, serving as a key link between machine predictions and human comprehension [Speith, 2022]. These methods are further subdivided into architecture-agnostic and architecture-specific approaches.

Architecture-agnostic interpretation methods work toward creating simple, interpretable approximations of predictive models. They can be applied across many machine learning models without specific prerequisites. These methods focus on understanding a model’s behavior and pinpointing the key factors that contribute to their predictions. An example is LIME [Haghifar et al., 2022], which offers interpretability but is limited to only linear interactions [Bramhall et al., 2020]. Another method, SHAP [Rashed-Al-Mahfuz et al., 2021], assigns each feature an importance value for a given prediction, providing a unified measure of feature importance. However, it can be computationally intensive and time-consuming [Le et al., 2021]. Lastly, Graph-LIME [Huang et al., 2022] provides explanations based on graph theory but may generate unstable explanations [Nguyen et al., 2021].

Conversely, architecture-specific interpretation methods are custom-shaped to fit particular architectures. The latter leverages shared structural elements among a subset of models [Selvaraju et al., 2017]. For instance, Backpropagation (BP) [Apicella et al., 2022] is commonly utilized as an interpretation method, but it can succumb to saturation issues [Sujatha Ravindran & Contreras-Vidal, 2023]. Likewise, Guided Backpropagation (Guided-BP) [Tjoa & Guan, 2022] is another architecture-specific method but shares similar challenges with BP methods.

In addition, Class Activation Mapping (CAM) [Tjoa & Guan, 2020] aims to emphasize influential regions within classification tasks but struggles with precise localization [Lin et al., 2021]. The Gradient-weighted Class Activation Mapping (Grad-CAM) method [Vaquerizo-Villar et al., 2023] ameliorates this issue and provides decent location accuracy [Chaddad et al., 2023, Nunnari et al., 2021]. However, CAM-style methods, including Grad-CAM and its variants, are traditionally limited to a network’s final layers [Jiang et al., 2021a, Selvaraju et al., 2017, Chattopadhyay et al., 2018]. Gradient-based methods like Grad-CAM often yield noisy maps, limiting interpretability [Omeiza et al., 2019]. While perturbation-based techniques provide insights, they are computationally intensive [Fong & Vedaldi, 2017]. CAM methods offer single-sample visual explanations but are dependent on model architecture [Wang et al., 2020a]. Layer CAM, an enhancement of Grad-CAM, offers better input region localization but may overlook important information from earlier layers. [Jiang et al., 2021a]

On the other hand, quantitative explanations are essential for evaluating the performance of XAI models and comparing them numerically [Hassija et al., 2023]. Metrics such as Average drop %, increase confidence %, and win % can be used to measure the quality of explanations and identify the strengths and weaknesses of different XAI methods [Naidu et al., 2020]. This allows researchers to develop more accurate and interpretable XAI models that can be used to solve a wider range of problems.

In conclusion, each interpretability method has its strengths and limitations and provides unique insights into the inner workings of machine learning models [Rahate et al., 2022]. Therefore, the choice of an interpretability method should mirror each application’s specific requirements and constraints.

#### 1.3.4 Summary

In general, nonlinear methods for extracting indirect functional connectivity data from trial-based EEG signals achieve acceptable accuracy. However, their complexity makes them susceptible to noise, resulting in spurious connectivity. Meanwhile, DL methods capably handle artifacts and EEG representations, necessitating minimal preprocessing. Specifically, models based on CNNs are well-suited for automatic spatial, temporal, and spectral filtering. Furthermore, incorporating intrinsic and post-hoc interpretability strategies into these models could enhance their transparency and interpretability.

## 2 Aims

The analysis above leads us to the following general and specific objectives:

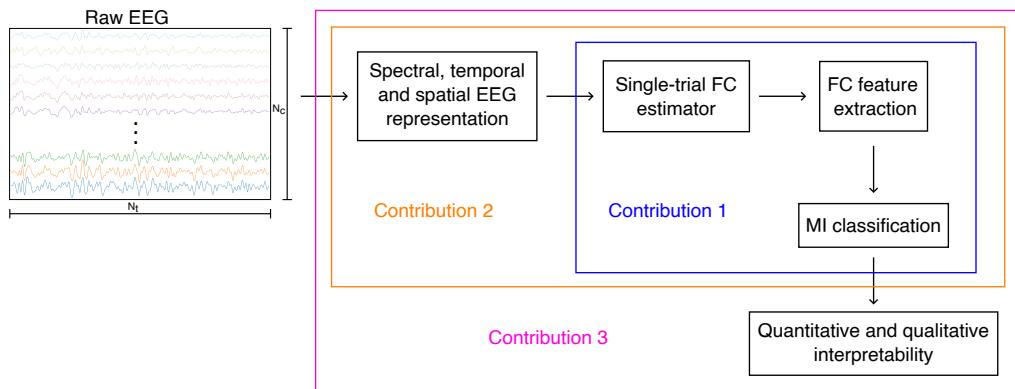
### 2.1 General Objective

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.

### 2.2 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, necessitating 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.

### 3 Outline and Contributions



**Figure 3-1:** Schematic display of the main thesis contributions, including the single-trial FC estimator and feature extraction, an end-to-end approach from EEG representations to MI classification, and quantitative and qualitative strategies for model interpretation

This chapter offers a clear introduction to the main contributions of this thesis, as illustrated in Figure 3-1. Firstly, it presents a novel single-trial FC estimator designed to improve MI classification tasks by addressing nonlinearities, intersubject variability, and spurious connections. Secondly, it introduces a comprehensive end-to-end approach aimed at extracting subject-specific spectral, temporal, and spatial EEG representations, which can aid in reducing the incidence of artifacts and establishing more relevant connections. Lastly, the chapter suggests qualitative and quantitative interpretability strategies for comparing DL EEG-based MI-BCI models, alongside a regularization approach to further mitigate spurious connections.

#### 3.1 Single-Trial FC for Enhancing MI-BCI Classification

As referenced in Section 1.2.1, implementing robust single-trial FC estimators and streamlined feature extraction methodologies is paramount to improving the MI-BCI model's accuracy without affecting its interpretability. More specifically, successful strategies must be devised to confront the complex issues related to spurious connectivities, non-linearity, non-stationarity, and inter-subject variability. Upon a comprehensive examination of the state-of-the-art in Section 1.3.1, our proposal addresses non-stationarity by implementing short sliding windows, which are presumed weakly stationary, at varying widths [Gaur et al., 2021b]. The problem of non-linearity integrating a Gaussian kernel that has proven effective against non-linear situations [Gunawardena et al., 2023]. Specifically, we propose an extension of Bochner's theorem [Bochner, 2020], named the Kernel-based Cross-Spectral FC (KCS-FC), to approximate the cross-spectral distribution of a stationary kernel function by using a weighted sum of Gaussian-based pairwise relationships within different sub-band filters and time windows. The challenge of inter-subject variability using a sub-band filtering tech-

nique [Mammone *et al.*, 2023]. Moreover, a vectorization representation is proposed for FC feature extraction, mainly because of its seamless integration with end-to-end models and its responses, which are closely similar to those in CSP under linear classifier scenarios [Meng *et al.*, 2023, Zoumpourlis & Patras, 2022]. To restrict spurious connectivities, we propose introducing an elastic-net regularization mechanism [Tay *et al.*, 2023]. Figure 5-1 shows a graphical depiction of this proposition.

## 3.2 Automatic Subject-Specific EEG Representation for MI-BCI

As referenced in Section 1.2.2, it is crucial to implement models that do not strongly rely on handcrafted subject-specific EEG representations. These kinds of models can help get rid of artifacts and reduce false connections, which will make MI-BCI more accurate, especially for people who do not do well with their BCI (BCI inefficiency) [Park *et al.*, 2023]. The following method is proposed based on the state-of-the-art exploration in Section 1.3.2. First, DL strategies are proposed due to their significant efficacy in dealing with artifact removal, further mitigating the problem of spurious connectivities[Altaheri *et al.*, 2023, Huang *et al.*, 2023, Hassanpour *et al.*, 2019]. Specifically, end-to-end DL models using raw EEG as inputs are preferred as they require little or no preprocessing and allow for light architectures [Dose *et al.*, 2018, Amin *et al.*, 2019]. Secondly, we propose using a CNN layer to automatically extract temporal and spectral features [Altaheri *et al.*, 2023, Craik *et al.*, 2019, Musallam *et al.*, 2021, Lawhern *et al.*, 2018]. Additionally, spatial extraction is performed by creating a custom layer incorporating the KCS-FC function from our contribution in Section 3.1. This contribution is illustrated in Figure 6-1.

## 3.3 Qualitative and Quantitative Post-Hoc and Intrinsic Explainability

As referenced in Section 1.2.3, transparency and interpretability strategies in MI-BCI models are critical in healthcare and medical decision-making scenarios [Miotto *et al.*, 2018, Xiao *et al.*, 2018]. Additionally, measuring the effectiveness of new MI-BCI models with qualitative and quantitative explainability from multiple domains is not easy. The following explainable AI strategy is proposed based on the state-of-the-art exploration in Section 1.3.3. Firstly, we propose qualitative interpretability using Layer-CAM post-hoc interpretability, which is based on the Grad-CAM method and can work with any CNN architecture [Chaddad *et al.*, 2023]. It presents a decent accuracy location [Nunnari *et al.*, 2021] and highlights spatiotemporal triggers in the raw EEG. Intrinsic interpretability is achieved by checking the last layer weights, resulting in channel importance that is represented as a topoplot. Secondly, for quantitative interpretation, we borrow the concept of average drop [Naidu *et al.*, 2020] from image processing to compare different MI-BCI models. **To compare results we include a regularization technique based on Renyi's entropy to analyze how interpretability is affected when the cross-information potential of the internal FC of the KCS-FCnet is maximized.** This leads to the proposed Interpretable Regularized Kernel Cross-Spectral Functional Connectivity Network (IRKCS-FCnet) depicted in Figure 7-1.

# 4 Materials and Methods

## 4.1 EEG-based Motor Imagery Datasets

We employed three motor-related public databases to appraise the performance of the proposed single-trial KCS-FC, automatic subject-specific EEG representations KCS-FCnet, and the qualitative and quantitative explainability IRKCS-FCnet. These databases are described in detail below.

### 4.1.1 BCI Competition IV Dataset IIa - DBI MI

#### Protocol Design

BCI Competition IV dataset IIa (DBI MI) is a public dataset<sup>1</sup> created by researchers in [Brunner *et al.*, 2008] that contains EEG recordings from nine subjects ( $M = 9$ ) who were instructed to perform four MI tasks ( $N_y = 4$ ) (left hand, right hand, feet, and tongue). Data were gathered in two days within six runs, yielding three runs per day. One run contains 12 trials per task, which were recorded by  $C = 22$  channels, for a total of  $R = 144$  trials of each label per subject, where each channel was sampled at **250Hz**. To estimate Electro-oculography (EOG) influence, EEG recording with opened, closed, and in-movement eyes lasting five minutes at the beginning of each session as illustrated in Figure 4-1.

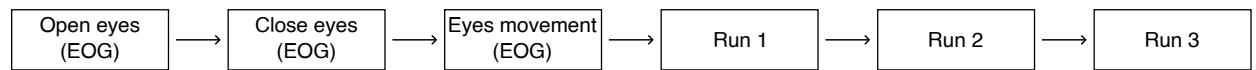
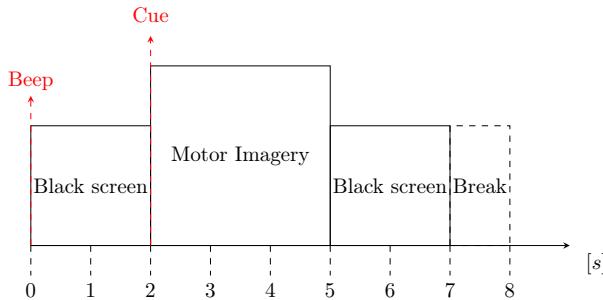


Figure 4-1: Schema representing a one-day session of EEG data collection for the DBI MI

Subjects were instructed to sit on a comfortable armchair in front of a computer screen. Figure 4-2 shows the seven-second MI protocol. The start was indicated by a short acoustic warning and a cross on a black screen for two seconds, then an arrow pointing left, right, down, or up indicated to the subjects to perform the desired MI task. Furthermore, the subjects were asked to continue the MI task until the cross disappeared **three** seconds later. Afterward, a black screen indicated a short break until the beep and the cross reappeared, and a new trial started.

---

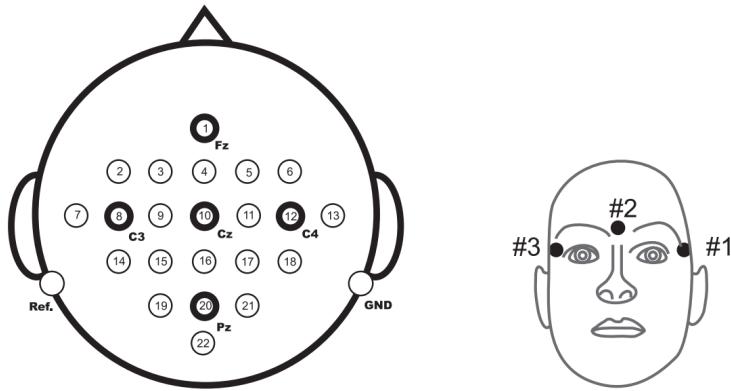
<sup>1</sup>[www.bbci.de/competition/iv](http://www.bbci.de/competition/iv)



**Figure 4-2:** Illustration of the seven-second DBI MI protocol used in the study. The protocol commences with a short acoustic warning and a cross shown on a screen for two seconds, followed by an arrow indicating the type of MI task to be performed. The black screen marks a brief break interval between tasks.

### Data Recording

Twenty-two Ag/AgCl electrodes with an inter-electrode distance of 3.5cm and an international 10 – 20 montage were used to acquire brain activity as shown in Figure 4-3. All signals were monopolar recorded with the left mastoid electrode serving as a reference and the right one as ground. Activity was sampled at 250s and bandpass filtered between 0.5Hz and 100Hz. The amplifier sensitivity was set to 100 $\mu$ V. Additionally, a notch filter centered at 50Hz was applied to diminish electrical line noise. Moreover, three monopolar EOG channels were also included with the same specifications but with amplifier sensitivity set to 1mV. This was done with the intention of subsequent artifact removal strategies. An expert visually inspected all EEG recordings to mark trials with high noise artifacts.



**Figure 4-3:** Illustration of the electrode montage including the left image adhering to the international 10-20 system used for acquiring brain activity and the right image showing the montage of the three monopolar EOG channels employed for subsequent artifact removal. **Source:** [Brunner *et al.*, 2008].

#### 4.1.2 Gamma Motor Execution Database—DBII ME

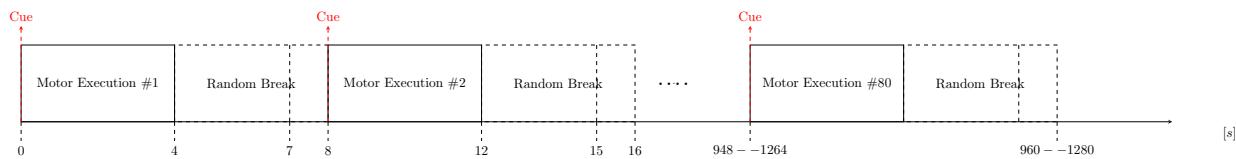
##### Protocol Design

We explore the data that are publicly available at (<https://gin.g-node.org/robintibor/high-gamma-dataset>); this collection is a dataset collected by authors in [Schirrmeister *et al.*, 2017], obtained from 14 healthy subjects (6 female, 2 left-handed, age  $27.2 \pm 3.6$ ) with four-second trials of executed movements divided into

#### 4.1. EEG-based Motor Imagery Datasets

#### 4. Materials and Methods

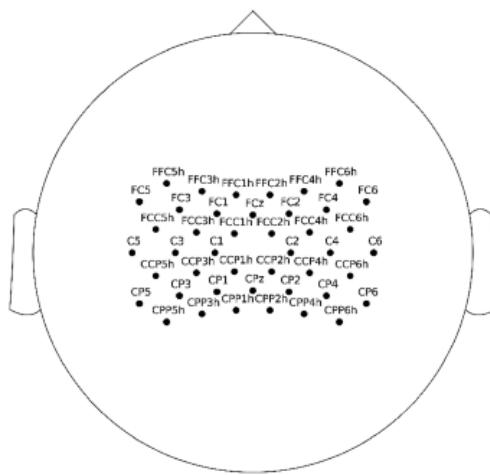
13 runs per subject and ( $R = 260$ ) trials were performed to collect each movement and rest. The four classes were movements of either the left hand, the right hand, both feet, and rest (no movement, but the same type of visual cue as the other classes). Visual cues were presented while using a monitor outside the cabin, which was visible through the shielded window. A fixation point was attached at the center of the screen. The subjects were instructed to relax, fixate on the mark, and keep as still as possible during the motor execution task. The tasks were as follows: Depending on the direction of a gray arrow that was shown on the black background. The subjects had to clench their toes (downward arrow), repetitively, perform sequential finger-tapping of their left (leftward arrow) or right (rightward arrow) hand, or relax (upward arrow). The movements were selected to require little proximal muscular activity while still being complex enough to keep subjects involved. Within the **4s** trials, the subjects performed the repetitive movements at their own pace, which had to be maintained as long as the arrow was showing. Per run, 80 arrows were displayed for **4s** each, with 3 to **4Hz** of continuous random inter-trial interval as shown in Figure 4-4



**Figure 4-4:** Schema representing a one-run session of EEG data collection for the DBII ME

#### Data Recording

The EEG setup for the experiment comprised active electromagnetic shielding, high-resolution low-noise amplifiers, an actively shielded 128-channel EEG cap, and full optical decoupling. All devices were battery-powered and communicated via optical fibers. Data were collected from 44 sensors covering the motor cortex as shown in Figure 4-5, and were sampled at a rate of **500Hz**. Subjects were accommodated comfortably in a dimly lit Faraday cabin where they received visual cues on a monitor approximately 1 meter away. A fixation point served as the focal point during the motor execution exercises. Subjects were instructed to stay still, relax, and limit body movements like blinking and swallowing to inter-trial intervals to ensure optimal data collection. The combined effect of the electromagnetic shielding and the comfortable, dimly lit environment helped maintain data coherence.



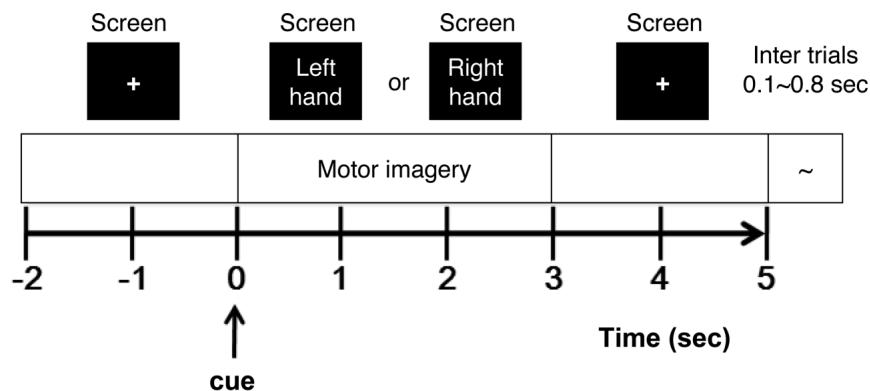
**Figure 4-5:** Positioning of the 44 sensors on the EEG cap, covering the motor cortex used for data collection in the motor execution experiment. Source: [Borra *et al.*, 2019]

### 4.1.3 MI BCI EEG Giga Science Database - DBIII MI

MI BCI EEG Giga Science database (DBIII MI) is an ideal choice for our validation as it has been widely used in the field of MI classification and has been shown to provide a robust benchmark for evaluating the performance of different models [Cho et al., 2017]. However, for the purpose of our study, we will only consider 50 subjects who met the minimum requirement of having at least 100 EEG trials recorded.

#### Protocol Design

This collection, publicly available at <sup>2</sup> and produced by researchers in [Cho et al., 2017] holds EEG recordings obtained from 52 subjects ( $M = 52$ ), holding 19 females, performing two MI tasks ( $N_y = 2$ ) (left and right hand). 50 were right-handed and 2 left-handed. Data were gathered in a laboratory with a noise level between 37 and 39 decibels during one of the next time slots, T1 (9:30–12:00), T2 (12:30–15:00), T3 (15:30–18:00), or T4 (19:00–21:30).  $R = 100$  trials of each label per subject were recorded by  $C = 64$  channels, where each channel was sampled at **512Hz**. Furthermore, EMG and EEG were recorded simultaneously with the same system and sampling rate to check actual hand movements.



**Figure 4-6:** Diagram illustrating the experiment protocol involving motor imagery tasks. The figure shows the sequence of prompting, performing, and break periods in a single run. Source: [Cho et al., 2017]

The experiment involved non-motor related and motor imagery tasks, starting with data collected under six different types of noise conditions: eye blinking, eye movements, head movement, chewing, and resting state. Each noise type was recorded twice for five seconds, except for the resting state condition, which was recorded for 60 seconds. Following the noise data collection, subjects performed MI tasks, which were prompted via a monitor and conducted as per a protocol explained in Figure 4-6. Each trial in this experiment stage started with a 2-second fixation on a cross displayed on a black screen. This was followed by a cue indicating a right or left-hand MI task. During the MI task, subjects were instructed to imagine the sensation of touching each finger of the designated hand with the thumb, beginning from the index finger and proceeding sequentially to the little finger. This emphasis on the kinesthetic rather than the visual aspect ensured that the subjects were indeed imagining the motor task. The MI task continued for 3 seconds, after which a blank screen signaled a rest period. This recovery interval lasted randomly between 2.1 and 2.8 seconds.

This task, break, and recovery cycle was repeated 20 times to form a complete run. Each subject performed between five and six such runs. After each run, a cognitive questionnaire was given to the subjects as an additional data collection point. Data quality was ensured by labeling and excluding trials with high voltage magnitude, classified as "bad trials", from the final analysis. Lastly, to maintain the engagement and

<sup>2</sup><http://gigadb.org/dataset/100295>

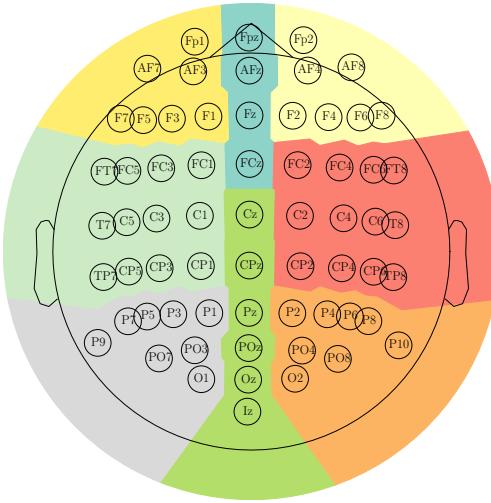
## 4.2. Theoretical Background

## 4. Materials and Methods

motivation of the subjects, they were provided with feedback on the accuracy of their task performance after each run.

### Data Recording

For this experiment, EEG data were collected using 64 active Ag/AgCl electrodes, according to a 64-channel montage based on the international 10-10 system and recorded at a sampling rate of **512Hz** as illustrated in Figure 4-7. The Biosemi ActiveTwo system was employed for EEG collection, with instructions for left or right-hand motor imagery being presented through the BCI2000 system 3.0.2. EMG was recorded concurrently with EEG to verify actual hand movements, with two electrodes attached to the flexor digitorum profundus and extensor digitorum on each arm. For every participant, the EEG channel locations were recorded with a 3D coordinate digitizer; positions were averaged from three measurements to accommodate for possible hand tremors.



**Figure 4-7:** Sensor positions in a 10–10 placement electrode system, containing 64 channels. Besides, it highlights in color the main parts of the brain ( Frontal left, Frontal, Frontal right, Central right, Posterior right, Posterior, Posterior left, Central left)

## 4.2 Theoretical Background

In this section, we introduce the mathematical formulation of relevant related concepts used throughout this study.

### 4.2.1 Motor Imagery Classification from EEG Signals

MI involves the neural simulation of a movement that, although not physically executed, activates the same areas of the brain as actual movements. EEGs capture these activations, and machine learning models can be built to interpret these signals, allowing the prediction of human actions purely from the EEGs.

From the mathematical perspective, let  $\{\mathbf{X}_r\}_{r=1}^R$  be a multi-channel EEG observation from trial  $r$ , where each element  $\mathbf{X}_r = \{\mathbf{x}_r^c \in \mathbb{R}^{N_t}\}_{c=1}^{N_c}$  contains  $N_t \in \mathbb{N}$  time instants and  $N_c \in \mathbb{N}$  number of channels. Moreover,

there exists a function  $\mathcal{F}$  (4-1) that exactly maps each trial  $\mathbf{X}_r$  into the label space  $y_r \in \{0, 1, \dots, N_y\}$  representing the type of motor imagery with  $N_y \in \mathbb{N}$  denoting the number of classes. The goal of EEG-MI classification is to find the best possible estimation function  $\hat{\mathcal{F}}(\cdot; \mathbf{w})$  that approximates the true function  $\mathcal{F}$  in (4-1), depending on a set of trainable parameters denoted as  $\mathbf{w}$

$$\mathcal{F} : \mathbf{X}_r \mapsto y_r \quad \forall r \in \{1, \dots, R\} \quad (4-1)$$

## Performance Measurement

Performance measurement is necessary to evaluate the effectiveness of an EEG signal MI classification model. There are several commonly used metrics to measure the performance of such models, including Accuracy, Kappa coefficient, and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC).

**Accuracy:** is the proportion of trials where a model correctly predicts the class label, represented as:

$$Acc = \frac{1}{R} \sum_{r=1}^R \delta_K(\hat{y}_r, y_r), \quad (4-2)$$

where  $\hat{y}_r$  and  $y_r$  are the predicted and actual class labels, respectively, for the  $r$ -th trial, and  $\delta_K(\cdot, \cdot)$  is the Kronecker delta function.

**Kappa Coefficient:** or Cohen's kappa, measures inter-rater agreement. It considers the agreement occurring by chance, hence, providing a more robust metric than simple accuracy. Given a confusion matrix, the kappa coefficient can be calculated as:

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad (4-3)$$

where  $p_o$  is the relative observed agreement, and  $p_e$  is the hypothetical probability of random chance agreement.

**AUC-ROC curve:** AUC represents the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

For instance, let  $\hat{p}_r$  be the predicted probability of the  $r$ -th trial belonging to the positive class. Then  $\hat{p}_r = \sigma(\hat{\mathcal{F}}(\mathbf{X}_r); \mathbf{w})$ , where  $\sigma(x) = 1/(1 + \exp(-x))$  is the logistic function that transforms the model output into a probability. Therefore, The TPR (or sensitivity) and FPR (or 1-specificity) can be defined as follows:

$$TPR = \frac{TP}{P} = \frac{\sum_{r=1}^R \delta_K(\hat{y}_r, y_r) \delta_K(y_r, 1)}{\sum_{r=1}^R \delta_K(y_r, 1)}, \quad (4-4)$$

$$FPR = \frac{FP}{N} = \frac{\sum_{r=1}^R \delta_K(\hat{y}_r, 1) \delta_K(y_r, 0)}{\sum_{r=1}^R \delta_K(y_r, 0)}, \quad (4-5)$$

where TP is the number of true positives, FP is the number of false positives, P is the number of positive instances, and N is the total number of negative ones. Thus, the AUC can be computed by integrating the curve over the interval  $[0, 1]$ . In practice, the integral is estimated by the trapezoidal rule since we only have a finite number of instances from which to construct the ROC curve. In essence, the AUC-ROC measures the model's ability to discriminate positive from negative cases.

### 4.2.2 Stochastic Processes

A stochastic process, in the context of MI-BCI systems, serves as a robust mathematical construct for characterizing the progression of underlying random systems over distinct mental events. More specifically,  $\chi_e^c$  outlines the behavior of the  $c$ -th channel at an event  $e$ . A trial or sample  $r$  of the event  $e$  can be drawn and be denoted as  $x_{e,r}^c(\tau)$ . This process characteristically encompasses a diverse array of random variables that mimic the dynamism inherent within an event. Hence, a set  $\{x_{e,r}^c(\tau)\}_{r=0,c=1}^{R,N_c}$  can effectively encapsulate the patterns within the stochastic process of the event  $e$ . For simplicity the notation  $x_r^c(\tau)$  is used, assuming a specific event.

#### Random Process Moments

In the study of random processes, 'moments' are vital quantitative measures that describe data set features, including mean, variance, skewness, and kurtosis. In this context, the focus is on the first and second moments. **The first moment is the mean**, providing a central data point while the second moment relates to the variance and correlation, indicating data dispersion and the relationship between variables, respectively. These moments give a comprehensive view of the dataset's structure and dynamics.

#### First Moment

The first moment, also known as the expectation, **of a random variable  $x_r^c(\tau)$** , represents the expected value or mean of the variable. It serves as a measure of the center of the data distribution. The expectation can be mathematically defined as follows:

$$\mathbb{E}_\tau\{x_r^c(\tau)\} = \mu_r^c = \int_{-\infty}^{+\infty} \tau P_r^c(\tau) d\tau, \quad (4-6)$$

where  $P_r^c(\tau)$  is the probability density function of the channel  $c$  at trial  $r$ .

#### Second Moment

The second moment of a random variable, often referred to as the variance, denoted as  $\tilde{\mathcal{R}}_r^c$ , serves as the measure of how much the potential outcomes of the variable deviate from its mean value. This quantifies the spread or dispersion of the data. The formulation for variance can be computed as follows.

$$\begin{aligned} \tilde{\mathcal{R}}_r^c(\tau) &= \mathbb{E}_\tau\{(x_r^c(\tau) - \mu_r^c)^2\} = \int_{-\infty}^{+\infty} (\tau - \mu_r^c)^2 P_r^c(\tau) d\tau \\ &= \mathbb{E}_\tau\{(x_r^c(\tau))^2\} - (\mathbb{E}\{x_r^c(\tau)\})^2 \end{aligned} \quad (4-7)$$

In the context of EEG-based MI-BCI systems, there is a common practice of centering each channel by eliminating the mean value. Consequently, when  $\mu_r^c = 0$ , the variance equation can be reformulated as:

$$\begin{aligned}\tilde{\mathcal{R}}_r^c(\tau) &= \mathbb{E}_\tau\{(x_r^c(\tau))^2\} = \int_{-\infty}^{+\infty} (\tau)^2 P_r^c(\tau) d\tau \\ &= \mathbb{E}_\tau\{(x_r^c(\tau))^2\}\end{aligned}\quad (4-8)$$

The Covariance, another important concept intertwined with variance, measures the joint variability or spread of two random variables. It determines how much the variables change together and quantifies their dependency. If two channels at trial  $r$   $x_r^c$  and  $x_r^{c'}$  are centered,  $\mu_r^c = 0$  and  $\mu_r^{c'} = 0$ , the covariance  $\tilde{\mathcal{R}}_r^{c,c'}$  can be computed as follows.

$$\tilde{\mathcal{R}}_r^{c,c'}(\tau) = \mathbb{E}_\tau [x_r^c(\tau)x_r^{c'}(\tau)] = \mathbb{E}_\tau [x_r^{c'}(\tau)x_r^c(\tau)] \quad (4-9)$$

Covariance is essential in many strategies, such as Common Spatial Patterns (CSP), used for MI-BCI, as they capture the relationships between different EEG channels  $c$ , and  $c'$  at trial  $r$ . These relationships contain relevant information regarding neuronal oscillations and synchronization, both of which are crucial aspects of MI tasks.

### Wide/Weak-Sense Stationarity Stochastic Processes

A stochastic process is deemed wide-sense stationary (WSS) or weak-sense stationary if it satisfies the following two conditions.

1. The mean function or the first moment of the process is constant. This implies that the expected value or the average value of the process should be constant over time and not rely on the underlying time. Mathematically, it is represented by the equation:

$$\mathbb{E}_\tau\{x_r^c(\tau)\} = \mu_r^c = \text{constant, for all } \tau, \quad (4-10)$$

where  $\mathbb{E}_\tau\{x_r^c(\tau)\}$  denotes the expected value of the random variable at any time  $\tau$  and  $\mu_r^c$  is the constant mean value.

2. The autocorrelation function or the second moment of the process depends only on the difference in time and not the actual time. This suggests that the correlation between two variables taken at different periods should only depend on the difference between those periods. It can be mathematically expressed as follows.

$$\mathbb{E}_\tau\{(x_r^c(\tau_1) - \mu_r^c)(x_r^c(\tau_2) - \mu_r^c)\} = \tilde{\mathcal{R}}_r^c(\tau_1 - \tau_2) = \tilde{\mathcal{R}}_r^c(\tau_\Delta), \quad (4-11)$$

Where the autocorrelation function  $\tilde{\mathcal{R}}_r^c(\tau_1 - \tau_2) = \tilde{\mathcal{R}}_r^c(\tau_\Delta)$  between two points at time instances  $\tau_1$  and  $\tau_2$  is only a function of their difference  $\tau_\Delta = (\tau_2 - \tau_1)$ .

### Wiener Khinchin Theorem

According to the Wiener Khinchin theorem, the autocorrelation function  $\tilde{\mathcal{R}}_r^c(\tau)$  of a WSS random process is a Fourier transform pair with its power spectral density  $\tilde{P}_r^c[\omega]$ . This means that the autocorrelation function in the time domain corresponds to the power spectral density in the frequency domain and vice versa. This can be mathematically represented as follows.

$$\tilde{\mathcal{R}}_r^c(\tau) = \int_{\mathbb{R}} \tilde{P}_r^c[\omega] e^{j2\pi f\tau} df \quad (4-12)$$

Likewise, the power spectral density is given by the Fourier transform of the autocorrelation function.

$$\tilde{P}_r^c[\omega] = \int_{\mathcal{U}} \tilde{\mathcal{R}}_r^c(\tau) e^{-j2\pi f\tau} d\tau \quad (4-13)$$

In the context of EEG-based MI-BCI systems, the Wiener-Khinchin theorem provides a useful tool for analyzing EEG signals. By transforming from the time domain to the frequency domain and vice versa, we gain insights into the spectral and temporal properties of the underlying stochastic processes, enabling efficient feature extraction and system identification.

### 4.2.3 EEG Classical Feature Extraction Techniques

The extraction of features from EEG signals is a crucial step towards decoding the information hidden within them. This step involves transforming the raw EEG data into a suitable set of features to represent the inherent properties that are key to differentiating different motor imagery classes. The random variable  $x_r^c(\tau)$  is being discretized, usually at 250Hz or 500Hz, creating a vector with  $N_t$  time samples  $\mathbf{x}_r^c \in \mathbb{R}^{N_t}$ . If all channels  $N_c$  are concatenated the matrix  $\mathbf{X}_r \in \mathbb{R}^{N_c, N_t}$  can be constructed. Let  $\mathcal{G}$  be a feature extraction function that transforms our raw EEG data into feature vectors as follows:

$$\mathbf{z}_r = \mathcal{G}(\mathbf{X}_r), \quad (4-14)$$

where  $\mathbf{z}_r \in \mathbb{R}^{N_p}$  represents the feature vector for trial  $r$  with  $N_p \in \mathbb{N}$  denoting the feature size. The feature extraction function  $\mathcal{G}$  can be defined in various ways based on different feature extraction methods. The objective here is to find an optimal feature extraction function  $\mathcal{G}$  such that the features  $\mathbf{z}_r = \mathcal{G}(\mathbf{X}_r)$  are discriminative enough for the classification of motor imagery EEG signals. After feature extraction step, the problem essentially becomes a feature-based classification problem, where we want to find a function  $\hat{\mathcal{F}}(\cdot, \mathbf{w})$  such that the estimated class label  $\hat{y}_r = \hat{\mathcal{F}}(\mathbf{z}_r; \mathbf{w})$  matches the true label  $y_r$  as accurately as possible.

### Common Spatial Patterns

Common Spatial Pattern (CSP) is a widely used technique for extracting features from EEG signals, especially for MI tasks. CSP helps in identifying spatial filters that maximize the difference of variance between classes. the core concept of CSP is the simultaneous diagonalization of two covariance matrices. CSP relies on Equation (4-9) to estimate the sample covariance matrix for each trial as follows:

$$\text{Cov}(\mathbf{X}_r) = \boldsymbol{\Sigma}_r = \frac{1}{N_t - 1} \mathbf{X}_r \mathbf{X}_r^T, \quad (4-15)$$

Where  $\text{Cov}(\cdot)$  is the covariance function,  $\boldsymbol{\Sigma}_r \in \mathbb{R}^{N_t \times N_c}$  is the covariance matrix for any trial  $r$ ,  $N_t$  is the number of time instances and  $\mathbf{X}_r^T$  is the transpose of the EEG data matrix for trial  $r$ . Next, the average class covariance matrices  $\boldsymbol{\Sigma}_{C1}$  and  $\boldsymbol{\Sigma}_{C2}$  are calculated taking the sum of all covariance matrices for a class divided by the total number of trials in that class. For instance, if we have a total of  $R_{C1}$  trials in the set  $C1$ , the average class matrix  $\boldsymbol{\Sigma}_{C1}$  can be computed as follows:

$$\boldsymbol{\Sigma}_{C1} = \frac{1}{R_{C1}} \sum_{r=1}^{R_{C1}} \boldsymbol{\Sigma}_r; \quad \forall r \in C1 \quad (4-16)$$

A similar process can be followed to obtain the average covariance matrix for class  $C2$ .

The goal of CSP is to find a set of projection vectors (spatial filters) that could maximize the variance for one class while minimizing it for the other class as defined in (4-17).

$$\mathbf{w}^* = \max_{\mathbf{w}} \frac{\mathbf{w}^T \boldsymbol{\Sigma}_{C1} \mathbf{w}}{\mathbf{w}^T \boldsymbol{\Sigma}_{C2} \mathbf{w}} \quad (4-17)$$

The solution to this problem is given by the joint eigenvectors of  $\boldsymbol{\Sigma}_{C1}$  and  $\boldsymbol{\Sigma}_{C2}$  corresponding to the maximum and minimum eigenvalues. The transformed EEG is not directly used in classification models. Instead, a common approach is to use the log variance of the transformed signals and often take from the top and bottom few filters. First, compute the transformed signal as:

$$\mathbf{S}_r = \mathbf{w}^* \mathbf{X}_r \quad (4-18)$$

To calculate the final features for the CSP method, we compute the log variance of these transformed signals as follows:

$$\mathbf{z}_r = \log \left( \frac{\text{diag}(\text{Cov}(\mathbf{S}_r))}{\text{Tr}(\text{Cov}(\mathbf{S}_r))} \right), \quad (4-19)$$

Where  $\text{Tr}(\cdot)$  stands for the trace operator,  $\text{diag}(\cdot)$  is the diagonal operator that extracts the elements in the principal diagonal, and  $\log(\cdot)$  represent the element wise logarithmic operator. For instance, if we're using CSP as our feature extraction method,  $\mathcal{G}_{CSP}$  will constitute the operations involved in (4-19), where the filters corresponding to the highest and lowest eigenvalues are selected, such as the from  $\mathbf{z}_r$  we take the first  $N_p/2$  and last  $N_p/2$  elements, where  $N_p$  is the final desired number of features. Finally, these features  $\mathbf{z}_r$  extracted by applying the CSP technique can now be fed to the classifier for further processing.

Sure, the requested sub-subsection would be as follows:

### Variations of Common Spatial Patterns

Over time, various adaptations of the CSP method have been developed to address specific issues and improve performance. These include the next regularizations techniques weighted by the  $\alpha$  parameter:

*Constant regularization:* Where a constant term is added to the sample variance for diminishing the risk of overfitting [Park et al., 2017].

$$\hat{\Sigma} = \{\Sigma + \alpha I\} \quad (4-20)$$

*Lasso regularization:* This method employs a penalty equal to the absolute value of the magnitude of the coefficients. It aims to simplify the used model by forcing some coefficients to zero in the spatial linear filters [Zhang et al., 2018a].

$$\mathbf{w}^* = \max_{\mathbf{w}} \left( \frac{\mathbf{w}^T \Sigma_{C1} \mathbf{w}}{\mathbf{w}^T \Sigma_{C2} \mathbf{w}} - \alpha \|\mathbf{w}\|_1 \right) \quad (4-21)$$

*Tikhonov regularization:* A squared magnitude of the spatial filter coefficients is employed as a penalty term, preventing coefficients from reaching very high values that could lead to overfitting [Fauzi et al., 2019]

$$\mathbf{w}^* = \max_{\mathbf{w}} \left( \frac{\mathbf{w}^T \Sigma_{C1} \mathbf{w}}{\mathbf{w}^T \Sigma_{C2} \mathbf{w}} - \alpha \|\mathbf{w}\|_2 \right) \quad (4-22)$$

*Elastic-net regularization:* This technique is a mix of the Lasso and Tikhonov regularizations, using both absolute and squared magnitudes, thus encompassing their benefits [Gu et al., 2021b].

$$\mathbf{w}^* = \max_{\mathbf{w}} \left( \frac{\mathbf{w}^T \Sigma_{C1} \mathbf{w}}{\mathbf{w}^T \Sigma_{C2} \mathbf{w}} - \alpha \|\mathbf{w}\|_1 \frac{1-\alpha}{2} \|\mathbf{w}\|_2 \right) \quad (4-23)$$

*Weighted regularization:* The diagonal of the sample covariance matrix is penalized with distinct weights to differentially shrink their effect on the spatial filters.[Deng et al., 2020]

$$\hat{\Sigma} = \{\Sigma + \alpha \text{diag}(\mathbf{a})\} \quad (4-24)$$

*Lq/p regularization:* A generalization of Lasso and Tikhonov regularizations. It is flexible with  $q$  and  $p$ , which can be any non-negative real numbers, providing more control over the model complexity [Cai et al., 2021].

$$\mathbf{w}^* = \max_{\mathbf{w}} \left( \frac{\mathbf{w}^T \Sigma_{C1} \mathbf{w}}{\mathbf{w}^T \Sigma_{C2} \mathbf{w}} - \alpha \|\mathbf{w}\|_{p,q} \right) \quad (4-25)$$

Each of these versions of CSP can be selected and implemented according to the specific demands and objectives of the dataset and problem at hand.

### Sub-band CSP

In practice, EEG signals not only differ in spatial patterns but also in frequency bands. Sub-band Common Spatial Pattern (SBCSP) is an extension that considers the frequency components of the EEG signals. First, it decomposes the original EEG signals into various frequency bands before applying the CSP method to each band individually. The generated sub-band CSP features are then concatenated to create a more robust feature vector that includes both spatial and spectral information, resulting in improved performance in classifying motor imagery tasks.

Suppose we define a band-pass filter  $\mathcal{H}_f$  that decomposes an EEG signal into the frequency band  $f \in \Omega$ , where  $\Omega$  is a set of filters. Thus, the filter operation can be represented as:

$$\tilde{\mathbf{X}}_{r,f} = \mathcal{H}_f(\mathbf{X}_r), \quad (4-26)$$

where  $\tilde{\mathbf{X}}_{r,f}$  represents the signal  $\mathbf{X}_r$  in the frequency band  $f$ . Then, the CSP transformation is performed over the sub-band signals, and the variance of the transformed signals is used to construct the sub-band CSP features as follows:

$$\tilde{\mathbf{z}}_{r,f} = \mathcal{G}_{CSP}(\tilde{\mathbf{X}}_{r,f}; \mathbf{w}_f^*), \quad (4-27)$$

where  $\mathcal{G}_{CSP}$  is the function encapsulating the CSP featurization described in (4-19), and  $\mathbf{w}_f^*$  is the spatial filter for the frequency band  $f$ . Therefore, to obtain the complete feature vector  $\mathbf{z}_r$  for trial  $r$ , the sub-band CSP features from all considered frequency bands are concatenated:

$$\mathbf{z}_r = [\tilde{\mathbf{z}}_{r,1}, \tilde{\mathbf{z}}_{r,2}, \dots, \tilde{\mathbf{z}}_{r,N_f}], \quad (4-28)$$

where  $N_f = |\Omega|$  is the total number of frequency bands and  $|\cdot|$  stands for the set cardinality.

This method improves classification performance over the traditional CSP by incorporating frequency information. It better captures distinct patterns across different frequency bands to differentiate between motor imagery tasks.

### Time Windowing and Sub-Band CSP

**EEG signals often exhibit different patterns across different time intervals and frequency bands during a predefined mental tasks.** Thus, along with the spatial and the frequency domain covered by sub-band CSP, it is also valuable to consider the time domain. We can incorporate information from the time dimension by applying a sliding time window over the continuous EEG signal and extracting sub-band CSP features from each time window. This way, we get a time-frequency (or t-f) distribution of CSP features, providing a more comprehensive representation of the signal for motor imagery classification tasks.

Suppose the raw EEG signal from trial  $r$  is given by  $\mathbf{X}_r$  and  $\Delta_t$  be the time window set. Thus, we can split  $\mathbf{X}_r$  into  $\tau = |\Delta_t|$  time windows, each containing  $N'_t$  time instants.  $\mathbf{X}_{r,t} \in \mathbb{R}^{N_t \times N'_t}$  denotes the signal in the  $t$ -th window. Next, we perform the sub-band CSP methods for each time-windowed signal. As a result, a sub-band CSP feature  $\tilde{\mathbf{z}}_{r,t,f}$  for the  $t$ -th time window and  $f$ -th frequency band is obtained. Finally, to obtain the complete feature vector that incorporates time, frequency, and spatial information, we concatenate all the sub-band CSP features across all time windows and frequency bands as follows:

$$\mathbf{z}_r = [\tilde{\mathbf{z}}_{r,1,1}, \tilde{\mathbf{z}}_{r,1,2}, \dots, \tilde{\mathbf{z}}_{r,\tau,N_f}]. \quad (4-29)$$

This method is essentially an enhancement of the Sub-band CSP that considers the information encoded in the time, frequency, and spatial domains simultaneously. By viewing the time dimension, it extracts more comprehensive features, which can lead to improved classification performance on motor imagery tasks.

#### 4.2.4 Functional Connectivity Estimators

In this section, we describe the most important functional connectivity estimators.

##### Correlation

Correlation (Corr) is a measure often used to represent the linear relationship between two EEG channels [Fagerholm *et al.*, 2020]. For two vectors  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  from two channels  $c$  and  $c'$  at trial  $r$  such that  $c \neq c'$ , the correlation is defined as:

$$\text{Corr}_r^{c,c'} = \frac{\mathbb{E}_t \left[ (x_r^c[t] - \mu_r^c)(x_r^{c'}[t] - \mu_r^{c'}) \right]}{\sigma_r^c \sigma_r^{c'}}, \quad (4-30)$$

where  $\mathbb{E}\{\cdot\}$  is the expected value,  $\mu_{X^c}$  and  $\mu_{X^{c'}}$  are the mean values and  $\sigma_{X^c}$  and  $\sigma_{X^{c'}}$  are the standard deviations of  $X^c$  and  $X^{c'}$  time series.

##### Cross-Correlation

Cross-correlation (Cross-corr) differs from Corr since it is a function with respect to a time lag  $\delta$ , which can be expressed as follows [Roy *et al.*, 2022].

$$\text{Cross-corr}_r^{c,c'}(\delta) = \frac{\mathbb{E}_t \left[ (x_r^c[t] - \mu^c)(x_r^{c'}[t + \delta] - \mu^{c'}) \right]}{\sigma^c \sigma^{c'}}. \quad (4-31)$$

##### Magnitude Square Coherence

The magnitude-squared coherence (MSC) is a measure that, as a function of frequency, estimates how well channel  $c$  relates linearly to another channel  $c'$  [Cattai *et al.*, 2021]. It is defined by the squared modulus of the cross-spectrum  $S_r^{c,c'}[\omega]$  of the two vectors,  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$ , divided by the product of their auto-spectral densities  $S_r^c[\omega]$  and  $S_r^{c'}[\omega]$  as:

$$\text{MSC}_r^{c,c'}[\omega] = \frac{|S_r^{c,c'}[\omega]|^2}{S_r^c[\omega] S_r^{c'}[\omega]}. \quad (4-32)$$

### Phase Locking Value

Phase Locking Value (PLV) is another technique commonly used in EEG analysis. PLV measures the consistency of the phase difference between a pair of channels, which can capture the synchronization activities between different brain regions. Suppose  $x_r^c[t]$  is the EEG signal from channel  $c$  at time  $t$  and trial  $r$ . The instantaneous phase  $\phi_{x^c}[t]$  can be extracted using the Hilbert transformation. According to authors in [Cattai et al., 2021], the phase difference between channel  $c$  and  $c'$  is calculated as follows:

$$\Delta\phi_{c,c'}[t] = \phi_{x^c}[t] - \phi_{x^{c'}}[t], \quad (4-33)$$

hence, the PLV between the two channels is then computed by averaging the phase difference over the time dimension and taking the absolute value:

$$PLV_r^{c,c'} = \left| \frac{1}{N_t} \sum_{t=1}^{N_t} e^{j\Delta\phi_{c,c'}[t]} \right|, \quad (4-34)$$

where  $j$  is the imaginary unit, values range from 0 to 1, with 1 indicating perfect phase locking (i.e., constant phase difference over time) and 0 indicating a random phase relationship.

### Phase Lag Index

The Phase Lag Index (PLI) captures the asymmetry of the distribution of phase differences between two signals and is calculated based on the relative phase difference between the two signals [Siviero et al., 2023].

$$PLI_r^{c,c'} = |\mathbb{E}_t[\text{sign}(\Delta\phi_{c,c'}[t])]|, \quad (4-35)$$

The resulting value lies in the interval  $[0, 1]$ , where a higher value indicates more phase synchrony.

### Weighted Phase Lag Index

The Weighted Phase Lag Index (WPLI) is a measure of the phase synchronization between two signals [Gonzalez-Astudillo et al., 2020]. It's a modification of PLI but takes the magnitude of the imaginary component of cross-spectrum into account; thus, it is less sensitive to noise. This measure can be expressed mathematically by extending the PLI definition.

$$WPLI_r^{c,c'} = \frac{\left| \mathbb{E}_f \left[ \left| \Im[S_r^{c,c'}[\omega]] \right| \text{sign} \left( \Im[S_r^{c,c'}[\omega]] \right) \right] \right|}{\mathbb{E} \left[ \left| \Im[S_{X^c, X^{c'}}] \right| \right]}, \quad (4-36)$$

where the  $S_r^{c,c'}[\omega]$  is the cross-spectrum of signals in channels  $c$  and  $c'$  at trial  $r$ ,  $\Im[\cdot]$  denotes the imaginary part of a complex number, and  $\text{sign}(\cdot)$  denotes the sign function:

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases} \quad (4-37)$$

### Imaginary Part of the Coherence

Imaginary Part of the Coherence (IPC) uses the imaginary part of the cross-spectrum, discarding any zero-lag interactions, often due to volume conduction, field spread, or common references [Cao *et al.*, 2022a]. The IPC function can be derived from the cross-spectrum of two signals as follows.

$$IPC_r^{c,c'}[\omega] = \frac{|\Im[S_r^{c,c'}[\omega]]|}{\sqrt{S_r^c[\omega]S_r^{c'}[\omega]}} \quad (4-38)$$

### Partial Coherence

Partial Coherence (PC) quantifies the unique linear relationship between two signals by removing the influence of other signals [Gonzalez-Astudillo *et al.*, 2020]. For multichannel signals, this can be understood as the coherence between two signals of interest after linearly regressing out the contributions from all other signals.

The PC between two signals  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  from two channels  $c$  and  $c'$  at trial  $r$  can be represented as:

$$PC_r^{c,c'}[\omega] = \frac{|S_r^{c,c'}[\omega]|^2 - |\sum_{k \neq c,c'} S_r^{c,k}[\omega]S_r^{c',k}[\omega]|^2}{(S_r^c[\omega]S_r^{c'}[\omega]) - (\sum_{k \neq c} S_r^{c,k}[\omega])(\sum_{k \neq c'} S_r^{c',k}[\omega])}, \quad (4-39)$$

where  $S_r^{c,c'}[\omega]$  is the cross-spectrum between two signals  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$ ,  $S_r^c[\omega]$  and  $S_r^{c'}[\omega]$  are the power spectral densities of signals  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  respectively. The sum is taken over all  $k$  channels, excluding the channels  $c$  and  $c'$ .

### Mutual Information

According to information theory, the Mutual Information, MI, of two random vectors  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  shows how a random vector is informative for the other one. Let  $P_r(\mathbf{x}_r^c)$  and  $P_r(\mathbf{x}_r^{c'})$  be the probability distributions of random vectors  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  at trial  $r$ , respectively. The entropy of  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  is defined as by [Gu *et al.*, 2023] as follows.

$$\begin{aligned} H(\mathbf{x}_r^c) &= -\sum_{t=0}^{N_t} P_r(x_r^c[t]) \log_b(P_r(x_r^c[t])), \\ H(\mathbf{x}_r^{c'}) &= -\sum_{t=0}^{N_t} P_r(x_r^{c'}[t]) \log_b(P_r(x_r^{c'}[t])), \end{aligned} \quad (4-40)$$

where  $t$  defines the time sample and  $H(\mathbf{x}_r^c | \mathbf{x}_r^{c'})$  and  $H(\mathbf{x}_r^c, \mathbf{x}_r^{c'})$  are conditional entropy and joint entropy between  $X^c$  and  $X^{c'}$ , defined respectively as follows.

$$\begin{aligned} H(\mathbf{x}_r^c, \mathbf{x}_r^{c'}) &= -\mathbb{E}_{\mathbf{x}_r^c} \left[ \mathbb{E}_{\mathbf{x}_r^{c'}} \left[ \log_b P_r(\mathbf{x}_r^c, \mathbf{x}_r^{c'}) \right] \right], \\ H(\mathbf{x}_r^c | \mathbf{x}_r^{c'}) &= -\mathbb{E}_{\mathbf{x}_r^c} \left[ \mathbb{E}_{\mathbf{x}_r^{c'}} \left[ \log_b P_r(\mathbf{x}_r^{c'} | \mathbf{x}_r^c) \right] \right], \end{aligned} \quad (4-41)$$

where  $\mathbb{E}$  is the expected value function. MI of two random vectors  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  is computed as

$$\text{MI}_r(\mathbf{x}_r^c, \mathbf{x}_r^{c'}) = H(\mathbf{x}_r^c) + H(\mathbf{x}_r^{c'}) - H(\mathbf{x}_r^c, \mathbf{x}_r^{c'}) = H(\mathbf{x}_r^{c'}) - H(\mathbf{x}_r^{c'} | \mathbf{x}_r^c) \quad (4-42)$$

$\text{MI}_r(\mathbf{x}_r^c, \mathbf{x}_r^{c'}) = 0$  if and only if random variables  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  are statistically independent.

## Granger Causality

Granger causality (GC) is a mathematical formalization of the Wiener's causality concept. For two stationary time series  $\mathbf{x}^c$  and  $\mathbf{x}^{c'}$ , the linear autoregressive model is defined.

$$\mathbf{x}_r^{c'}[t] = \sum_{k=0}^o a_k \mathbf{x}_r^{c'}[t-k] + \epsilon[t], \quad (4-43)$$

Where  $o$  is the model order,  $a_k$  is the model coefficients, and  $\epsilon[t]$  is the error at time  $t$ . Likewise, let define the bivariate autoregressive model as follows.

$$\mathbf{x}_r^{c'}[t] = \sum_{k=0}^o a_k \mathbf{x}_r^c[t-k] \sum_{k=0}^o b_k \mathbf{x}_r^{c'}[t-k] + \epsilon'[t], \quad (4-44)$$

where  $o$  is the model order,  $a_k$  and  $b_k$  are the model coefficients, and  $\epsilon'[t]$  is the error at time  $t$ .

The idea behind GC is that each value  $\mathbf{x}_r^{c'}[t]$  will be better defined including information from  $\mathbf{x}_r^c[t]$ , that means  $|\epsilon'| < |\epsilon|$ . Thus, the GC can be defined as follows [Rezaei & Shalbaf, 2023].

$$\text{GC}(\mathbf{x}_r^c \rightarrow \mathbf{x}_r^{c'}) = \log \left( \frac{\text{var}\{\epsilon\}}{\text{var}\{\epsilon'\}} \right), \quad (4-45)$$

where  $\epsilon, \epsilon' \in \mathbb{R}^{N_t-o}$  are vectors holding the prediction errors, and  $\text{var}\{\cdot\}$  stands for the variance operator. If the past of  $\mathbf{x}_r^c$  does not improve the prediction of  $\mathbf{x}_r^{c'}$ , then  $\text{var}\{\epsilon\} \approx \text{var}\{\epsilon'\}$  and  $\text{GC}(\mathbf{x}_r^c \rightarrow \mathbf{x}_r^{c'}) \rightarrow 0$ ; if the prediction improves, then  $\text{var}\{\epsilon\} \gg \text{var}\{\epsilon'\}$  and  $\text{GC}(\mathbf{x}_r^c \rightarrow \mathbf{x}_r^{c'}) \gg 0$ .

## Directed Transfer Function

The Directed Transfer Function (DTF) is based on the concept of Granger causality and estimated using multivariate autoregressive (MVAR) models, which uses all signals simultaneously [Rezaei & Shalbaf, 2023]. First, the MVAR is calculated as follows:

$$\mathbf{X}_r[t] = \sum_{t'=1}^{\tilde{p}} (\mathbf{A}_r[t'] \mathbf{X}_r[t-t']) + \epsilon_r[t], \quad (4-46)$$

where  $\mathbf{X}_r[t]$  is the matrix containing all channels at trial  $r$  and time sample  $t$ ,  $\mathbf{A}_r[t']$  is the coefficient matrix at lag  $t'$ ,  $\tilde{p}$  is the model order, and  $\epsilon_r[t]$  is the residual matrix at time  $t$  and trial  $r$ . Second, the MVAR model

## 4.2. Theoretical Background

## 4. Materials and Methods

can be transformed into its frequency domain version using the Fourier transform to obtain the transfer function matrix  $\mathbf{H}_r[\omega]$  as follows:

$$\mathbf{H}_r[\omega] = \left[ I - \sum_{t'=1}^{\tilde{p}} \mathbf{A}_r[t'] \mathcal{F}(\mathbf{X}_r[t-t']) \right]^{-1}, \quad (4-47)$$

where  $\mathbf{H}_r[\omega]$  represents the transfer matrix,  $\omega$  is the frequency and  $\mathcal{F}$  is the Fourier transform. Finally, the influence from channel  $c$  to channel  $c'$  can be calculated as follows:

$$\text{DTF}_{r,c'}^c[\omega] = \frac{|h_r^{c,c'}[\omega]|^2}{\sum_{k=0}^{N_c} |h_r^{c,k}[\omega]|^2}, \quad (4-48)$$

where  $h_r^{c,c'}[\omega]$  is the element of the transfer matrix  $H_r[\omega]$  from channel  $c$  to channel  $c'$ , and  $N_c$  is the total number of channels. The square of the magnitude of  $h_r^{c,c'}[\omega]$  is normalized by the sum of squares of the magnitudes of all elements in channel  $c$  of  $H_r[\omega]$ .

## Partial Directed Coherence

Partial Directed Coherence (PDC) is a frequency-domain measure based on MVAR models, designed to examine directed interactions between any pair of channels  $c$  and  $c'$ . Similar to the DTF, the PDC is derived using Granger causality principles, but the normalization method differs, allowing the differentiation between direct and indirect interactions. As in DFT the first step is to fit the MVAR model to the data as follows:

$$\mathbf{X}_r[t] = \sum_{t'=1}^{\tilde{p}} (\mathbf{A}_r[t'] \mathbf{X}_r[t-t']) + \boldsymbol{\epsilon}_r[t], \quad (4-49)$$

where  $\mathbf{X}_r[t]$  is the matrix containing all channels at trial  $r$  and time sample  $t$ ,  $\mathbf{A}_r[t']$  is the coefficient matrix at lag  $t'$ ,  $\tilde{p}$  is the model order, and  $\boldsymbol{\epsilon}_r[t]$  is the residual matrix at time  $t$  and trial  $r$ . Second, the MVAR model can be transformed into its frequency domain version using the Fourier transform to obtain the matrix of coefficient in the frequency domain  $\tilde{\mathbf{A}}_r[\omega]$  as follows:

$$\begin{aligned} \tilde{a}_r^{c,c'}[\omega] &= \sum_{t'=0}^{\tilde{p}} \mathcal{F}(a_r^{c,c'}[t']), \\ &= \sum_{t'=0}^{\tilde{p}} a_r^{c,c'}[t'] \exp^{j2\pi\omega\tau t'}, \end{aligned} \quad (4-50)$$

where  $\tilde{a}_r^{c,c'}[0] = 1$ ,  $\mathbf{H}_r[\omega]$  represents the transfer matrix,  $\omega$  is the frequency and  $\mathcal{F}$  is the Fourier transform. Finally, the PDC from channel  $c$  to channel  $c'$  at frequency  $\omega$  and trial  $r$  can be written as follows:

$$\text{PDC}_r^{c,c'}[\omega] = \frac{|\tilde{a}_r^{c,c'}[\omega]|}{\sqrt{\left| \sum_{k=0}^{N_c} \tilde{a}_r^{c,k}[\omega] \right|^2}}. \quad (4-51)$$

PDC ranges from 0 to 1, providing the degree of a direct influence of one channel on another in the frequency domain [Gaxiola-Tirado et al., 2017].

## Transfer Entropy

Transfer Entropy (TE) is an alternative measure of direct functional connectivity based on information theory. TE does not require a model of the interaction and is inherently nonlinear. TE for two observed

## 4. Materials and Methods

## 4.2. Theoretical Background

channels  $\mathbf{x}_r^c$  and  $\mathbf{x}_r^{c'}$  can be written as:

$$\text{TE}(\mathbf{x}_r^c \rightarrow \mathbf{x}_r^{c'}) = \sum_{x_r^{c'}[t+u], \mathbf{x}^{c'}[t]^{d_{\mathbf{x}^{c'}}}, \mathbf{x}_r^c[t]^{d_{\mathbf{x}_r^c}}} P(x_r^{c'}[t+u], x_r^{c'}[t]^{d_{\mathbf{x}_r^{c'}}}, \mathbf{x}_r^c[t]^{d_{\mathbf{x}_r^c}}) \log \left( \frac{P(x_r^{c'}[t+u] | x_r^{c'}[t]^{d_{\mathbf{x}_r^{c'}}}, \mathbf{x}_r^c[t]^{d_{\mathbf{x}_r^c}})}{P(x_r^{c'}[t+u] | x_r^{c'}[t]^{d_{\mathbf{x}_r^{c'}}})} \right), \quad (4-52)$$

where  $t$  is a time-index and  $u$  denotes the prediction time.  $x_r^{c'}[t]^{d_{\mathbf{x}_r^{c'}}}$  and  $x_r^c[t]^{d_{\mathbf{x}_r^c}}$  are  $d_{\mathbf{x}_r^{c'}}-$  and  $d_{\mathbf{x}_r^c}-$ dimensional delay vectors [Rezaei & Shalbaf, 2023]. Note that the probability density functions is unknown. Hence, the first step is to estimate it using methods like histograms [Choi & Kim, 2021], kernel estimates [Kuang, 2021], or K-nearest neighbors [Martínez-Cancino et al., 2020].

### Synchronization Likelihood

The Synchronization Likelihood (SL) method is based on the concept of generalized synchronization and can be applied to both linear and nonlinear systems [Gonzalez-Astudillo et al., 2021]. From each time sample  $x_r^c[t]$  the embedded vectors  $\mathbf{q}_r^c$  are calculated as follows:

$$\mathbf{q}_r^c[t] = [x_r^c[t], x_r^c[t+l], x_r^c[t+2l], \dots, x_r^c[t+(m-1)l]], \quad (4-53)$$

where  $m$  is the embedding dimension and  $l$  is the lag. Then, for each vector  $\mathbf{q}_r^c[t]$  a time-lag simmilarity probability  $p_r^c[t]$  is computed using a predefined distance  $\varepsilon$  as follows:

$$p_r^c[t] = \frac{1}{2(\tau_{\Delta 2} - \tau_{\Delta 1})} \sum_{t'=1}^{N_t} \tilde{\Theta}(\varepsilon - \langle \mathbf{q}_r^c[t], \mathbf{q}_r^c[t'] \rangle_2), \quad (4-54)$$

where  $\langle \cdot \rangle_2$  is the Euclidean distance,  $\tilde{\Theta}(\cdot)$  is the Heaviside step function described in Equation (4-55),  $\tau_{\Delta 1}$  is the Theiler correction to avoid autocorrealtion noise, and  $\tau_{\Delta 2}$  is the length of the time window of interest.

$$\tilde{\Theta}(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1, & \text{if } x > 0. \end{cases} \quad (4-55)$$

Each channel  $c$  at time  $t$  has a critical distance  $\varepsilon_r^c[t]$ , that can be determined assumming  $p_r^c[t] = p_{\text{ref}}$  and  $p_{\text{ref}} \ll 1$ . Hence, the number of channels where the embedded vectors  $\mathbf{q}_r^c[t]$  and  $\mathbf{q}_r^c[t']$  are closer than the critical distance can be computed as follows:

$$h_r[t, t'] = \sum_{c=1}^{N_c} \tilde{\Theta}(\varepsilon_r^c[t] - \langle \mathbf{q}_r^c[t], \mathbf{q}_r^c[t'] \rangle_2), \quad (4-56)$$

where  $h_r[t, t']$  shows the number of embedded vectors that are alike and is bounded by  $0 \geq h_r[t, t'] \leq N_c$ . We can now defined the synchronization likelihood  $SL_r^c[t, t']$  for each channel  $c$  at trial  $r$  and at any time point pair  $[t, t']$  as:

$$SL_r^c[t, t'] = \begin{cases} \frac{h_r[t, t'] - 1}{N_c - 1}, & \text{if } \langle \mathbf{q}_r^c[t], \mathbf{q}_r^c[t'] \rangle_2 < \varepsilon_r^c[t] \\ 0, & \text{if } \langle \mathbf{q}_r^c[t], \mathbf{q}_r^c[t'] \rangle_2 \geq \varepsilon_r^c[t], \end{cases} \quad (4-57)$$

we can now take the average over all  $t'$  time samples to obtain  $SL_r^c[t]$ , which describes how strongly the channel  $c$  at time  $t$  on trial  $r$  is synchronized to all other  $N_c - 1$  channels. SL is bounded by  $p_{\text{ref}} \geq SL_r^c[t] \leq 1$ . If channel  $c$  is uncorrelated with the other channels, then  $SL_r^c[t] = p_{\text{ref}}$ ; whereas if it achieves maximal synchronization with all channels, then  $SL_r^c[t] = 1$  [Gonzalez-Astudillo et al., 2021]. Note that to compute the pair-wise connectivity, the SL analysis should be conducted for each pair of channels independently.

### 4.2.5 Reproducing Kernel Hilbert Spaces for Machine Learning

The central idea behind Reproducing Kernel Hilbert Spaces (RKHS) is to use the kernel function to map the input data into a high-dimensional feature space where learning problems become more tractable. In the context of EEG signal classification for motor imagery tasks, the raw EEG data often resides in a high-dimensional and complex space. Traditional linear models might struggle to capture the intricate relationships hidden within accurately. By mapping the raw data into an RKHS via a suitable kernel function, we can transform the complex problem into a simpler one, enabling more effective learning and prediction. Moreover, many kernel methods project data into RKHS implicitly, avoiding the actual high-dimensional mapping and the accompanying computational burden (referred to as the "kernel trick").

#### Reproducing Kernel Hilbert Spaces

Let  $\kappa(\cdot, \cdot)$  be a real-valued positive definite kernel,  $\mathcal{H}$  a nonempty space, and  $\phi(\cdot)$  be a function that maps any element in  $\mathcal{X}$  to the  $\mathcal{H}$  space. One can define a vector space by taking a linear combination of the form

$$\phi(\cdot) = \sum_{i=1}^N \alpha_i \kappa(\cdot, x_i) \quad (4-58)$$

where  $N \in \mathbb{N}$ ,  $\alpha_i \in \mathbb{R}$ , and  $\{x_i \in \mathcal{X}; \forall i\}$ . Thus, the dot product between  $\phi$  and another function  $\phi'$ , defined as  $\phi'(\cdot) = \sum_{j=1}^{N'} \beta_j \kappa(\cdot, x_j)$ , can be defined as follows:

$$\begin{aligned} \langle \phi, \phi' \rangle &= \sum_{i=1}^N \sum_{j=1}^{N'} \alpha_i \beta_j \kappa(x_i, x_j) . \\ &= \sum_{j=1}^{N'} \beta_j \phi(x_j) . \\ &= \sum_{i=1}^N \alpha_i \phi'(x_i) . \end{aligned} \quad (4-59)$$

The dot product  $\langle \cdot, \cdot \rangle$  is bilinear and symmetric, as  $\langle \phi, \phi' \rangle = \langle \phi', \phi \rangle$ . Furthermore, it is positive definite as for any function  $\phi$  we have:

$$\langle \phi, \phi \rangle \sum_{i,j=1}^N \alpha_i \alpha_j \kappa(x_i, x_j) \geq 0. \quad (4-60)$$

Finally, the dot product is equivalent to the  $\kappa(\cdot, \cdot)$  function as  $\langle \kappa(\cdot, x), \phi \rangle = f(x)$  and  $\langle \kappa(\cdot, x), \kappa(\cdot, x') \rangle = \kappa(x, x')$ . By virtue of these properties, positive definite kernels  $\kappa$  are also called RKHS. In general, an RKHS can be defined as follows:

**Definition 4.2.1 (Reproducing Kernel Hilbert Space)** *Let  $\mathcal{X}$  be a nonempty set and  $\mathcal{H}$  a Hilbert space of functions  $\phi : \mathcal{X} \rightarrow \mathcal{H}$ . Then,  $\mathcal{H}$  is called a reproducing kernel Hilbert space endowed with the dot product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ , and the norm  $|\phi|_{\mathcal{H}} = \sqrt{\langle \phi, \phi \rangle_{\mathcal{H}}}$ , if there exists a bilinear function  $\kappa : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  with the following properties.*

- $\kappa$  has the reproducing property  $\langle \kappa(\cdot, x), \phi \rangle = f(x) \quad \forall f \in \mathcal{H}$ , in particular  $\langle \kappa(\cdot, x), \kappa(\cdot, x') \rangle = \kappa(x, x')$ .
- $\kappa$  spans  $\mathcal{H}$ , i.e.  $\mathcal{H} = \overline{\text{span}\{\kappa(x, \cdot) | x \in \mathcal{X}\}}$ , where  $\overline{\mathcal{X}}$  denotes the completion of the set  $\mathcal{X}$ .
- RKHS uniquely determines  $\kappa$ , since the symmetry property is not fulfilled as  $\langle \kappa(\cdot, x), \kappa'(\cdot, x') \rangle = \kappa(x, x')$  but  $\langle \kappa'(\cdot, x), \kappa(\cdot, x') \rangle = \kappa'(x, x')$ , then  $\kappa(x, x') \neq \kappa'(x, x')$

### Bochner's Theorem

Bochner's theorem relates a time stationary kernel  $\kappa_x(\tau_\Delta) = \kappa(x(\tau_1), x(\tau_2))$  with its cross-spectral counterpart  $S_x(\varpi)$  via the Fourier Transform as defined in Equation (4-61).

$$\kappa_x(\tau_\Delta) = \int_{\mathcal{F}} \exp(j2\pi\tau_\Delta\varpi) S_x(\varpi) d\varpi, \quad (4-61)$$

**Theorem 4.2.1** Let  $\kappa(x, y) = q(x - y)$  be a translation invariant kernel, real-valued and symmetric. Assume that  $q$  is continuous. Then the two following are equivalent:

(i)  $\kappa$  is positive definite.

(ii) There exists a positive and finite measure  $\mu$  on  $\mathbb{R}^p$  such that  $q$  is the Fourier Transform of  $\mu$ , i.e. for all  $x \in \mathbb{R}^p$

$$q(x) = \int_{\mathbb{R}^p} \exp(-iu^\top x) \mu(u) du \quad (4-62)$$

In the realm of EEG-based MI-BCI systems, Bochner's theorem is a powerful analytical tool. It allows for the transformation of kernels from the time domain to the frequency domain when scrutinizing EEG signals. By providing a profound understanding of the spectral and temporal properties of the underlying stochastic processes, this theorem plays a significant role in unraveling the intricate patterns in EEG data for efficient feature extraction and precise identification of cognitive processes.

#### 4.2.6 Information Theoretic Learning from Kernel Matrices

Information Theoretic Learning (ITL) is a learning framework that harnesses information theory for supervised and unsupervised learning algorithms. Unlike traditional models that use Shannon-based entropy, ITL utilizes Renyi's  $\alpha$ -order entropy, a more generalized concept [Li & Principe, 2020]. Shannon entropy is defined as the expected information of the outcomes of a random variable. For a continuous random variable  $X$ , and using the linear averaging operator, Shannon entropy can be represented as  $H(X) = \mathbb{E}\{I(X)\} = \int p(x)I(x)dx$ , where  $I(x) = -\log(p(x))$ . However, the linear mean is only a specific instance of the averaging operator. In general, the expected value connected with a monotonous function  $g(x)$ , with inverse  $g^{-1}(x)$ , is denoted as  $\mathbb{E}\{x\} = g^{-1}(\int p(x)g(x)dx)$ . Thanks to the additivity postulate for independent events, the possible options for  $g(x)$  fall into either  $g(x) = cx$  or  $g(x) = c2^{(1-\alpha)x}$ , as demonstrated in [Rényi, 1961]. The former leads to the linear mean and consequently the Shannon entropy, while the latter results in:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left( \int p(x)^\alpha dx \right) \quad (4-63)$$

This equation relates to Renyi's  $\alpha$  entropy where  $\alpha \neq 1$  and  $\alpha \geq 0$  [Principe, 2010, Rényi, 1961]. It includes Shannon's entropy definition when  $\alpha \rightarrow 1$ . In cases when one must estimate entropy from discrete data, the probability density function of a discrete random variable  $X$ ,  $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ , the Parzen density estimator can approximate as  $\hat{p}(x) = \frac{1}{n} \sum_{i=1}^n \kappa(x, x_i)$ , where  $\kappa(\cdot, \cdot) \in \mathbb{R}$  is a positive definite kernel function. For the case of  $\alpha = 2$ , the Parzen approximation provides:

$$\hat{H}_2(X) = -\log \left( \frac{1}{n^2} \sum_{i,j=1}^n \kappa_{\sqrt{2}\sigma}(x_i, x_j) \right), \quad (4-64)$$

the information potential, the quantity inside the log function, correspond to the squared norm  $\frac{1}{n} \sum_{i=1}^n \phi(x_i)$ , which by the law of large numbers converges to  $|\mathbb{E}\{X\}|^2$ . The empirical estimation used in Equation (4-64) links both Hilbert space representations and it can be rewritten using the terms of a Gram matrix  $\mathbf{K}_{ij} = \kappa_{2\sigma}(x_i, x_j)$  as follows.

$$\hat{H}_2(X) = -\log \left( \frac{1}{n^2} \|\mathbf{K}\|_F^2 \right) + C(\sigma), \quad (4-65)$$

where  $C(\sigma)$  is the normalization factor of the Parzen window, the Frobenius norm of the Gram matrix  $\mathbf{K}$ , defined as  $\|\mathbf{K}\|_F^2 = \text{tr}(\mathbf{K}\mathbf{K})$ , and  $\text{tr}(\cdot)$  is the matrix trace. The cross-information potential seeks to quantify the information contribution of each sample involved in a Parzen approximation and is directly related to Renyi's entropy by a strictly monotonic function [Giraldo et al., 2014]. Moreover, Equation (4-65) is an entropy-like value possessing properties similar to Renyi's entropy without having to estimate probability distributions. Given a Gram matrix  $\mathbf{K}$  with elements  $k_{ij} = \kappa(x_i, x_j)$ , a kernel-based formulation of Renyi's  $\alpha$ -order entropy can be outlined as follows:

$$H_\alpha(\mathbf{K}) = \frac{1}{1-\alpha} \log (\text{tr}(\mathbf{K}^\alpha)), \quad (4-66)$$

which holds that  $\text{tr}(\mathbf{K}) = 1$ . The power  $\alpha$  of  $\mathbf{K}$  can be obtained using the spectral theorem [Giraldo et al., 2014].

#### 4.2.7 EEG Deep Learning Techniques

DL has brought remarkable developments in various fields, and EEG signal classification is no exception. It offers the advantage of automatic feature extraction, reducing the need for manual intervention, and often outperforms traditional machine learning classifiers. DL consists of multiple interconnected layers, that work together to predict a given input sample. They can be divided into discriminative, representative, and generative models.

##### Discriminative Models in Deep Learning

These models play a vital role in numerous machine-learning applications, largely due to their extraordinary effectiveness in predicting class labels given the input data. Generally, discriminative models concentrate on learning the conditional probability,  $P(y|x)$ . In other words, they aim to model the decision boundary between classes. We can categorize discriminative models based on their architecture in the following manner.

**Multilayer Perceptron** Traditional multilayer perceptrons (MLPs), also known as feed-forward neural networks, are the simplest type of artificial neural network. Comprising input, hidden, and output layers, their main goal is to transform the input data into meaningful outputs by adjusting the weights associated with the nodes via a learning algorithm [Amin et al., 2019]. Mathematically, the output of each neuron in the neural network can be represented as follows:

$$h_i = \sigma \left( \sum_{j=1}^{N_h} w_{ji} x_j + b_i \right), \quad (4-67)$$

where  $x_j$  represents the inputs,  $w_{ji}$  the weights,  $b_i$  the bias,  $h_i$  the  $i$ -th layer, and  $N_h$  is the number of layers. The function  $\sigma$  is a nonlinear activation function, which introduces non-linearity into the output of a neuron. This assists in learning complex non-linear patterns within the data. Specifically in the field of MI-BCI classification,  $h_0$  is the flattened representation of  $\mathbf{X}_r$ , while  $h_{N_h}$  is the final layer representing  $\mathbf{z}_r$ .

The output  $\mathbf{z}_r$ , serves as the input for a classification layer that generates predictions. For binary classification, a function such as a sigmoid could be utilized to constrain the output between 0 and 1. In situations involving multi-class classification problems, a softmax function is generally employed.

**Convolutional Neural Networks** A Convolutional Neural Network (CNN) is a specialized architecture that is particularly useful for handling grid-like data, such as a time series or an image. It often consists of convolutional, pooling, and fully connected layers. Its core concept is to apply a set of learnable filters or kernels to the input samples [Zhang et al., 2021c]. For instance, a filter  $\mathbf{W}_f$  can be defined for a raw EEG input data  $\mathbf{X}_r$  at trial  $r$  to get its filtered version  $\hat{\mathbf{X}}_r$  as follows:

$$\hat{\mathbf{X}}_{r,f} = \varphi(\mathbf{X}_r; \mathbf{W}_f), \quad (4-68)$$

where  $\varphi(\cdot, \mathbf{W}_f)$  is a convolutional layer and  $\mathbf{W}_f$  are trainable weights. In addition, if  $\mathbf{W}_f$  is set, the CNN is called pooling layers and can perform down-sampling operations to provide an abstracted form of representation and to reduce the computational cost for subsequent layers.

CNNs can learn to extract hierarchically organized features, where initial layers capture local and generic features (like edges in an image or short-term changes in an EEG signal), and deeper layers capture more global and task-specific features. Through backpropagation and gradient descent, a CNN is trained to optimize the weights of the kernels in the convolutional layers and the weights in the fully connected layers to improve classification performance. In the context of EEG signal classification, convolutional layers can be used to automatically and adaptively extract local and shift-invariant features, and the fully connected layers can learn to differentiate between different classes of motor imagery tasks.

**Recurrent Neural Networks** Recurrent Neural Networks (RNNs) are designed to use sequential information by performing the same task for every element in a sequence, with the output depending on the previous computations. Each neuron or unit in an RNN takes the current and previously received inputs into account to generate the output. Mathematically, the hidden state  $h_t$  at time  $t$  can be calculated as:

$$h_t = \sigma(\mathbf{w}_h h_{t-1} + \mathbf{w}_x \mathbf{x}_{r,t} + \mathbf{b}_h) \quad (4-69)$$

The output at time  $t$  can be represented as:

$$y_t = \psi(\mathbf{w}_y h_t + \mathbf{b}_y) \quad (4-70)$$

Here,  $\mathbf{w}_h, \mathbf{w}_x, \mathbf{w}_y$  represent the weight matrices, while  $b_t$  denotes the bias terms.  $\sigma$  refers to the activation function, while  $\psi$  typically represents a softmax function utilized in multi-class classification problems. Due to the consideration of the temporal dynamic behavior of inputs by RNNs, they are extensively used in the processing of sequential data. Furthermore, RNN variants are discussed below.

**Gated Recurrent Unit (GRU)** is a variant of the standard RNN that attempts to solve the vanishing gradients problem, making it more efficient in handling longer sequences. It introduces two types of gates:

## 4.2. Theoretical Background

## 4. Materials and Methods

update gates and reset gates. The update gates determine the degree to which the previous state information is kept for the current state, while the reset gates control the amount of past information that is forgotten [Luo et al., 2018]. Mathematically, the GRU computations are implemented with the following equations:

$$u_t = \sigma(\mathbf{w}_{hu}\mathbf{h}_{t-1} + \mathbf{w}_{xu}\mathbf{x}_t + \mathbf{b}_u) \quad (4-71)$$

$$r_t = \sigma(\mathbf{w}_{hr}\mathbf{h}_{t-1} + \mathbf{w}_{xr}\mathbf{x}_t + \mathbf{b}_r) \quad (4-72)$$

$$\tilde{h}_t = \tanh(\mathbf{w}_h\mathbf{h}_{t-1} \odot r_t + \mathbf{w}_x\mathbf{x}_t + \mathbf{b}_h) \quad (4-73)$$

$$h_t = (1 - u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t \quad (4-74)$$

In these equations,  $u_t$  is the update gate,  $r_t$  is the reset gate, and  $\tilde{h}_t$  is the candidate activation.

Like the GRU, **Long Short-Term Memory (LSTM)** is designed to address the vanishing gradient problem encountered in standard RNNs [Kumar et al., 2019]. However, LSTMs have a more complex structure. An LSTM introduces a cell state and three gates: an input gate, a forget gate, and an output gate. This additional complexity allows for the modeling of longer context dependencies. The following equations mathematically represent the LSTM computations:

$$f_t = \sigma(\mathbf{w}_{hf}\mathbf{h}_{t-1} + \mathbf{w}_{xf}\mathbf{x}_t + \mathbf{b}_f) \quad (4-75)$$

$$i_t = \sigma(\mathbf{w}_{hi}\mathbf{h}_{t-1} + \mathbf{w}_{xi}\mathbf{x}_t + \mathbf{b}_i) \quad (4-76)$$

$$\tilde{C}_t = \tanh(\mathbf{w}_{hC}\mathbf{h}_{t-1} + \mathbf{w}_{xC}\mathbf{x}_t + \mathbf{b}_C) \quad (4-77)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4-78)$$

$$o_t = \sigma(\mathbf{w}_{ho}\mathbf{h}_{t-1} + \mathbf{w}_{xo}\mathbf{x}_t + \mathbf{b}_o) \quad (4-79)$$

$$h_t = o_t \odot \tanh(C_t) \quad (4-80)$$

In these equations,  $f_t$ ,  $i_t$ , and  $o_t$  correspond to the forget gate, input gate, and output gate, respectively, while  $C_t$  denotes the cell state and  $h_t$  is the hidden state.

**Graph Neural Networks** Graph Neural Networks (GNNs) represent an innovative shift in neural network architectures, emphasizing the relationships between data rather than treating each data point individually or with a set grid. GNNs are specifically designed for processing graphical data, providing an insightful approach to problems where data are represented as graphs, such as social networks, molecular chemistry, and brain connectivity networks [Ju & Guan, 2023]. The fundamental idea in GNNs is the propagation and transformation of features through the vertices of a graph, where the feature of a given vertex is iteratively updated by aggregating features from its neighbors, similar to CNNs. The aggregation function is a critical design choice and can vary significantly per the specific GNN type. The aggregated features are then combined with the original features through a combining function, and a non-linear transformation is typically applied.

Let us consider Graph Convolutional Networks (GCNs) as an example. In a GCN, the feature representation of each node,  $\mathbf{H}^{(l)} \in \mathbb{R}^{n \times d}$ , is updated in every layer  $l \in \{0, \dots, L\}$  by passing through the following propagation rule:

$$\mathbf{H}^{(l+1)} = \sigma \left( \tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2} \mathbf{H}^{(l)} \mathbf{W}^{(l)} \right), \quad (4-81)$$

Where  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_N$  is the adjacency matrix of the undirected graph  $G$  with added self-connections by including the identity matrix  $\mathbf{I}_N$ .  $\tilde{\mathbf{D}}$  is the diagonal node degree matrix of  $\tilde{\mathbf{A}}$ .  $\mathbf{W}^{(l)}$  is the weight matrix for the  $l$ -th neural network layer, and  $\sigma(\cdot)$  represents a non-linear activation function such as the ReLU. Thus, each node aggregates the features from its neighboring nodes defined by  $\tilde{\mathbf{A}}$ . This creates a form of information exchange among adjacent nodes. Then, the linear transformation  $\mathbf{W}^{(l)}$  is applied, following which it undergoes a non-linear activation function [Raeisi et al., 2022].

Other types of GNNs include Graph Attention Networks (GATs), which introduce an attention mechanism allowing different weights for different nodes in the neighborhood according to their relative importance [Demir et al., 2022]. While Graph Neural Networks are capable of managing complex graph-structured data, they can be computationally demanding due to the necessity of processing all nodes and their neighbors within the graph. However, ongoing advancements in optimization techniques are continually improving their efficiency [Gao et al., 2024].

## EEG-based Deep Learning MI Classification Models

**EEGnet:** is a compact CNN specifically designed for EEG-based BCIs proposed in [Lawhern et al., 2018]. It is versatile across various BCI paradigms, can be trained with limited data, and produces interpretable features. The network involves sequenced convolutional steps and utilizes depthwise separable convolutions to reduce the number of parameters that need training, thereby lowering the network's complexity. EEGNet effectively extracts temporal and spatial features and applies batch normalization before the Exponential Linear Unit (ELU) nonlinearity [Qiumei et al., 2019]. Dropout and average pooling techniques are used to prevent over-fitting and dimension reduction, respectively. Interestingly, the model eliminates using a dense layer prior to the softmax classification to reduce free parameters; the architecture can be seen in Table 4-1. This innovative approach ensures EEGNet's efficacy in capturing significant and diverse time-scale data and enhances the network's accuracy in EEG signal classification.

**Shallowconvnet:** is a convolutional neural network (ConvNet) architecture designed specifically for EEG decoding tasks [Schirrmeister et al., 2017]. It comprises two main convolutional layers where the first layer is intended to bandpass each EEG channel. Bandpass filtering is a critical process in EEG signal processing that serves to filter unwanted frequency components and retain frequencies of interest. The second layer of the ShallowConvNet operates similar to spatial filters that process the signals in the spatial domain. An average

**Table 4-1:** EEGnet arquitecture

<b>Layer</b>	<b>Filters</b>	<b>Kernel size</b>	<b>Stride</b>	<b>Extra parameters</b>
Conv2D	8	(1,32)	(1,1)	bias=False, padding=same momentum=0.99, epsilon=0.001
BatchNormalization	.	.	.	.
DepthwiseConv2D	16	( $N_c$ ,1)	(1,1)	bias=False, $D_m = 2$ , max norm = 1 momentum=0.99, epsilon=0.001
BatchNormalization	.	.	.	.
Elu	.	.	.	.
AveragePooling2D	.	(1,4)	(1,1)	.
Dropout	.	.	.	$D_p = 0.5$
SeparableConv2D	16	(1,16)	(1,1)	bias=False, padding=same momentum=0.99, epsilon=0.001
BatchNormalization	.	.	.	.
Elu	.	.	.	.
AveragePooling2D	.	(1,8)	(1,1)	.
Dropout	.	.	.	$D_p = 0.5$
Flatten	.	.	.	.
Dense	.	.	.	units= $N_y$
Softmax	.	.	.	.

pooling layer is applied in the ShallowConvNet to reduce spatial data and dimensionality, thereby helping curb overfitting. To further mimic the filter bank common spatial pattern (FBCSP) strategy [Ang *et al.*, 2008], a broadly accepted approach in building spatial filters for EEG signal classification, a fully connected layer is added to the ShallowConvNet. The activation function used in this model is both square and logarithmic in nature, which aims to introduce non-linearity to the model and help it learn complex patterns. The architecture can be seen in Table 4-2

**Table 4-2:** Shallowconvnet architecture

<b>Layer</b>	<b>Filters</b>	<b>Kernel size</b>	<b>Stride</b>	<b>Extra parameters</b>
Conv2D	40	(1,13)	(1,1)	max norm = 2
Conv2D	40	( $N_c$ ,1)	(1,1)	max norm = 2
BatchNormalization	.	.	.	momentum=0.99, epsilon=0.001
Square activation	.	.	.	.
AveragePooling2D	.	(1,35)	(1,7)	.
log activation	.	.	.	.
Dropout	.	.	.	$D_p = 0.5$
Flatten	.	.	.	.
Dense	.	.	.	units= $N_y$
Softmax	.	.	.	.

*Deepconvnet:* also introduced by [Schirrmeyer *et al.*, 2017], is built to glean deeper, more intricate features by implementing additional convolutional layers. This model goes beyond the two-layer structure of the ShallowConvNet, incorporating three more convolutional layers, thus creating a five-layer network structure. While this model's complexity possibly aids in unearthing subtle patterns or features not discernible with less sophisticated models, it also poses potential risks. Due to its complexity, DeepConvNet is more susceptible to overfitting, where it could end up learning not the underlying patterns for different neural states but rather, specific variability present in individual trials. As such, when utilizing DeepConvNet, particular care must be taken to monitor and prevent overfitting, which might involve techniques such as regularization or early stopping. The architecture can be seen in Table 4-3

**Table 4-3:** Deepconvnet architecture

<b>Layer</b>	<b>Filters</b>	<b>Kernel size</b>	<b>Stride</b>	<b>Extra parameters</b>
Conv2D	25	(1,5)	(1,1)	max norm = 2
Conv2D	25	( $N_c$ ,1)	(1,1)	max norm = 2
BatchNormalization	.	.	.	momentum=0.1, epsilon=1e-5
Elu	.	.	.	.
MaxPooling2D	.	(1,2)	(1,2)	.
Dropout	.	.	.	$D_p = 0.5$
Conv2D	50	(1,5)	(1,1)	max norm = 2
BatchNormalization	.	.	.	momentum=0.1, epsilon=1e-5
Elu	.	.	.	.
MaxPooling2D	.	(1,2)	(1,2)	.
Dropout	.	.	.	$D_p = 0.5$
Conv2D	100	(1,5)	(1,1)	max norm = 2
BatchNormalization	.	.	.	momentum=0.1, epsilon=1e-5
Elu	.	.	.	.
MaxPooling2D	.	(1,2)	(1,2)	.
Dropout	.	.	.	$D_p = 0.5$
Conv2D	200	(1,5)	(1,1)	max norm = 2
BatchNormalization	.	.	.	momentum=0.1, epsilon=1e-5
Elu	.	.	.	.
MaxPooling2D	.	(1,2)	(1,2)	.
Dropout	.	.	.	$D_p = 0.5$
Flatten	.	.	.	.
Dense	.	.	.	units= $N_y$
Softmax	.	.	.	.

*TCNet-Fusion:* is a more recent approach proposed by [Musallam *et al.*, 2021] that comprises three main parts. Like EEGnet, it includes a temporal component that allows for learning different bandpass frequencies and a depth-wise separable convolution to extract spatial features for each temporal filter. A Temporal Convolutional Network (TCN) block is also applied to extract temporal patterns, which are then concatenated with the output of a separable convolution to alleviate feature loss. These outputs are then flattened and concatenated, followed by a separable convolution. A final dense layer with softmax activation is used to classify the concatenated features into MI classes. Similarly, adversarial learning has been successfully applied in many DL applications; these architectures rely on training a generative model that enforces invariance and generalization. Regardless, adversarial architectures are affected by overfitting, especially with subjects that do not have a similar pattern in each session [Özdenizci *et al.*, 2019]. The architecture can be seen in Table 4-4

#### 4.2.8 Explainable AI in MI-BCI

Explainable Artificial Intelligence (XAI) is a rapidly evolving field that aims to make the decision-making process of complex AI models understandable. In the context of MI-BCI, XAI can provide valuable insights into how predictions are made, thereby enhancing trustworthiness, causality, transferability, confidence, fairness, accessibility, and interactivity [Arrieta *et al.*, 2020]. XAI techniques can be categorized based on interpretation types, model specificity, explanation scopes, and explanation forms.

Regarding interpretation types, we have intrinsic and post-hoc techniques. Intrinsic techniques provide an explanation based on the structure or components of the model itself. For instance, in linear regression models, the trained weights can provide an intrinsic interpretation, indicating the contribution of each feature.

**Table 4-4:** TCNet-Fusion architecture

<b>Layer</b>	<b>Filters</b>	<b>Kernel size</b>	<b>Stride</b>	<b>Extra parameters</b>
Permute	.	.	.	(2,1,3)
Conv2D	24	(32,1)	(1,1)	bias=False, padding=same
BatchNormalization	.	.	.	momentum=0.99, epsilon=0.001
DepthwiseConv2D	16	(1, $N_c$ )	(1,1)	bias=False, $D_m = 2$ , max norm = 1
BatchNormalization	.	.	.	momentum=0.99, epsilon=0.001
Elu	.	.	.	.
AveragePooling2D	.	(8,1)	(1,1)	.
Dropout	.	.	.	$D_p = 0.5$
SeparableConv2D	48	(16,1)	(1,1)	bias=False, padding=same
BatchNormalization	.	.	.	momentum=0.99, epsilon=0.001
Elu	.	.	.	.
AveragePooling2D	.	(8,1)	(1,1)	.
Dropout-2	.	.	.	$D_p = 0.5$
TCN-Blocks	.	.	.	layers=2, kernel-s=4, filt=12 $D_p=0.3$ , N-residuals=2
Concatenate	.	.	.	[TCN-blocks, Dropout-2]
Flatten-1	.	.	.	[Dropout-2]
Flatten-2	.	.	.	[Concatenate]
Concatenate	.	.	.	[Flatten-1, Flatten-2]
Dense	.	.	.	units= $N_y$
Softmax	.	.	.	.

On the other hand, post-hoc techniques refer to external methods that analyze a trained model, such as Class Activation Maps (CAMs).

In terms of model specificity, we have model-specific and model-agnostic techniques. Model-specific techniques require a unique model structure or property. Therefore, all intrinsic models fall into this category as they rely on the model structure [Tjoa & Guan, 2020]. Model-agnostic techniques, on the other hand, do not have any special requirements for the model, making them widely used in XAI models.

Explanation scope can be either global or local. Local explanations focus on the variables that contribute to the decision for a single input, providing insights into the main characteristics associated with that input. Global explanations aim to give insights into the model itself, explaining the interactions of the model parameters and answering the question of how the model makes predictions.

Finally, explanation forms can be divided into feature-based and example-based explanations. Feature-based explanations use gradient or feature map values to approximate the input importance, highlighting the parts that impact the final decision the most. An overlay map over the original input sample commonly represents them. Example-based explanations, on the other hand, try to explain models by showing one or more examples similar to the given input, usually extracted at the training step.

## Class Activation Maps

Class Activation Maps (CAMs) are powerful visualization tools that highlight the input regions important for making predictions in Convolutional Neural Networks (CNNs). They provide a heatmap visualization that indicates the areas in the input image that the model focuses on when making a prediction. CAMs are based on the concept that the spatial information is encoded in the feature maps of convolutional layers in CNNs. By projecting the weights of the output layer back onto the convolutional feature maps, we can generate a coarse heatmap for each class in the output layer. This heatmap, or CAM, indicates the discriminative

image regions CNN uses to identify a specific class.

The CAM for a particular class  $c$  can be computed as follows:

$$L_{CAM} = \sum_k w_{k,c} F_k, \quad (4-82)$$

where  $F_k$  are the spatial feature maps of the last convolutional layer, and  $w_{k,c}$  are the weights corresponding to class  $c$  in the output layer. The CAM  $M_c$  is a weighted combination of the feature maps, where the weights are determined by the importance of the feature map for class  $c$ .

### Gradient Class Activation Maps

Gradient Class Activation Maps (Grad-CAMs) are an extension of CAMs that aim to provide visual explanations of decisions made by any Convolutional Neural Network (CNN) architecture without requiring any re-training or modification to the original model. Grad-CAMs use the gradient information flowing into the last convolutional layer of the CNN to understand each neuron for a decision of interest. These gradients, flowing backward from the specific class at the output, are global-average-pooled to obtain the neuron importance weights. These are then used to create a coarse localization map highlighting the important regions in the image for predicting the concept.

The Grad-CAM for a particular class  $c$  is computed as follows:

$$L_{Grad-CAM}^c = \text{ReLU} \left( \sum_k \alpha_k^c F_k \right), \quad (4-83)$$

where  $F_k$  are the spatial feature maps of the last convolutional layer, and  $\alpha_k^c$  are the neuron importance weights for class  $c$ , computed as the global average of the gradients of the class score with respect to the feature map. The  $\text{ReLU}(\cdot)$  function ensures that only the features that positively influence the class of interest are visualized.

### Layer Class Activation Maps

Layer Class Activation Maps (Layer CAM) is a recent advancement in the field of explainable AI, providing a method to visualize the regions of input that are important for making predictions in Convolutional Neural Networks (CNNs). Layer CAM extends the concept of CAMs and Grad-CAMs by offering a more precise localization of the important regions. Layer CAM leverages the gradient information of the last convolutional layer to generate a heatmap that highlights the discriminative regions in the input image. Unlike Grad-CAM, which uses global average pooling of gradients, Layer CAM directly multiplies the feature maps with the corresponding gradients, followed by a  $\text{ReLU}$  operation and a channel-wise summation to generate the final heatmap.

The Layer CAM for a particular class  $c$  is computed as follows:

$$L_{Layer-CAM}^c = \text{ReLU} \left( \sum_k F_k \cdot \frac{\partial y^c}{\partial F_k} \right), \quad (4-84)$$

where  $F_k$  are the spatial feature maps of the last convolutional layer, and  $\frac{\partial y^c}{\partial F_k}$  are the gradients of the class score concerning the feature map. The  $ReLU(\cdot)$  function ensures that only the features that positively influence the class of interest are visualized.

## 4.3 Computing Resources

The experiments were carried out on Python-based libraries, including the well-known Scikitlearn, TensorFlow, and Keras. We have used the free tier of Google Colaboratory and Kaggle cloud computing services with the following specifications.

- **Google Colaboratory**<sup>3</sup>: provides 2 CPU cores, approximately 12 Gb of RAM, and 12/16 Gb on GPU
- **Kaggle**<sup>4</sup>: provides 4 CPU cores, 16Gb RAM and 16Gb on GPU

In addition, the research group has as a regional ally the Centro de Bioinformática y Biología Computacional - BIOS, which provides a computer specialized in machine learning with a 32-core processor, 500Gb of RAM, and a 16Gb GPU.

---

<sup>3</sup><https://colab.research.google.com/>

<sup>4</sup><https://www.kaggle.com/>

# 5 Single-Trial Kernel-based Functional Connectivity for Enhanced Feature Extraction in EEG-based MI-BCI

Here, we propose a single-trial kernel-based functional connectivity measure as an EEG-based feature extraction method to deal with inter/intra-subject variability in motor-related tasks as depicted in Figure 5-1. To this end, from spatio-temporal-frequency patterns, we extract the functional connectivity between the EEG channels through a novel kernel-based cross-spectral distribution estimator. Namely, the Gaussian kernel is used to compute the pairwise kernel-based channel dependencies because of its universal approximating ability. Further, we optimize the spectral combination weights within a sparse-based  $\ell_2$ -norm as a feature selection framework matching the motor-related labels that perform a dimensionality reduction of the extracted Gaussian Functional Connectivity features. We also interpret the extracted functional connectivity patterns by clustering them into ensembles of individuals with common behavior. Although our proposal is a data-driven approach, as other well-known motor-related feature extraction methods, i.e., CSP [Rongrong et al., 2020], it does not require the full trial set to compute discriminative EEG channel dependencies. The latter can benefit the implementation of real-time brain-machine interfaces. We contrast the proposed Gaussian functional connectivity measure with the baseline CSP method, two commonly used single-trial FC measures (Cross-Correlation Coefficient and Phase Lag Value), and more elaborated feature extraction approaches in the state-of-the-art. The validation results that were accomplished in three EEG databases (two with MI and one with ME) show that the proposed connectivity measure allows for reaching very competitive classifier performance with less affected values by feature extraction parameters, like the tuning of sliding time window. Besides, the introduced connectivity measure does not demand a prior linear spatial filtering as a preprocessing procedure. Additionally, the interpretation of subject clusters using our functional connectivity patterns benefits in understanding the inter/intra-subject variability in motor-related tasks.

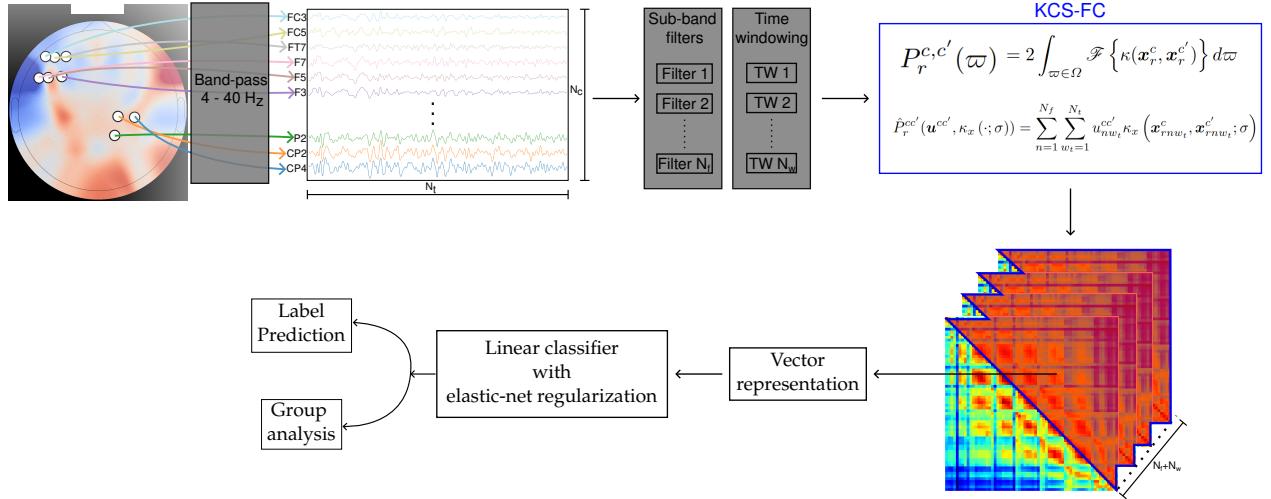
## 5.1 Kernel-based Covariance Function

The Wiener-Khinchin's theorem states that a real-valued auto-correlation function,  $R_r^c(\tau)$ , of a weak-sense stationary stochastic process  $x_r^c(\tau)$  can be defined as [Cohen, 1998]:

$$R_r^c(\tau) = \int_{\varpi \in \Omega} \exp(j2\pi\tau\varpi) dP_r^c(\varpi), \quad (5-1)$$

where  $P_r^c(\varpi) \in \mathbb{R}[0, 1]$  is a monotonic spectral distribution function absolutely continuous and differentiable over frequency  $\varpi \in \Omega$ . According to Bochner's theorem the univariable relationship in Equation (5-1) a stationary positive-definite kernel  $\kappa_r^{cc'}(\Delta_x) = \kappa(x_r^c, x_r^{c'})$  can be related with its cross-spectral counterpart

## 5.2. Gaussian Functional Connectivity from Kernel-based Spectral Distribution



**Figure 5-1:** Illustration of the proposed Kernel-based Cross-Spectral FC for enhancing the MI-BCI model’s accuracy. Key features include robust FC estimators using a Gaussian kernel for non-linearity, short sliding windows for addressing non-stationarity, sub-band filtering for managing inter-subject variability, vectorization representation for FC extraction and elastic-net regularization to curb spurious connectivities

$S_r^{cc'}(\varpi)$  if and only if the following assumption holds between both spectral representations [Bochner, 2020]:

$$\kappa_r^{cc'}(\Delta_x) = \int_{\varpi \in \Omega} \exp(j2\pi\Delta_x\varpi) S_r^{cc'}(\varpi) d\varpi, \quad (5-2)$$

where  $\Delta_x = \mathbf{x}_r^c - \mathbf{x}_r^{c'}$  is the vector delay,  $\varpi \subseteq \Omega$  is the frequency domain that contains the bandwidth set of analysis  $\Omega$ , and  $S_r^{cc'}(\varpi)$  is the cross-spectral density that preserves the following equality:  $S_r^{cc'}(\varpi) = dP_r^{cc'}(\varpi)/d\varpi$ , with  $P_r^{cc'}(\varpi) \in \mathbb{R}[0, 1]$  being the cross-spectral distribution that is related to the mapping kernel,  $\kappa : \mathbb{R}^{N_t} \times \mathbb{R}^{N_t} \rightarrow \mathbb{R}$ . As regards  $\kappa(\mathbf{x}_r^c, \mathbf{x}_r^{c'}) = \langle \phi(\mathbf{x}_r^c), \phi(\mathbf{x}_r^{c'}) \rangle_{\mathcal{H}}$ , it is a positive-definite stationary kernel inducing the nonlinear feature mapping  $\phi(\cdot)$  to a Reproducing Kernel Hilbert Space,  $\mathcal{H}$ . Notation  $\langle \cdot, \cdot \rangle$  represents the dot product. Therefore, we compute the cross-spectral distribution  $P_r^{cc'}(\varpi)$  within a bandwidth  $\Omega$ , as below:

$$\begin{aligned} P_r^{cc'}(\varpi) &= 2 \int_{\varpi \in \Omega} S_r^{cc'}(\varpi) d\varpi, \\ &= 2 \int_{\varpi \in \Omega} \mathcal{F}\{\kappa(\mathbf{x}_r^c, \mathbf{x}_r^{c'})\} d\varpi, \end{aligned} \quad (5-3)$$

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

As a result, Equation (5-3) preserves the frequency-based interpretation of the kernel-based pairwise dependencies that were estimated between vectors of random functions. Furthermore, the imposed stationary kernel favors the extraction of nonlinear data dependencies [Bishop, 2006]

## 5.2 Gaussian Functional Connectivity from Kernel-based Spectral Distribution

Let  $\mathbf{x}_{rnw_t}^c$  be an EEG channel  $c$  at trial  $r$  within filter  $n$  and time window  $w_t$ , and  $\{\mathbf{x}_{rnw_t}^c \in \mathbb{R}^{N_t}, y_r \in \mathbb{R}[0, 1]\}_{c,r,n,w_t=1}^{N_c, R, N_f, N_t}$  be an time segmented filtered EEG set holding all channels  $N_c$ ,  $R$  trials,  $N_f$  frequency

bandwidths,  $N_t$  time windows,  $y_r$  is the class or label assigned to the  $r$ -th trial, and the notation  $|\cdot|$  stands for set cardinality. Since the Gaussian kernel is preferred in pattern classification because of its universal approximating ability and mathematical tractability [Álvarez-Meza *et al.*, 2014a], we propose the Gaussian function to compute the pairwise kernel-based dependencies between channels  $c$  and  $c'$ , without the case  $c \neq c'$  in Equation (5-2), termed Gaussian Functional Connectivity (KCS-FC), defined as follows:

$$\kappa_x(\mathbf{x}_{rnw_t}^c, \mathbf{x}_{rnw_t}^{c'}; \sigma) = \exp\left(-\|\mathbf{x}_{rnw_t}^c - \mathbf{x}_{rnw_t}^{c'}\|_2^2 / 2\sigma^2\right), \quad (5-4)$$

where  $\sigma \in \mathbb{R}^+$  is the length scale hyperparameter. Notation  $\|\cdot\|_q$  stands for  $\ell_q$ -norm.

With the aim to implement the single-trial extraction, we estimate the cross-spectral distribution  $\hat{\mathbf{P}}_r^{cc'}$  between a pair of channels  $c$  and  $c'$  at trial  $r$  as follows:

$$\hat{\mathbf{P}}_r^{cc'}(\mathbf{u}^{cc'}, \kappa_x(\cdot; \sigma)) = \sum_{n=1}^{N_f} \sum_{w_t=1}^{N_t} u_{nw_t}^{cc'} \kappa_x(\mathbf{x}_{rnw_t}^c, \mathbf{x}_{rnw_t}^{c'}; \sigma), \quad (5-5)$$

where  $\mathbf{u}^{cc'} \in \mathbb{R}^{N_f N_t}$  is the spatio-temporal-frequency (textits-t-f) relevance vector that codes the pairwise undirected dependency between channels and contains the values  $u_{nw_t}^{cc'}$  holding the relevance value estimated at  $n$ -th frequency and  $w_t$  window.

Using the single-trial kernel-based spectral distribution representation in Equation (5-5), we propose extracting the sparse functional connectivity, using the Elastic Net regularization [Tay *et al.*, 2023], aiming to finding discriminative and interpretable brain activity patterns.

$$\mathbf{u}^* = \arg \min_{\mathbf{u}} \mathbb{E} \left\{ \sum_{r=1}^R \left\| \sum_{c,c'=1}^{N_c} \hat{\mathbf{P}}_r^{cc'}(\mathbf{u}^{cc'}, \kappa_x(\cdot; \sigma)) - y_r \right\|_2^2 \right\} + \alpha \sum_{c,c'=1}^{N_c} \|\mathbf{u}^{cc'}\|_1 + \frac{1-\alpha}{2} \sum_{c,c'=1}^{N_c} \|\mathbf{u}^{cc'}\|_2 \quad : \forall c < c', \quad (5-6)$$

where  $\alpha \in \mathbb{R}^+$  is the regularization hyperparameters,  $\|\cdot\|_q$  is the  $\ell_q$ -norm.

Without loss of generallity, equation (5-6) can be rewritten as:

$$\tilde{\mathbf{u}}^* = \arg \min_{\tilde{\mathbf{u}}} \mathbb{E} \left\{ \sum_{r=1}^R \|\tilde{\mathbf{u}} \tilde{\mathbf{p}}_r^\top - y_r\|_2^2 \right\} + \alpha \|\tilde{\mathbf{u}}\|_1 + \frac{1-\alpha}{2} \|\tilde{\mathbf{u}}\|_2, \quad (5-7)$$

where  $\top$  is the vector transpose,  $\tilde{\mathbf{u}}$  and  $\tilde{\mathbf{p}}$  stand for the spatio-temporal-frequency relevance matrix and cross-spectral distribution vector concatenation obtained as:

$$\begin{aligned} \tilde{\mathbf{u}} &= \left[ u_{11}^{12}, u_{11}^{13}, \dots, u_{11}^{1N_c}, u_{11}^{23}, \dots, u_{11}^{(N_c-1)N_c}, u_{12}^{12}, \dots, u_{N_f N_t}^{(N_c-1)N_c} \right], \\ \tilde{\mathbf{p}}_r &= \left[ \hat{P}_{r11}^{12}, \hat{P}_{r11}^{13}, \dots, \hat{P}_{r11}^{1N_c}, \hat{P}_{r11}^{23}, \dots, \hat{P}_{r11}^{(N_c-1)N_c}, \hat{P}_{r12}^{12}, \dots, \hat{P}_{rN_f N_t}^{(N_c-1)N_c} \right], \end{aligned} \quad (5-8)$$

Note that for simplicity the term  $\hat{\mathbf{P}}_r^{cc'}(\mathbf{U}_r^{cc'}, \kappa_x(\cdot; \sigma))$  is rewritten as  $\hat{\mathbf{P}}_{rnw_t}^{cc'}$ .

Consequently, the optimization framework in Equation (5-7) allows for computing the spectral combination weights from the extracted filter-bank representations that match the MI class-probability set, resulting in a sparse, relevant kernel-based functional connectivity between pairs of EEG channels.

### 5.3 Experimental Set-Up

The evaluation of the considered functional connectivity measures for enhanced feature extraction is performed according to the following stages:

(i) Preprocessing and trial-based extraction of s-t-f representations. For extracting the subject EEG dynamics over time accurately, the sliding window length of feature extraction is fixed to the next values:  $\tau = [0.5, 1.0, 1.5, 2.0]$  s, having an overlap of 75%. Additionally, similar to authors in [Ang *et al.*, 2008] the frequency bands of interest are selected with the lower frequency limit set at 4 Hz, the upper limit set to 40 Hz, and an overlap of 50%.

(ii) The estimation of the single-trial functional connectivity from the extracted s-t-f features. For comparison, the proposed KCS-FC is contrasted with two commonly used single-trial FC measures. To facilitate this comparison, the KCS-FC in Equation (5-6) is interchanged with new single-trial FC measures. These measures can be estimated from a frequency  $n$ , a time window  $w_t$ , and a pair of channels  $cc'$  (with  $c \neq c'$ , for all  $cc' \in C$ ) as described in [Rodrigues *et al.*, 2019a].

$$\rho(\mathbf{x}_{rnw_t}^c, \mathbf{x}_{rnw_t}^{c'}) = \langle \mathbf{x}_{rnw_t}^c, \mathbf{x}_{rnw_t}^{c'} \rangle \quad (5-9)$$

$$\Delta\phi(\mathbf{x}_{rnw_t}^c, \mathbf{x}_{rnw_t}^{c'}) = |\exp(j(\phi_{rnw_t}^c - \phi_{rnw_t}^{c'}))| \quad (5-10)$$

where  $\mathbf{x}_{rnw_t}^c$  and  $\mathbf{x}_{rnw_t}^{c'}$  is the real-valued EEG data from the corresponding electrodes at trial  $r$  (within temporal window  $w_t$  and frequency band  $n$ ) with instantaneous phases  $\phi_n^{ct}$  and  $\phi_n^{c't}$ , respectively. The pairwise relationships in Equations (5-9) and (5-10) are referred as Cross-Correlation Coefficient(CCF) and Phase Lag Value (PLV), respectively.

For feeding the classifier procedure, we evaluate two feature representation approaches that are extracted from each tested FC measure: Firstly, the feature representation is the CSP's pattern vector (linear projection of the sample covariance), computed as detailed in [Velasquez-Martinez *et al.*, 2020c]; in particular, the well-known CSP's eigendecomposition is applied by replacing the covariance matrix with the corresponding FC measure. Secondly, the concatenated (vectorized) triangular representation of the symmetrical FC matrices is extracted, as detailed in [Barachant *et al.*, 2013]. Both of the approaches comprise a filter-bank-based concatenation strategy through a fixed sliding window size and overlap values; that is, the temporal dependencies between windows are not directly modeled.

(iii) Sparse feature selection and classifier performance computation. Here, for implementation purposes, the optimizing problem is solved through the well-known Elastic-net algorithm to deal appropriately with redundant information. It is worth noting that we prefer Elastic-net over the Lasso-based regularization, because, when the number of features is greater than the number of training samples, Lasso behaves erratically [Géron, 2022].

For comparative purposes, we also perform feature extraction using the baseline CSP-based spatial filtering that is widely used for filtering measures of synchronization [Haddad *et al.*, 2019, Kumar *et al.*, 2018]. To this end, we perform the CSP feature extraction, adjusting the sliding time window length at each evaluated value of  $\tau$  and fixing the variance of the surrogate space to the first three eigenvectors of the spatial filtering matrix, as suggested in [Lotte *et al.*, 2018]. It is worth mentioning that we evaluate the FC measures in Equations (5-9) and (5-10), as well the CSP-based feature extraction scheme, through the sparse model presented in Equation (5-6) using a vector concatenation approach through temporal windows and frequency bands.

(iv) Clustering of subject inefficiency and interpretation analysis. For enhancing the interpretive analysis, we cluster the differences in neural responses that depend on the users' motor skills. The intra/inter-subject variability affects the FC estimators' robustness properties, becoming stronger for the worst-performing individuals.

Regarding the performance measure, the classifier accuracy  $a_c \in [0, 1]$  is computed, as follows:

$$a_c = (T_P + T_N)/(T_P + T_N + F_P + F_N),$$

where  $T_P$ ,  $T_N$ ,  $F_P$ , and  $F_N$  are true-positives, true-negatives, false-positives, and false-negatives, respectively. A stratified 10-fold cross-validation strategy randomly partitions the training and testing sets.

## 5.4 Results and Discussion

### 5.4.1 Influencing FC Parameters on Accuracy Estimation

#### Impact of Prior CSP Filtering

For interpretation purposes, Figure 5-2 presents the subjects' accuracy in decreasing order by all evaluated EEG data collections. Each data set is tested in two configurations to feed the sparse feature selector: using feature concatenation (odd rows) and linearly mapping through CSP (even rows). The FC measures' accuracy is estimated at each  $\tau$  (PLV—left column, CCF—middle column, KCS-FC—right column), adjusting the parameter  $\sigma$  to the median averaged over the input distances for kernel-based functional connectivity computation.

The top plots display the results that were performed by the databases with a similar amount of subjects: DMBI MI (first and second rows) and DB DBII ME (third and fourth rows). As seen, PLV yields the most scattered outcomes, with a significant portion accomplishing low accuracy. A similar situation holds for DBIII MI with a more visible subject variability: PLV discriminates the worst, followed by CCF, and KCS-FC performs a bit better, still outperforming the other FC when compared and having the accuracy values less spread throughout the subject set. However, the accuracy estimates are significantly more dispersed across the individuals than in DBII ME, regardless of the FC measure. Several factors may account for the higher variability of MI data. The EEG montage of MI protocols is more complex (64 MI channels vs. 44 ME channels). Besides, the ME signals are reported as having a higher degree of regularity than MI practicing, at least in terms of motor-related discriminability [Shamsi *et al.*, 2020a]. Consequently, even if the DBIII MI collection doubles the number of trials compared with DBII ME, the subject variability of MI data remains high enough to limit the CSP effectiveness (see the bottom row). Moreover, numerous subjects achieve accuracy values that are below 60%, which is, under the BCI-inefficiency level reported [Sannelli *et al.*, 2019].

#### Influence of Sliding Windows

Figure 5-3 summarizes **impact of the sliding window's length** on the classifier discrimination ability using each evaluated measure of FC. In the case of DBI MI, the first row shows that PLV and CCF both obtain very close performance if the CSP filtering is not involved at all tested window values. Otherwise, the latter measure notably improves in performance. Nevertheless, the proposed KCS-FC outperforms, and this behavior remains valid in the second row that presents the accuracy of DBII ME with more subject variability, as shown by the average bars. However, some fluctuations in the PLV and CCF performance are more evident after applying CSP, depending on the window. Instead, the KCS-FC measure presents performance nearly invariant across the range of considered sliding window lengths, being slightly enhanced by the CSP filtering algorithm.

**Additionally**, the bottom row presents the classifier performance that is obtained by MI activity (DBIII MI) with the most challenging subject variability, showing that the accuracy of all FC measures drops, particularly in the case of PLV. As seen, the PLV and CCF accuracy values are more sensitive to the sliding window influence, while the proposed KCS-FC assessments remain constant over the range of  $\tau$ . One more aspect to highlight is that the CSP-based filtering effectiveness becomes more dependent on the specific sliding window selected. Furthermore, under increased subject variability, this spatial filtering method's application noticeably decreases the KCS-FC performance.

### 5.4.2 Estimated Classifier Accuracy of Individuals

The following consideration is linked to the quality of individual FC assessments because of their diverse variability. For this purpose, Figure 5-4 illustrates the individual classifier performance at each window length, for which the subjects are displayed on the horizontal axis in decreasing order of the CCF accuracy achieved at  $\tau = 2$  s (the continuous line outlined in blue color). The rationale for choosing this specific FC estimation case as the baseline is that the cross-correlation coefficient that is presented in Equation (5-9) can be directly associated with the conventional Pearson estimate with the simplest interpretation regarding pairwise relationships [Georgiadis *et al.*, 2019].

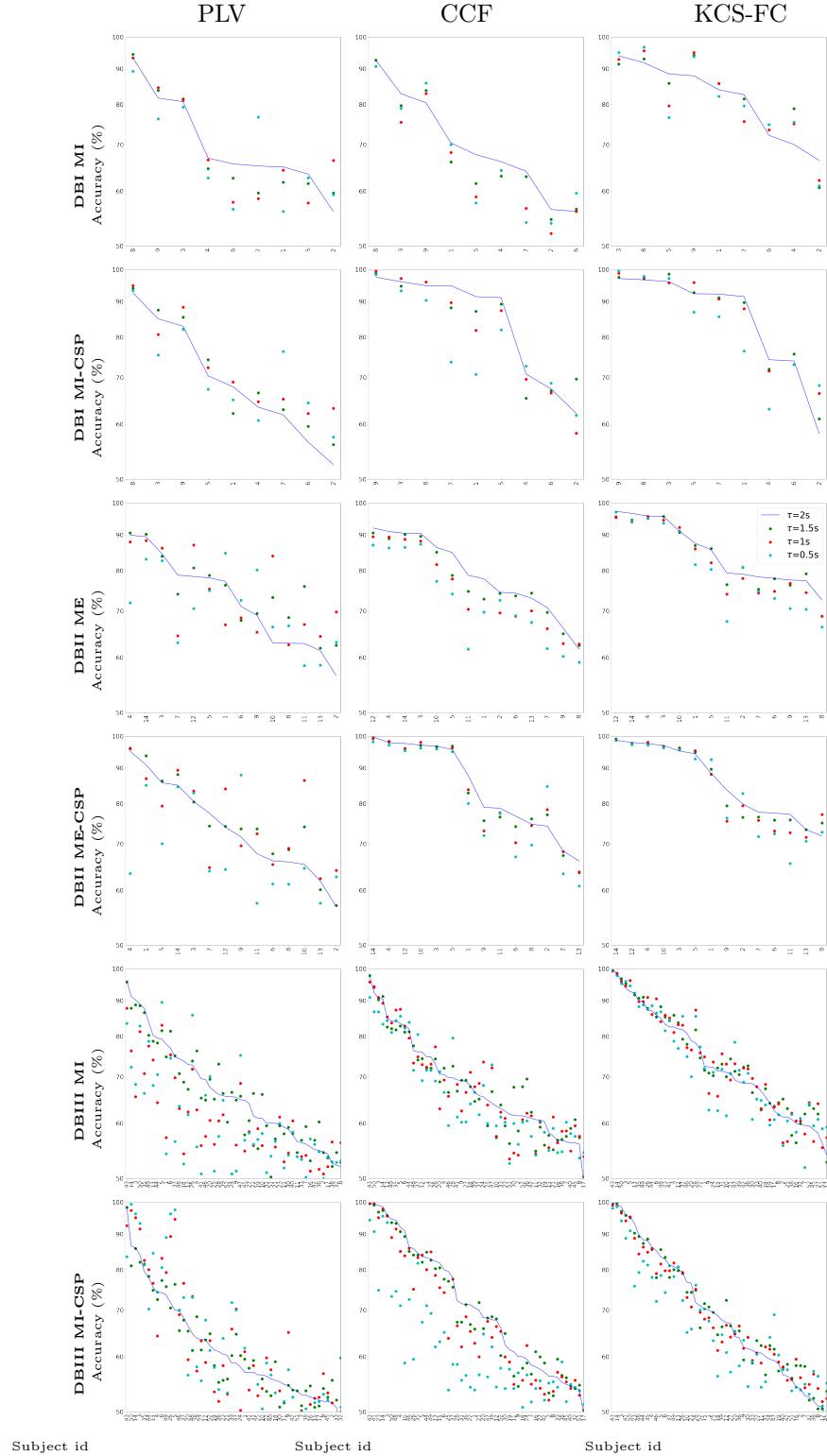
At first sight, besides being distinctive from the CCF and GCF, the variability of accuracy estimates that are achieved by PLV (green line) becomes visible across the subject set of DBII ME: the shorter the window length, the more variable the subject estimates. Rather, the CCF and KCS-FC measures produce accuracy values following a similar order of subjects, for which the enhancing effect of CSP-based filtering can be observed. In the case of DBIII MI, the variability in classifier performance that is provided by the FC metrics grows noticeably, regardless of applying the CSP filtering. Thus, the order of subjects provided by each FC measure is particular and dissimilar to each other. Nevertheless, when compared with the baseline CCF-based accuracy (blue line), the proposed GCF measure without CSP allows for enhancing almost every single subject's discrimination ability, regardless of the window length applied. Moreover, the kernel-based FC measure ensures that several subjects exceed the BCI-inefficiency level, as previously assessed using CCF. It is noteworthy that using the CSP algorithm appreciably reduces the KCS-FC effectiveness for dealing with the subject variability of DBIII MI.

### 5.4.3 Interpretation of Subject Clusters Using Functional Connectivity Patterns

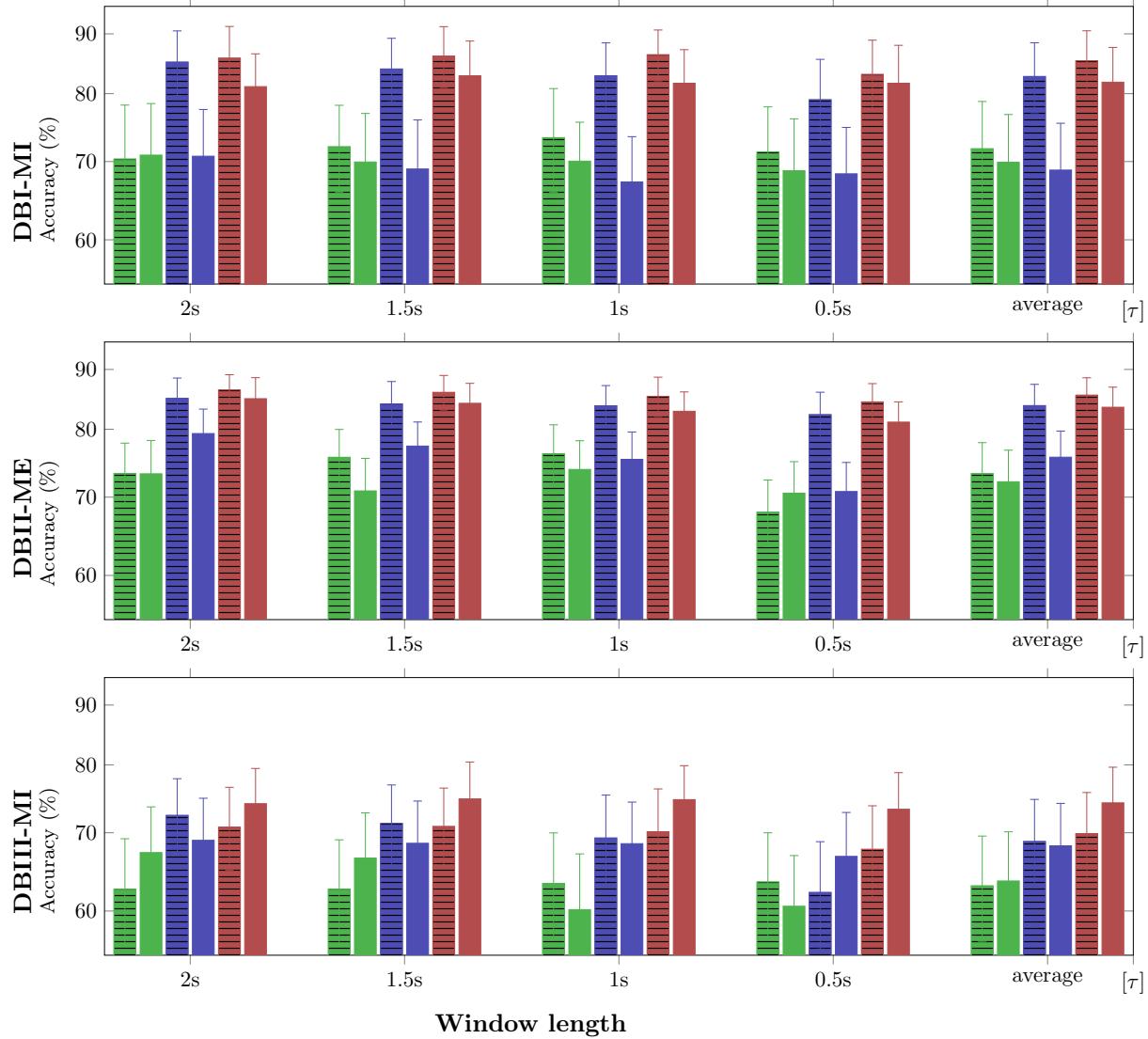
The main goal can be framed as the grouping of subjects having similar spatial networks involved in the motor-related tasks under consideration, aiming to properly interpret the evaluated functional connectivity measures. To this end, we cluster the extracted FC sets into partitions, each one with resembling subjects in terms of FC dynamics, being coded in the relevance vector  $v$  of Equation (5-6), and the motor-related accuracy. However, because of the poor clustering performance in high dimensional spaces, we accomplish a previous dimensional reduction stage on the normalized subject's relevance vector set (fixing a 2D low-dimensionality) through t-Distributed Stochastic Neighbor Embedding (t-SNE), which preserves the spatial relationships in the higher space (nearest-neighbors) [Linderman & Steinerberger, 2019]. Subsequently, we concatenate the t-SNE low-dimensional representation with the corresponding individual accuracy to perform a k-means-based clustering. At each window length, the latter approach is carried out, as proposed in [Kim *et al.*, 2019].

One concern is how a subject may switch between the assigned clusters when accounting for the influence of extracted FC measures at each window length. To clarify this aspect, Figure 5-5 displays the clustering results achieved by each functional connectivity measure using a matrix that illustrates the cells colored according to the individual group assessed, ranking the assessed groups in decreasing order of accuracy averaged over the corresponding subset. The three groups are colored, as follows: Group I, giving the best accuracy I (in yellow), Group II with regular accuracy (purple), and Group III, performing the worst accuracy (Turquoise).

As seen for DBII ME, the intragroup similarity remains comparable in groups I and III over the values of  $\tau$ , while Group II has noticeable volatile behavior. This result may be explained because of the low number of database subjects (only 14), making any exchange seriously affect the intergroup distribution. In DBIII MI, the high subject variability that is reflected by the functional connectivity extracted yields high changeability between groups, at least for the PLV and CCF measures. Rather, the proposed KCS-FC measure is less affected, as seen at window values of 2 and 1.5 s. Further, groups II and III become scrambled, meaning that the subject variability extracted at the shorter window lengths remains hard to overcome.



**Figure 5-2:** Subject accuracy at each window length  $\tau$  performed by the evaluated FC measures: PLV, CCF, and KCS-FC. Subjects are displayed on the horizontal axis in decreasing order of each FC accuracy at  $\tau = 2\text{s}$ . Notation CSP stands for the accuracy estimated after the spatial CSP filtering.

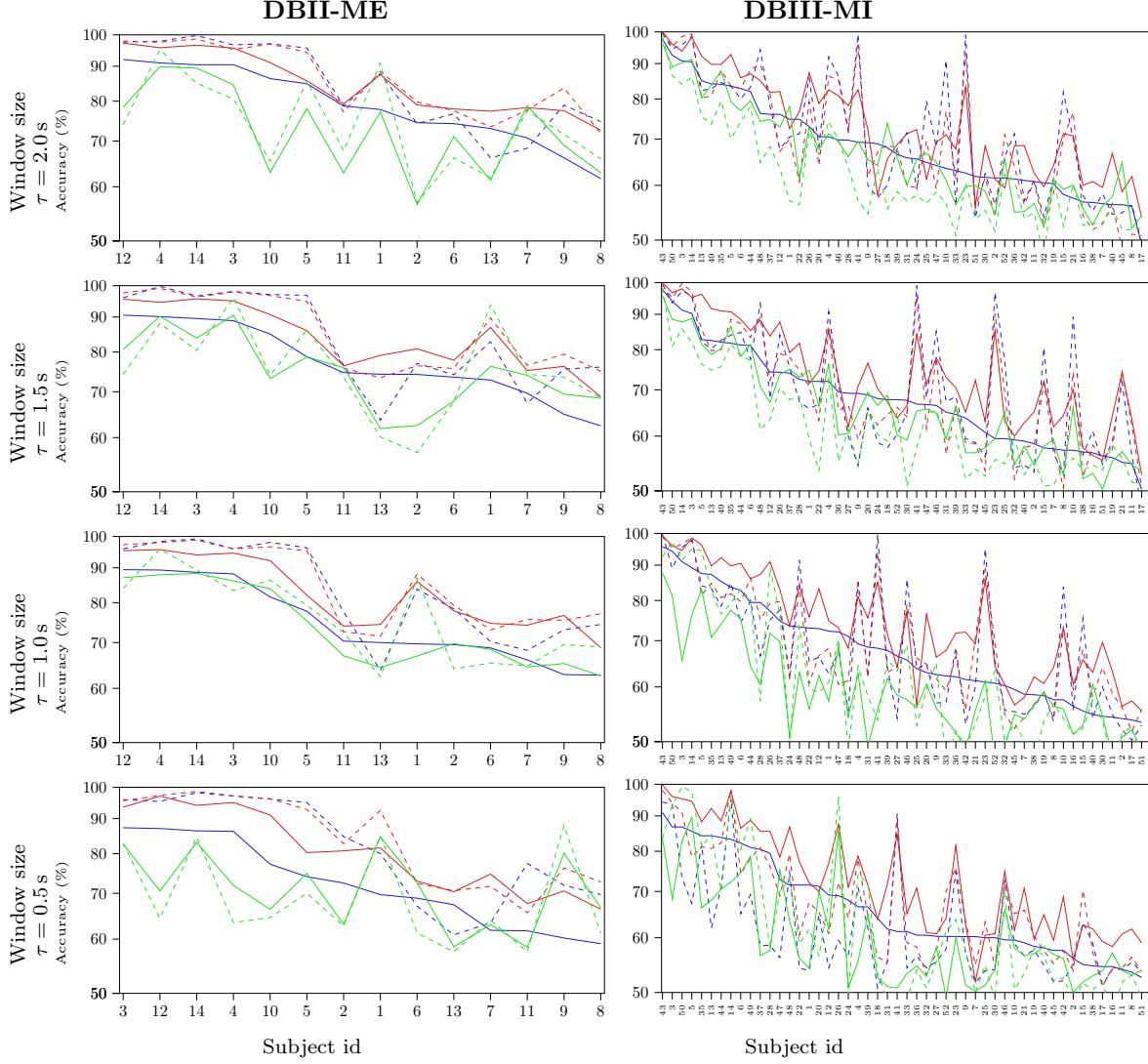


**Figure 5-3:** Accuracy performed at each  $\tau$  by the evaluated measures of FC: PLV (green), CCF (blue), and KCS-FC (red). The full-filled colored bars stand for concatenate representation, while the bars with horizontal lines present the results after CSP filtering.

The last aspect of the study is the interpretability of motor-related tasks as reliable sources of neural responses in the assessed clusters of dynamic behavior (the groups are denoted as I, II, and III).

Figure 5-6 presents the topographic channel plots (topoplots) that contain the relevant activated scalp areas that are inferred from the extracted FC measures, principally contributing to the discriminability between labeled motor-related tasks. We also include the most relevant pairwise functional connectivity links between electrodes that are estimated to fulfill the 99 percentile of the normalized relevance weights values (between 0 to 1) computed from Equation (5-6). The background stands for the accumulated relevance that is mapped to the channel positions. In DBII ME, only the SMR area is depicted according to the electrode set configuration above-described.

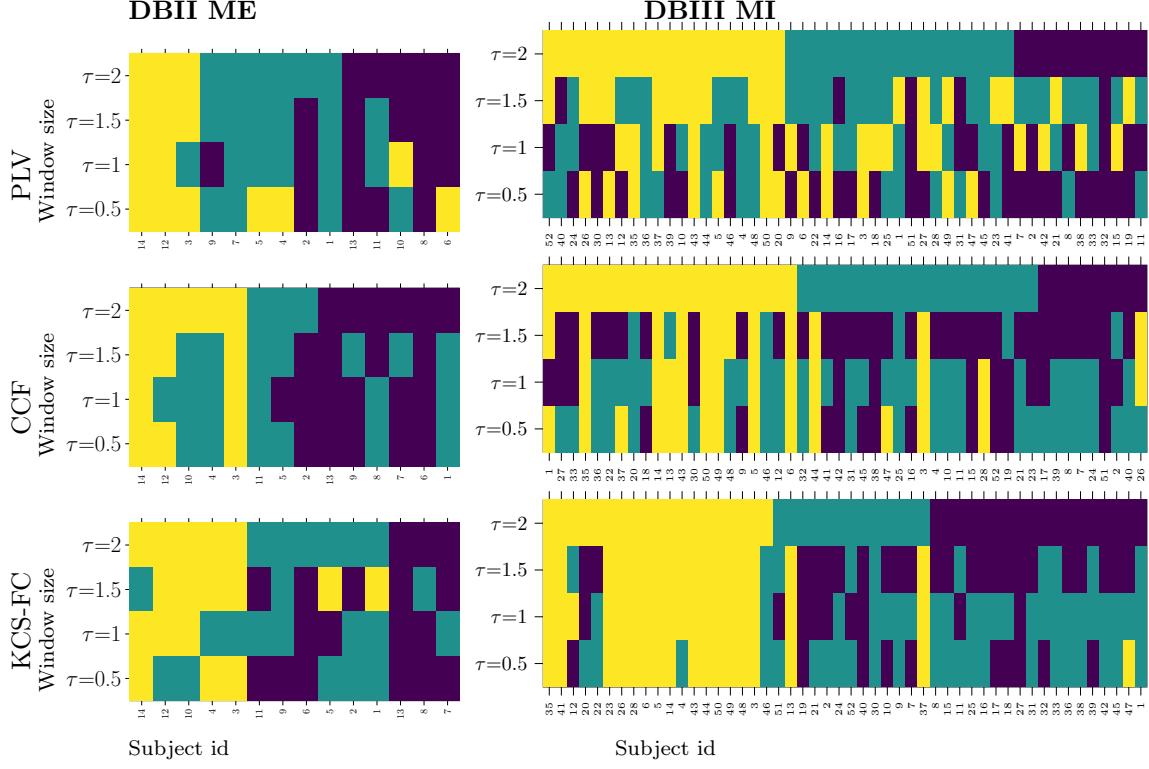
PLV performs in DBII ME the first three plots in the top row. The topoplots reflect high evoked response amplitudes that are densely spread over the SMR area, which means that the PLV variability hampers the sparse feature selection framework to obtain a reduced set of extracted FC features. Still, a few links



**Figure 5-4:** Individual classifier performance at different window lengths using FC: PLV (green), CCF (blue), and KCS-FC (red). The accuracy assessments performed without CSP are outlined with a continuous line, after CSP filtering—with a dashed line. The subjects are displayed on the horizontal axis in decreasing order of their CCF-based accuracy.

surpass the 99 percentile, which enables some information regarding the relevant electrodes. Contrary to this, the three topoplots that are delivered by CCF (middle row) and KCS-FC (bottom row) show low amplitudes of neural activity in Group I, slightly increased activity of Group II, and increased amplitudes of the worst-performing group of individuals (III). However, the latter FC measure clearly shows two foci that are symmetrically located over each hemisphere, as expected in motor-related tasks, at least for GI and, to a less extent, for GII. Besides, the amount of relevant links estimated is low. Consequently, the sparse-based  $\ell_2$ -norm framework enables the effective dimensionality reduction of the extracted KCS-FC features. Moreover, the accuracy using KCS-FC assessed by every group of individuals outperforms the other compared functional connectivity metrics: PLV and CCF.

In DBIII ME, the full electrode montage is employed and, thus, the estimated neural responses are all spread over the scalp, as seen in the three last topoplots of every row. Likewise, the number of relevant links increases. Nonetheless, the compared FC measures' performance has some similarities with DBII MI: PLV presents high background activity, CCF has amplitudes with fewer amplitudes, and GCF achieves a very



**Figure 5-5:** Clustering variability of individuals that belong to Group I (cells in yellow), Group II (Turquoise), and Group III (purple), depending on the extracting window length  $\tau$ .

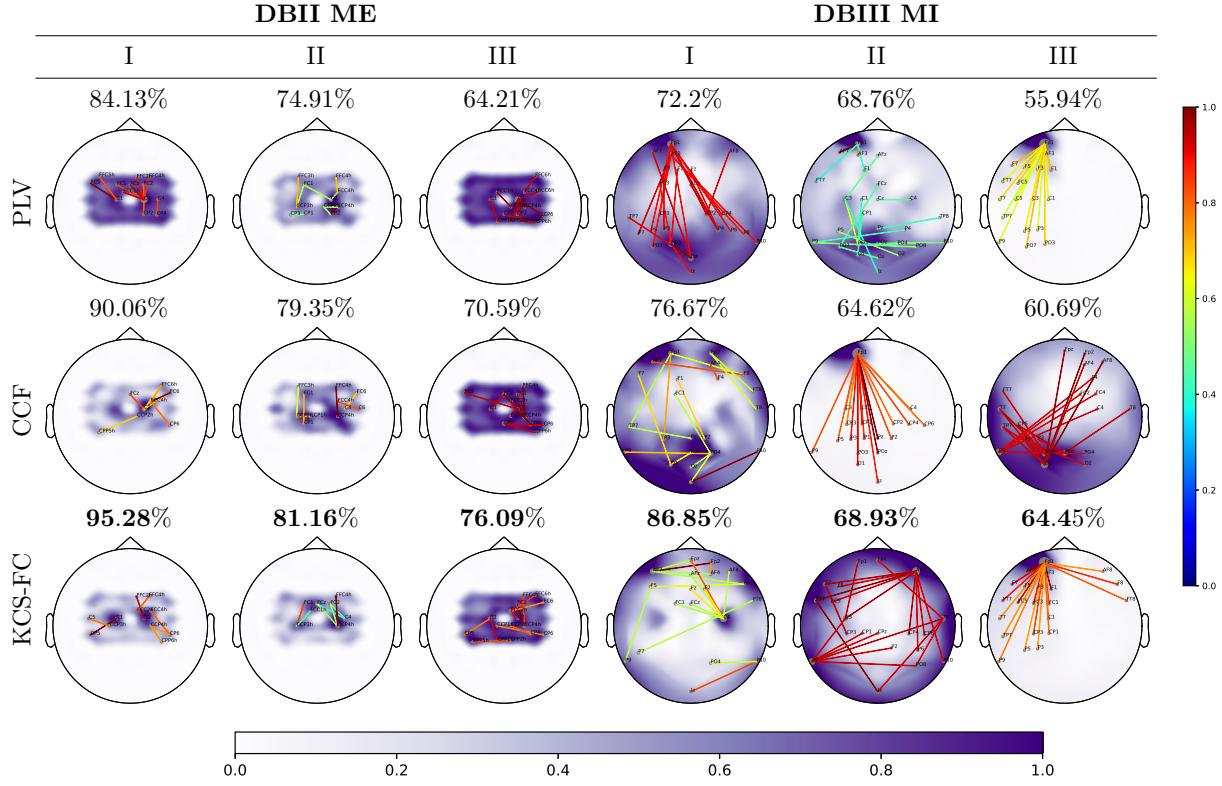
focalized activity over the SMR hemispheres, which play a critical role in MI tasks. Once again, the proposed KCS-FC measure outperforms the values of accuracy attained by other compared FC metrics. However, G III produces highly increased response amplitudes abnormally confined over the frontal zone (outside the SMR zone) and numerous links going to different electrodes. This issue may be explained because of acquisition artifacts, which may introduce noise and distortions, severely damaging the EEG data quality.

#### 5.4.4 Selecting the Sliding Time Window

The window length selected for extracting the EEG dynamics over time is a pivotal parameter. From the results obtained, both contrasted single-trial FC measures (Cross-Correlation Coefficient and Phase Lag Value) show several fluctuations over the evaluated values of  $\tau$  that diminish the performed accuracy, becoming worse in databases with increased subject variability, as is the case of DBIII MI. On the other hand, the proposed Gaussian Functional Connectivity measure enables accuracy estimates that are less affected by the sliding time window and, thus, KCS-FC allows for extracting the subject's dynamics more accurately within wider ranges of  $\tau$ .

#### 5.4.5 Prior Spatial Filtering Versus to Feed the Concatenation of FC Feature Sets

Because of the poor signal-to-noise ratio of scalp EEG measurements, the baseline CSP-based spatial filtering is very frequently accomplished. Meanwhile, the evaluating results indicate that the CSP effectiveness degrades noticeably as the inter/intrasubject variability increases. Thus, a big partition of subjects in



**Figure 5-6:** Spatial relevance mapped to topoplots estimated for each group of individuals from the FC measures (regarding the relevance weights  $v$  and the classification accuracy). The normalized relevance weights values (between 0 to 1) are computed from Equation (5-6). The background stands for the accumulated spatial relevance mapped to the EEG channel positions. The colored links represent the normalized FC relevance weights that hold a value higher than the 99 percentile of vector  $v$ . Above each plot, the average accuracy is shown over the corresponding group.

**Table 5-1:** Classifier accuracy comparison of approaches using functional connectivity features recently reported against the Gaussian FC performance in discriminating MI tasks. Notation TSGSP is temporally constrained sparse group spatial pattern, STR is space-time recurrence, and OPTICAL is Optimized CSP with long short term memory (LSTM). The best value performed for each database is marked in bold.

Data	Time Window	Filter Band	Interpretation	Feature Extraction	Accuracy (%)
DBI-MI	✓	✓	✓	TSGSP [Zhang et al., 2018c]	<b>82.50 ± 12.2</b>
	-	-	✓	STR connectivity [Rodrigues et al., 2019a]	69.56±15.02
	✓	-	✓	Renyi's $\alpha$ -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., 2019a]	62.00 ± 13.00
	✓	✓	✓	Proposed KCS-FC	<b>74.12 ± 12.13</b>

DBIII MI turns out to be below the BCI-inefficiency level. These findings are according to the reported CSP variability from trial-to-trial [Gaur et al., 2021a], and the subject-dependent choice of its extraction window [Velasquez-Martinez et al., 2020b].

One more restriction is that CSP is more oriented to power-based features [Kumar et al., 2018], so that its validity of the combination with PLV is questionable, as the achieved results definitely indicate. The combination of CSP with CCF and KCS-FC allows for enhancing the performed accuracy, but in DBI MI and DBII ME with a relatively moderate number of subjects. The testing of DBIII MI with the largest considered

intra/inter-subject variability tends to nullify CSP filtering before CCF feature extraction. Moreover, the proposed KCS-FC measures' performance improves if avoiding using the spatial filtering algorithm, which means that FC feature sets' straightforward concatenation is enough to feed the sparse-based  $\ell_2$ -norm feature selection framework. Table 5-1 displays the evaluated MI data's accuracy (i.e., DBI MI, and DBIII MI) using functional connectivity measures that have been reported recently, noting that the proposed KCS-FC approach provides very competitive classifier performance values.

## 5.5 Summary

We have introduced a novel method for single-trial kernel cross-spectral functional connectivity, known as KCS-FC, which addresses prevalent challenges inherent to EEG-based MI-BCI systems, identified in Section 1.2.1. These include issues related to non-stationarity, non-linearity, inter-subject variability, and spurious connectivities. To achieve this, we applied an extension of Bochner's theorem to approximate the cross-spectral distribution as a weighted sum of Gaussian-kernel-based pairwise relationships between channels, considering various time windows and frequency sub-bands. This weighted sum is managed by a linear regression model combined with an elastic net regularization that aligns the value with the MI label and differentiates spurious connectivities. In testing our method, we compared it with two prevalent single-trial FC estimators and respective CSP versions using two publicly available MI datasets. The results reveal that our KCS-FC strategy shows competitive accuracy in performance while maintaining interpretable FC matrices.

However, our approach heavily relies on handcrafted feature extraction processes, including time windowing and sub-band selection. This reliance on handcrafted features, as discussed in Section 1.2.2, does not significantly assist in artifact removal and only marginally increases accuracy for lower-performing individuals. Future research, as suggested in Section 1.3.2, should focus on adopting end-to-end DL models that can autonomously extract EEG representations. This would potentially be more effective in removing artifacts and could potentially enhance the performance of low-performing individuals.

# 6 KCS-FCnet: Kernel Cross-Spectral Functional Connectivity Network for Automatic EEG Representation in MI-BCI

Here we propose an end-to-end technique for classifying MI using EEG signals, termed Kernel Cross-Spectral Functional Connectivity Network (KCS-FCnet), as depicted in Figure 6-1. Our approach overcomes current DL limitations by introducing a cross-spectral Gaussian functional connectivity data-driven estimator to classify MI tasks from raw data. KCS-FCnet utilizes 1D convolutional layers to extract temporal-frequency features from input channels and a cross-spectral distribution estimation that codes relevant temporal and spectral MI patterns. It also includes a functional connectivity feature map, which improves the interpretability of the model by extracting meaningful pairwise channel relationships. Our approach is evaluated on a publicly available dataset and achieves state-of-the-art results for EEG-based MI classification. Furthermore, it demonstrated robustness to different experimental settings and individual differences. Lastly, our results suggest that the KCS-FCnet architecture is a highly effective method for EEG-based MI classification and can potentially be applied in real-world BCI.

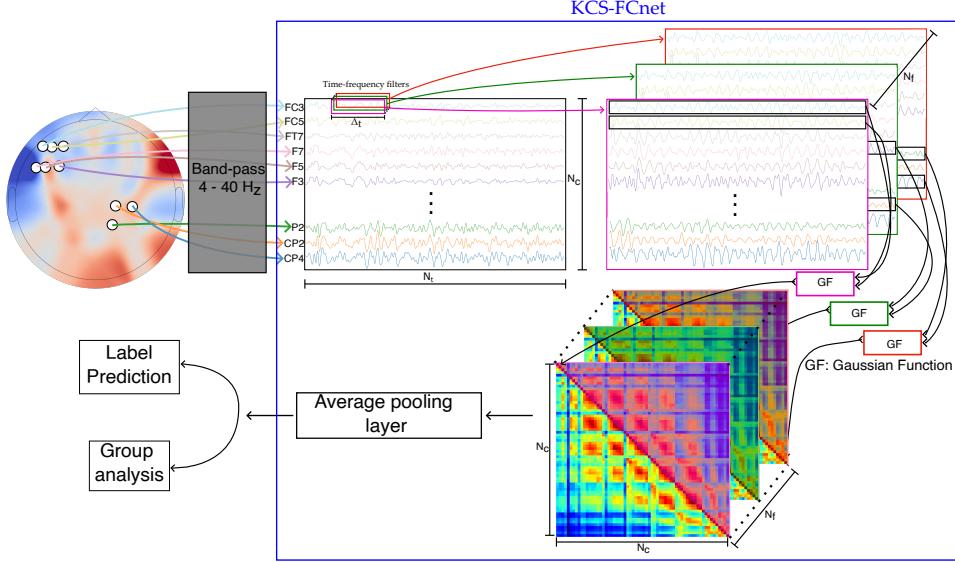
## 6.1 Kernel Cross-Spectral Functional Connectivity Network

The input-output EEG dataset,  $\{\mathbf{X}_r \in \mathbb{R}^{N_c \times N_t}, \mathbf{y}_r \in \{0, 1\}^{N_y}\}_{r=1}^R$ , comprises  $R$  trials,  $N_t$  time instants,  $N_c$  channels, and  $N_y$  classes. To enhance the most informative EEG spatial-temporal-spectral patterns from  $\mathbf{X}_r$  and reduce noise for improved MI class prediction, we propose to estimate the cross-spectral distribution among channels using a function composition. This approach gathers 1-D convolutional-based feature layers for extracting time-frequency patterns within each EEG channel and a Gaussian kernel-based pairwise similarity, as follows:

$$\hat{\mathbf{P}}_r(\mathbf{w}_f) = \tilde{K}(\cdot; \sigma) \circ \varphi(\mathbf{X}_r; \mathbf{w}_f), \quad (6-1)$$

where  $\hat{\mathbf{P}}_r(\mathbf{w}_f) \in [0, 1]^{N_c \times N_c \times N_f}$ ,  $N_f$  is the number of convolutional filters, notation  $\circ$  stands for function composition,  $\varphi(\cdot; \mathbf{w}_f)$  is a 1-D convolutional layer that can be used to automatically extract frequency patterns ruled by the weight vector  $\mathbf{w}_f \in \mathbb{R}^{\Delta_t}$ , with  $\Delta_t < N_t$ . Of note, in Equation (6-1) function  $\tilde{K}(\cdot; \sigma)$  is the convolutional filter concatenation of all pair-wise values  $\kappa_x(\mathbf{x}_{rf}^c, \mathbf{x}_{rf}^{c'}; \sigma)$  and is obtained as:

$$\tilde{K}(\tilde{\mathbf{X}}_r; \sigma) = [\mathbf{K}_{r1}, \mathbf{K}_{r2}, \dots, \mathbf{K}_{rf}, \dots, \mathbf{K}_{rN_f}], \quad (6-2)$$



**Figure 6-1:** Schematic diagram illustrating the proposed Kernel Cross-Spectral FC Network (KCS-FCnet), including the use of DL strategies and a CNN layer for artifact removal and spectral and temporal feature extraction. The diagram further outlines the process of spatial extraction via a custom layer incorporating the kernel cross-spectral function

where  $\mathbf{K}_{rf} \in \mathbb{R}^{N_c \times N_c}$  is the kernel matrix for a trial  $r$  at a convolutional filter  $f$  and it is calculated as follows:

$$\mathbf{K}_{rf} = \begin{bmatrix} \kappa_x(\mathbf{x}_{rf}^1, \mathbf{x}_{rf}^1; \sigma) & \kappa_x(\mathbf{x}_{rf}^1, \mathbf{x}_{rf}^2; \sigma) & \dots & \kappa_x(\mathbf{x}_{rf}^1, \mathbf{x}_{rf}^{N_c}; \sigma) \\ \kappa_x(\mathbf{x}_{rf}^2, \mathbf{x}_{rf}^1; \sigma) & \kappa_x(\mathbf{x}_{rf}^2, \mathbf{x}_{rf}^2; \sigma) & \dots & \kappa_x(\mathbf{x}_{rf}^2, \mathbf{x}_{rf}^{N_c}; \sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa_x(\mathbf{x}_{rf}^{N_c}, \mathbf{x}_{rf}^1; \sigma) & \kappa_x(\mathbf{x}_{rf}^{N_c}, \mathbf{x}_{rf}^2; \sigma) & \dots & \kappa_x(\mathbf{x}_{rf}^{N_c}, \mathbf{x}_{rf}^{N_c}; \sigma). \end{bmatrix} \quad (6-3)$$

We compute the average functional connectivity measure  $\tilde{\mathbf{P}}_r \in \mathbb{R}^{N_c \times N_c}$  over convolutional filters, as follows:

$$\tilde{\mathbf{P}}_r = \text{AvgPooling}_f (\hat{\mathbf{P}}_r(\mathbf{w}_f)), \quad (6-4)$$

where  $\mathbf{w}_f$  is the  $f$ -th convolutional filter,  $N_f$  is the number of convolutional filters. This measure provides a way to analyze how different frequency bands of a single-trial EEG relate to each other across channels. After computing the average functional connectivity measure and taking advantage of the symmetric property of the Gaussian functional connectivity, the vectorized version of  $\tilde{\mathbf{P}}_r$  is calculated as:

$$\bar{\mathbf{p}}_r = [\tilde{p}_r^{12}, \tilde{p}_r^{13}, \dots, \tilde{p}_r^{cc'}, \dots, \tilde{p}_r^{(N_c-1)N_c}]; \forall c < c', \quad (6-5)$$

where  $\bar{\mathbf{p}}_r \in \mathbb{R}^{N_c(N_c-1)/2}$ . Next, a softmax-based output layer is applied over vector  $\bar{\mathbf{p}}_r$  to obtain the MI class probability membership  $\hat{\mathbf{y}}_r \in [0, 1]^{N_y}$  as:

$$\hat{\mathbf{y}}_r = \text{softmax} (\mathbf{V}\bar{\mathbf{p}}_r + \mathbf{b}), \quad (6-6)$$

where  $\mathbf{V} \in \mathbb{R}^{N_c(N_c-1)/2 \times N_y}$ ,  $\mathbf{b} \in \mathbb{R}^{N_y}$ . In addition, a gradient descent-based framework using back-propagation is employed to optimize the parameter set  $\Theta = \{\mathbf{w}_f, \mathbf{V}, \mathbf{b}, \sigma; \forall f \in \{1, 2, \dots, N_f\}\}$ , as follows [Zhang et al., 2021a]:

$$\Theta^* = \arg \min_{\Theta} \mathbb{E}_r \{ \mathcal{L}(\mathbf{y}_r, \hat{\mathbf{y}}_r | \Theta); \forall r \in \{1, 2, \dots, R\} \}, \quad (6-7)$$

being  $\mathcal{L}\{\cdot\}$  a given loss function, i.e., cross-entropy. The optimization problem outlined in Equation (6-7) enables the training of our Kernel Cross-Spectral Functional Connectivity Network (KCS-FCnet) for the classification of MI tasks.

## 6.2 Experimental Set-Up

### 6.2.1 KCS-FCnet Implementation Details

In this study, we evaluate the efficacy of our proposed method for extracting subject-specific functional connectivity matrices from the KCS-FCnet that predicts MI output labels from EEG records. To accomplish this, we have developed a pipeline consisting of the following steps, which were tested on the Giga dataset (as detailed in Section 4.1):

- Raw EEG Preprocessing: First, we load subject recordings using a **custom database loader** module (<https://github.com/UN-GCPDS/python-gcpds.databases> (accessed on 27 January 2023)). Next, we downsample each signal from 512 Hz to 128 Hz using the Fourier method provided by the SciPy signal resample function (<https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.resample.html>). Then each time series trial was filtered between [4, 40] Hz, using a fifth-order Butterworth bandpass filter. In addition, we clipped the records from 0.5 s to 2.5 s post cue onset, retaining only information from the motor imagery task. Preprocessing step resembles the one provided by authors in [Lawhern *et al.*, 2018]. Note that since we are analyzing only the MI time segment, we assume the signal to be stationary. Our straightforward preprocessing aims to investigate five distinct brain rhythms within the 4 to 40 Hz range, including theta, alpha, and three beta waves. Theta waves (4–8 Hz), located in the hippocampus and various cortical structures, are believed to indicate an “online state” and are associated with sensorimotor and mnemonic functions, as stated by authors in [Abhang *et al.*, 2016]. In contrast, sensory stimulation and movements suppress alpha-band activity (8–13 Hz). It is modulated by attention, working memory, and mental tasks, potentially serving as a marker for higher motor control functions. Besides, tested preprocessing also comprises three types of beta waves: Low beta waves (12–15 Hz) or “beta one” waves, mainly associated with focused and introverted concentration. Second, mid-range beta waves (15–20 Hz), or “beta two” waves, are linked to increased energy, anxiety, and performance. Third, high beta waves (18–40 Hz), or “beta three” waves, are associated with significant stress, anxiety, paranoia, high energy, and high arousal.
- KCS-FCnet Training: We split trials within each subject data using the standard 5-fold 80–20 scheme. That means shuffling the data and taking 80% of it to train (training set), holding out the remaining 20% to validate trained models (testing set), and repeating the process five times [Schirrmeyer *et al.*, 2017]. For the sake of comparison, we calculate the accuracy, Cohen’s kappa, and the area under the ROC curve to compare performance between models [Warrens, 2015, Géron, 2022]. It is worth noting that we rescale the kernel length according to the new sampling frequency as in [Lawhern *et al.*, 2018]. The GridSearchCV class from SKlearn is used to find the best hyperparameter combination of our KCS-FCnet. The number of filters  $N_f$  is searched within the set {2, 3, 4}.
- Group-Level Analysis: We build a scoring matrix that contains as many rows as subjects in the dataset, 50 for Giga, and six columns, including accuracy, Cohen’s kappa, and the area under the ROC curve scores, along with their respective standard deviation. To keep the intuition of the higher, the better, and constrain all columns to be between [0, 1] in the score matrix, we replace the standard deviation with its complement and normalize the Cohen’s kappa by adding to it the unit and dividing by two. Then, using the score matrix and the k-means clustering algorithm [Géron, 2022], with  $k$  set as three, we trained a model to cluster subjects’ results based on the baseline model EEGnet [Lawhern *et al.*, 2018] in one of three groups: best, intermediate, and worst performing subjects. Next, we order each subject

based on a projected vector obtained from the first component of the well-known Principal Component Analysis (PCA) algorithm applied to the score matrix. Next, with the trained  $k$ -means, the subjects analyzed by our KCS-FCnet were clustered using the score matrix. The aim is to compare and check how subjects change between EEGnet and KCS-FCnet-based groups.

A KCS-FCnet sketch can be visualized in [Figure 6-1](#). The detailed KCS-FCnet architecture is summarized in [Table 6-1](#). All experiments were carried out in Python 3.8, with the Tensorflow 2.4.1 API, on Google Colaboratory and Kaggle environments. The fine-tuning process for the model's parameters begins by utilizing the training set for optimization. To evaluate the model's performance, the test set is employed solely for reporting scores. The categorical cross-entropy loss function is applied, and no additional callbacks are utilized. The training phase involves passing the entire batch of samples. Additionally, to support further analysis and experimentation, the model weights and performance scores are systematically saved for future reference.

**Table 6-1:** Detailed KCS-FCnet architecture for MI classification.

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

## 6.2.2 Functional Connectivity Pruning and Visualization

To compare functional connectivity between the groups mentioned above, first, we have to check which connections are relevant for class separability. It is worth noting that a high correlation in the functional connectivity matrix does not guarantee a higher class separability. Therefore, we use the two-sample Kolmogorov–Smirnov (2KS) test to overcome this issue and select only relevant connections as in [Gu *et al.*, 2020b]. The null hypothesis is that both samples are drawn from the same unknown distribution. Thus, we group the trials of each connection for a subject according to the label to build the samples “right” and “left”. Then, every pair is passed through the 2KS test, and connections holding a  $p$ -value equal to or lower than 0.05 are kept. Hence, we can state that both samples came from different distributions and the classes are distinguishable. Next, we build a  $p$ -value matrix containing the information on whether a connection is relevant. To visualize how each  $p$ -value matrix changes across subjects and groups, we plot each  $p$ -value matrix on a 2D visualization, where both dimensions are calculated using the well-known  $t$ -distributed Stochastic Neighbor Embedding ( $t$ -SNE) algorithm [Van der Maaten & Hinton, 2008], from the SKlearn library, over the EEGnet score matrix. It is noteworthy that the perplexity parameter has been specifically set to a value of ten, while all other parameters have been retained at their default settings.

Next, to effectively depict the connections between various regions of the brain, we employ a specialized connectivity visualizer (<https://github.com/UN-GCPDS/python-gcpds.visualizations>) which utilizes the Circos plot technique to display only the most significant connections, specifically those that fall within the 99-th percentile. To further enhance the analysis, we have chosen to plot the subject closest to the centroid of each group, thereby allowing for a detailed examination of one individual from each group.

### 6.2.3 Method Comparison

We compare the proposed KCS-FCnet with four end-to-end DL models that have been reported recently for effectively extracting relevantly explainable information from raw EEG. As with our proposal, the contrasted architectures are selected because they benefit from convolutional layers to extract temporal-frequency features for improving MI classification performance. Namely, (i) the EEGnet architecture in [Lawhern *et al.*, 2018] operates depthwise separable convolutions to reduce the number of training parameters, extracting temporal and spatial convolution features from each channel of a previous feature map; (ii) Shallowconvnet in [Schirrmeister *et al.*, 2017] incorporates two convolution layers (for sequential bandpass and spatial filtering of the EEG channels) followed by a square and log activation function, an average pooling layer, and a fully connected layer to emulate the baseline strategy of Filter Bank Common Spatial Patterns [Ang *et al.*, 2008]; (iii) Deepconvnet proposed by [Schirrmeister *et al.*, 2017] employs three convolutional layers to extract DL features; and (iv) TCNet-Fusion comprises three filtering stages to extract temporal, bandpass spectral, and spatial features, as explained in detail in [Musallam *et al.*, 2021].

For concrete testing, individual subject accuracy and standard deviation scores are only compared between the EEGnet and our KCS-FCnet due to their similarity in architecture and the number of parameters. For all provided approaches, the average classification performance along the 50 subjects in Giga is computed. Every architecture is implemented using TensorFlow2 and the SciKeras library, which allows wrapping any deep learning model as a SKlearn classifier. For the EEGnet, Shallowconvnet, Deepconvnet, and TCNet-Fusion, we use the hyperparameters that each work reported as the best combination. The complete codes for training, validating, and saving the model are publicly available (EEGnet<sup>1</sup>, Shallowconvnet<sup>2</sup>, Deepconvnet<sup>3</sup>, TCNet-Fusion<sup>4</sup>, and KCS-FCnet<sup>5</sup>).

## 6.3 Results and Discussion

### 6.3.1 Subject Dependent and Group Analysis Results

The proposed KCS-FCnet architecture is closely compared to the EEGnet architecture in this study, with a focus on subject-specific accuracy scores and their standard deviation. The comparison is illustrated in Figure **6-2**, where the dotted orange line represents the EEGnet and the dotted blue line represents the proposed KCS-FCnet. The blue and red bars indicate whether a specific subject's accuracy improves or decreases when using the KCS-FCnet, respectively. Additionally, the background of the figure includes low-opacity green, yellow, and red bars to indicate the group belongingness of the subjects (best, intermediate, and worst-performing clusters). The X-axis of the figure displays the subjects sorted based on their maximum score values as determined by the EEGnet results. The average accuracy for EEGnet and KCS-FCnet is 69.0

<sup>1</sup><https://www.kaggle.com/dggarciam94/eegnet-11-11-2022-version>

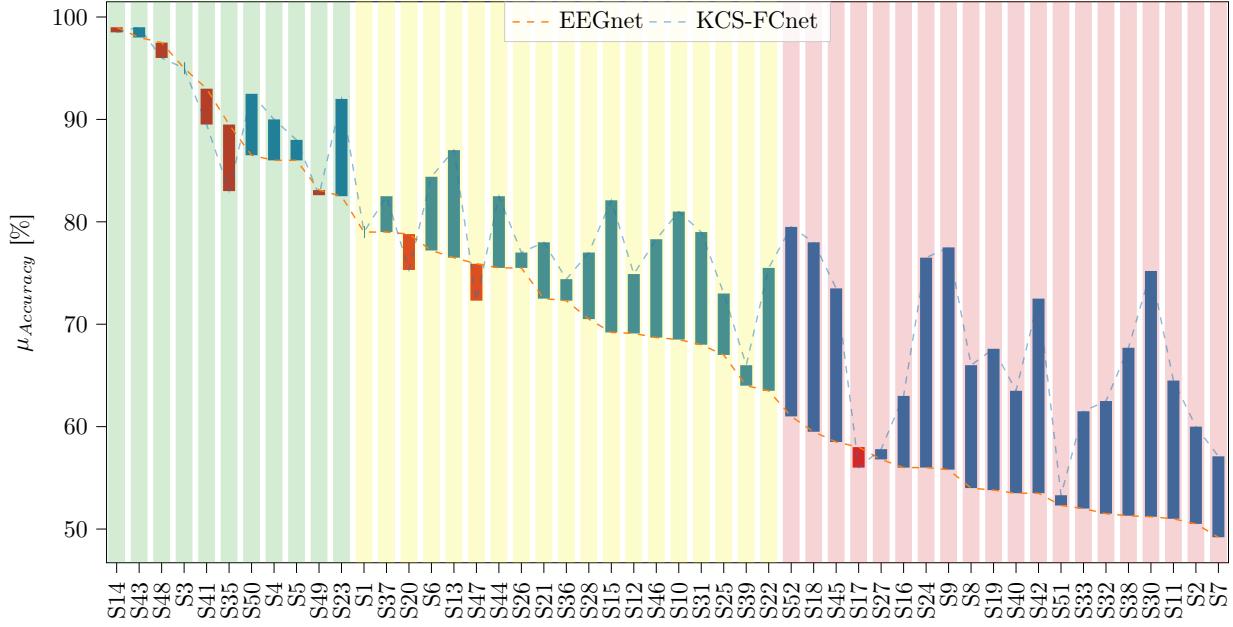
<sup>2</sup><https://www.kaggle.com/dggarciam94/shallownet-11-11-2022-version>

<sup>3</sup><https://www.kaggle.com/dggarciam94/deepconvnet-11-11-2022-version>

<sup>4</sup><https://www.kaggle.com/dggarciam94/tcnet-fusion-11-11-2022-version>

<sup>5</sup><https://www.kaggle.com/code/dggarciam94/gfcnet-11-11-2022-version>

and 76.4, respectively, resulting in an incremental of 7.4 for our proposal. Overall, it is demonstrated that KCS-FCnet can effectively classify motor imagery tasks using raw EEG as input data.

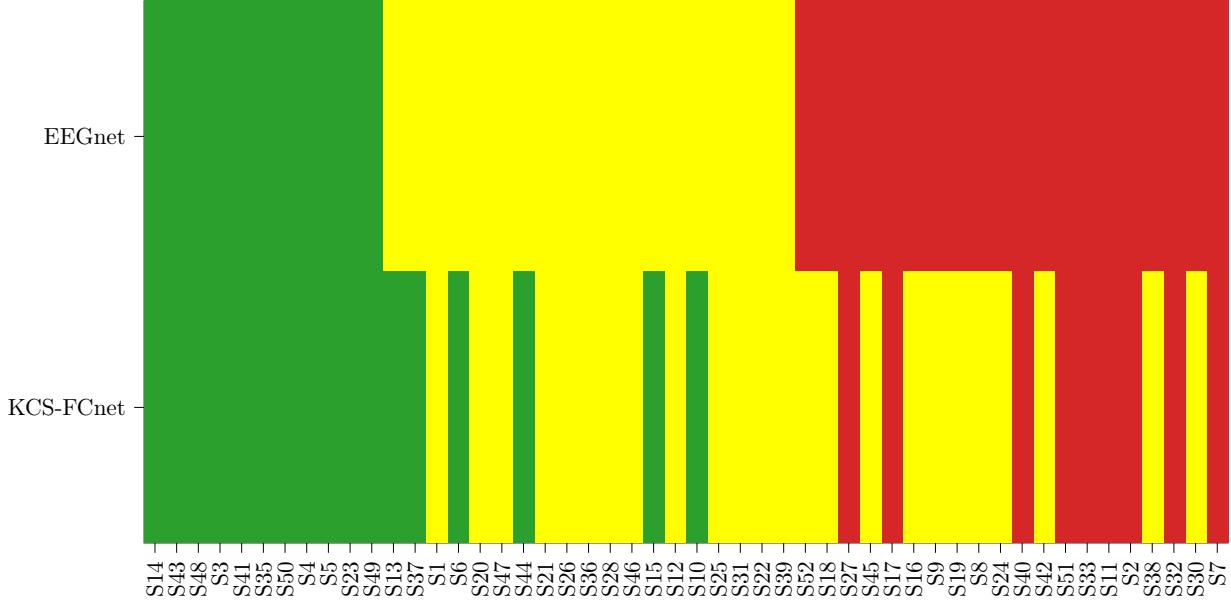


**Figure 6-2:** Subject specific results. EEGnet and KCS-FCnet average accuracies are depicted, with subjects being sorted based on their performance using EEGnet. The blue bars represent an improvement in performance using the KCS-FCnet, while the red bars indicate a decrease in performance. The background codes the group membership (best—G I, medium—GII, and worst—GIII performance clusters).

Moreover, our proposed method, KCS-FCnet, demonstrates mixed results in terms of accuracy for the subjects studied. On the one hand, seven subjects experienced a decrease in accuracy, of which only four experienced a reduction of three points or more. On the other hand, the remaining subjects experience an increase in accuracy, with the majority experiencing an increase of more than five points. Notably, our approach has a particularly strong impact on subjects in the third group, resulting in only one case where KCS-FCnet fails to surpass the baseline performance and two cases with less than one point of increase. Additionally, our data-driven functional connectivity method proves effective in extracting relevant feature maps for subjects in the second group, with over ten subjects experiencing an accuracy increment of at least three points. The first group, consisting of subjects with good performance, does not see remarkable results from KCS-FCnet, with only one subject experiencing a decrease in performance by five points and one subject experiencing an increase of more than five points. In general, subjects with the best performance appear to have a limitation when trying to include more relevant feature maps, yet our network is able to preserve their classification performance in most cases. In contrast, poor-performance subjects have more room for enhancement in the feature map, which is why we see a more significant increment in the third group.

Figure 6-3 illustrates the subject group belongingness and the impact of the KCS-FCnet method on group classification. The first row shows the subjects organized based on the EEGnet results, while the bottom row shows how each subject changes or maintains their group based on the KCS-FCnet results. For example, in the red group on the EEGnet row, subjects starting from S52, when we look at the new grouping based on KCS-FCnet for the same subset of subjects, it is evident that a total of eleven subjects significantly improved their performance, moving to the yellow cluster, while only nine remained in the red one. Additionally, six subjects had a major performance increase and were promoted to the best group (green), which demonstrates the effectiveness of the proposed framework. Furthermore, the subjects that were originally in the best group maintained their status, indicating that the best-performing subjects are less likely to improve. Then, our approach achieves better MI discrimination compared to EEGnet, particularly for bad and

medium-performing subjects, which is important as it highlights the model's capacity to handle challenging cases.



**Figure 6-3:** KCS-FCnet subject group enhancement regarding the EEGnet performance. Note that green, yellow and red represent best, medium, and worst performance regarding the average accuracy along subjects. First row: Subjects organized based on the EEGnet classification. Second row: Subject membership changes based on the KCS-FCnet results.

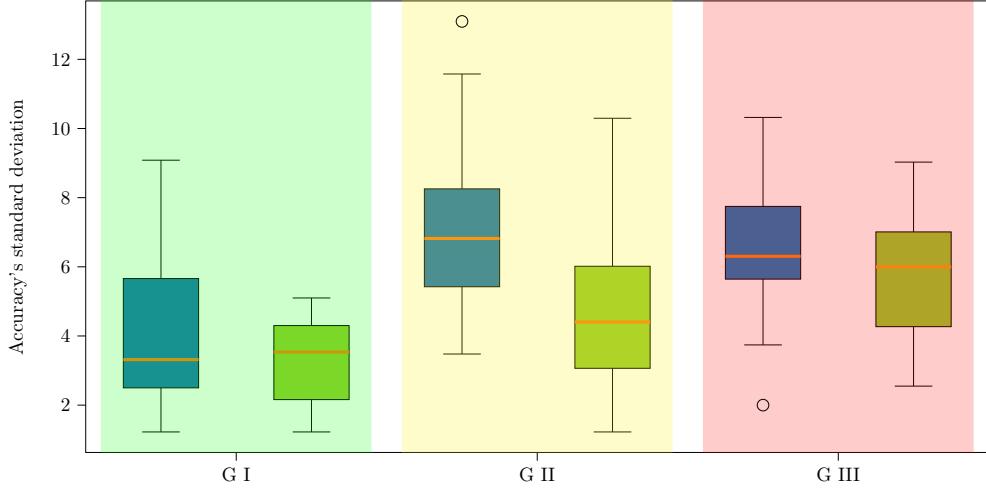
Next, Table 6-2 shows the accuracy results for each group for EEGnet and KCS-FCnet. It is important to note that while the difference in the first group is insignificant, with only 0.9 points, there is a notable improvement in the second group, with an increment of 5.6 accuracy points. Additionally, the third group shows a considerable increment of 12.4 points. A similar trend is observed in the standard deviation, where the second group has the most reduction of 2.6 points. Hence, our proposal not only outperforms EEGnet in terms of accuracy but also reduces the variability for all clusters.

**Table 6-2:** Group-based accuracy results for EEGnet and KCS-FCnet. The average accuracy for the best, medium, and worst-grouped subjects is depicted. The KCS-FCnet average increase for each cluster is also reported.

Approach	Group	Accuracy	KCS-FCnet Gain
EEGnet	G I	$90.6 \pm 4.3$	.
	G II	$72.2 \pm 7.3$	.
	G III	$54.3 \pm 6.6$	.
KCS-FCnet	G I	$91.5 \pm 3.3$	<b>0.9</b>
	G II	$77.8 \pm 4.7$	<b>5.6</b>
	G III	$66.7 \pm 5.6$	<b>12.4</b>

Figure 6-4 compares the accuracy standard deviation for EEGnet and KCS-FCnet. The background boxes indicate the group membership. For the first group, we can see an improvement in the variability scores for our proposed method, with a difference of four points between the maximum values. The second group shows a slight reduction in all standard deviation values for our method; however, the variability proportion remains almost the same. For the last group, there is a similar behavior for both methods. Overall, our proposed strategy reduces the variability and maintains a similar average accuracy score among subjects in the best group while increasing the average accuracy and maintaining the variability for the second and

third clusters.

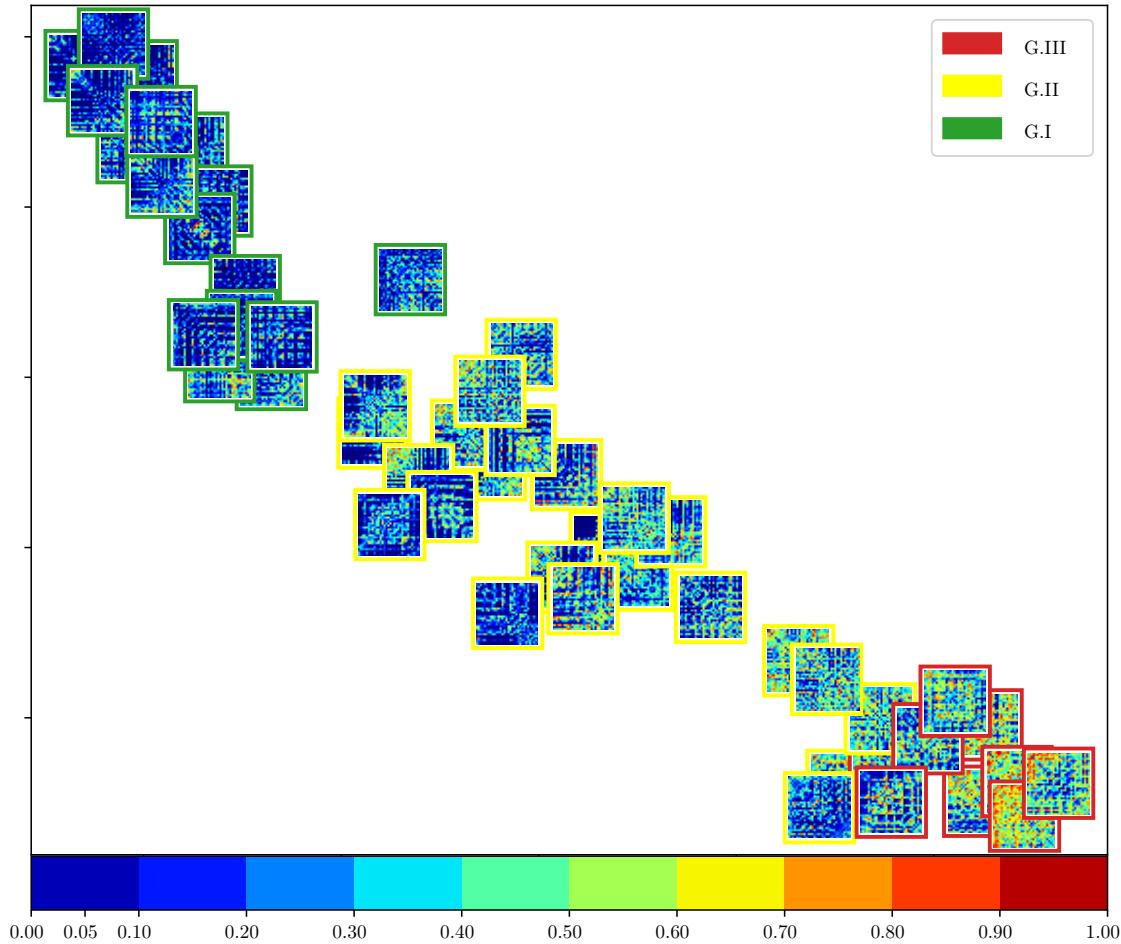


**Figure 6-4:** Group comparison between EEGnet (blue boxplots) and KCS-FCnet (green boxplots) concerning the accuracy's standard deviation. The background codes the group membership (best—GI, medium—GII, and worst—GIII performance clusters).

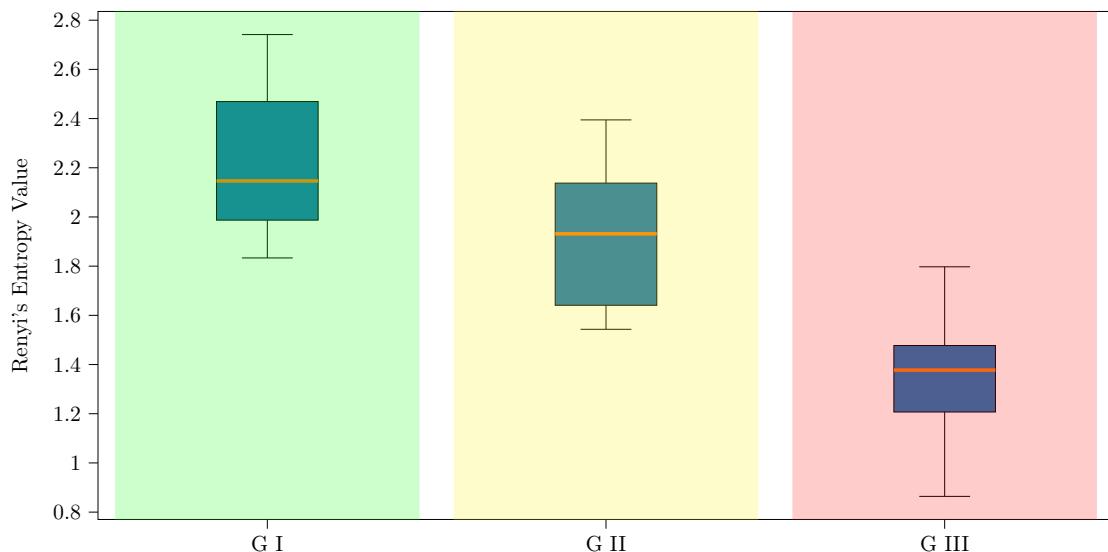
### 6.3.2 Estimated Functional Connectivity Results

In this study, we employed the two-sample Kolmogorov–Smirnov test to calculate a functional connectivity matrix for each subject. The matrix includes information about the separability of the MI classes. The null hypothesis asserts that the distribution of connection pairs for classes 0 and 1 is identical. A  $p$ -value is calculated, and we reject the null hypothesis only if the  $p$ -value is less than 5%. In other words, connection pairs with lower  $p$ -values indicate higher class separability and are considered more informative for the classification task. Figure 6-5 depicts the results of the test in the form of  $p$ -value-based matrices for each subject, which are plotted in a 2D projection using the t-SNE algorithm to reduce the dimensionality of the score matrix. Each matrix has a colored outer square that indicates the group membership. The matrices in the top left corner (first group) have the lowest  $p$ -values for every connection pair, indicating that almost every pair has a different class distribution, resulting in high accuracy scores, e.g., more than 90%. Conversely, matrices in the bottom right corner (third group) have the least significant  $p$ -values, indicating that only a few connection pairs can reject the null hypothesis; then, the class probability distribution for almost every pair cannot be distinguished. There is a gradual transition between matrices from the highest  $p$ -values in the bottom right corner to the lowest in the top left corner. Additionally, each group keeps an intra-subject  $p$ -value similarity for similar EEG channel connections.

Furthermore, Figure 6-6 details the amount of information preserved within each subject representative connectivity matrix. We utilize the widely used quadratic Rényi's entropy [Bromiley *et al.*, 2004] to quantify the interpretability performance from pruned functional connectivity matrices. Namely, a higher entropy value indicates a higher interpretability of that particular group of subjects concerning both relevant pairwise channel relationships and MI discrimination capability. The background boxes represent the group membership, and the box-and-whisker plots depict each cluster's distribution of Rényi's entropy values. The first group displays the most significant values, indicating that most connections discriminate highly between classes. In contrast, the third group has the lowest values, suggesting poor class discriminability. As expected, the groups that perform better show higher retention of information by the KSC-FCnet-based functional connectivity matrix.



**Figure 6-5:**  $t$ -SNE 2D projection of pruned functional connectivity matrices based on KSC-FCnet and two-sample Kolmogorov–Smirnov test. The color bar depicts the  $p$ -value of every connection for each subject matrix, where deep blue means more class separability. Therefore, the bluer the matrix, the better the discriminability. Outer boxes indicate subject group belongingness: green G I, yellow G II, and red G III.  $p$ -values below 5% are taken as significant.



**Figure 6-6:** Rényi's entropy-based retained information within the estimated functional connectivity matrices using KSC-FCnet. The background codes the group membership (best, medium, and worst performance clusters). Boxplot representation is used to present the retained information within each group.

Further, the Circos plot is a valuable tool to visualize which EEG channels are most important for each subject's experiment. Figure 6-7 shows that relevant channel dependencies are kept mainly for the best performance group. Note that all connections are normalized between the three subjects (connectivities above the 99th percentile are shown). For the G I case, the most robust connections are found between the frontal, central left, and right areas, with a few connections in the posterior region. This pattern is consistent with a good-performing subject who presents the most relevant information in the sensorimotor area (central left and right). Conversely, G II shows significant connections between the center-right and frontal areas, with fewer robust connections in the central left. It is worth noting that EEG noise may be present in the connectivity feature map around the central left region. Notably, G III has no significant connections, indicating that the model could not extract noise-free and discriminative connectivities.

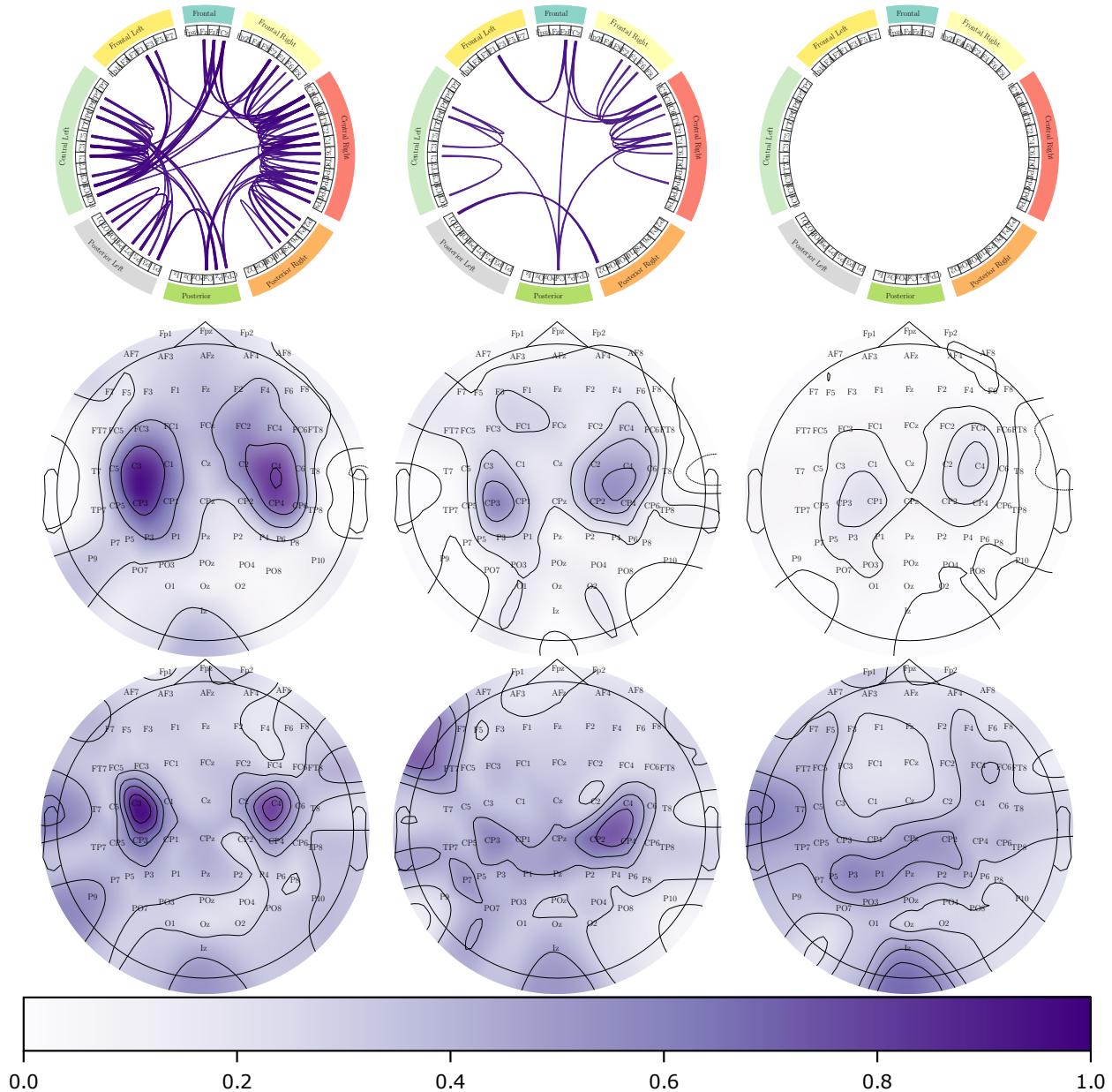
The second and third rows in Figure 6-7 depict the most significant brain areas using 2KS and the weights of the last layer in the KCS-FCnet. In G I, similar results are observed, highlighting the sensorimotor area; however, the results in the third row are more concentrated around C4 and C3. It suggests that subjects in the best-performing group do not exhibit much noise and the MI task can be completed using only a few sensors. For G II, there is a slight difference between the results. In particular, a high activation in the left frontal area is observed, while the information is more focused on the sensorimotor area. Finally, for the last group, the most significant difference is observed. While in the 2KS test, some importance (below 0.4) is observed around C4 and C3, in the weight-based approach, there is no clear pattern, indicating that our DL approach can not find relevant information in the sensorimotor area.

Table 6-3 summarizes the results of KCS-FCnet and the contrasted end-to-end architectures of Convolutional neural network models. Deepconvnet performs the poorest, making it unsuitable for handling high intra-class variability. By contrast, Shallowconvnet and TCNet-Fusion have values of quality measures that are very close to each other, being more competitive. Despite this, KCS-FCnet achieves the highest scores, outperforming the other models.

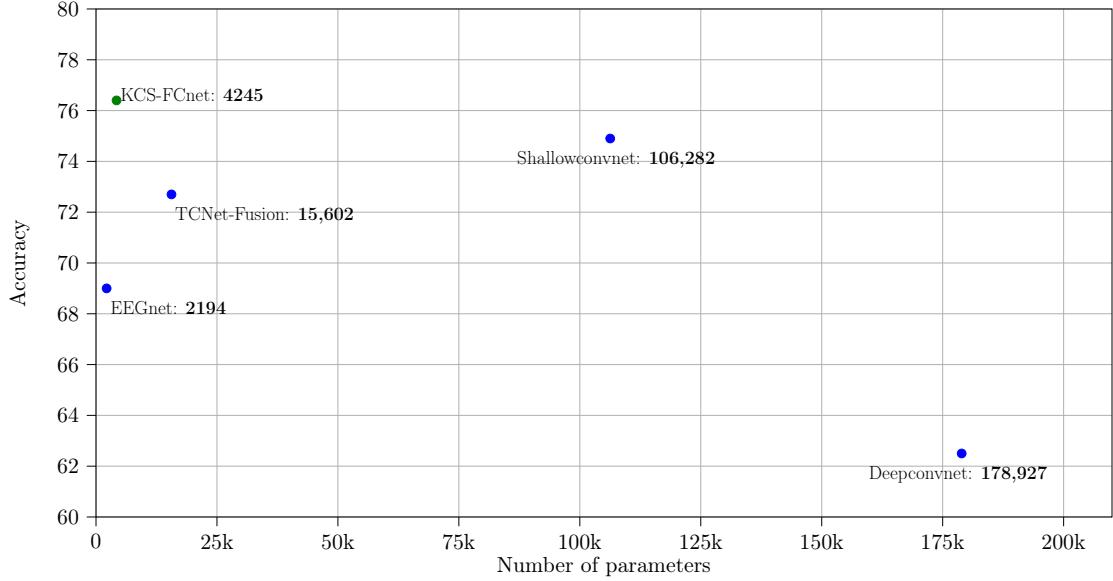
Another essential aspect to quantify the model performance is the number of trainable parameters. Figure 6-8 presents the required number of trainable parameters vs. the attained MI classification performance for each studied DL approach. As seen, a higher number of trainable parameters does not necessarily imply a higher classification accuracy. In fact, two clusters are evident: models holding less than 20k trainable parameters and algorithms requiring more than 100k parameters. Notably, the EEGnet gathers 2.194 trainable parameters and got a 69% accuracy score. Then, the Deepconvnet has the most significant number of trainable parameters (178.927) but only achieves a 62.5% accuracy score. The overfitting issue can explain the latter, especially when dealing with a high intra-class variability MI dataset. As previously stated, the Shallowconvnet, TCNet-Fusion, and KCS-FCnet have the highest accuracy scores. However, the Shallowconvnet has more than 100k trainable parameters, and the TCNet-Fusion has more than 25k. Conversely, our KCS-FCnet, not only outperforms these architectures in terms of accuracy but also requires the lowest complexity to achieve competitive discrimination results.

**Table 6-3:** Method Comparison results regarding the average MI classification for the Giga database.

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



**Figure 6-7:** KSC-FCnet functional connectivity results (Circos plots and topoplots). The first row illustrates the 99th percentile of the most significant functional connections across the centroid subjects concerning the studied best, medium, and worst performance clusters (see Figure 4-7), with the opacity representing the strength of the connectivity. The second and third rows display topoplots of the two-sampled Kolmogorov–Smirnov test and the weights of the classification layer on the KCS-FCnet, respectively. The purer the purple color, the more important the brain area is. The left to right column represents each group from G I to G III.



**Figure 6-8:** Method comparison results: number of trainable parameters vs. average motor imagery classification accuracy for the Giga database.

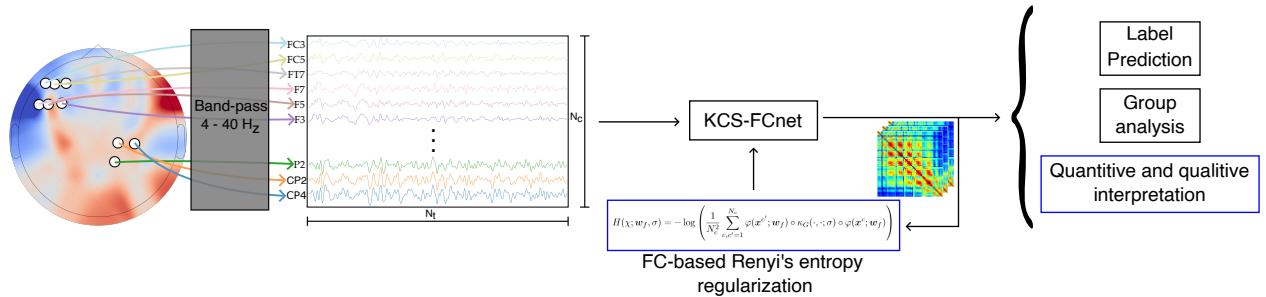
## 6.4 Summary

We have designed a novel end-to-end DL EEG-based MI classification method, termed the Kernel Cross-Spectral Functional Connectivity Network (KCS-FCnet). This strategy directly addresses fundamental issues like the elimination of handcrafted feature extraction steps and artifact removal that often generate false connectivities, as elaborated in Section 1.2.2. Our methodology combines cross-spectral analysis with a data-driven KCS-FC block to model the non-linear connections between different channels and a CNN layer to automatically refine temporal, spectral, and spatial representations from EEGs. We appraised the efficacy of our approach alongside other end-to-end DL techniques using an extensively utilized public dataset. The results demonstrated that our KCS-FCnet approach consistently outperforms prevalent state-of-the-art methods in terms of both EEG-based MI classification and spatio-temporal-frequency interpretability. This system also minimizes the number of trainable parameters, thus reducing architectural complexity.

Although our approach outperforms other DL solutions when it comes to accuracy, it remains a challenge to conclusively affirm if our system offers enhanced interpretability as discussed in Section 1.2.3. To ensure a comprehensive analysis of the interpreted results, it is essential to employ both qualitative and quantitative methodologies. Future studies, as suggested in Section 1.3.3, should aim to incorporate post-hoc and intrinsic strategies to develop MI-BCI models that are more transparent and interpretable.

# 7 IRKCS-FCnet: Interpretable Regularized Kernel Cross-Spectral Functional Connectivity Network with Qualitative and Quantitative Post-Hoc and Intrinsic Explainability

In this chapter, we introduce an Interpretable Regularized Kernel Cross-Spectral FC network, named IRKCS-FCnet, as depicted in Figure 7-1. This approach minimizes spurious connectivities by optimizing Renyi's entropy of KCS-FC matrices. Both IRKCS-FCnet and its plain version, KCS-FCnet, deploy a CNN layer to extract temporal-spectral features for enhanced FC estimation. The performance of our proposed approach is assessed employing both post-hoc and intrinsic quantitative and qualitative strategies, and it is compared against EEGnet, the recognized standard. While achieving notable accuracy, the IRKCS-FCnet also produces improved interpretable FC matrices. Overall, our results demonstrate that our model architecture achieves comparable MI classification while improving the transparency and interoperability.



**Figure 7-1:** Illustration of the proposed Interpretable Regularized Kernel Cross-Spectral Functional Connectivity Network (IRKCS-FCnet) model, with both qualitative and quantitative strategies for improved MI-BCI model explainability

## 7.1 Cross-Information Potential Regularization

Let the set  $\{\mathbf{X}^c \in \mathbb{R}^{N_t}\}_{c=1}^{N_c}$  be i.i.d realizations drawn from the random variable  $\chi$  with density function  $\zeta(\chi)$ . We can estimate the density function using the Parzen density estimation strategy as follow.

$$\hat{\zeta}(\chi) = \frac{1}{N_c} \sum_{c=1}^{N_c} \kappa(\chi, x^c; \sigma) \quad (7-1)$$

Since the cross-information potential is a special case of the definite kernel  $\kappa(\cdot, \cdot; \sigma)$  and it is related to the Renyi's entropy  $\alpha = 2$  by a monotonic function as show in Equation (7-2). We can maximize either the Renyi's entropy or the cross-information potential.

$$\begin{aligned} H_2(\chi) &= -\log \left( \int_{\chi} \hat{\zeta}(\chi)^2 \right) d\zeta \\ &= -\log \left( \frac{1}{N_c^2} \sum_{c,c'=1}^{N_c} \kappa(x^c, x^{c'}) \right) \end{aligned} \quad (7-2)$$

This concept can be extended to the function composition proposed in Chapter 6 Equation (6-1) as follows.

$$\hat{\zeta}(\chi) = \frac{1}{N_c} \sum_{c=1}^{N_c} \kappa_x(\chi, \cdot; \sigma) \circ \varphi(x^c; w_f), \quad (7-3)$$

Where  $\circ$  stands for function composition,  $\varphi(\cdot; w_f)$  is a 1-D convolutional layer that can be used to automatically extract frequency patterns ruled by the weight vector  $w_f \in \mathbb{R}^{\Delta_t}$ , with  $\Delta_t < N_t$ . Thus, the Renyi's entropy can be calculated as follow.

$$H(\chi; w_f, \sigma) = -\log \left( \frac{1}{N_c^2} \sum_{c,c'=1}^{N_c} \kappa_x(\cdot, \cdot; \sigma) \circ (\varphi(x^{c'}; w_f), \varphi(x^c; w_f)) \right) \quad (7-4)$$

Note that the maximal potential information is obtained when each sample provides a unique information about the Parzen estimation of the density function. Therefore, maximizing the cross-information potential will induce the diagonal  $\mathbf{K}$  matrix. Since the main objective is to enhance the interpretability while maintaining the results obtained in Chapter 6, we include the  $H(\chi; w_f, \sigma)$  into the optimization problem as follows:

$$\Theta^* = \arg \min_{\Theta} \mathbb{E}_r \left\{ \mathcal{L}(y_r, \hat{y}_r | \Theta) - H(\tilde{P}_r; w_f, \sigma); \forall r \in \{1, 2, \dots, R\} \right\}, \quad (7-5)$$

where  $\tilde{P}_r$  is the average functional connectivity measure calculated in Equation (6-4) and  $\hat{y}$  is the class probability vector obtained in Equation (6-6). Here, the objective is akin to the information bottleneck principle, where we aim to maximize the cross-information potential while maintaining the accuracy obtained in Chapter 6. In this context, the cross-information potential serves as a regularization technique on the typical cross-entropy optimization problem.

## 7.2 Experimental Set-Up

- **Raw EEG Preprocessing:** The preprocessing stage starts with the loading of subject recordings using a **custom database loader** module (<https://github.com/UN-GCPDS/python-gcpds.databases>.

Proceeding with the presumption that 2.5s windows are stationary, a Fourier method is applied to down-sample each signal from 512 Hz to 128 Hz, as provided by the SciPy library with its signal resample function <https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.resample.html>). Each time series trial is filtered within the [4 – 40] Hz range using a fifth-order Butterworth bandpass filter. Subsequently, we construct an array of time-based records, utilizing 2.5s windows and a 50% overlap, starting from 0s and ending at 7s. This enables us to analyze windows that contain no MI information, partial MI information, and solely MI information sequentially. It's important to note that the cue onset happens 2s post the start of the recording; this preprocessing tries to mirror the one in [Lawhern et al., 2018]. The filter [4 – 40] Hz aimed at probing five distinct brain rhythms encompassing theta, alpha, and three beta waves [Abhang et al., 2016]. Where theta waves [4 – 8] Hz, typically observed in the hippocampus and various cortical structures, are theorized to signify an “online state” and are linked with sensorimotor and mnemonic functionalities. On the contrary, alpha-band activity [8 – 13] Hz is suppressed by sensory stimulation and movements, potentially acting as a marker for higher motor control functions given its modulation by attention, working memory, and mental tasks. The preprocessing also incorporates three beta wave types: “beta one” or low beta waves [12 – 15] Hz typically correlate with focused and introverted focus; “beta two” or mid-range beta waves [15 – 20] Hz are associated with augmented energy, anxiety, and performance; “beta three” or high beta waves [18 – 40] Hz are indicative of considerable stress, anxiety, paranoia, high energy, and high arousal.

- **Training:** Post-preprocessing, we are left with five datasets, each encapsulating 2.5s EEG information, all characterized by the same label set. Each dataset is individually trained as per the subsequent steps: Firstly, trials from each subject's data are split based on a standard 5-fold 80-20 scheme. In other words, the data is shuffled, and 80% of it is utilized for training (forming the training set), withholding the remaining 20% to validate the trained models (forming the testing set). This sequence is repeated five times in total. Subsequently, the plain model is trained with one of the training sets, searching for the hyperparameter  $N_f$  within the set {2, 3, 4} and  $\Delta_t$  within the set {50, 25}. It's important to note that the kernel length scale set corresponds with the new sampling frequency [Lawhern et al., 2018]. **The ADAM algorithm [Kingma, 2014] is employed to tune the model parameters with a learning rate of 0.1.** To locate the optimal combination of hyperparameters for our model, the GridSearchCV class from SKlearn is utilized. Given the need for performance comparison, we compute the accuracy, Cohen's kappa, and the area under the ROC curve to assess performance between models [Warrens, 2015, Géron, 2022]. Following the conclusion of this procedure, the best performance metrics and hyperparameters are saved. For the sake of comparison, the optimal hyperparameters secured from the plain trained model are loaded onto the regularized version. The splitting and training steps are replicated similarly, except only for the new hyperparameter that modifies the regularization weight  $\Upsilon$  on the cost function, and it is explored within the set {0, 0.2, 0.5, 0.7, 1, 5, 10}. Here, we aspire to compare the performance of the regularized version using the same hyperparameters over the identical time window. The underlying hypothesis is that the regularized version performs at least equivalently, or superior, to the plain version. We also aim to gain insights into how regularization impacts the connectivity matrix. Concluding the training process, we have ten models for each subject: five corresponding to the plain model and five to its regularized version.
- **Model Level Analysis:** By leveraging the benchmark EEGnet model, we contrast the performances of both plain and regularized models. Subjects are sorted based on their respective EEGnet accuracy score. In addition, we examine the accuracy of each of the five windows mentioned above, analyzing how performance attenuates when the provision of MI information is constrained. For instance, we start when subjects see a blank screen, followed by moments entirely consumed by MI tasks, and finally reach situations where subjects behold a blank screen immediately after executing an MI task. In this final instance, the brain might continue engaging with the MI task for several seconds more. This comparative analysis promises to provide insights into the effects of varying MI information availability on model performance.
- **Group Level Analysis:** Focusing on the window demonstrating the highest performance, we use the k-means clustering algorithm [Géron, 2022] to cluster the accuracy scores of each subject. Setting  $k$

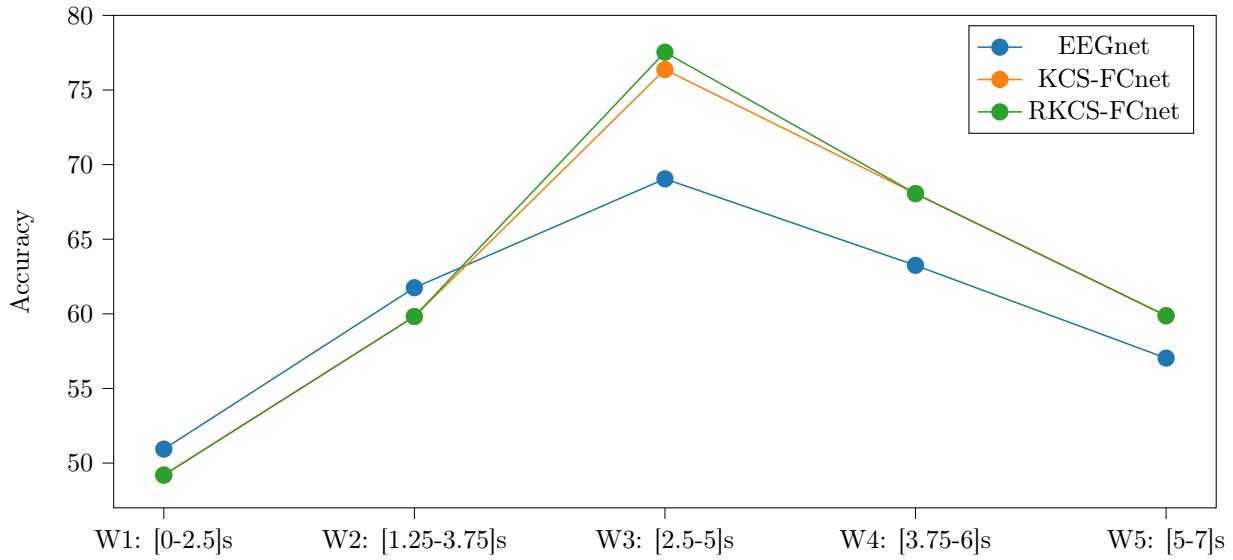
to three, we create a model categorizing subjects' performance based on their results in the EEGnet model [Lawhern *et al.*, 2018]. This separates the subjects into three distinct performance groups: best (G I), intermediate (G II), and worst (G III) performers. Applying this trained k-means model, we then cluster the data from the plain and regularized models. The aim here is to identify shifts in group assignments between the baseline performance and the results from plain and regularized models. To further our analysis, we calculate Renyi's entropy for each identified group, where higher Renyi's entropy implies a sparser connectivity matrix and higher cross-information potential.

- **Post-hoc interpretability:** We also examine the quantitative differences between the plain and regularized model versions, specifically looking at average and random feature drops. Using a paired t-test, we validate that dropping relevant features impacts accuracy differently than dropping random features. This involves creating scenarios where 5%, 10%, 15%, 20%, and 25% of features are dropped based on relevance and randomly. The impact on accuracy in each scenario is then compared to the performance when all features are considered. By concatenating these calculated accuracy drop values and running a paired t-test, we verify their statistical difference. Given that our regularized version is designed to extract more relevant features, we expect to observe higher accuracy drop values with each drop percentage increase compared with its plain version.
- **Intrinsic interpretability:** We also took the weights of the last layer to understand which channels were the most important. To do so, we calculated the values of the average weights for each group. Then, we summed up the absolute values for each branch class's average weights, sorted them, and took the top ten channels. Next, we represented them in a topoplot-like illustration. The hypothesis is that the best group will have all or most of the most important channels near the sensorimotor cortex area. While the noise affects the subjects with lower accuracy performance, the most important channels can be found in different areas. Furthermore, we extracted the Power Spectral Density (PSD) with the Welch method of the CNN layer output to analyze how the temporal and spectral filters modify the signals. This was done for each subject and then averaged by groups. Here, we are specifically interested in exploring the alpha wave, particularly the mu rhythm, oscillating between 8 and **13Hz**. Activation of this rhythm, found primarily in motor-related brain areas, is hypothesized to be activated in the highest performing group [Hobson & Bishop, 2017, Llanos *et al.*, 2013]. For other groups, a more widespread activation should be seen. In addition, the average of the last branch class weights is used to show the overall importance of each channel. This is shown by a 2D scalp representation, which aids in seeing which brain areas are the most important for each group. Here, we hypothesize that the sensorimotor areas are the most important for the best groups, while more areas are included as we approach the worst-performing group.
- **Individual Subject Analysis:** For a deep-dive into each group's characteristics, we select the subjects closest to their respective group's centroid for analysis. We initially calculate the PSD using the Welch method over cross and direct-label class normalized activation maps with the modified LayerCAM, which return one map per filter. The primary goal is to examine the relative importance of each filter coupled with the frequencies they extracted. We hypothesize that information will mostly be retained by the direct class, where each filter ideally should contribute equally towards the classification objective.

## 7.3 Results and Discussion

### 7.3.1 Model Level Analysis

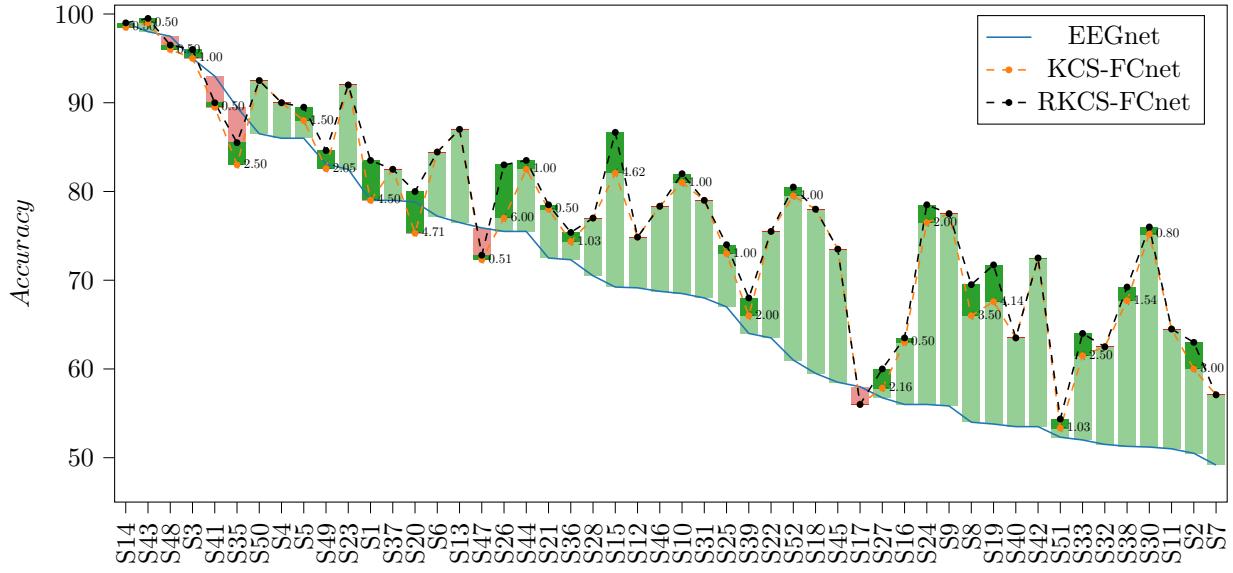
Figure 7-2 shows the comparison of performance between different windows, each containing various amounts of MI information. The first window has no MI information, the second one contains half of the MI task, the third 0.5 seconds after the cue along with the complete span of MI information, the fourth comprises half MI and half post MI, and finally, the last window exclusively captures post-MI information.



**Figure 7-2:** Performance comparison across EEGnet (—), KCS-FCnet (—), and RKCS-FCnet (—) models at different time windows.

The figure reveals three distinct patterns. Firstly, all three models exhibit an overall accuracy lower than 65% when operating with none or insufficient MI information. Also, it is important to highlight that for these initial two windows, the EEGnet model outperforms the others, demonstrating a higher accuracy score. This can be attributed to the fact that FC-based models compare full-time windows to derive the connectivity matrix, making them more susceptible to windows conveying information from varied cognitive tasks. Moreover, as EEGnet continues to add more convolution layers, the impact of non-relevant time segments becomes minor. Secondly, all models hit their peak accuracy in the [2.5 – 5]s window. Within this window, we observe roughly equivalent average accuracy between the two FC-based models, but there is a significant difference from the EEGnet. Thirdly, the phase of windows carrying post-MI information still shows greater accuracy performance than the first two windows, even though it contains post-task information. There could be residual information, as the brain may continue engaging with the task. Interestingly, a noticeable difference in accuracy score for the FC-based models is seen only in the best-performing window. This could occur because the regularization works optimally when the whole matrix represents the time segment holding the MI class under analysis. Henceforth, our focus will be on the best-performing window [2.5 – 5]s for deeper analysis.

Figure 7-3 illustrates a comparative analysis of the accuracy rates achieved by the EEGnet, KCS-FCNet, and RKCS-FCnet models over the best performing window [0.5 – 2.5]s. It provides a visual representation of the average accuracy for individual subjects. Subjects have been ordered based on the accuracy performance of EEGnet, depicted as a blue line, and the dotted orange and black lines, respectively denote the KCS-FCNet and RKCS-FCnet models. Changes in accuracy are represented by color coding, where green indicates an

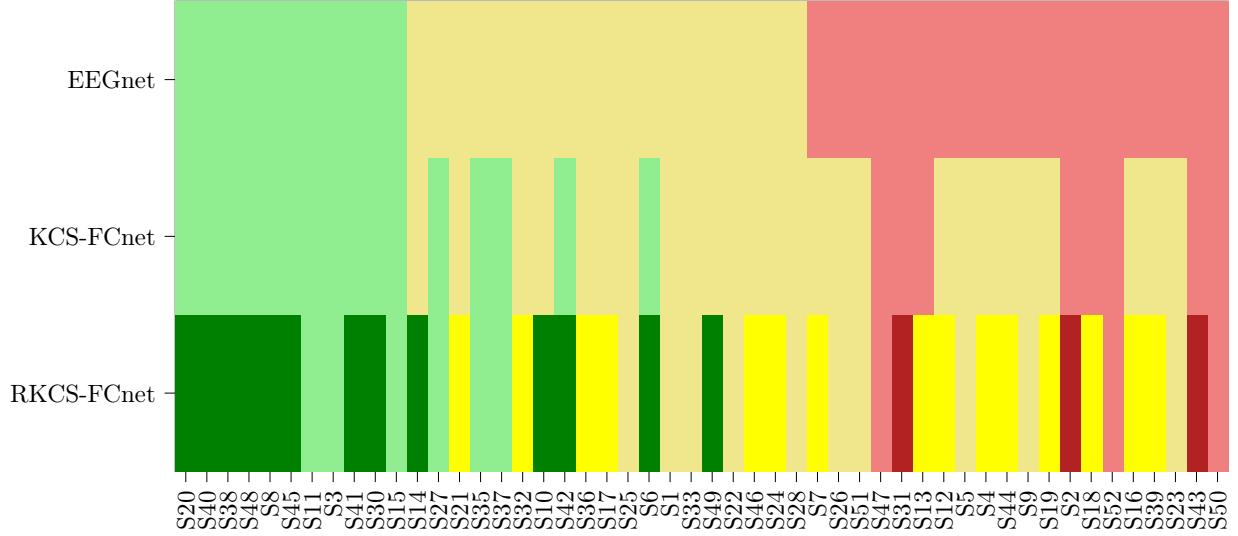


**Figure 7-3:** Comparison of subject-specific average accuracy across EEGnet, KCS-FCnet, and RKCS-FCnet models. Subjects are arranged in descending order based on the accuracy of the EEGnet model. The color-coded bars illustrate performance shifts, with dark green(■) indicating an increase in accuracy on the RKCS-FCnet over KCS-FCnet, light green(□) indicating an increase in accuracy on KCS-FCnet over EEGnet, and while red(■) indicates an accuracy reduction.

increase in accuracy, whereas red signifies a decrease. Additionally, the differential accuracy between the plain and regularized versions is presented with solid green bars and numerical values for comparative evaluation, highlighting how much advancement the regularized version has achieved over its plain counterpart. For instance, it can be seen in subject 26, with a notable increase of six percentile points. Overall, both the KCS-FCnet and RKCS-FCnet models outperform the benchmark model, with only a few exceptions where the accuracy drops. It's interesting to note that the regularized version consistently performs at least as well as, or better than, the plain version, with an accuracy boost of more than four percentage points observed for subjects 1, 20, 26, and 15. It is worth noting that increases in accuracy were not significant in subjects already achieving more than 80% accuracy, suggesting strong predictive power offered by either approach for those subjects. However, subjects falling below 80% FC-based methods demonstrated the most marked improvements, particularly under the 70% threshold, where the nature of FC effectively prevents noisy connections. Lastly, the overall comparison illustrates that subject 17 yields identical performance across all three models.

### 7.3.2 Group Level Analysis

Figure 7-4 illustrates the distribution of subjects across various performance groups for each approach. We use three colors (light green, light yellow, and light red) to differentiate the groups: best performance, intermediate performance, and worst performance, respectively. In the first row, we present the distribution of performance for the EEGnet model. This will serve as our benchmark for comparing distributions. The second row reveals a shift in group belonging when using the KCS-FCnet model. Specifically, five subjects from the intermediate group have moved to the best performing group, and twelve subjects advanced from the worst to the intermediate group. Note the significant concentration of subjects in the intermediate performance group. The same trend is observed in the third row, showcasing the group distribution for the RKCS-FCnet. Light-colored symbols mark subjects who remain in their initial group with same accuracy performance, while those with solid colors depict subjects with improved accuracy rates. For instance,



**Figure 7-4:** Group transitions for subjects across EEGnet, KCS-FCnet, and RKCS-FCnet models. The color representation, green, yellow, and red, correspondingly signifies the best, intermediate, and worst average accuracy performance. First row: Initial group allocation of subjects based on EEGnet model results. Second row: Subject group transitions resulting from implementing the KCS-FCnet model. Third row: Subject group shifts according to the outcomes of the RKCS-FCnet model.

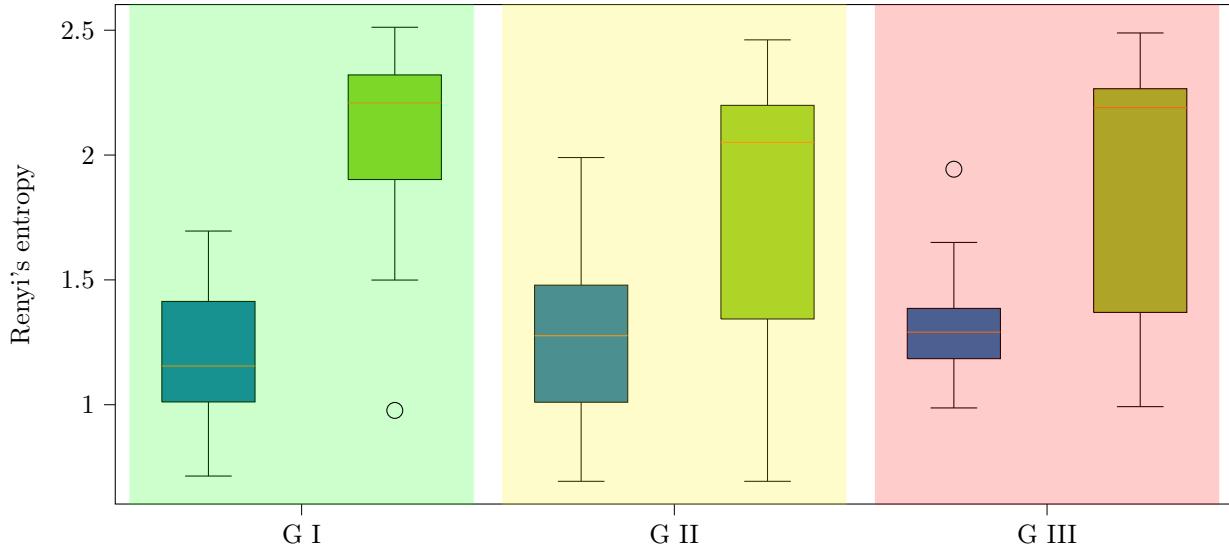
**Table 7-1:** Comparison of accuracy performance across groups and models

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	<b><math>92.28 \pm 4.79</math></b>	<b><math>79.26 \pm 4.93</math></b>	<b><math>67.77 \pm 7.83</math></b>

subjects 14, 10, and 49 are promoted from the intermediate to the best group, while subjects 13 and 18 moved up from the worst to the intermediate group. It is worth noting that the regularized model maintains or enhances subject group standings and accuracy. For the RKCS-FCnet model only six subjects are left in the worst performance group. This marks an improvement compared to the eight subjects under the KCS-FCnet model and a significant advance from the twenty in the same group under our baseline EEGnet model. In fact, performance was found to be improved for 31 subjects as compared to the KCS-FCnet model. Thus, the incorporation of FC information offers notable advantages, especially in aiding the subjects from the worst performance group, more so when combined with Renyi’s regularization.

Table 7-1 provides a comparative perspective on the accuracy performance of three models (EEGnet, KCS-FCnet, IRKCS-FCnet) across the three distinct groups, where bold numbers represent the highest group score. Each group varies based on performance—from the highest performing (Group I) to the lowest (Group III). Note that The regularized version (IRKCS-FCnet) reaches the highest accuracy scores. In Group I, the accuracy difference is marginal at about one percentage point between the EEGnet baseline and KCS-FCnet. Yet, the improvement becomes more marked with the IRKCS-FCnet, which outpaces the baseline by nearly two percentage points. Interestingly, the standard deviation for the regularized version also witnesses a decrease of more than one percentage point, suggesting enhanced consistency in performance for the regularized version. Group II exemplifies the most dramatic difference in accuracy, with an uplift from 72.15% in EEGnet to almost 80% in RKCS-FCnet. Despite the nearly constant standard deviation across all models in this group, the number of subjects included in this group increases significantly for the FC-based models compared to the baseline model, indicating greater stability. Finally, in Group III, the FC-based models perform similarly, outperforming the baseline model by over 12 percentage points. The most substantial difference between FC-based and baseline models is seen in this group; a possible

explanation for this substantial improvement is that FC information plays a critical role in enhancing the model's performance, particularly for subjects with potentially noisier EEG recordings, as evident in Figure 7-3.



**Figure 7-5:** Box and whisker plot comparing Renyi's entropy across groups for KCS-FCnet (blue boxplots) and RKCS-FCnet (green boxplots) models. The background codes the group membership (best, intermediate, and worst performance clusters)

As demonstrated in Figure 7-5, the regularized model consistently displays higher entropy values across all groups. Remember that within the cost function, along with cross-entropy, the pursuit of maximizing this specific value is prioritized. In this scenario, higher entropy values suggest that there is a greater focus on obtaining independent rows in the FC matrix, which in turn indicates the inclusion of less noisy connections. Group I shows the most significant difference. Except for one sample, all other samples exceed 1.5. This trend can be attributed to the fact that best-performing subjects might rely on fewer connections to achieve high accuracy, allowing for more effective regularization. A similar pattern is observed in Groups II and III, with increased entropy values seen in the regularized model's results in both cases. However, Group II showcases a wider spread of values. This could be explained by the larger number of subjects included in this group for both plain and regularized versions. The overlap seen in this box and whisker plot, becoming more extensive in the intermediate group, led us to perform a paired t-test. This statistical test helps ascertain whether the observed values can be considered as belonging to different sets, thus establishing the statistical difference between them. The assessed p-values for Groups I, II, and III were 0.00006, 0.00175, and 0.00005, respectively. All these p-values fall below the significance level of 0.05, with Groups I and III exhibiting exceptionally low p-values below 0.001. Consequently, this statistically significant variance reiterates that the values derived from the regularized model are distinctively different across all groups.

### Post-Hoc Interpretability

**Table 7-2:** p-values comparing accuracy drop of KCS-FCnet and RKCS-FCnet models vs Random (\* indicates p-value < 0.01)

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

We quantified the average drop for both models to validate that IRKCS-FCnet is better at extracting relevant features than its plain version. Here, we progressively eliminate features from the most relevant and monitor the resultant decrease in accuracy performance. We introduced a random version for comparison to ensure valuable feature learning in both models. In this context, all features hold an equal likelihood of being dropped. As shown in Table 7-2 the calculated p-values confirm that the drop in accuracy upon feature removal in both models is statistically different from the random drop across all performance groups (G I, G II, and G III).

**Table 7-3:** Average accuracy drop across feature removal percentages for KCS-FCnet and RKCS-FCnet, with higher difference drops indicated by (green) for GI, (yellow) for GII, and (red) for GIII.

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

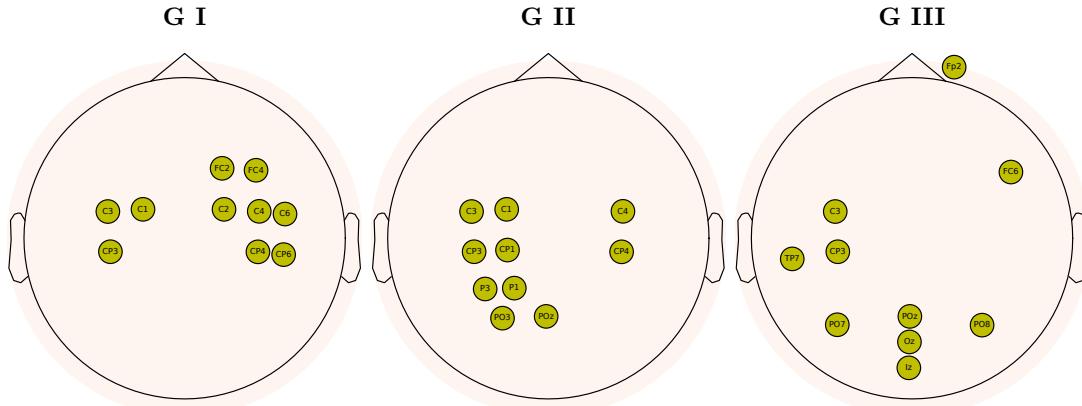
Table 7-3 illustrates the decrease in accuracy as more important features are removed, with higher drop values indicating the removal of more critical features. Notably, in both models, groups that initially achieved higher accuracy scores suffer larger accuracy drops, especially when contrasted with other groups. This effect is particularly evident when only 5% of the features are removed. Across all groups and every percentage of feature drops, barring one case for GII, the IRKCS-FCnet outperforms in terms of feature importance as it shows higher accuracy drops. Additionally, the most substantial differences are represented by (green) with a 5.34 percentage points for GI, (yellow) with a 1.32 percentage points gain for the plain model in the case of GII, and (red) with a 1.82 percentage points for GIII.

For instance, in Group I, the plain model suffers a drop of 17.12 accuracy points, whereas the regularized model experiences a more significant drop of 22.46 points, marking a difference of over 5 points. Similar but less pronounced differences are observed for Groups II and III. However, as we increase the feature removal amount to higher percentages, the discrepancy between the plain and regularized versions' drop values gradually diminishes. We can attribute this trend to the finite set of critical features. After their removal, eliminating additional features might not have the same level of impact. Moreover, Group I, as the top performing group, exhibits the largest drops until 15% of removed features. It is worth noting that as we progress from removing 15% to 25% of features, all drop values seemingly plateau, approximating an accuracy drop of almost 40%. Recalling that the MI task explored holds a binary classification, any dummy model would still be expected to achieve an average accuracy of about 50%. Therefore, this observation of the models standing near to 40% drop in accuracy as more features are removed underscores the relative significance of each feature in our models.

## Intrinsic Interpretability

The top ten channels for each group can be seen in Figure 7-6. For Group I, the best-performing, we can see that most of the channels reside in the sensorimotor cortex (CP3, CP4, CP6), central sulcus (C1, C2, C3, C4, C6), and the motor cortex (FC2, FC4). As stated in various studies about individuals performing motor tasks, we can observe significant activation differences in contiguous structures such as the anterior and posterior areas of the central sulcus, motor cortex, and sensorimotor [Eichert et al., 2021]. Moving on to Group II, the first notable difference is that the channels from the motor cortex are not among the top ten, with 60% of the electrodes located in the motor-related brain area. The remaining 40% are found on the parietal lobe (P1, P3, PO3, POz), a region primarily responsible for fine sensation related to sensory perception and integration. In this instance, this area may be activated due to the visuospatial navigation

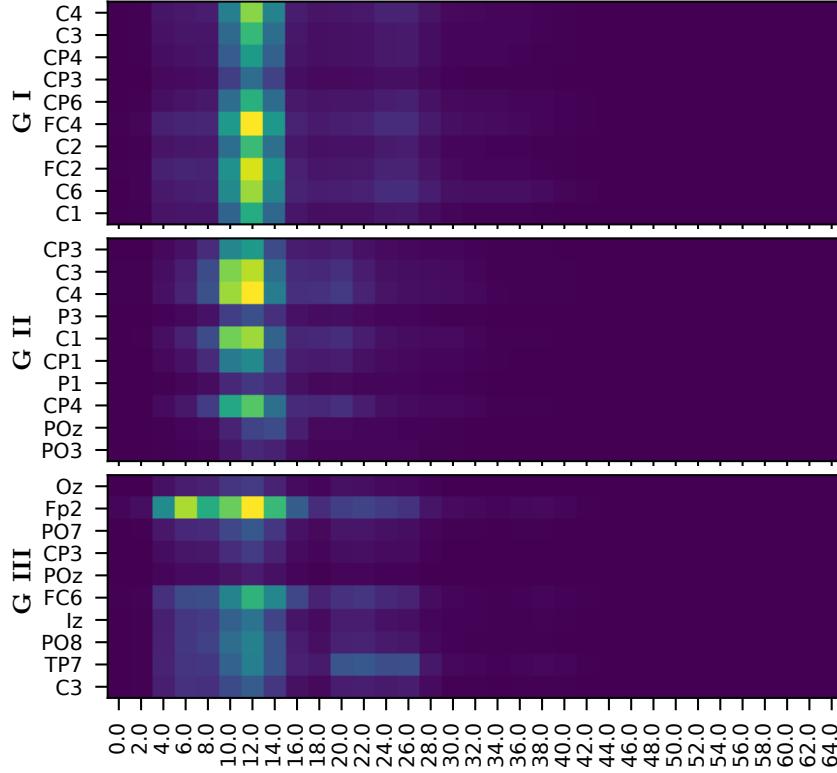
and reasoning when viewing the screen instructing to perform a specific Motor Imagery task. It is logical to infer visual stimuli interpretation is occurring since half of these channels are located in the parietal-occipital area [Phunruangsakao *et al.*, 2023]. Lastly, in Group III, comprising the worst-performing individuals, we observe that only 30% of the top ten channels are situated in the motor-related area (FC6, C3, CP3). The remaining 70% consist of electrodes from temporal (TP7), parietal-occipital (PO7, POz, PO8), occipital (Oz, Iz - the most posterior occipital electrode), and prefrontal (Fp2) regions. The frontal and occipital lobes suggest that visual information remains influential in brain network organization. Specifically, the Fp2 channels may suggest the presence of high ocular artifacts [Han *et al.*, 2023].



**Figure 7-6:** Visualization of the top ten channels in each group. The channels for Group I (left) are primarily concentrated in the sensorimotor cortex, central sulcus, and motor cortex. In Group II (middle), the majority of the channels are located in the motor-related brain area and parietal lobe, with visual stimuli interpretation inferred from the number of channels in the parietal-occipital area. For Group III (right), the top channels are predominantly located in non-motor-related areas, with a particular influence from visual information in the frontal and occipital lobes and possible high ocular artifacts suggested by the presence of Fp2 channels.

In order to gain a deeper understanding of the spectral filters produced by the CNN layer, Figure 7-7 presents the average PSD for each group on the top ten channels, ranked in order of importance. In Group I, we observe a relatively stable spectrum, with the spectral power mainly concentrated between 10 and **14Hz**, exhibiting higher spectral power centered around **12Hz**. This observation aligns with various studies noting that motor imagery-related tasks predominantly influence the Mu rhythm, oscillating between 8 and **13Hz** [Al-Saegh *et al.*, 2021, Hobson & Bishop, 2017, Llanos *et al.*, 2013]. Group II demonstrates a similar pattern; however, the spectral power appears slightly skewed to the left. Nonetheless, it still falls within the range of the Mu oscillations associated with motor imagery in motor-related brain areas [Chen *et al.*, 2021]. As for the last group, the pre-frontal Fp2 channel captures most of the spectral power, with significant influences from  $\alpha$  and  $\theta$  waves. With Oz and Fp2 emerging as the two most important channels and Fp2 commanding the majority of the spectral power, it can be inferred that a substantial amount of artifacts are introduced [Han *et al.*, 2023]. These artifacts directly impact this group's task performance accuracy as we showed in Table 7-1.

Figure Figure 7-8 offers a 2D illustration of scalp spatial relevance for each group and each class (Right and Left hand), enhancing our understanding of spatial interpretability. Clear activation is noted in the sensorimotor region for the first two groups, primarily for the left hand, with positive significance in the contralateral region and negative significance in the ipsilateral [Van der Lubbe *et al.*, 2021]. An inverse yet similar pattern is observed for the right hand. Group II displays a spread of importance that extends towards the parietal and occipital lobes yet retains a significant concentration around the sensorimotor cortex. For Group I, channels C3 and C4 are the most important, and channel CP3 is also highlighted for Group II. Group III shows an intriguing behavior with a faint contralateral arrangement for the left class and ipsilateral for the right. Nevertheless, numerous activations in non-motor-related external channels, like Fp2, Oz, PO8, and PO7, suggest a possibility of artifact interference [Collazos-Huertas *et al.*, 2023]. Intense

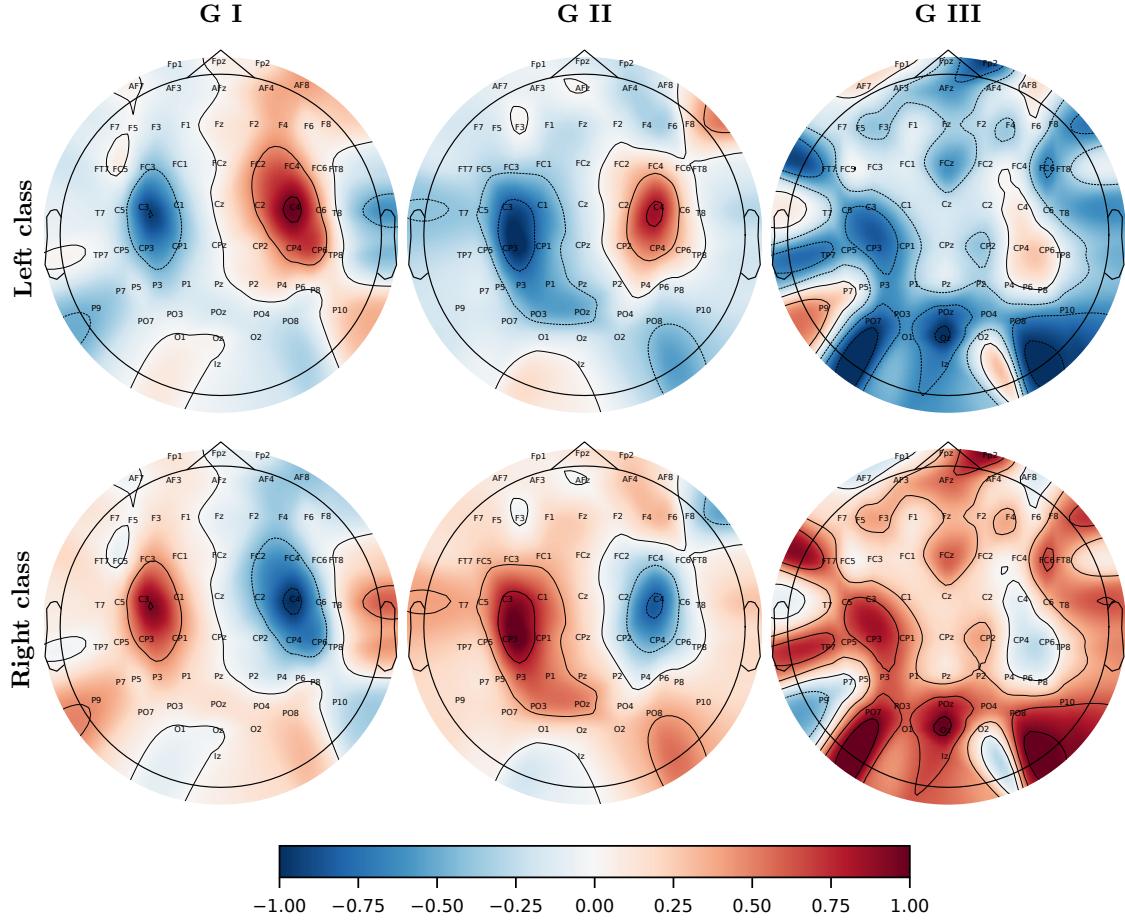


**Figure 7-7:** Illustration of the average PSD for each group plotted across the ten most important channels, categorized by importance from top to bottom. Relatively stable spectrum is visible within Group I (top), with predominant spectral power between 10 and **14Hz**. Group II (middle) exhibits a similar pattern, with a slightly left-skewed power distribution, yet still within the range of Mu oscillations. On the contrary, Group III (bottom) displays an increase in spectral power at the pre-frontal Fp2 channel, significantly influenced by  $\alpha$  and  $\theta$  waves, indicating the introduction of substantial artifacts.

activation observed in the occipital lobe could indicate the likely impact of substantial visual stimuli on processing motor-related information. These dispersed activations align with the modest accuracy level of around 68% achieved in the IRKCS-FCnet. The proposed regularization method allows for efficient handling of artifacts and effective capture of crucial information around the motor-related brain area. However, in extreme situations, challenges persist. Even though an increase is observed in the accuracy level, nearing the 70% threshold for BCI inefficiency, noise and artifacts continue to hinder the performance of the MI task [Vidaurre *et al.*, 2021].

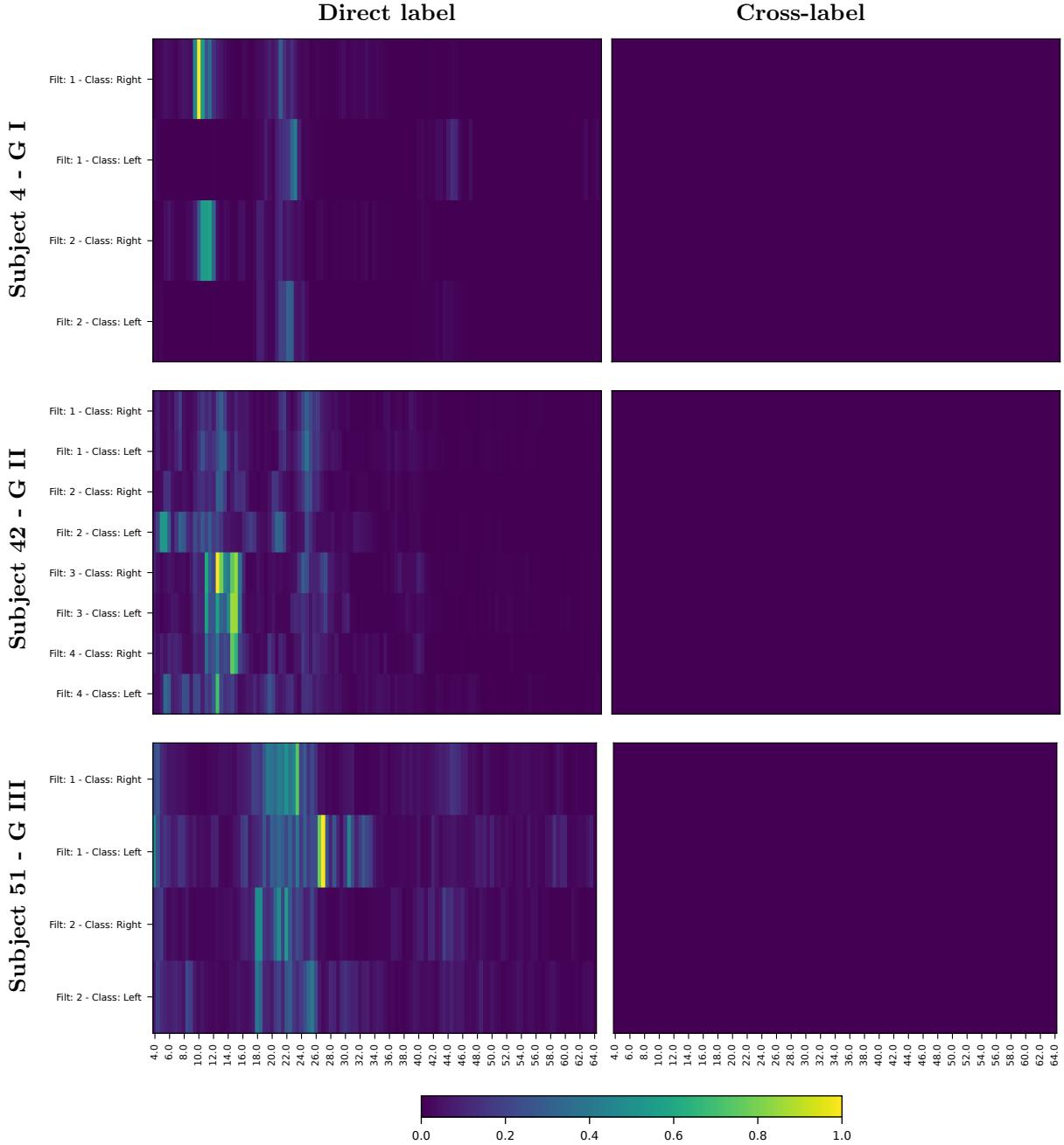
### 7.3.3 Individual Subject Analysis

Figure Figure 7-9 displays individual subjects from Group I, Group II, and Group III, specifically subjects 4, 42, and 51, respectively. Across all cases, the cross-label importance is zero, thereby indicating no model misdirection towards the contrary class. While this is a positive aspect, the overall zero cross-label importance suggests potential overfitting in our model. For subject 4 of Group I, the direct label PSD shows that the  $\alpha$  brain wave is primarily relevant for classifying the right-hand movement, as per the expectations for motor-related tasks. Interestingly, the left-hand movement assigns greater importance to higher frequency bands (around 20 to **26Hz**) and no importance to motor-related frequencies. This pattern suggests a model bias towards the right class, indicating a potential increase in false positives (samples incorrectly classified as right-hand movement). In the case of subject 42 from Group II, who uses four filters, the critical frequencies



**Figure 7-8:** Spatial relevance for each group derived from the final layer of the IRKCS-FCnet model. The background represents the accumulated spatial relevance mapped to EEG channel positions. The first row represents the left class, while the second represents the right class. The columns are dedicated to each group: the left column for Group I, the middle column for Group II, and the right column for Group III. The figure visibly presents activation in sensorimotor regions across groups, potential artifact interferences, and alignment with the low accuracy level achieved in G III.

seem somewhat dispersed yet primarily concentrated within the 10 to **18Hz** range. Filters three and four, which highlight the motor-related frequencies, are most important. In contrast, the less critical filters extend across the entire available bandwidth. This scenario does not show any clear preferences between right and left classes. On the other hand, subject 51 from Group III, one of the individuals with the lowest performance, presents a distinct pattern. The most important frequencies, predominantly for the left-hand class, appear above motor-related frequencies, specifically around **28Hz**. The right-hand class does not exhibit comparable crucial filters or frequencies. Interestingly, for this class, most of the motor-related frequencies are considered less or non important.



**Figure 7-9:** Illustration of the PSD across subjects 4, 42, and 51 from groups I, II, and III, respectively. While Groups I and II display localized and specific frequency band importance for motor tasks, Group III suggests more spread frequencies. Zero cross-label importance across all subjects indicates that the model effectively prevents misclassification but suggests potential overfitting.

## 7.4 Summary

We have developed an interpretable, regularized, end-to-end DL architecture for EEG-based MI-BCI systems called the Interpretable Regularized Kernel Cross-Spectral Functional Connectivity Network (IRKCS-

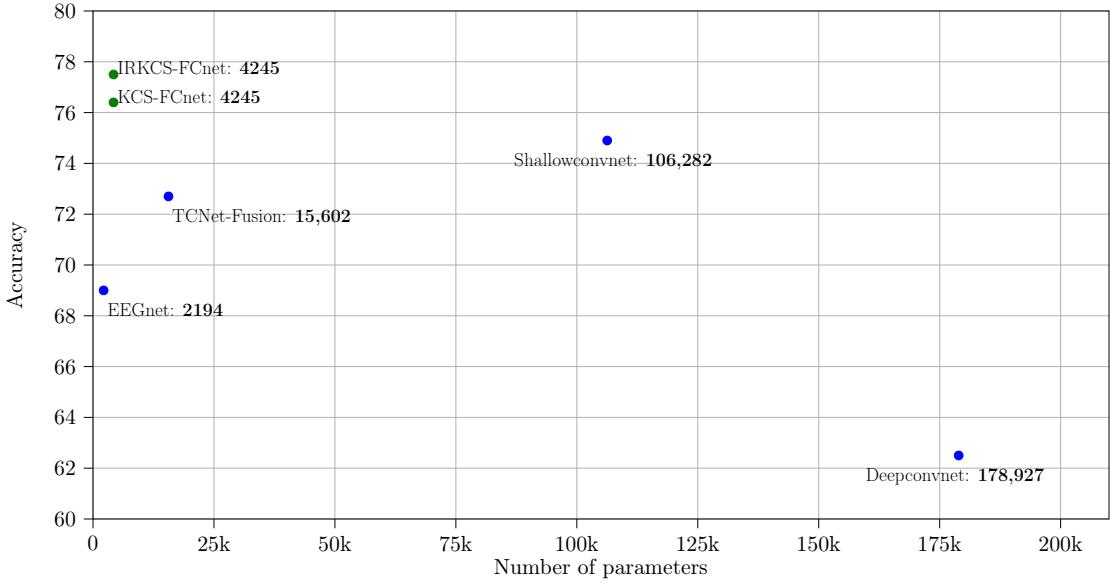
FCnet). This architecture, when combined with a post-hoc and intrinsic interpretability strategy, addresses transparency and interpretation challenges in MI-BCI models, as detailed in Section 1.2.3. We employ the proposed architecture to derive an intrinsic interpretation directly from the weights of the final layer. At the same time, we enhance the cross-information potential in the KCS-FC block to achieve sparser FC connectivities. This effectively removes spurious and nonrelevant connectivities related to the MI task. To facilitate this, we have integrated Renyi's entropy  $\alpha = 2$  into the cross-entropy loss function, emphasizing the maximization of cross-information potential and increasing the probability of the correct label. Moreover, by employing the average drop score, an effective method to measure interpretability results based on the hierarchical removal of critical features, we can compare regular and plain versions of KCS-FC-based DL architectures. Notably, our IRKCS-FCnet has revealed the contralateral concept of MI in both the right and left-hand classes. This comprehensive interpretability method has deepened our understanding of the operation of the KCS-FC block and has provided helpful insight into the potential further enhancement of the proposed architectures.

# 8 Final Remarks

## 8.1 Conclusions

In this thesis, we have effectively demonstrated the power of the kernel cross-spectral approach in approximating the cross-spectral distribution based on the extension of Bochner's theorem, which has significant utility as a single-trial FC estimator. Furthermore, we encapsulated this concept within a custom DL layer, facilitating the integration of KCS-FC into end-to-end DL models KCS-FCnet and posterior IRKCS-FCnet. This required minimal preprocessing steps while achieving superior accuracy compared to other end-to-end DL architectures. Beyond this, we developed a data-driven regularization approach to mitigate the number of connectivities over the adjacency matrix. This approach integrated the concept of cross-information potential from the field of information-theoretic learning, utilizing kernel matrices and the  $\alpha = 2$  Renyi's entropy. As such, we successfully avoided the need to obtain the data's probability distribution. Moreover, we employed both qualitative and quantitative interpretability strategies to advocate for improved model transparency, aiding the search for a suitable measure of a model's interpretability. This was accomplished by comparing average drop scores, thereby gaining insights into the interpretability of competing models. Figure 8-1 presents a comparison between the state-of-the-art end-to-end EEG-based DL architectures (depicted in blue), and the two newly proposed architectures, KCS-FCnet and IRKCS-FCnet (depicted in green). Both of the proposed solutions outperform the existing architectures in terms of accuracy and are ranked second in terms of parameter quantity. This displays the remarkable capacity of KCS-FC to decode EEG signals. The main conclusions of this study can be summarized as follows.

- The validation results obtained from three EEG databases—two with MI and one with ME—demonstrate that the proposed single-trial KCS-FC provides highly competitive classifier performance and is less influenced by handcrafted feature extraction steps, such as the tuning of sliding time window length. Moreover, the benefits of interpreting PLV sets appear limited due to their susceptibility to subject variability, resulting in indistinct topographic channel representations, and PLV provides the poorest classifier performance. Conversely, applying the CCF method results in topoplots showing increased activity over the SMR electrodes, in line with expectations for motor-related tasks. Furthermore, the proposed KCS-FC distinctly highlights two symmetrically located channels over each hemisphere's SMR electrodes, while the accuracy using KCS-FC surpasses other comparable FC estimators. Based on these findings, implementing single-trial KCS-FC for decoding motor skills presents promising potential for MI-BCI systems, enabling high accuracy while maintaining the well-established interpretability of the FC matrices.
- KCS-FCnet excels at automatically estimating EEG representations, clearly outperforming other end-to-end architectures. Compared to even more complex methods, KCS-FCnet surpasses them in terms of performance. Although slightly more complex than the EEGnet, it has greater accuracy, especially in contrast to architectures with 25 times more parameters, marking it as a highly efficient model. Further, our analysis of results from three distinct natural groups highlighted that high-performance subjects retain more critical connections than their counterparts. It also addresses the issue of BCI



**Figure 8-1:** Comparison of state-of-the-art end-to-end EEG-based DL architectures (in blue) with the proposed KCS-FCnet and IRKCS-FCnet architectures (in green) in terms of accuracy and number of parameters.

inefficiency by enhancing the accuracy of low-performing individuals and only marginally reducing the performance of one low-performing subject. Our findings underscore the KCS-FCnet architecture as a promising DL strategy for EEG-based MI classification. Given its lightweight setup in terms of parameters, this method is well suited to practical BCI applications, suggesting substantial potential for significant advancements in this field.

- The IRKCS-FCnet has proven to be an intrinsic interpretable model, providing valuable insight into spatial connections for MI tasks. More specifically, it highlights the contralateral concept of motor-related neurological states that are deeply rooted within the complex brain network. Furthermore, an additional tool, the average drop score, measures the relevancy of the interpretative results. This aspect can facilitate the identification of transparent models by offering a broader comparative framework that encapsulates performance and interpretive analysis. Importantly, we demonstrated that maximizing the cross-information potential regulates the cross-entropy loss function, mitigating spurious connectivities and reducing the number of true connections that fail to distinguish between MI tasks.

## 8.2 Future Work

While our contributions reflect high accuracy and interpretability, additional intricate aspects of the proposed method may warrant further analysis. For instance, the feature extraction process could potentially be refined by utilizing more advanced methods for measuring multivariate similarity, such as centered kernel alignment [Brockmeier *et al.*, 2014, Alvarez-Meza *et al.*, 2017].

Similarly, extracting more information from the time lag, rather than relying solely on an instantaneous relationship, could help handle volume conductive issues as they primarily apply to instantaneous links. This approach may also serve to draw out more information regarding how past times influence the selected channel [Uribe *et al.*, 2019, Bakhshali *et al.*, 2020, De La Pava Panche *et al.*, 2019]. The current method's reliance on pairwise relationships might also overlook the influence of a third channel on the interaction between two electrodes, thereby hinting at the potential benefit of implementing more complex

solutions, such as multivariate approaches, capable of handling the information carried by other electrodes [Vidaurre *et al.*, 2023].

This work sets the stage for more intricate studies in the DL domain, such as cross-subject transfer entropy, where gaining information from different subjects could enhance the decoding of MI tasks embedded in EEG recordings, particularly for low-performing individuals [Wei *et al.*, 2023]. Another promising approach would incorporate Riemannian geometry to encapsulate the internal relationships between different channels in the FC matrix's feature extraction process [Carrara & Papadopoulou, 2023]. Given the recent advancements in the field of graph neural networks, it is conceivable to integrate this concept, given our work's heavy reliance on FC matrices, which naturally generate a dynamic graph. This direction can lead to even more nuanced and sophisticated EEG analysis frameworks [Ma *et al.*, 2023].

### 8.3 Academic Products

- 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)
- 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)
- 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 (Enviado)
- Github repository containing all notebooks used in this study [https://github.com/dggarciam/PhD\\_code\\_reposotory](https://github.com/dggarciam/PhD_code_reposotory)

# References

- [Abdulkarim & Al-Faiz, 2021] **Abdulkarim, H. & Al-Faiz, M. Z.**: , 2021; Online multiclass EEG feature extraction and recognition using modified convolutional neural network method; *International Journal of Electrical and Computer Engineering (IJECE)*; **11** (5): 4016–4026.
- [Abhang et al., 2016] **Abhang, P. A.; Gawali, B. W. & Mehrotra, S. C.**: , 2016; Chapter 3 - technical aspects of brain rhythms and speech parameters; en *Introduction to EEG- and Speech-Based Emotion Recognition* (Editado por **Abhang, P. A.; Gawali, B. W. & Mehrotra, S. C.**); Academic Press; ISBN 978-0-12-804490-2; págs. 51–79; doi:<https://doi.org/10.1016/B978-0-12-804490-2.00003-8>; URL <https://www.sciencedirect.com/science/article/pii/B9780128044902000038>.
- [Abiri et al., 2019] **Abiri, R.; Borhani, S.; Sellers, E. W.; Jiang, Y. & Zhao, X.**: , 2019; A comprehensive review of EEG-based brain-computer interface paradigms; *Journal of Neural Engineering*; **16** (1): 011001.
- [Ahani et al., 2014] **Ahani, A.; Wahbeh, H.; Nezamfar, H.; Miller, M.; Erdogmus, D. & Oken, B.**: , 2014; Quantitative change of EEG and respiration signals during mindfulness meditation; *Journal of neuroengineering and rehabilitation*; **11** (1): 1–11.
- [Ahirwal et al., 2014] **Ahirwal, M. K.; Kumar, A. & Singh, G. K.**: , 2014; Adaptive filtering of EEG/ERP through bounded range artificial bee colony (BR-ABC) algorithm; *Digital Signal Processing*; **25**: 164–172.
- [Ahn et al., 2014] **Ahn, M.; Lee, M.; Choi, J. & 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 et al., 2019] **Ai, Q.; Chen, A.; Chen, K.; Liu, Q.; Zhou, T.; Xin, S. & Ji, Z.**: , 2019; Feature extraction of four-class motor imagery EEG signals based on functional brain network; *Journal of Neural Engineering*; **16** (2): 026032.
- [Akella et al., 2021] **Akella, A.; Singh, A. K.; Leong, D.; Lal, S.; Newton, P.; Clifton-Bligh, R.; McLachlan, C. S.; Gustin, S. M.; Maharaj, S.; Lees, T. et al.**: , 2021; Classifying multi-level stress responses from brain cortical EEG in nurses and non-health professionals using machine learning auto encoder; *IEEE Journal of Translational Engineering in Health and Medicine*; **9**: 1–9.
- [Akuthota et al., 2023] **Akuthota, S.; Rajkumar, K. & Ravichander, J.**: , 2023; Eeg based motor imagery bci using four class iterative filtering & four class filter bank common spatial pattern; en *2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS)*; IEEE; págs. 429–434.
- [Al-Fahad et al., 2020] **Al-Fahad, R.; Yeasin, M. & Bidelman, G. M.**: , 2020; Decoding of single-trial EEG reveals unique states of functional brain connectivity that drive rapid speech categorization decisions; *Journal of Neural Engineering*; **17** (1): 016045.
- [Al-Nafjan et al., 2017] **Al-Nafjan, A.; Hosny, M.; Al-Ohal, Y. & Al-Wabil, A.**: , 2017; Review and classification of emotion recognition based on EEG brain-computer interface system research: a systematic review; *Applied Sciences*; **7** (12): 1239.
- [Al Nazi et al., 2021] **Al Nazi, Z.; Hossain, A. A. & Islam, M. M.**: , 2021; Motor imagery EEG classification using random subspace ensemble network with variable length features; *International Journal Bioautomation*; **25** (1): 13.
- [Al-Saegh et al., 2021] **Al-Saegh, A.; Dawwd, S. A. & Abdul-Jabbar, J. M.**: , 2021; Deep learning for motor imagery EEG-based classification: A review; *Biomedical Signal Processing and Control*; **63**: 102172.
- [Al-Shargie et al., 2020] **Al-Shargie, F. M.; Hassanin, O.; Tariq, U. & Al-Nashash, H.**: , 2020; EEG-based semantic vigilance level classification using directed connectivity patterns and graph theory analysis; *IEEE Access*; **8**: 115941–115956.
- [Alazrai et al., 2019] **Alazrai, R.; Abuhijleh, M.; Alwanni, H. & Daoud, M. I.**: , 2019; A deep learning framework for decoding motor imagery tasks of the same hand using eeg signals; *IEEE Access*; **7**: 109612–109627.
- [Aldayel et al., 2020] **Aldayel, M.; Ykhlef, M. & Al-Nafjan, A.**: , 2020; Deep learning for EEG-based preference classification in neuromarketing; *Applied Sciences*; **10** (4): 1525.
- [Ali et al., 2018] **Ali, H. T.; Kammoun, A. & Couillet, R.**: , 2018; Random matrix-improved kernels for large dimensional spectral clustering; en *2018 IEEE Statistical Signal Processing Workshop (SSP)*; IEEE; págs. 453–457.
- [Alizadeh et al., 2023] **Alizadeh, N.; Afrakhteh, S. & Mosavi, M.**: , 2023; Multi-task eeg signal classification using correlation-based imf selection and multi-class csp; *IEEE Access*.
- [Alsharif et al., 2020] **Alsharif, A.; Salleh, N.; Baharun, R. & Safaei, M.**: , 2020; Neuromarketing approach: An overview and future research directions; *Journal of Theoretical and Applied Information Technology*; **98** (7): 991–1001.

## REFERENCES

## REFERENCES

- [Altaheri *et al.*, 2023] Altaheri, H.; Muhammad, G.; Alsulaiman, M.; Amin, S. U.; Altwaijri, G. A.; Abdul, W.; Bencherif, M. A. & 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.
- [Álvarez-Meza *et al.*, 2014a] Álvarez-Meza, A.; Cárdenas-Peña, D. & Castellanos-Dominguez, G.: , 2014a; Unsupervised kernel function building using maximization of information potential variability; en *Iberoamerican Congress on Pattern Recognition*; Springer; págs. 335–342.
- [Álvarez-Meza *et al.*, 2017] Álvarez-Meza, A.; Orozco-Gutierrez, A. & Castellanos-Dominguez, G.: , 2017; Kernel-based relevance analysis with enhanced interpretability for detection of brain activity patterns; *Frontiers in neuroscience*; **11**: 550.
- [Álvarez-Meza *et al.*, 2014b] Álvarez-Meza, A. M.; Cárdenas-Peña, D. & Castellanos-Dominguez, G.: , 2014b; Unsupervised kernel function building using maximization of information potential variability; en *Iberoamerican Congress on Pattern Recognition*; Springer; págs. 335–342.
- [Alwasiti *et al.*, 2020] Alwasiti, H.; Yusoff, M. Z. & Raza, K.: , 2020; Motor imagery classification for brain computer interface using deep metric learning; *IEEE Access*; **8**: 109949–109963.
- [Amin *et al.*, 2019] Amin, S. U.; Alsulaiman, M.; Muhammad, G.; Mekhtiche, M. A. & Hossain, M. S.: , 2019; Deep learning for eeg motor imagery classification based on multi-layer cnns feature fusion; *Future Generation computer systems*; **101**: 542–554.
- [Amir *et al.*, 2020] Amir, J.; Amir, R.; Ednaldo Birgante, P. & Md Kafiul, I.: , 2020; Classification of emotions induced by horror and relaxing movies using single-channel EEG recordings.
- [Ang *et al.*, 2008] Ang, K. K.; Chin, Z. Y.; Zhang, H. & Guan, C.: , 2008; Filter bank common spatial pattern (fbcsp) in brain-computer interface; en *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*; IEEE; págs. 2390–2397.
- [Anowar *et al.*, 2021] Anowar, F.; Sadaoui, S. & Selim, B.: , 2021; Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE); *Computer Science Review*; **40**: 100378.
- [Antonakakis *et al.*, 2020] Antonakakis, M.; Schrader, S.; Aydin, Ü.; Khan, A.; Gross, J.; Zervakis, M.; Rampp, S. & Wolters, C. H.: , 2020; Inter-subject variability of skull conductivity and thickness in calibrated realistic head models; *Neuroimage*; **223**: 117353.
- [Anupallavi & MohanBabu, 2020] Anupallavi, S. & MohanBabu, G.: , 2020; A novel approach based on BSPCI for quantifying functional connectivity pattern of the brain's region for the classification of epileptic seizure; *Journal of Ambient Intelligence and Humanized Computing*; 1–11.
- [Apicella *et al.*, 2021] Apicella, A.; Arpaia, P.; Frosolone, M. & Moccaldi, N.: , 2021; High-wearable EEG-based distraction detection in motor rehabilitation; *Scientific Reports*; **11** (1): 1–9.
- [Apicella *et al.*, 2022] Apicella, A.; Isgrò, F.; Pollastro, A. & Prevete, R.: , 2022; Toward the application of xai methods in eeg-based systems; *arXiv preprint arXiv:2210.06554*.
- [Arrieta *et al.*, 2020] Arrieta, A. B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R. *et al.*: , 2020; Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai; *Information fusion*; **58**: 82–115.
- [Arvaneh *et al.*, 2011] Arvaneh, M.; Guan, C.; Ang, K. K. & Quek, C.: , 2011; Optimizing the channel selection and classification accuracy in EEG-based BCI; *IEEE Transactions on Biomedical Engineering*; **58** (6): 1865–1873.
- [Bakhshali *et al.*, 2020] Bakhshali, M.; Khademi, M.; Ebrahimi, A. & Moghimi, S.: , 2020; Eeg signal classification of imagined speech based on riemannian distance of correntropy spectral density; *Biomedical Signal Processing and Control*; **59**: 101899.
- [Bakhshayesh *et al.*, 2019] Bakhshayesh, H.; Fitzgibbon, S. P.; Janani, A. S.; Grummelt, T. S. & Pope, K. J.: , 2019; Detecting synchrony in eeg: A comparative study of functional connectivity measures; *Computers in biology and medicine*; **105**: 1–15.
- [Bang & Lee, 2022] Bang, J.-S. & Lee, S.-W.: , 2022; Interpretable convolutional neural networks for subject-independent motor imagery classification; en *2022 10th International Winter Conference on Brain-Computer Interface (BCI)*; IEEE; págs. 1–5.
- [Baniqued *et al.*, 2021] Baniqued, P. D. E.; Stanyer, E. C.; Awais, M.; Alazmani, A.; Jackson, A. E.; Mon-Williams, M. A.; Mushtaq, F. & Holt, R. J.: , 2021; Brain-computer interface robotics for hand rehabilitation after stroke: a systematic review; *Journal of NeuroEngineering and Rehabilitation*; **18** (1): 1–25.
- [Barachant *et al.*, 2013] Barachant, A.; Bonnet, S.; Congedo, M. & Jutten, C.: , 2013; Classification of covariance matrices using a riemannian-based kernel for BCI applications; *Neurocomputing*; **112**: 172–178.
- [Barrios *et al.*, 2019] Barrios, J. A.; Ezquerro, S.; Bertomeu-Motos, A.; Nann, M.; Badesa, F. J.; Fernandez, E.; Soekadar, S. R. & 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.
- [Bastos & Schoffelen, 2016] Bastos, A. M. & Schoffelen, J.-M.: , 2016; A tutorial review of functional connectivity analysis methods and their interpretational pitfalls; *Frontiers in systems neuroscience*; **9**: 175.
- [Batres-Mendoza *et al.*, 2016] Batres-Mendoza, P.; Montoro-Sanjose, C. R.; Guerra-Hernandez, E. I.; Almanza-Ojeda, D. L.; Rostro-Gonzalez, H.; Romero-Troncoso, R. J. & Ibarra-Manzano, M. A.: , 2016; Quaternion-based signal analysis for motor imagery classification from electroencephalographic signals; *Sensors*; **16** (3): 336.

- [Battaglia & Brovelli, 2020] **Battaglia, D. & Brovelli, A.**: , 2020; Functional connectivity and neuronal dynamics: insights from computational methods; *The Cognitive Neurosciences*.
- [Belwafi et al., 2020] **Belwafi, K.; Gannouni, S. & Aboalsamh, H.**: , 2020; An effective zeros-time windowing strategy to detect sensorimotor rhythms related to motor imagery eeg signals; *IEEE Access*; **8**: 152669–152679.
- [Benzy et al., 2020] **Benzy, V.; Vinod, A.; Subasree, R.; Alladi, S. & Raghavendra, K.**: , 2020; Motor imagery hand movement direction decoding using brain computer interface to aid stroke recovery and rehabilitation; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **28** (12): 3051–3062.
- [Betzel et al., 2019] **Betzel, R. F.; Bertolero, M. A.; Gordon, E. M.; Gratton, C.; Dosenbach, N. U. & Bassett, D. S.**: , 2019; The community structure of functional brain networks exhibits scale-specific patterns of inter-and intra-subject variability; *Neuroimage*; **202**: 115990.
- [Billinger et al., 2013] **Billinger, M.; Brunner, C. & Müller-Putz, G. R.**: , 2013; Single-trial connectivity estimation for classification of motor imagery data; *Journal of neural engineering*; **10** (4): 046006.
- [Bishop, 2006] **Bishop, C. M.**: , 2006; *Pattern recognition and machine learning*; springer.
- [Bochner, 2020] **Bochner, S.**: , 2020; *Harmonic analysis and the theory of probability*; University of California press.
- [Bonci et al., 2021] **Bonci, A.; Fiori, S.; Higashi, H.; Tanaka, T. & Verdini, F.**: , 2021; An introductory tutorial on brain–computer interfaces and their applications; *Electronics*; **10** (5): 560.
- [Borra et al., 2019] **Borra, D.; Fantozzi, S. & Magosso, E.**: , 2019; Eeg motor execution decoding via interpretable sinc-convolutional neural networks; en *Mediterranean Conference on Medical and Biological Engineering and Computing*; Springer; págs. 1113–1122.
- [Borra et al., 2020] **Borra, D.; Fantozzi, S. & Magosso, E.**: , 2020; Interpretable and lightweight convolutional neural network for EEG decoding: application to movement execution and imagination; *Neural Networks*; **129**: 55–74.
- [Braga-Neto & Dougherty, 2004] **Braga-Neto, U. M. & Dougherty, E. R.**: , 2004; Is cross-validation valid for small-sample microarray classification?; *Bioinformatics*; **20** (3): 374–380.
- [Bramhall et al., 2020] **Bramhall, S.; Horn, H.; Tieu, M. & Lohia, N.**: , 2020; Qlime-a quadratic local interpretable model-agnostic explanation approach; *SMU Data Science Review*; **3** (1): 4.
- [Brockmeier et al., 2014] **Brockmeier, A.; Choi, J.; Kriminger, E.; Francis, J. & Principe, J.**: , 2014; Neural decoding with kernel-based metric learning; *Neural computation*; **26** (6): 1080–1107.
- [Bromiley et al., 2004] **Bromiley, P.; Thacker, N. & Bouhova-Thacker, E.**: , 2004; Shannon entropy, renyi entropy, and information; statistics and inf; *Series (2004-004)*.
- [Brunner et al., 2008] **Brunner, C.; Leeb, R.; Müller-Putz, G.; Schlögl, A. & Pfurtscheller, G.**: , 2008; Bci competition 2008-graz data set a; *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*; **16**: 1–6.
- [Brunner et al., 2015] **Brunner, C.; Birbaumer, N.; Blankertz, B.; Guger, C.; Kübler, A.; Mattia, D.; Millán, J. d. R.; Miralles, F.; Nijholt, A.; Opisso, E. et al.**: , 2015; Bnci horizon 2020: towards a roadmap for the bci community; *Brain-computer interfaces*; **2** (1): 1–10.
- [Brusini et al., 2021] **Brusini, L.; Stival, F.; Setti, F.; Menegatti, E.; Menegaz, G. & Storti, S. F.**: , 2021; A systematic review on motor-imagery brain-connectivity-based computer interfaces; *IEEE Transactions on Human-Machine Systems*; **51** (6): 725–733.
- [Buchanan et al., 2021] **Buchanan, D. M.; Ros, T. & Nahas, R.**: , 2021; Elevated and slowed EEG oscillations in patients with post-concussive syndrome and chronic pain following a motor vehicle collision; *Brain Sciences*; **11** (5): 537.
- [Bullmore & Sporns, 2009] **Bullmore, E. & Sporns, O.**: , 2009; Complex brain networks: graph theoretical analysis of structural and functional systems; *Nature reviews neuroscience*; **10** (3): 186–198.
- [Cai et al., 2021] **Cai, Q.; Gong, W.; Deng, Y. & Wang, H.**: , 2021; Single-trial EEG classification via common spatial patterns with mixed lp-and lq-norms; *Mathematical Problems in Engineering*; **2021**.
- [Caicedo-Acosta et al., 2021] **Caicedo-Acosta, J.; Castaño, G. A.; Acosta-Medina, C.; Alvarez-Meza, A. & Castellanos-Dominguez, G.**: , 2021; Deep neural regression prediction of motor imagery skills using EEG functional connectivity indicators; *Sensors*; **21** (6): 1932.
- [Cao et al., 2022a] **Cao, J.; Zhao, Y.; Shan, X.; Wei, H.-l.; Guo, Y.; Chen, L.; Erkoyuncu, J. A. & Sarrigiannis, P. G.**: , 2022a; Brain functional and effective connectivity based on electroencephalography recordings: A review; *Human brain mapping*; **43** (2): 860–879.
- [Cao et al., 2022b] **Cao, L.; Wang, W.; Huang, C.; Xu, Z.; Wang, H.; Jia, J.; Chen, S.; Dong, Y.; Fan, C. & 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.
- [Cárdenas-Peña et al., 2017] **Cárdenas-Peña, D.; Collazos-Huertas, D. & Castellanos-Dominguez, G.**: , 2017; Enhanced data representation by kernel metric learning for dementia diagnosis; *Frontiers in neuroscience*; **11**: 413.
- [Carrara & Papadopoulou, 2023] **Carrara, I. & Papadopoulou, T.**: , 2023; Classification of bci-eeg based on augmented covariance matrix; *arXiv preprint arXiv:2302.04508*.

## REFERENCES

## REFERENCES

- [Cattai *et al.*, 2021] Cattai, T.; Colonnese, S.; Corsi, M.-C.; Bassett, D. S.; Scarano, G. & Fallani, F. D. V.: , 2021; Phase/amplitude synchronization of brain signals during motor imagery bci tasks; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **29**: 1168–1177.
- [Cattan *et al.*, 2018] Cattan, G.; Mendoza, C.; Andreev, A. & Congedo, M.: , 2018; Recommendations for integrating a p300-based brain computer interface in virtual reality environments for gaming; *Computers*; **7** (2): 34.
- [Chaddad *et al.*, 2023] Chaddad, A.; Peng, J.; Xu, J. & Bouridane, A.: , 2023; Survey of explainable ai techniques in healthcare; *Sensors*; **23** (2): 634.
- [Chakraborty *et al.*, 2017] Chakraborty, S.; Tomsett, R.; Raghavendra, R.; Harborne, D.; Alzantot, M.; Cerutti, F.; Srivastava, M.; Preece, A.; Julier, S.; Rao, R. M. *et al.*: , 2017; Interpretability of deep learning models: A survey of results; en *2017 IEEE smartworld, ubiquitous intelligence & computing, advanced & trusted computed, scalable computing & communications, cloud & big data computing, Internet of people and smart city innovation (smartworld/SCAL-COM/UIC/ATC/CBDcom/IOP/SCI)*; IEEE; págs. 1–6.
- [Chamola *et al.*, 2020] Chamola, V.; Vineet, A.; Nayyar, A. & Hossain, E.: , 2020; Brain-computer interface-based humanoid control: A review; *Sensors*; **20** (13): 3620.
- [Chao & Liu, 2020] Chao, H. & Liu, Y.: , 2020; Emotion recognition from multi-channel EEG signals by exploiting the deep belief-conditional random field framework; *IEEE Access*; **8**: 33002–33012.
- [Chattpadhyay *et al.*, 2018] Chattpadhyay, A.; Sarkar, A.; Howlader, P. & Balasubramanian, V. N.: , 2018; Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks; en *2018 IEEE winter conference on applications of computer vision (WACV)*; IEEE; págs. 839–847.
- [Chen *et al.*, 2020] Chen, Y.; Hang, W.; Liang, S.; Liu, X.; Li, G.; Wang, Q.; Qin, J. & Choi, K.-S.: , 2020; A novel transfer support matrix machine for motor imagery-based brain computer interface; *Frontiers in Neuroscience*; **14**.
- [Chen *et al.*, 2021] Chen, Y. Y.; Lambert, K. J.; Madan, C. R. & Singhal, A.: , 2021; Mu oscillations and motor imagery performance: A reflection of intra-individual success, not inter-individual ability; *Human Movement Science*; **78**: 102819.
- [Chen *et al.*, 2019] Chen, Z.; Ji, J. & Liang, Y.: , 2019; Convolutional neural network with an element-wise filter to classify dynamic functional connectivity; en *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; IEEE; págs. 643–646.
- [Chiariion *et al.*, 2023] Chiariion, G.; Sparacino, L.; Antonacci, Y.; Faes, L. & 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 *et al.*, 2017] Cho, H.; Ahn, M.; Ahn, S.; Kwon, M. & Jun, S.: , 2017; EEG datasets for motor imagery brain–computer interface; *GigaScience*; **6** (7): gix034.
- [Choi & Kim, 2021] Choi, I. & Kim, W. C.: , 2021; Detecting and analyzing politically-themed stocks using text mining techniques and transfer entropy—focus on the republic of korea’s case; *Entropy*; **23** (6): 734.
- [Clark *et al.*, 2020] Clark, S. V.; King, T. Z. & Turner, J. A.: , 2020; Cerebellar contributions to proactive and reactive control in the stop signal task: A systematic review and meta-analysis of functional magnetic resonance imaging studies; *Neuropsychology review*; 1–24.
- [Coelho *et al.*, 2012] Coelho, G. P.; Barbante, C. C.; Boccato, L.; Attux, R. R.; Oliveira, J. R. & Von Zuben, F. J.: , 2012; Automatic feature selection for BCI: an analysis using the davies-bouldin index and extreme learning machines; en *The 2012 International Joint Conference on Neural Networks (IJCNN)*; IEEE; págs. 1–8.
- [Cohen, 1998] Cohen, L.: , 1998; The generalization of the wiener-khinchin theorem; en *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, tomo 3; IEEE; págs. 1577–1580.
- [Collazos-Huertas *et al.*, 2020] Collazos-Huertas, D.; Álvarez-Meza, A.; Acosta-Medina, C.; Castaño-Duque, G. & Castellanos-Domínguez, G.: , 2020; CNN-based framework using spatial dropping for enhanced interpretation of neural activity in motor imagery classification; *Brain Informatics*; **7** (1): 1–13.
- [Collazos-Huertas *et al.*, 2021] Collazos-Huertas, D. F.; Velasquez-Martinez, L. F.; Perez-Nastar, H. D.; Alvarez-Meza, A. M. & Castellanos-Domínguez, G.: , 2021; Deep and wide transfer learning with kernel matching for pooling data from electroencephalography and psychological questionnaires; *Sensors*; **21** (15): 5105.
- [Collazos-Huertas *et al.*, 2023] Collazos-Huertas, D. F.; Álvarez-Meza, A. M.; Cárdenas-Peña, D. A.; Castaño-Duque, G. A. & Castellanos-Domínguez, C. G.: , 2023; Posthoc interpretability of neural responses by grouping subject motor imagery skills using cnn-based connectivity; *Sensors*; **23** (5): 2750.
- [Congedo *et al.*, 2017a] Congedo, M.; Barachant, A. & Bhatia, R.: , 2017a; Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review; *Brain-Computer Interfaces*; **4** (3): 155–174.
- [Congedo *et al.*, 2017b] Congedo, M.; Barachant, A. & Koopaei, E. K.: , 2017b; Fixed point algorithms for estimating power means of positive definite matrices; *IEEE Transactions on Signal Processing*; **65** (9): 2211–2220.
- [Congedo *et al.*, 2017c] Congedo, M.; Rodrigues, P. L. C.; Bouchard, F.; Barachant, A. & Jutten, C.: , 2017c; A closed-form unsupervised geometry-aware dimensionality reduction method in the riemannian manifold of SPD matrices; en *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; IEEE; págs. 3198–3201.
- [Craik *et al.*, 2019] Craik, A.; He, Y. & Contreras-Vidal, J. L.: , 2019; Deep learning for electroencephalogram (EEG) classification tasks: a review; *Journal of Neural Engineering*; **16** (3): 031001.

## REFERENCES

## REFERENCES

- [Dai et al., 2020] **Dai, H.; Su, S.; Zhang, Y. & Jian, W.**: , 2020; Effect of spatial filtering and channel selection on motor imagery BCI; en *Proceedings of the 2020 Conference on Artificial Intelligence and Healthcare*; págs. 270–274.
- [Dai et al., 2019] **Dai, M.; Zheng, D.; Na, R.; Wang, S. & Zhang, S.**: , 2019; Eeg classification of motor imagery using a novel deep learning framework; *Sensors*; **19** (3): 551.
- [Daly et al., 2012] **Daly, I.; Nasuto, S. J. & Warwick, K.**: , 2012; Brain computer interface control via functional connectivity dynamics; *Pattern recognition*; **45** (6): 2123–2136.
- [Darvish Ghanbar et al., 2021] **Darvish Ghanbar, K.; Yousefi Rezaii, T.; Farzamnia, A. & Saad, I.**: , 2021; Correlation-based common spatial pattern (ccsp): A novel extension of csp for classification of motor imagery signal; *Plos one*; **16** (3): e0248511.
- [De La Pava Panche et al., 2019] **De La Pava Panche, I.; Alvarez-Meza, A. & Orozco-Gutierrez, A.**: , 2019; A data-driven measure of effective connectivity based on renyi's  $\alpha$ -entropy; *Frontiers in neuroscience*; **13**: 1277.
- [Demir et al., 2022] **Demir, A.; Koike-Akino, T.; Wang, Y. & Erdoğmus, D.**: , 2022; Eeg-gat: graph attention networks for classification of electroencephalogram (eeg) signals; en *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*; IEEE; págs. 30–35.
- [Deng et al., 2020] **Deng, Y.; Li, Z.; Wang, H.; Lu, X. & Fan, H.**: , 2020; Local temporal joint recurrence common spatial patterns for MI-based BCI; en *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, tomo 1; IEEE; págs. 813–816.
- [Desbois et al., 2021] **Desbois, A.; Cattai, T.; Corsi, M.-C. & de Vico Fallani, F.**: , 2021; Functional connectivity for bci: Openvibe implementation; en *JJC-ICON'2021-Journée Jeunes Chercheurs en Interfaces Cerveau-Ordinateur et Neurofeedback*.
- [Ding et al., 2023] **Ding, X.; Yang, L. & Li, C.**: , 2023; Study of mi-bci classification method based on the riemannian transform of personalized eeg spatiotemporal features; *Mathematical Biosciences and Engineering*; **20** (7): 12454–12471.
- [Dose et al., 2018] **Dose, H.; Møller, J. S.; Iversen, H. K. & Puthusserypady, S.**: , 2018; An end-to-end deep learning approach to mi-eeg signal classification for bcis; *Expert Systems with Applications*; **114**: 532–542.
- [Duan et al., 2014] **Duan, L.; Zhong, H.; Miao, J.; Yang, Z.; Ma, W. & Zhang, X.**: , 2014; A voting optimized strategy based on ELM for improving classification of motor imagery BCI data; *Cognitive Computation*; **6** (3): 477–483.
- [Duan et al., 2020] **Duan, L.; Duan, H.; Qiao, Y.; Sha, S.; Qi, S.; Zhang, X.; Huang, J.; Huang, X. & Wang, C.**: , 2020; Machine learning approaches for MDD detection and emotion decoding using EEG signals; *Frontiers in Human Neuroscience*; **14**.
- [Duan et al., 2021] **Duan, S.; Yu, S. & Príncipe, J. C.**: , 2021; Modularizing deep learning via pairwise learning with kernels; *IEEE Transactions on Neural Networks and Learning Systems*.
- [Duclos et al., 2021] **Duclos, C.; Maschke, C.; Mahdid, Y.; Berkun, K.; da Silva Castanheira, J.; Tarnal, V.; Picton, P.; Vanini, G.; Golmirzaie, G.; Janke, E. et al.**: , 2021; Differential classification of states of consciousness using envelope-and phase-based functional connectivity; *NeuroImage*; **237**: 118171.
- [Eichert et al., 2021] **Eichert, N.; Watkins, K. E.; Mars, R. B. & Petrides, M.**: , 2021; Morphological and functional variability in central and subcentral motor cortex of the human brain; *Brain Structure and Function*; **226** (1): 263–279.
- [EPMoghaddam et al., 2022] **EPMoghaddam, D.; Sheth, S. A.; Haneef, Z.; Gavvala, J. & Aazhang, B.**: , 2022; Epileptic seizure prediction using spectral width of the covariance matrix; *Journal of Neural Engineering*; **19** (2): 026029.
- [Fagerholm et al., 2020] **Fagerholm, E. D.; Moran, R. J.; Violante, I. R.; Leech, R. & Friston, K. J.**: , 2020; Dynamic causal modelling of phase-amplitude interactions; *NeuroImage*; **208**: 116452.
- [Fahimi et al., 2020] **Fahimi, F.; Dosen, S.; Ang, K. K.; Mrachacz-Kersting, N. & Guan, C.**: , 2020; Generative adversarial networks-based data augmentation for brain-computer interface; *IEEE transactions on neural networks and learning systems*; **32** (9): 4039–4051.
- [Fan et al., 2021] **Fan, F.-L.; Xiong, J.; Li, M. & Wang, G.**: , 2021; On interpretability of artificial neural networks: A survey; *IEEE Transactions on Radiation and Plasma Medical Sciences*; **5** (6): 741–760.
- [Fauzi et al., 2019] **Fauzi, H.; Azzam, M. A.; Shapiai, M. I.; Kyoso, M.; Khairuddin, U. & Komura, T.**: , 2019; Energy extraction method for EEG channel selection; *Telkomnika*; **17** (5): 2561–2571.
- [Feng et al., 2020] **Feng, Z.; Qian, L.; Hu, H. & Sun, Y.**: , 2020; Functional connectivity for motor imaginary recognition in brain-computer interface; en *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*; IEEE; págs. 3678–3682.
- [Fine & Scheinberg, 2001] **Fine, S. & Scheinberg, K.**: , 2001; Efficient SVM training using low-rank kernel representations; *Journal of Machine Learning Research*; **2** (Dec): 243–264.
- [Fogelson et al., 2013] **Fogelson, N.; Li, L.; Li, Y.; Fernandez-del Olmo, M.; Santos-Garcia, D. & Peled, A.**: , 2013; Functional connectivity abnormalities during contextual processing in schizophrenia and in parkinson's disease; *Brain and cognition*; **82** (3): 243–253.
- [Fong & Vedaldi, 2017] **Fong, R. C. & Vedaldi, A.**: , 2017; Interpretable explanations of black boxes by meaningful perturbation; en *Proceedings of the IEEE international conference on computer vision*; págs. 3429–3437.
- [Frau-Pascual et al., 2019] **Frau-Pascual, A.; Fogarty, M.; Fischl, B.; Yendiki, A.; Aganj, I.; Initiative, A. D. N. et al.**: , 2019; Quantification of structural brain connectivity via a conductance model; *NeuroImage*; **189**: 485–496.

- [Freer *et al.*, 2019] **Freer, D.; Deligianni, F. & Yang, G.-Z.**: , 2019; Adaptive riemannian bci for enhanced motor imagery training protocols; en *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*; IEEE; págs. 1–4.
- [Friston, 2002] **Friston, K.**: , 2002; Functional integration and inference in the brain; *Progress in neurobiology*; **68** (2): 113–143.
- [Friston *et al.*, 2013] **Friston, K.; Moran, R. & Seth, A. K.**: , 2013; Analysing connectivity with granger causality and dynamic causal modelling; *Current opinion in neurobiology*; **23** (2): 172–178.
- [Friston, 2011] **Friston, K. J.**: , 2011; Functional and effective connectivity: a review; *Brain connectivity*; **1** (1): 13–36.
- [Frolov *et al.*, 2019] **Frolov, N.; Maksimenko, V.; Lüttjohann, A.; Koronovskii, A. & Hramov, A.**: , 2019; Feed-forward artificial neural network provides data-driven inference of functional connectivity; *Chaos: An Interdisciplinary Journal of Nonlinear Science*; **29** (9): 091101.
- [Fu *et al.*, 2020] **Fu, R.; Han, M.; Tian, Y. & Shi, P.**: , 2020; Improvement motor imagery EEG classification based on sparse common spatial pattern and regularized discriminant analysis; *Journal of Neuroscience Methods*; **343**: 108833.
- [Ganin *et al.*, 2013] **Ganin, I. P.; Shishkin, S. L. & Kaplan, A. Y.**: , 2013; A p300-based brain-computer interface with stimuli on moving objects: four-session single-trial and triple-trial tests with a game-like task design; *PLoS one*; **8** (10): e77755.
- [Gao *et al.*, 2024] **Gao, H.; Wang, X.; Chen, Z.; Wu, M.; Cai, Z.; Zhao, L.; Li, J. & Liu, C.**: , 2024; Graph convolutional network with connectivity uncertainty for eeg-based emotion recognition; *IEEE Journal of Biomedical and Health Informatics*.
- [Gao *et al.*, 2022] **Gao, S.; Yang, J.; Shen, T. & Jiang, W.**: , 2022; A parallel feature fusion network combining gru and cnn for motor imagery eeg decoding; *Brain Sciences*; **12** (9): 1233.
- [Gao *et al.*, 2021] **Gao, Z.; Dang, W.; Wang, X.; Hong, X.; Hou, L.; Ma, K. & Perc, M.**: , 2021; Complex networks and deep learning for EEG signal analysis; *Cognitive Neurodynamics*; **15** (3): 369–388.
- [García-Murillo *et al.*, 2021] **García-Murillo, D. G.; Alvarez-Meza, A. & Castellanos-Dominguez, G.**: , 2021; Single-trial kernel-based functional connectivity for enhanced feature extraction in motor-related tasks; *Sensors*; **21** (8): 2750.
- [Garg *et al.*, 2023] **Garg, D.; Verma, G. K. & Singh, A. K.**: , 2023; A review of deep learning based methods for affect analysis using physiological signals; *Multimedia Tools and Applications*; 1–46.
- [Gaur *et al.*, 2021a] **Gaur, P.; Gupta, H.; Chowdhury, A.; McCreadie, K.; Pachori, R. & Wang, H.**: , 2021a; A sliding window common spatial pattern for enhancing motor imagery classification in eeg-bci; *IEEE Transactions on Instrumentation and Measurement*; **70**: 1–9.
- [Gaur *et al.*, 2021b] **Gaur, P.; Gupta, H.; Chowdhury, A.; McCreadie, K.; Pachori, R. B. & Wang, H.**: , 2021b; A sliding window common spatial pattern for enhancing motor imagery classification in eeg-bci; *IEEE Transactions on Instrumentation and Measurement*; **70**: 1–9.
- [Gaur *et al.*, 2021c] **Gaur, P.; McCreadie, K.; Pachori, R. B.; Wang, H. & Prasad, G.**: , 2021c; An automatic subject specific channel selection method for enhancing motor imagery classification in EEG-BCI using correlation; *Biomedical Signal Processing and Control*; **68**: 102574.
- [Gaxiola-Tirado *et al.*, 2017] **Gaxiola-Tirado, J. A.; Salazar-Varas, R. & Gutiérrez, D.**: , 2017; Using the partial directed coherence to assess functional connectivity in electroencephalography data for brain–computer interfaces; *IEEE Transactions on Cognitive and Developmental Systems*; **10** (3): 776–783.
- [Georgiadis *et al.*, 2018] **Georgiadis, K.; Laskaris, N.; Nikolopoulos, S. & Kompatsiaris, I.**: , 2018; Exploiting the heightened phase synchrony in patients with neuromuscular disease for the establishment of efficient motor imagery bcis; *Journal of neuroengineering and rehabilitation*; **15** (1): 1–18.
- [Georgiadis *et al.*, 2019] **Georgiadis, K.; Laskaris, N.; Nikolopoulos, S. & Kompatsiaris, I.**: , 2019; Connectivity steered graph fourier transform for motor imagery bci decoding; *Journal of neural engineering*; **16** (5): 056021.
- [Géron, 2022] **Géron, A.**: , 2022; *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*; " O'Reilly Media, Inc.".
- [Ghanbar *et al.*, 2021] **Ghanbar, K. D.; Rezaii, T. Y.; Farzamnia, A. & Saad, I.**: , 2021; Correlation-based common spatial pattern (ccsp): A novel extension of csp for classification of motor imagery signal; *PLoS One*; **16** (3): 1–18.
- [Ghane & Hossain, 2020] **Ghane, P. & Hossain, G.**: , 2020; Learning patterns in imaginary vowels for an intelligent brain computer interface (BCI) design; *arXiv preprint arXiv:2010.12066*.
- [Giraldo *et al.*, 2014] **Giraldo, L. G. S.; Rao, M. & Principe, J. C.**: , 2014; Measures of entropy from data using infinitely divisible kernels; *IEEE Transactions on Information Theory*; **61** (1): 535–548.
- [Golugula *et al.*, 2011] **Golugula, A.; Lee, G. & Madabhushi, A.**: , 2011; Evaluating feature selection strategies for high dimensional, small sample size datasets; en *2011 Annual International conference of the IEEE engineering in medicine and biology society*; IEEE; págs. 949–952.
- [Gonuguntla *et al.*, 2016] **Gonuguntla, V.; Wang, Y. & Veluvolu, K. C.**: , 2016; Event-related functional network identification: application to EEG classification; *IEEE journal of selected topics in signal processing*; **10** (7): 1284–1294.
- [Gonzalez-Astudillo *et al.*, 2020] **Gonzalez-Astudillo, J.; Cattai, T.; Bassignana, G.; Corsi, M.-C. & Fallani, F. D. V.**: , 2020; Network-based brain computer interfaces: principles and applications; *Journal of Neural Engineering*.
- [Gonzalez-Astudillo *et al.*, 2021] **Gonzalez-Astudillo, J.; Cattai, T.; Bassignana, G.; Corsi, M.-C. & Fallani, F. D. V.**: , 2021; Network-based brain–computer interfaces: principles and applications; *Journal of neural engineering*; **18** (1): 011001.

- [Gramfort & Clerc, 2007] Gramfort, A. & Clerc, M.: , 2007; Low dimensional representations of MEG/EEG data using laplacian eigenmaps; en *2007 Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging*; IEEE; págs. 169–172.
- [Graña & Morais-Quilez, 2023] Graña, M. & Morais-Quilez, I.: , 2023; A review of graph neural networks for electroencephalography data analysis; *Neurocomputing*: 126901.
- [Grigorev et al., 2021] Grigorev, N. A.; Savosenkov, A. O.; Lukoyanov, M. V.; Udotatina, A.; Shusharina, N. N.; Kaplan, A. Y.; Hramov, A. E.; Kazantsev, V. B. & 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.
- [Grilo et al., 2019] Grilo, M.; Ribeiro, L.; Moraes, C.; Melo, C.; Fantinato, D.; Sampaio, L.; Neves, A. & Ramos, R.: , 2019; Artifact removal in EEG based emotional signals through linear and nonlinear methods; en *2019 E-Health and Bioengineering Conference (EHB)*; IEEE; págs. 1–4.
- [Gu et al., 2023] Gu, H.; Wang, J. & Han, Y.: , 2023; Decoding of brain functional connections underlying natural grasp task using time-frequency cross mutual information; *IEEE Access*.
- [Gu et al., 2021a] Gu, J.; Wei, M.; Guo, Y. & Wang, H.: , 2021a; Common spatial pattern with l21-norm; *Neural Processing Letters*: 1–20.
- [Gu et al., 2020a] Gu, L.; Yu, Z.; Ma, T.; Wang, H.; Li, Z. & Fan, H.: , 2020a; EEG-based classification of lower limb motor imagery with brain network analysis; *Neuroscience*; **436**: 93–109.
- [Gu et al., 2020b] Gu, L.; Yu, Z.; Ma, T.; Wang, H.; Li, Z. & Fan, H.: , 2020b; Random matrix theory for analysing the brain functional network in lower limb motor imagery; en *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*; IEEE; págs. 506–509.
- [Gu et al., 2021b] Gu, X.; Cao, Z.; Jolfaei, A.; Xu, P.; Wu, D.; Jung, T.-P. & Lin, C.-T.: , 2021b; Eeg-based brain-computer interfaces (bcis): A survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications; *IEEE/ACM transactions on computational biology and bioinformatics*; **18** (5): 1645–1666.
- [Guillot & Debarnot, 2019] Guillot, A. & Debarnot, U.: , 2019; Benefits of motor imagery for human space flight: a brief review of current knowledge and future applications; *Frontiers in Physiology*; **10**: 396.
- [Gunawardena et al., 2023] Gunawardena, R.; Sarrigiannis, P. G.; Blackburn, D. J. & He, F.: , 2023; Kernel-based nonlinear manifold learning for eeg-based functional connectivity analysis and channel selection with application to alzheimer's disease; *Neuroscience*.
- [Guo et al., 2020] Guo, Y.; Zhang, Y.; Chen, Z.; Liu, Y. & Chen, W.: , 2020; Eeg classification by filter band component regularized common spatial pattern for motor imagery; *Biomedical Signal Processing and Control*; **59**: 101917.
- [Gupta et al., 2015] Gupta, A.; Agrawal, R. & Kaur, B.: , 2015; Performance enhancement of mental task classification using EEG signal: a study of multivariate feature selection methods; *Soft Computing*; **19** (10): 2799–2812.
- [Haddad et al., 2019] Haddad, A.; Shamsi, F.; Ghovanloo, M. & Najafizadeh, L.: , 2019; Early decoding of tongue-hand movement from EEG recordings using dynamic functional connectivity graphs; en *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*; IEEE; págs. 373–376.
- [Haghaniifar et al., 2022] Haghaniifar, A.; Majdabadi, M. M.; Choi, Y.; Deivalakshmi, S. & Ko, S.: , 2022; Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning; *Multimedia Tools and Applications*; **81** (21): 30615–30645.
- [Hamedi et al., 2014] Hamedi, M.; Salleh, S.-H.; Noor, A. M. & Mohammad-Rezazadeh, I.: , 2014; Neural network-based three-class motor imagery classification using time-domain features for bci applications; en *2014 IEEE region 10 symposium*; IEEE; págs. 204–207.
- [Hamedi et al., 2016] Hamedi, M.; Salleh, S.-H. & Noor, A. M.: , 2016; Electroencephalographic motor imagery brain connectivity analysis for BCI: a review; *Neural computation*; **28** (6): 999–1041.
- [Han et al., 2023] Han, S.; Zhang, C.; Lei, J.; Han, Q.; Du, Y.; Wang, A.; Bai, S. & Zhang, M.: , 2023; Cepstral analysis based artifact detection, recognition and removal for prefrontal eeg; *IEEE Transactions on Circuits and Systems II: Express Briefs*.
- [Hang et al., 2020] Hang, W.; Feng, W.; Liang, S.; Wang, Q.; Liu, X. & Choi, K.-S.: , 2020; Deep stacked support matrix machine based representation learning for motor imagery EEG classification; *Computer methods and programs in biomedicine*; **193**: 105466.
- [Hassanpour et al., 2019] Hassanpour, A.; Moradikia, M.; Adeli, H.; Khayami, S. R. & Shamsinejadbabaki, P.: , 2019; A novel end-to-end deep learning scheme for classifying multi-class motor imagery electroencephalography signals; *Expert Systems*; **36** (6): e12494.
- [Hassija et al., 2023] Hassija, V.; Chamola, V.; Mahapatra, A.; Singal, A.; Goel, D.; Huang, K.; Scardapane, S.; Spinelli, I.; Mahmud, M. & Hussain, A.: , 2023; Interpreting black-box models: a review on explainable artificial intelligence; *Cognitive Computation*: 1–30.
- [Hata et al., 2016] Hata, M.; Kazui, H.; Tanaka, T.; Ishii, R.; Canuet, L.; Pascual-Marqui, R. D.; Aoki, Y.; Ikeda, S.; Kanemoto, H.; Yoshiyama, K. et al.: , 2016; Functional connectivity assessed by resting state eeg correlates with cognitive decline of alzheimer's disease—an eloreta study; *Clinical Neurophysiology*; **127** (2): 1269–1278.
- [He et al., 2019] He, B.; Astolfi, L.; Valdés-Sosa, P. A.; Marinazzo, D.; Palva, S. O.; Bénar, C.-G.; Michel, C. M. & Koenig, T.: , 2019; Electrophysiological brain connectivity: theory and implementation; *IEEE Transactions on Biomedical Engineering*; **66** (7): 2115–2137.

- [He *et al.*, 2013] He, L.; Hu, Y.; Li, Y. & Li, D.: , 2013; Channel selection by rayleigh coefficient maximization based genetic algorithm for classifying single-trial motor imagery EEG; *Neurocomputing*; **121**: 423–433.
- [Hejazi & Motie Nasrabadi, 2019] Hejazi, M. & Motie Nasrabadi, A.: , 2019; Prediction of epilepsy seizure from multi-channel electroencephalogram by effective connectivity analysis using granger causality and directed transfer function methods; *Cognitive neurodynamics*; **13**: 461–473.
- [Herranz-Gómez *et al.*, 2020] Herranz-Gómez, A.; Gaudiosi, C.; Angulo-Díaz-Parreño, S.; Suso-Martí, L.; La Touche, R. & Cuena-Martínez, F.: , 2020; Effectiveness of motor imagery and action observation on functional variables: An umbrella and mapping review with meta-meta-analysis; *Neuroscience & Biobehavioral Reviews*.
- [Hobson & Bishop, 2017] Hobson, H. M. & 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.
- [Hochberg *et al.*, 2012] Hochberg, L. R.; Bacher, D.; Jarosiewicz, B.; Masse, N. Y.; Simeral, J. D.; Vogel, J.; Haddadin, S.; Liu, J.; Cash, S. S.; Van Der Smagt, P. *et al.*: , 2012; Reach and grasp by people with tetraplegia using a neurally controlled robotic arm; *Nature*; **485** (7398): 372–375.
- [Hossain *et al.*, 2023] Hossain, K. M.; Islam, M. A.; Hossain, S.; Nijholt, A. & Ahad, M. A. R.: , 2023; Status of deep learning for eeg-based brain–computer interface applications; *Frontiers in computational neuroscience*; **16**: 1006763.
- [Hosseini *et al.*, 2020a] Hosseini, M.; Powell, M.; Collins, J.; Callahan-Flinton, C.; Jones, W.; Bowman, H. & Wyble, B.: , 2020a; I tried a bunch of things: The dangers of unexpected overfitting in classification of brain data; *Neuroscience & Biobehavioral Reviews*.
- [Hosseini *et al.*, 2020b] Hosseini, M.-P.; Hosseini, A. & Ahi, K.: , 2020b; A review on machine learning for EEG signal processing in bioengineering; *IEEE reviews in biomedical engineering*.
- [Hrisca-Eva & Lazar, 2021] Hrisca-Eva, O.-D. & Lazar, A. M.: , 2021; Multi-sessions outcome for EEG feature extraction and classification methods in a motor imagery task.; *Traitement du Signal*; **38** (2).
- [Hsu & Chen, 2020] Hsu, L. & Chen, Y.-J.: , 2020; Neuromarketing, subliminal advertising, and hotel selection: An EEG study; *Australasian Marketing Journal (AMJ)*; **28** (4): 200–208.
- [Huang *et al.*, 2023] Huang, G.; Zhao, Z.; Zhang, S.; Hu, Z.; Fan, J.; Fu, M.; Chen, J.; Xiao, Y.; Wang, J. & Dan, G.: , 2023; Discrepancy between inter-and intra-subject variability in eeg-based motor imagery brain-computer interface: Evidence from multiple perspectives; *Frontiers in Neuroscience*; **17**: 1122661.
- [Huang *et al.*, 2022] Huang, Q.; Yamada, M.; Tian, Y.; Singh, D. & Chang, Y.: , 2022; Graphlime: Local interpretable model explanations for graph neural networks; *IEEE Transactions on Knowledge and Data Engineering*.
- [Hurtado-Rincón *et al.*, 2016] Hurtado-Rincón, J. V.; Martínez-Vargas, J. D.; Rojas-Jaramillo, S.; Giraldo, E. & Castellanos-Domínguez, G.: , 2016; Identification of relevant inter-channel EEG connectivity patterns: a kernel-based supervised approach; en *International Conference on Brain Informatics*; Springer; págs. 14–23.
- [Idowu *et al.*, 2021] Idowu, O. P.; Ilesanmi, A. E.; Li, X.; Samuel, O. W.; Fang, P. & Li, G.: , 2021; An integrated deep learning model for motor intention recognition of multi-class EEG signals in upper limb amputees; *Computer Methods and Programs in Biomedicine*; **206**: 106121.
- [Ismail & Karwowski, 2020] Ismail, L. E. & Karwowski, W.: , 2020; A graph theory-based modeling of functional brain connectivity based on EEG: A systematic review in the context of neuroergonomics; *IEEE Access*; **8**: 155103–155135.
- [Iturralde *et al.*, 2012] Iturralde, P. A.; Patrone, M.; Lecumberry, F. & Fernández, A.: , 2012; Motor intention recognition in EEG: In pursuit of a relevant feature set; en *Iberoamerican Congress on Pattern Recognition*; Springer; págs. 551–558.
- [Jahromy *et al.*, 2019] Jahromy, F. Z.; Bajoulevand, A. & Daliri, M. R.: , 2019; Statistical algorithms for emotion classification via functional connectivity.; *Journal of Integrative Neuroscience*; **18** (3): 293–297; doi:10.31083/j.jin.2019.03.601; URL <https://app.dimensions.ai/details/publication/pub.1121421951andhttps://jin.imrpress.com/EN/article/downloadArticleFile.do?attachType=PDF&id=1604>.
- [Jain & Zongker, 1997] Jain, A. & Zongker, D.: , 1997; Feature selection: Evaluation, application, and small sample performance; *IEEE transactions on pattern analysis and machine intelligence*; **19** (2): 153–158.
- [Janapati *et al.*, 2023] Janapati, R.; Dalal, V. & Sengupta, R.: , 2023; Advances in modern eeg-bci signal processing: A review; *Materials Today: Proceedings*; **80**: 2563–2566.
- [Jiang *et al.*, 2021a] Jiang, P.-T.; Zhang, C.-B.; Hou, Q.; Cheng, M.-M. & Wei, Y.: , 2021a; Layercam: Exploring hierarchical class activation maps for localization; *IEEE Transactions on Image Processing*; **30**: 5875–5888.
- [Jiang *et al.*, 2021b] Jiang, Y.; Chen, W.; Li, M.; Zhang, T. & You, Y.: , 2021b; Synchroextracting chirplet transform-based epileptic seizures detection using EEG; *Biomedical Signal Processing and Control*; **68**: 102699.
- [Jin *et al.*, 2021] Jin, J.; Sun, H.; Daly, I.; Li, S.; Liu, C.; Wang, X. & Cichocki, A.: , 2021; A novel classification framework using the graph representations of electroencephalogram for motor imagery based brain-computer interface; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **30**: 20–29.
- [Johnson *et al.*, 2018] Johnson, N. N.; Carey, J.; Edelman, B. J.; Doud, A.; Grande, A.; Lakshminarayanan, K. & He, B.: , 2018; Combined rtms and virtual reality brain–computer interface training for motor recovery after stroke; *Journal of neural engineering*; **15** (1): 016009.
- [Ju & Guan, 2023] Ju, C. & Guan, C.: , 2023; Graph neural networks on spd manifolds for motor imagery classification: A perspective from the time–frequency analysis; *IEEE Transactions on Neural Networks and Learning Systems*.

- [Kant et al., 2020] **Kant, P.; Laskar, S. H.; Hazarika, J. & Mahamune, R.**: , 2020; Cwt based transfer learning for motor imagery classification for brain computer interfaces; *Journal of Neuroscience Methods*; **345**: 108886.
- [Kaper et al., 2004] **Kaper, M.; Meinicke, P.; Grossekathofer, U.; Lingner, T. & Ritter, H.**: , 2004; BCI competition 2003-data set IIB: support vector machines for the p300 speller paradigm; *IEEE Transactions on biomedical Engineering*; **51** (6): 1073–1076.
- [Kawanabe et al., 2006] **Kawanabe, M.; Krauledat, M. & Blankertz, B.**: , 2006; *A bayesian approach for adaptive BCI classification*; Citeseer.
- [Khademi et al., 2023] **Khademi, Z.; Ebrahimi, F. & Kordy, H. M.**: , 2023; A review of critical challenges in mi-bci: From conventional to deep learning methods; *Journal of Neuroscience Methods*; **383**: 109736.
- [Khan et al., 2021] **Khan, D. M.; Yahya, N.; Kamel, N. & Faye, I.**: , 2021; Effective connectivity in default mode network for alcoholism diagnosis; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **29**: 796–808.
- [Khan et al., 2020] **Khan, M. A.; Das, R.; Iversen, H. K. & 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*; **103843**.
- [Khosla et al., 2020] **Khosla, A.; Khandnor, P. & Chand, T.**: , 2020; A comparative analysis of signal processing and classification methods for different applications based on EEG signals; *Biocybernetics and Biomedical Engineering*; **40** (2): 649–690.
- [Kim et al., 2022] **Kim, S.-J.; Lee, D.-H. & Lee, S.-W.**: , 2022; Rethinking cnn architecture for enhancing decoding performance of motor imagery-based eeg signals; *IEEE Access*; **10**: 96984–96996.
- [Kim et al., 2019] **Kim, Y.; Lee, S.; Kim, H.; Lee, S.; Lee, S. & Kim, D.**: , 2019; Reduced burden of individual calibration process in brain-computer interface by clustering the subjects based on brain activation; en *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*; IEEE; págs. 2139–2143.
- [Kingma, 2014] **Kingma, D. P.**: , 2014; Adam: A method for stochastic optimization; *arXiv preprint arXiv:1412.6980*.
- [Knapen, 2021] **Knapen, T.**: , 2021; Topographic connectivity reveals task-dependent retinotopic processing throughout the human brain; *Proceedings of the National Academy of Sciences*; **118** (2).
- [Köllöd et al., 2023] **Köllöd, C. M.; Adolf, A.; Iván, K.; Márton, G. & Ulbert, I.**: , 2023; Deep comparisons of neural networks from the eegnet family; *Electronics*; **12** (12): 2743.
- [Kondratyev et al., 2020] **Kondratyev, A.; Schwarz, C. & Horvath, B.**: , 2020; Data anonymisation, outlier detection and fighting overfitting with restricted boltzmann machines; *Outlier Detection and Fighting Overfitting with Restricted Boltzmann Machines* (January 27, 2020).
- [Kostiukovich et al., 2021] **Kostiukovich, K.; Stirenko, S.; Gordienko, N.; Rokovyi, O.; Alienin, O. & Gordienko, Y.**: , 2021; Convolutional and recurrent neural networks for physical action forecasting by brain-computer interface; en *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, tomo 2; IEEE; págs. 973–978.
- [Kotte & Dabbakuti, 2020] **Kotte, S. & Dabbakuti, J. K.**: , 2020; Methods for removal of artifacts from eeg signal: A review; en *Journal of Physics: Conference Series*, tomo 1706; IOP Publishing; pág. 012093.
- [Kraeutner et al., 2016] **Kraeutner, S. N.; MacKenzie, L. A.; Westwood, D. A. & Boe, S. G.**: , 2016; Characterizing skill acquisition through motor imagery with no prior physical practice.; *Journal of Experimental Psychology: Human Perception and Performance*; **42** (2): 257.
- [Kuang, 2021] **Kuang, P.-C.**: , 2021; Measuring information flow among international stock markets: An approach of entropy-based networks on multi time-scales; *Physica A: Statistical Mechanics and its Applications*; **577**: 126068.
- [Kübler, 2020] **Kübler, A.**: , 2020; The history of BCI: From a vision for the future to real support for personhood in people with locked-in syndrome; *Neuroethics*; **13** (2): 163–180.
- [Kulatilleke, 2023] **Kulatilleke, G.**: , 2023; Towards efficient graph neural networks for optimizing illicit dark web interventions.
- [Kulkarni et al., 2019] **Kulkarni, P. M.; Xiao, Z.; Robinson, E. J.; Jami, A. S.; Zhang, J.; Zhou, H.; Henin, S. E.; Liu, A. A.; Osorio, R. S.; Wang, J. et al.**: , 2019; A deep learning approach for real-time detection of sleep spindles; *Journal of neural engineering*; **16** (3): 036004.
- [Kumar et al., 2018] **Kumar, S.; Reddy, T. & Behera, L.**: , 2018; Eeg based motor imagery classification using instantaneous phase difference sequence; en *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*; IEEE; págs. 499–504.
- [Kumar et al., 2019] **Kumar, S.; Sharma, A. & Tsunoda, T.**: , 2019; Brain wave classification using long short-term memory network based OPTICAL predictor; *Scientific reports*; **9** (1): 1–13.
- [Kumar et al., 2021] **Kumar, S.; Sharma, R. & Sharma, A.**: , 2021; Optical+: a frequency-based deep learning scheme for recognizing brain wave signals; *Peerj Computer Science*; **7**: e375.
- [Kwon et al., 2021] **Kwon, B.-H.; Jeong, J.-H. & Lee, S.-W.**: , 2021; Visual motion imagery classification with deep neural network based on functional connectivity; *arXiv preprint arXiv:2103.02851*.
- [Kwon et al., 2019] **Kwon, O.-Y.; Lee, M.-H.; Guan, C. & Lee, S.-W.**: , 2019; Subject-independent brain–computer interfaces based on deep convolutional neural networks; *IEEE transactions on neural networks and learning systems*; **31** (10): 3839–3852.

- [Ladda et al., 2021] Ladda, A. M.; Lebon, F. & Lotze, M.: , 2021; Using motor imagery practice for improving motor performance—a review; *Brain and cognition*; **150**: 105705.
- [Lahane et al., 2019] Lahane, P.; Jagtap, J.; Inamdar, A.; Karne, N. & Dev, R.: , 2019; A review of recent trends in eeg based brain-computer interface; en *2019 International Conference on Computational Intelligence in Data Science (ICCIDIS)*; IEEE; págs. 1–6.
- [Lakshmi et al., 2014] Lakshmi, M. R.; Prasad, T. & Prakash, D. V. C.: , 2014; Survey on EEG signal processing methods; *International Journal of Advanced Research in Computer Science and Software Engineering*; **4** (1).
- [Lawhern et al., 2018] Lawhern, V. J.; Solon, A. J.; Waytowich, N. R.; Gordon, S. M.; Hung, C. P. & Lance, B. J.: , 2018; Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces; *Journal of neural engineering*; **15** (5): 056013.
- [Le et al., 2021] Le, N. Q. K.; Kha, Q. H.; Nguyen, V. H.; Chen, Y.-C.; Cheng, S.-J. & Chen, C.-Y.: , 2021; Machine learning-based radiomics signatures for egfr and kras mutations prediction in non-small-cell lung cancer; *International journal of molecular sciences*; **22** (17): 9254.
- [Lebedev & Nicolelis, 2006] Lebedev, M. A. & Nicolelis, M. A.: , 2006; Brain–machine interfaces: past, present and future; *TRENDS in Neurosciences*; **29** (9): 536–546.
- [Ledoit & Wolf, 2004] Ledoit, O. & Wolf, M.: , 2004; A well-conditioned estimator for large-dimensional covariance matrices; *Journal of multivariate analysis*; **88** (2): 365–411.
- [Lee et al., 2004] Lee, F.; Scherer, R.; Leeb, R.; Schlägl, A.; Bischof, H. & Pfurtscheller, G.: , 2004; Feature mapping using PCA, locally linear embedding and isometric feature mapping for EEG-based brain computer interface; Citeseer.
- [Lee & Choi, 2019] Lee, H. K. & Choi, Y.-S.: , 2019; Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain-computer interface; *Entropy*; **21** (12): 1199.
- [Lee et al., 2020] Lee, M.; Yoon, J.-G. & Lee, S.-W.: , 2020; Predicting motor imagery performance from resting-state eeg using dynamic causal modeling; *Frontiers in human neuroscience*; **14**: 321.
- [Leeuwis et al., 2021] Leeuwis, N.; Yoon, S. & Alimardani, M.: , 2021; Functional connectivity analysis in motor-imagery brain computer interfaces; *Frontiers in Human Neuroscience*; **15**: 732946.
- [Li et al., 2021a] Li, B.; Cheng, T. & Guo, Z.: , 2021a; A review of eeg acquisition, processing and application; en *Journal of Physics: Conference Series*, tomo 1907; IOP Publishing; pág. 012045.
- [Li et al., 2015] Li, J.; Struzik, Z.; Zhang, L. & Cichocki, A.: , 2015; Feature learning from incomplete eeg with denoising autoencoder; *Neurocomputing*; **165**: 23–31.
- [Li et al., 2019a] Li, J. W.; Barma, S.; Mak, P. U.; Pun, S. H. & Vai, M. I.: , 2019a; Brain rhythm sequencing using eeg signals: A case study on seizure detection; *IEEE Access*; **7**: 160112–160124.
- [Li & Principe, 2020] Li, K. & Principe, J. C.: , 2020; Fast estimation of information theoretic learning descriptors using explicit inner product spaces; *arXiv preprint arXiv:2001.00265*.
- [Li & Chen, 2021] Li, M. & Chen, W.: , 2021; FFT-based deep feature learning method for EEG classification; *Biomedical Signal Processing and Control*; **66**: 102492.
- [Li et al., 2016] Li, M.; Luo, X.; Yang, J. & Sun, Y.: , 2016; Applying a locally linear embedding algorithm for feature extraction and visualization of MI-EEG; *Journal of Sensors*; **2016**.
- [Li et al., 2019b] Li, M.-A.; Han, J.-F. & Duan, L.-J.: , 2019b; A novel mi-eeg imaging with the location information of electrodes; *IEEE Access*; **8**: 3197–3211.
- [Li, 2022] Li, P.: , 2022; Bayesian networks for brain-computer interfaces: A survey; *arXiv preprint arXiv:2206.07487*.
- [Li et al., 2019c] Li, S.; Xie, X.; Gu, Z.; Yu, Z. L. & Li, Y.: , 2019c; Motor imagery classification based on local isometric embedding of riemannian manifold; en *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*; IEEE; págs. 2368–2372.
- [Li et al., 2019d] Li, Y.; Zhang, X.-R.; Zhang, B.; Lei, M.-Y.; Cui, W.-G. & Guo, Y.-Z.: , 2019d; A channel-projection mixed-scale convolutional neural network for motor imagery EEG decoding; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **27** (6): 1170–1180.
- [Li et al., 2020] Li, Y.; Liu, Y.; Cui, W.-G.; Guo, Y.-Z.; Huang, H. & Hu, Z.-Y.: , 2020; Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **28** (4): 782–794.
- [Li et al., 2021b] Li, Y.; Fu, B.; Li, F.; Shi, G. & Zheng, W.: , 2021b; A novel transferability attention neural network model for EEG emotion recognition; *Neurocomputing*; **447**: 92–101.
- [Lim et al., 2021] Lim, J.-S.; Lee, J.-J. & Woo, C.-W.: , 2021; Post-stroke cognitive impairment: pathophysiological insights into brain disconnectome from advanced neuroimaging analysis techniques; *Journal of Stroke*; **23** (3): 297–311.
- [Lin et al., 2021] Lin, Y.-S.; Lee, W.-C. & Celik, Z. B.: , 2021; What do you see? evaluation of explainable artificial intelligence (xai) interpretability through neural backdoors; en *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*; págs. 1027–1035.

## REFERENCES

## REFERENCES

- [Linderman & Steinerberger, 2019] Linderman, G. & Steinerberger, S.: , 2019; Clustering with t-sne, provably; *SIAM Journal on Mathematics of Data Science*; **1** (2): 313–332.
- [Liu & Gillies, 2016] Liu, R. & Gillies, D. F.: , 2016; Overfitting in linear feature extraction for classification of high-dimensional image data; *Pattern Recognition*; **53**: 73–86.
- [Liu & Yang, 2021] Liu, T. & Yang, D.: , 2021; A densely connected multi-branch 3d convolutional neural network for motor imagery eeg decoding; *Brain Sciences*; **11** (2): 197.
- [Liu et al., 2024] Liu, W.; Guo, C. & Gao, C.: , 2024; A cross-session motor imagery classification method based on riemannian geometry and deep domain adaptation; *Expert Systems with Applications*; **237**: 121612.
- [Llanos et al., 2013] Llanos, C.; Rodriguez, M.; Rodriguez-Sabate, C.; Morales, I. & Sabate, M.: , 2013; Mu-rhythm changes during the planning of motor and motor imagery actions; *Neuropsychologia*; **51** (6): 1019–1026.
- [Lopez et al., 2020] Lopez, C. A. F.; Li, G. & Zhang, D.: , 2020; Beyond technologies of electroencephalography-based brain-computer interfaces: A systematic review from commercial and ethical aspects; *Frontiers in Neuroscience*; **14**.
- [Lotte et al., 2018] Lotte, F.; Bougrain, L.; Cichocki, A.; Clerc, M.; Congedo, M.; Rakotomamonjy, A. & Yger, F.: , 2018; A review of classification algorithms for eeg-based brain-computer interfaces: a 10 year update; *Journal of neural engineering*; **15** (3): 031005.
- [Lu et al., 2011] Lu, C.-F.; Teng, S.; Hung, C.-I.; Tseng, P.-J.; Lin, L.-T.; Lee, P.-L. & Wu, Y.-T.: , 2011; Reorganization of functional connectivity during the motor task using EEG time-frequency cross mutual information analysis; *Clinical Neurophysiology*; **122** (8): 1569–1579.
- [Luo et al., 2019] Luo, J.; Wang, J.; Xu, R. & Xu, K.: , 2019; Class discrepancy-guided sub-band filter-based common spatial pattern for motor imagery classification; *Journal of neuroscience methods*; **323**: 98–107.
- [Luo et al., 2020] Luo, J.; Gao, X.; Zhu, X.; Wang, B.; Lu, N. & Wang, J.: , 2020; Motor imagery eeg classification based on ensemble support vector learning; *Computer methods and programs in biomedicine*; **193**: 105464.
- [Luo et al., 2018] Luo, T.-j.; Zhou, C.-l. & Chao, F.: , 2018; Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network; *BMC bioinformatics*; **19** (1): 1–18.
- [Luo et al., 2021] Luo, Z.; Jin, R.; Shi, H. & Lu, X.: , 2021; Research on recognition of motor imagination based on connectivity features of brain functional network; *Neural plasticity*; **2021**.
- [Ma et al., 2023] Ma, W.; Wang, C.; Sun, X.; Lin, X. & Wang, Y.: , 2023; A double-branch graph convolutional network based on individual differences weakening for motor imagery eeg classification; *Biomedical Signal Processing and Control*; **84**: 104684.
- [Ma et al., 2020] Ma, X.; Wang, D.; Liu, D. & Yang, J.: , 2020; Dwt and cnn based multi-class motor imagery electroencephalographic signal recognition; *Journal of neural engineering*; **17** (1): 016073.
- [Maiseli et al., 2023] Maiseli, B.; Abdalla, A. T.; Massawe, L. V.; Mbise, M.; Mkocha, K.; Nassor, N. A.; Ismail, M.; Michael, J. & Kimambo, S.: , 2023; Brain-computer interface: trend, challenges, and threats; *Brain Informatics*; **10** (1): 20.
- [Maksimenko et al., 2017] Maksimenko, V. A.; Lüttjohann, A.; Makarov, V. V.; Goremyko, M. V.; Koronovskii, A. A.; Nedaivozov, V.; Runnova, A. E.; van Luijtelaar, G.; Hramov, A. E. & Boccaletti, S.: , 2017; Macroscopic and microscopic spectral properties of brain networks during local and global synchronization; *Physical Review E*; **96** (1): 012316.
- [Mammone et al., 2023] Mammone, N.; Ieracitano, C.; Adeli, H. & Morabito, F. C.: , 2023; Autoencoder filter bank common spatial patterns to decode motor imagery from eeg; *IEEE Journal of Biomedical and Health Informatics*.
- [Martínez-Cancino et al., 2020] Martínez-Cancino, R.; Delorme, A.; Wagner, J.; Kreutz-Delgado, K.; Sotero, R. C. & Makeig, S.: , 2020; What can local transfer entropy tell us about phase-amplitude coupling in electrophysiological signals?; *Entropy*; **22** (11): 1262.
- [Matemba et al., 2020] Matemba, E. D.; Li, G.; Gogan, I. C. W. & Maiseli, B. J.: , 2020; Technology acceptance model: recent developments, future directions, and proposal for hypothetical extensions; *International Journal of Technology Intelligence and Planning*; **12** (4): 315–348.
- [McFarland & Wolpaw, 2011] McFarland, D. J. & Wolpaw, J. R.: , 2011; Brain-computer interfaces for communication and control; *Communications of the ACM*; **54** (5): 60–66.
- [Meanti et al., 2020] Meanti, G.; Carratino, L.; Rosasco, L. & Rudi, A.: , 2020; Kernel methods through the roof: handling billions of points efficiently; *arXiv preprint arXiv:2006.10350*.
- [Meers et al., 2020] Meers, R.; Nuttall, H. E. & Vogt, S.: , 2020; Motor imagery alone drives corticospinal excitability during concurrent action observation and motor imagery; *Cortex*; **126**: 322–333.
- [Meng et al., 2023] Meng, J.; Zhao, Y.; Wang, K.; Sun, J.; Yi, W.; Xu, F.; Xu, M. & Ming, D.: , 2023; Rhythmic temporal prediction enhances neural representations of movement intention for brain-computer interface; *Journal of Neural Engineering*; **20** (6): 066004.
- [Miah et al., 2020] Miah, A. S. M.; Rahim, M. A. & Shin, J.: , 2020; Motor-imagery classification using riemannian geometry with median absolute deviation; *Electronics*; **9** (10): 1584.
- [Miao et al., 2020] Miao, M.; Hu, W.; Yin, H. & Zhang, K.: , 2020; Spatial-frequency feature learning and classification of motor imagery eeg based on deep convolution neural network; *Computational and mathematical methods in medicine*; **2020**.

- [Midha *et al.*, 2021] Midha, S.; Maior, H. A.; Wilson, M. L. & Sharples, S.: , 2021; Measuring mental workload variations in office work tasks using fmris; *International Journal of Human-Computer Studies*; **147**: 102580.
- [Miladinović *et al.*, 2021] Miladinović, A.; Ajčević, M.; Jarmolowska, J.; Marusic, U.; Colussi, M.; Silveri, G.; Battaglini, P. P. & Accardo, A.: , 2021; Effect of power feature covariance shift on BCI spatial-filtering techniques: A comparative study; *Computer Methods and Programs in Biomedicine*; **198**: 105808.
- [Miljević *et al.*, 2022] Miljević, A.; Bailey, N. W.; Vila-Rodriguez, F.; Herring, S. E. & Fitzgerald, P. B.: , 2022; Electroencephalographic connectivity: a fundamental guide and checklist for optimal study design and evaluation; *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*; **7** (6): 546–554.
- [Millan & Mourino, 2003] Millan, J. R. & Mourino, J.: , 2003; Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project; *IEEE transactions on neural systems and rehabilitation engineering*; **11** (2): 159–161.
- [Miotto *et al.*, 2018] Miotto, R.; Wang, F.; Wang, S.; Jiang, X. & Dudley, J. T.: , 2018; Deep learning for healthcare: review, opportunities and challenges; *Briefings in bioinformatics*; **19** (6): 1236–1246.
- [Mirzaei & Ghasemi, 2021] Mirzaei, S. & Ghasemi, P.: , 2021; EEG motor imagery classification using dynamic connectivity patterns and convolutional autoencoder; *Biomedical Signal Processing and Control*; **68**: 102584.
- [Mishra *et al.*, 2018] Mishra, P. K.; Jagadish, B.; Kiran, M.; Rajalakshmi, P. & Reddy, D. S.: , 2018; A novel classification for eeg based four class motor imagery using kullback-leibler regularized riemannian manifold; en *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*; IEEE; págs. 1–5.
- [Modares-Haghghi *et al.*, 2021] Modares-Haghghi, P.; Boostani, R.; Nami, M. & Sanei, S.: , 2021; Quantification of pain severity using EEG-based functional connectivity; *Biomedical Signal Processing and Control*; **69**: 102840.
- [Mohammadi *et al.*, 2021] Mohammadi, L.; Einalou, Z.; Hosseinzadeh, H. & Dadgostar, M.: , 2021; Cursor movement detection in brain-computer-interface systems using the k-means clustering method and LSVM; *Journal of Big Data*; **8** (1): 1–15.
- [Molina-Giraldo *et al.*, 2015] Molina-Giraldo, S.; Carvajal-González, J.; Álvarez-Meza, A. M. & Castellanos-Domínguez, G.: , 2015; Video segmentation framework based on multi-kernel representations and feature relevance analysis for object classification; en *Pattern recognition applications and methods*; Springer; págs. 273–283.
- [Molnar *et al.*, 2020] Molnar, C.; Casalicchio, G. & Bischl, B.: , 2020; Interpretable machine learning—a brief history, state-of-the-art and challenges; en *Joint European conference on machine learning and knowledge discovery in databases*; Springer; págs. 417–431.
- [Moran *et al.*, 2012] Moran, A.; Guillot, A.; MacIntyre, T. & Collet, C.: , 2012; Re-imagining motor imagery: Building bridges between cognitive neuroscience and sport psychology; *British Journal of Psychology*; **103** (2): 224–247.
- [Mridha *et al.*, 2021] Mridha, M. F.; Das, S. C.; Kabir, M. M.; Lima, A. A.; Islam, M. R. & Watanobe, Y.: , 2021; Brain-computer interface: Advancement and challenges; *Sensors*; **21** (17): 5746.
- [Mumtaz *et al.*, 2018] Mumtaz, W.; Kamel, N.; Ali, S. S. A.; Malik, A. S. *et al.*: , 2018; An EEG-based functional connectivity measure for automatic detection of alcohol use disorder; *Artificial intelligence in medicine*; **84**: 79–89.
- [Musallam *et al.*, 2021] Musallam, Y. K.; AlFassam, N. I.; Muhammad, G.; Amin, S. U.; Alsulaiman, M.; Abdul, W.; Altaheri, H.; Bencherif, M. A. & Algabri, M.: , 2021; Electroencephalography-based motor imagery classification using temporal convolutional network fusion; *Biomedical Signal Processing and Control*; **69**: 102826.
- [Naidu *et al.*, 2020] Naidu, R.; Ghosh, A.; Maurya, Y.; Kundu, S. S. *et al.*: , 2020; Is-cam: Integrated score-cam for axiomatic-based explanations; *arXiv preprint arXiv:2010.03023*.
- [Nentwich *et al.*, 2020] Nentwich, M.; Ai, L.; Madsen, J.; Telesford, Q. K.; Haufe, S.; Milham, M. P. & Parra, L. C.: , 2020; Functional connectivity of eeg is subject-specific, associated with phenotype, and different from fmri; *NeuroImage*; **218**: 117001.
- [Nguyen *et al.*, 2021] Nguyen, H. T. T.; Cao, H. Q.; Nguyen, K. V. T. & Pham, N. D. K.: , 2021; Evaluation of explainable artificial intelligence: Shap, lime, and cam; en *Proceedings of the FPT AI Conference*; págs. 1–6.
- [Nicolini *et al.*, 2020] Nicolini, C.; Forcellini, G.; Minati, L. & Bifone, A.: , 2020; Scale-resolved analysis of brain functional connectivity networks with spectral entropy; *NeuroImage*; **211**: 116603.
- [Nisar *et al.*, 2018] Nisar, H.; Thee, K. W.; Lim, S. H.; Yap, V. V.; Teh, P. C.; Nor, N. M. & Chow, C. M.: , 2018; Brain functional connectivity analysis using single trial EEG for understanding individual mechanisms; en *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)*; IEEE; págs. 209–214.
- [Nkengfack *et al.*, 2020] Nkengfack, L. C. D.; Tchiotsop, D.; Atangana, R.; Louis-Door, V. & Wolf, D.: , 2020; EEG signals analysis for epileptic seizures detection using polynomial transforms, linear discriminant analysis and support vector machines; *Biomedical Signal Processing and Control*; **62**: 102141.
- [Noshadi *et al.*, 2014] Noshadi, S.; Abootalebi, V.; Sadeghi, M. T. & Shahvazian, M. S.: , 2014; Selection of an efficient feature space for EEG-based mental task discrimination; *Biocybernetics and Biomedical Engineering*; **34** (3): 159–168.
- [Nunnari *et al.*, 2021] Nunnari, F.; Kadir, M. A. & Sonntag, D.: , 2021; On the overlap between grad-cam saliency maps and explainable visual features in skin cancer images; en *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*; Springer; págs. 241–253.
- [Oikonomou *et al.*, 2017] Oikonomou, V. P.; Georgiadis, K.; Liaros, G.; Nikolopoulos, S. & Kompatsiaris, I.: , 2017; A comparison study on eeg signal processing techniques using motor imagery eeg data; en *2017 IEEE 30th international symposium on computer-based medical systems (CBMS)*; IEEE; págs. 781–786.

- [Omeiza *et al.*, 2019] Omeiza, D.; Speakman, S.; Cintas, C. & Weldermariam, K.: , 2019; Smooth grad-CAM++: An enhanced inference level visualization technique for deep convolutional neural network models; *arXiv preprint arXiv:1908.01224*.
- [O'Reilly & Chanmittakul, 2021] O'Reilly, J. A. & Chanmittakul, W.: , 2021; L1 regularization-based selection of EEG spectral power and ECG features for classification of cognitive state; en *2021 9th International Electrical Engineering Congress (iEECON)*; IEEE; págs. 365–368.
- [Ortiz-Echeverri *et al.*, 2019] Ortiz-Echeverri, C. J.; Salazar-Colores, S.; Rodríguez-Reséndiz, J. & Gómez-Loenzo, R. A.: , 2019; A new approach for motor imagery classification based on sorted blind source separation, continuous wavelet transform, and convolutional neural network; *Sensors*; **19** (20): 4541.
- [Özdenizci *et al.*, 2019] Özdenizci, O.; Wang, Y.; Koike-Akino, T. & Erdoğmuş, D.: , 2019; Adversarial deep learning in EEG biometrics; *IEEE signal processing letters*; **26** (5): 710–714.
- [Padfield *et al.*, 2019] Padfield, N.; Zabalza, J.; Zhao, H.; Masero, V. & Ren, J.: , 2019; Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges; *Sensors*; **19** (6): 1423.
- [Pandey & Miyapuram, 2021] Pandey, P. & Miyapuram, K. P.: , 2021; BRAIN2DEPTH: Lightweight CNN model for classification of cognitive states from EEG recordings; *arXiv preprint arXiv:2106.06688*.
- [Parhi & Tewfik, 2021] Parhi, M. & Tewfik, A. H.: , 2021; Classifying imaginary vowels from frontal lobe EEG via deep learning; en *2020 28th European Signal Processing Conference (EUSIPCO)*; IEEE; págs. 1195–1199.
- [Park *et al.*, 2023] Park, S.; Ha, J. & Kim, L.: , 2023; Improving performance of motor imagery-based brain–computer interface in poorly performing subjects using a hybrid-imagery method utilizing combined motor and somatosensory activity; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **31**: 1064–1074.
- [Park *et al.*, 2017] Park, S.-H.; Lee, D. & Lee, S.-G.: , 2017; Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **26** (2): 498–505.
- [Paz-Linares *et al.*, 2017] Paz-Linares, D.; Vega-Hernandez, M.; Rojas-Lopez, P. A.; Valdes-Hernandez, P. A.; Martinez-Montes, E. & Valdes-Sosa, P. A.: , 2017; Spatio temporal EEG source imaging with the hierarchical bayesian elastic net and elitist lasso models; *Frontiers in neuroscience*; **11**: 635.
- [Perdikis *et al.*, 2018] Perdikis, S.; Tonin, L.; Saeedi, S.; Schneider, C. & Millán, J. d. R.: , 2018; The cybathlon bci race: Successful longitudinal mutual learning with two tetraplegic users; *PLoS biology*; **16** (5): e2003787.
- [Peterson *et al.*, 2019] Peterson, V.; Wyser, D.; Lambercy, O.; Spies, R. & Gassert, R.: , 2019; A penalized time-frequency band feature selection and classification procedure for improved motor intention decoding in multichannel EEG; *Journal of Neural Engineering*; **16** (1): 016019.
- [Philip *et al.*, 2022] Philip, B. S.; Prasad, G. & Hemanth, D. J.: , 2022; Non-stationarity removal techniques in meg data: A review; *Procedia Computer Science*; **215**: 824–833.
- [Phunruangsakao *et al.*, 2023] Phunruangsakao, C.; Achancaray, D.; Bhattacharyya, S.; Izumi, S.-I. & Hayashibe, M.: , 2023; Effects of visual-electrotactile stimulation feedback on brain functional connectivity during motor imagery practice; *Scientific Reports*; **13** (1): 17752.
- [Pope *et al.*, 2019] Pope, P. E.; Kolouri, S.; Rostami, M.; Martin, C. E. & Hoffmann, H.: , 2019; Explainability methods for graph convolutional neural networks; en *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; págs. 10772–10781.
- [Prakash *et al.*, ] Prakash, E. I.; Shrikumar, A. & Kundaje, A.: , ; Towards more realistic simulated datasets for benchmarking deep learning models in regulatory genomics; en *Machine Learning in Computational Biology*; PMLR; págs. 58–77.
- [Principe, 2010] Principe, J. C.: , 2010; *Information theoretic learning: Renyi's entropy and kernel perspectives*; Springer Science & Business Media.
- [Pulgarin-Giraldo *et al.*, 2017] Pulgarin-Giraldo, J. D.; Ruales-Torres, A.; Álvarez-Meza, A. M. & Castellanos-Dominguez, G.: , 2017; Relevant kinematic feature selection to support human action recognition in MoCap data; en *International Work-Conference on the Interplay Between Natural and Artificial Computation*; Springer; págs. 501–509.
- [Qian *et al.*, 2018] Qian, X.; Loo, B. R. Y.; Castellanos, F. X.; Liu, S.; Koh, H. L.; Poh, X. W. W.; Krishnan, R.; Fung, D.; Chee, M. W.; Guan, C. *et al.*: , 2018; Brain-computer-interface-based intervention re-normalizes brain functional network topology in children with attention deficit/hyperactivity disorder; *Translational psychiatry*; **8** (1): 149.
- [Qiao *et al.*, 2023] Qiao, R.; Zhang, H. & Tian, Y.: , 2023; Eeg cortical network reveals the temporo-spatial mechanism of visual search; *Brain Research Bulletin*; **203**: 110758.
- [Qiumei *et al.*, 2019] Qiumei, Z.; Dan, T. & Fenghua, W.: , 2019; Improved convolutional neural network based on fast exponentially linear unit activation function; *Ieee Access*; **7**: 151359–151367.
- [Raeisi *et al.*, 2020] Raeisi, K.; Mohebbi, M.; Khazaei, M.; Seraji, M. & Yoonessi, A.: , 2020; Phase-synchrony evaluation of EEG signals for multiple sclerosis diagnosis based on bivariate empirical mode decomposition during a visual task; *Computers in biology and medicine*; **117**: 103596.
- [Raeisi *et al.*, 2022] Raeisi, K.; Khazaei, M.; Croce, P.; Tamburro, G.; Comani, S. & Zappasodi, F.: , 2022; A graph convolutional neural network for the automated detection of seizures in the neonatal eeg; *Computer methods and programs in biomedicine*; **222**: 106950.

- [Rahate et al., 2022] **Rahate, A.; Walambe, R.; Ramanna, S. & Kotecha, K.**: , 2022; Multimodal co-learning: challenges, applications with datasets, recent advances and future directions; *Information Fusion*; **81**: 203–239.
- [Rahimi et al., 2007] **Rahimi, A.; Recht, B. et al.**: , 2007; Random features for large-scale kernel machines.; en *NIPS*, tomo 3; Citeseer; pág. 5.
- [Ramadan & Vasilakos, 2017] **Ramadan, R. A. & Vasilakos, A. V.**: , 2017; Brain computer interface: control signals review; *Neurocomputing*; **223**: 26–44.
- [Ramadan et al., 2015] **Ramadan, R. A.; Refat, S.; Elshahed, M. A. & Ali, R. A.**: , 2015; Basics of brain computer interface; en *Brain-Computer Interfaces*; Springer; págs. 31–50.
- [Ras et al., 2018] **Ras, G.; van Gerven, M. & Haselager, P.**: , 2018; Explanation methods in deep learning: Users, values, concerns and challenges; *Explainable and interpretable models in computer vision and machine learning*: 19–36.
- [Rashed-Al-Mahfuz et al., 2021] **Rashed-Al-Mahfuz, M.; Moni, M. A.; Uddin, S.; Alyami, S. A.; Summers, M. A. & Eapen, V.**: , 2021; A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (eeg) data; *IEEE journal of translational engineering in health and medicine*; **9**: 1–12.
- [Rasheed et al., 2021] **Rasheed, M. A.; Chand, P.; Ahmed, S.; Sharif, H.; Hoodbhoy, Z.; Siddiqui, A. & Hasan, B. S.**: , 2021; Use of artificial intelligence on electroencephalogram (EEG) waveforms to predict failure in early school grades in children from a rural cohort in pakistan; *Plos one*; **16** (2): e0246236.
- [Rashid et al., 2020] **Rashid, M.; Sulaiman, N.; PP Abdul Majeed, A.; Musa, R. M.; Ab Nasir, A. F.; Bari, B. S. & Khatun, S.**: , 2020; Current status, challenges, and possible solutions of eeg-based brain-computer interface: a comprehensive review; *Frontiers in neurorobotics*; **25**.
- [Rashkov et al., 2019] **Rashkov, G.; Bobe, A.; Fastovets, D. & Komarova, M.**: , 2019; Natural image reconstruction from brain waves: a novel visual bci system with native feedback; *BioRxiv*: 787101.
- [Rathee et al., 2017] **Rathee, D.; Cecotti, H. & Prasad, G.**: , 2017; Single-trial effective brain connectivity patterns enhance discriminability of mental imagery tasks; *Journal of neural engineering*; **14** (5): 056005.
- [Ravaioli et al., 2023] **Ravaioli, G.; Domingos, T. & Teixeira, R. F.**: , 2023; A framework for data-driven agent-based modelling of agricultural land use; *Land*; **12** (4): 756.
- [Rejer & Lorenz, 2013] **Rejer, I. & Lorenz, K.**: , 2013; Genetic algorithm and forward method for feature selection in EEG feature space; *Journal of Theoretical and Applied Computer Science*; **7** (2): 72–82.
- [Rényi, 1961] **Rényi, A.**: , 1961; On measures of entropy and information; en *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, tomo 4; University of California Press; págs. 547–562.
- [Reuderink et al., 2011] **Reuderink, B.; Farquhar, J.; Poel, M. & Nijholt, A.**: , 2011; A subject-independent brain-computer interface based on smoothed, second-order baselining; en *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; IEEE; págs. 4600–4604.
- [Rezaei & Shalbaf, 2023] **Rezaei, E. & Shalbaf, A.**: , 2023; Classification of right/left hand motor imagery by effective connectivity based on transfer entropy in electroencephalogram signal; *Basic and Clinical Neuroscience*; **14** (2): 213.
- [Riahi et al., 2020] **Riahi, N.; Vakorin, V. A. & Menon, C.**: , 2020; Estimating fugl-meyer upper extremity motor score from functional-connectivity measures; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **28** (4): 860–868.
- [Rimbert et al., 2020] **Rimbert, S.; Bougrain, L. & Fleck, S.**: , 2020; Learning how to generate kinesthetic motor imagery using a bci-based learning environment: A comparative study based on guided or trial-and-error approaches; en *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*; IEEE; págs. 2483–2498.
- [Rodrigues et al., 2019a] **Rodrigues, P.; Stefano, C.; Attux, R.; Castellano, G. & Soriano, D.**: , 2019a; Space-time recurrences for functional connectivity evaluation and feature extraction in motor imagery brain-computer interfaces; *Medical & biological engineering & computing*; **57** (8): 1709–1725.
- [Rodrigues et al., 2020] **Rodrigues, P.; Fim-Neto, A.; Sato, J.; Soriano, D. & Nasuto, S.**: , 2020; Single-trial functional connectivity dynamics of event-related desynchronization for motor imagery eeg-based brain-computer interfaces; en *Brazilian Congress on Biomedical Engineering*; Springer; págs. 1887–1893.
- [Rodrigues et al., 2019b] **Rodrigues, P. G.; Stefano Filho, C. A.; Attux, R.; Castellano, G. & Soriano, D. C.**: , 2019b; Space-time recurrences for functional connectivity evaluation and feature extraction in motor imagery brain-computer interfaces; *Medical & biological engineering & computing*; **57** (8): 1709–1725.
- [Rodrigues et al., 2022] **Rodrigues, P. G.; Filho, C. A. S.; Takahata, A. K.; Suyama, R.; Attux, R.; Castellano, G.; Sato, J. R.; Nasuto, S. J. & Soriano, D. C.**: , 2022; Can dynamic functional connectivity be used to distinguish between resting-state and motor imagery in eeg-bcis?; en *Complex Networks & Their Applications X: Volume 2, Proceedings of the Tenth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2021 10*; Springer; págs. 688–699.
- [Rongrong et al., 2020] **Rongrong, F.; Mengmeng, H.; Yongsheng, T. & Peiming, S.**: , 2020; Improvement motor imagery eeg classification based on sparse common spatial pattern and regularized discriminant analysis; *Journal of Neuroscience Methods*; **343**: 108833.
- [Roweis & Ghahramani, 1999] **Roweis, S. & Ghahramani, Z.**: , 1999; A unifying review of linear gaussian models; *Neural computation*; **11** (2): 305–345.

- [Roy *et al.*, 2022] Roy, G.; Bhoi, A. & Bhaumik, S.: , 2022; A comparative approach for mi-based eeg signals classification using energy, power and entropy; *IRBM*; **43** (5): 434–446.
- [Rubinov & Sporns, 2010] Rubinov, M. & Sporns, O.: , 2010; Complex network measures of brain connectivity: uses and interpretations; *Neuroimage*; **52** (3): 1059–1069.
- [Ruffino *et al.*, 2017] Ruffino, C.; Papaxanthis, C. & Lebon, F.: , 2017; Neural plasticity during motor learning with motor imagery practice: Review and perspectives; *Neuroscience*; **341**: 61–78.
- [Ruiz-Gómez *et al.*, 2020] Ruiz-Gómez, S. J.; Gómez, C.; Poza, J.; Revilla-Vallejo, M.; Gutiérrez-de Pablo, V.; Rodríguez-González, V.; Maturana-Candela, A. & Hornero, R.: , 2020; Volume conduction effects on connectivity metrics: Application of network parameters to characterize alzheimer's disease continuum; en *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*; IEEE; págs. 30–33.
- [Saha & Baumert, 2020] Saha, S. & Baumert, M.: , 2020; Intra-and inter-subject variability in EEG-based sensorimotor brain computer interface: a review; *Frontiers in computational neuroscience*; **13**: 87.
- [Said *et al.*, 2017] Said, S.; Bombrun, L.; Berthoumieu, Y. & Manton, J. H.: , 2017; Riemannian gaussian distributions on the space of symmetric positive definite matrices; *IEEE Transactions on Information Theory*; **63** (4): 2153–2170.
- [Sakkalis, 2011] Sakkalis, V.: , 2011; Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG; *Computers in biology and medicine*; **41** (12): 1110–1117.
- [Samuel *et al.*, 2017] Samuel, O. W.; Geng, Y.; Li, X. & 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.
- [Sannelli *et al.*, 2019] Sannelli, C.; Vidaurre, C.; Müller, K. & Blankertz, B.: , 2019; A large scale screening study with a smr-based bci: Categorization of bci users and differences in their smr activity; *PLoS One*; **14** (1): e0207351.
- [Sarin *et al.*, 2020] Sarin, M.; Verma, A.; Mehta, D. H.; Shukla, P. K. & Verma, S.: , 2020; Automated ocular artifacts identification and removal from EEG data using hybrid machine learning methods; en *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*; IEEE; págs. 1054–1059.
- [Sarraf, 2017] 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.
- [Schirrmeyer *et al.*, 2017] Schirrmeyer, R. T.; Springenberg, J. T.; Fiederer, L. D. J.; Glasstetter, M.; Eggensperger, K.; Tangermann, M.; Hutter, F.; Burgard, W. & Ball, T.: , 2017; Deep learning with convolutional neural networks for EEG decoding and visualization; *Human brain mapping*; **38** (11): 5391–5420.
- [Seghier & Price, 2018] Seghier, M. L. & Price, C. J.: , 2018; Interpreting and utilising intersubject variability in brain function; *Trends in cognitive sciences*; **22** (6): 517–530.
- [Selvaraju *et al.*, 2016] Selvaraju, R. R.; Das, A.; Vedantam, R.; Cogswell, M.; Parikh, D. & Batra, D.: , 2016; Grad-CAM: Why did you say that?; *arXiv preprint arXiv:1611.07450*.
- [Selvaraju *et al.*, 2017] Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D. & Batra, D.: , 2017; Grad-cam: Visual explanations from deep networks via gradient-based localization; en *Proceedings of the IEEE international conference on computer vision*; págs. 618–626.
- [Shali & Setarehdan, 2020] Shali, R. K. & Setarehdan, S. K.: , 2020; The impact of electrode reduction in the diagnosis of dyslexia; en *2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME)*; IEEE; págs. 118–125.
- [Shamsi *et al.*, 2020a] Shamsi, F.; Haddad, A. & Najafizadeh, L.: , 2020a; Early classification of motor tasks using dynamic functional connectivity graphs from eeg; *Journal of Neural Engineering*.
- [Shamsi *et al.*, 2020b] Shamsi, F.; Haddad, A. *et al.*: , 2020b; Recognizing pain in motor imagery EEG recordings using dynamic functional connectivity graphs; en *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*; IEEE; págs. 2869–2872.
- [Shamsi *et al.*, 2021] Shamsi, F.; Haddad, A. & Najafizadeh, L.: , 2021; Early classification of motor tasks using dynamic functional connectivity graphs from EEG; *Journal of Neural Engineering*; **18** (1): 016015.
- [Shao *et al.*, 2020] Shao, L.; Zhang, L.; Belkacem, A. N.; Zhang, Y.; Chen, X.; Li, J. & Liu, H.: , 2020; EEG-controlled wall-crawling cleaning robot using SSVEP-based brain-computer interface; *Journal of healthcare engineering*; **2020**.
- [Sharma *et al.*, 2023] Sharma, N.; Sharma, M.; Singhal, A.; Vyas, R.; Malik, H.; Afthanorhan, A. & Hossaini, M. A.: , 2023; Recent trends in eeg based motor imagery signal analysis and recognition: A comprehensive review.; *IEEE Access*.
- [Shrikumar *et al.*, 2017] Shrikumar, A.; Greenside, P. & Kundaje, A.: , 2017; Learning important features through propagating activation differences; en *International conference on machine learning*; PMLR; págs. 3145–3153.
- [Si *et al.*, 2020] Si, Y.; Li, F.; Duan, K.; Tao, Q.; Li, C.; Cao, Z.; Zhang, Y.; Biswal, B.; Li, P.; Yao, D. *et al.*: , 2020; Predicting individual decision-making responses based on single-trial eeg; *NeuroImage*; **206**: 116333.
- [Simon *et al.*, 2013] Simon, N.; Friedman, J.; Hastie, T. & Tibshirani, R.: , 2013; A sparse-group lasso; *Journal of computational and graphical statistics*; **22** (2): 231–245.
- [Simonyan *et al.*, 2013] Simonyan, K.; Vedaldi, A. & Zisserman, A.: , 2013; Deep inside convolutional networks: Visualising image classification models and saliency maps; *arXiv preprint arXiv:1312.6034*.

- [Singh *et al.*, 2021] Singh, A.; Hussain, A. A.; Lal, S. & 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.
- [Sisti *et al.*, 2022] Sisti, H. M.; Beebe, A.; Bishop, M. & Gabrielsson, E.: , 2022; A brief review of motor imagery and bimanual coordination; *Frontiers in Human Neuroscience*; **16**: 1037410.
- [Sitaram *et al.*, 2017] 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.
- [Siviero *et al.*, 2023] Siviero, I.; Menegaz, G. & Storti, S. F.: , 2023; Functional connectivity and feature fusion enhance multiclass motor-imagery brain-computer interface performance; *Sensors*; **23** (17): 7520.
- [Smith *et al.*, 2011] Smith, S. M.; Miller, K. L.; Salimi-Khorshidi, G.; Webster, M.; Beckmann, C. F.; Nichols, T. E.; Ramsey, J. D. & Woolrich, M. W.: , 2011; Network modelling methods for fmri; *Neuroimage*; **54** (2): 875–891.
- [Smola & Schölkopf, 2000] Smola, A. J. & Schölkopf, B.: , 2000; Sparse greedy matrix approximation for machine learning.
- [Smuha, 2019] Smuha, N. A.: , 2019; The eu approach to ethics guidelines for trustworthy artificial intelligence; *Computer Law Review International*; **20** (4): 97–106.
- [Somers *et al.*, 2018] Somers, B.; Francart, T. & Bertrand, A.: , 2018; A generic eeg artifact removal algorithm based on the multi-channel wiener filter; *Journal of neural engineering*; **15** (3): 036007.
- [Sorger & Goebel, 2020] Sorger, B. & Goebel, R.: , 2020; Real-time fmri for brain-computer interfacing; *Handbook of clinical neurology*; **168**: 289–302.
- [Souto *et al.*, 2020] Souto, D. O.; Cruz, T. K. F.; Coutinho, K.; Julio-Costa, A.; Fontes, P. L. B. & Haase, V. G.: , 2020; Effect of motor imagery combined with physical practice on upper limb rehabilitation in children with hemiplegic cerebral palsy; *NeuroRehabilitation*; **46** (1): 53–63.
- [Speith, 2022] Speith, T.: , 2022; A review of taxonomies of explainable artificial intelligence (xai) methods; en *2022 ACM Conference on Fairness, Accountability, and Transparency*; págs. 2239–2250.
- [Springenberg *et al.*, 2014] Springenberg, J. T.; Dosovitskiy, A.; Brox, T. & Riedmiller, M.: , 2014; Striving for simplicity: The all convolutional net; *arXiv preprint arXiv:1412.6806*.
- [Stefano Filho *et al.*, 2018] Stefano Filho, C. A.; Attux, R. & Castellano, G.: , 2018; Can graph metrics be used for EEG-BCIs based on hand motor imagery?; *Biomedical Signal Processing and Control*; **40**: 359–365.
- [Stefano Filho *et al.*, 2021] Stefano Filho, C. A.; Ignacio Serrano, J.; Attux, R.; Castellano, G.; Rocon, E. & del Castillo, M. D.: , 2021; Reorganization of resting-state EEG functional connectivity patterns in children with cerebral palsy following a motor imagery virtual-reality intervention; *Applied Sciences*; **11** (5): 2372.
- [Stephan *et al.*, 2009] Stephan, K.; Friston, K. & Squire, L.: , 2009; Functional connectivity.
- [Stergiadis *et al.*, 2022] Stergiadis, C.; Kostaridou, V.-D. & Klados, M. A.: , 2022; Which bss method separates better the eeg signals? a comparison of five different algorithms; *Biomedical Signal Processing and Control*; **72**: 103292.
- [Sturm *et al.*, 2016] Sturm, I.; Lapuschkin, S.; Samek, W. & Müller, K.-R.: , 2016; Interpretable deep neural networks for single-trial EEG classification; *Journal of neuroscience methods*; **274**: 141–145.
- [Subasi & Gursoy, 2010] Subasi, A. & Gursoy, M. I.: , 2010; EEG signal classification using PCA, ICA, LDA and support vector machines; *Expert systems with applications*; **37** (12): 8659–8666.
- [Subasi *et al.*, 2021] Subasi, A.; Tuncer, T.; Dogan, S.; Tanko, D. & Sakoglu, U.: , 2021; EEG-based emotion recognition using tunable Q wavelet transform and rotation forest ensemble classifier; *Biomedical Signal Processing and Control*; **68**: 102648.
- [Sujatha Ravindran & Contreras-Vidal, 2023] Sujatha Ravindran, A. & Contreras-Vidal, J.: , 2023; An empirical comparison of deep learning explainability approaches for eeg using simulated ground truth; *Scientific Reports*; **13** (1): 17709.
- [Sun *et al.*, 2020a] Sun, C.; Lin, K.; Huang, Y.-T.; Ching, E. S.; Lai, P.-Y. & Chan, C.: , 2020a; Directed effective connectivity and synaptic weights of in vitro neuronal cultures revealed from high-density multielectrode array recordings; *bioRxiv*.
- [Sun *et al.*, 2020b] Sun, X.; Hu, B.; Zheng, X.; Yin, Y. & Ji, C.: , 2020b; Emotion classification based on brain functional connectivity network; en *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; IEEE; págs. 2082–2089.
- [Suto & Oniga, 2018] Suto, J. & Oniga, S.: , 2018; Music stimuli recognition in electroencephalogram signal; *Elektronika ir Elektrotechnika*; **24** (4): 68–71.
- [Šverko *et al.*, 2022] Šverko, Z.; Vrankić, M.; Vlahinić, S. & Rogelj, P.: , 2022; Complex pearson correlation coefficient for eeg connectivity analysis; *Sensors*; **22** (4): 1477.
- [Tabar & Halici, 2016] Tabar, Y. R. & Halici, U.: , 2016; A novel deep learning approach for classification of eeg motor imagery signals; *Journal of neural engineering*; **14** (1): 016003.
- [Tafreshi *et al.*, 2019] Tafreshi, T. F.; Daliri, M. R. & Ghodousi, M.: , 2019; Functional and effective connectivity based features of eeg signals for object recognition; *Cognitive neurodynamics*; **13**: 555–566.

- [Talukdar *et al.*, 2019] **Talukdar, U.; Hazarika, S. M. & Gan, J. Q.**: , 2019; Motor imagery and mental fatigue: inter-relationship and eeg based estimation; *Journal of computational neuroscience*; **46**: 55–76.
- [Tang *et al.*, 2014] **Tang, Q.; Wang, J. & Wang, H.**: , 2014; L1-norm based discriminative spatial pattern for single-trial EEG classification; *Biomedical Signal Processing and Control*; **10**: 313–321.
- [Tang *et al.*, 2020] **Tang, X.; Li, W.; Li, X.; Ma, W. & Dang, X.**: , 2020; Motor imagery eeg recognition based on conditional optimization empirical mode decomposition and multi-scale convolutional neural network; *Expert Systems with Applications*; **149**: 113285.
- [Tang *et al.*, 2023] **Tang, X.; Wang, S.; Deng, X.; Liu, K.; Tian, Y.; Wang, H. & Gao, X.**: , 2023; Graph-based information separator and area convolutional network for eeg-based intention decoding; *IEEE Transactions on Cognitive and Developmental Systems*.
- [Tay *et al.*, 2023] **Tay, J. K.; Narasimhan, B. & Hastie, T.**: , 2023; Elastic net regularization paths for all generalized linear models; *Journal of statistical software*; **106**.
- [Tayeb *et al.*, 2019] **Tayeb, Z.; Fedjaev, J.; Ghaboosi, N.; Richter, C.; Everding, L.; Qu, X.; Wu, Y.; Cheng, G. & Conradt, J.**: , 2019; Validating deep neural networks for online decoding of motor imagery movements from eeg signals; *Sensors*; **19** (1): 210.
- [Teo & Chew, 2014] **Teo, W.-P. & Chew, E.**: , 2014; Is motor-imagery brain-computer interface feasible in stroke rehabilitation?; *PM&R*; **6** (8): 723–728.
- [Thakur *et al.*, 2016] **Thakur, K. T.; Albanese, E.; Giannakopoulos, P.; Jette, N.; Linde, M.; Prince, M. J.; Steiner, T. J. & Dua, T.**: , 2016; Neurological disorders.
- [Thiebaut de Schotten *et al.*, 2020] **Thiebaut de Schotten, M.; Foulon, C. & Nachev, P.**: , 2020; Brain disconnections link structural connectivity with function and behaviour; *Nature communications*; **11** (1): 5094.
- [Tibrewal *et al.*, 2022] **Tibrewal, N.; Leeuwis, N. & Alimardani, M.**: , 2022; Classification of motor imagery eeg using deep learning increases performance in inefficient bci users; *Plos one*; **17** (7): e0268880.
- [Tibshirani, 1996] **Tibshirani, R.**: , 1996; Regression shrinkage and selection via the lasso; *Journal of the Royal Statistical Society: Series B (Methodological)*; **58** (1): 267–288.
- [Tjoa & Guan, 2020] **Tjoa, E. & Guan, C.**: , 2020; A survey on explainable artificial intelligence (xai): Toward medical xai; *IEEE transactions on neural networks and learning systems*; **32** (11): 4793–4813.
- [Tjoa & Guan, 2022] **Tjoa, E. & Guan, C.**: , 2022; Evaluating weakly supervised object localization methods right? a study on heatmap-based xai and neural backed decision tree; *A Study on Heatmap-Based Xai and Neural Backed Decision Tree (November 27, 2022)*.
- [Tortora *et al.*, 2020] **Tortora, S.; Ghidoni, S.; Chisari, C.; Micera, S. & Artoni, F.**: , 2020; Deep learning-based BCI for gait decoding from EEG with LSTM recurrent neural network; *Journal of Neural Engineering*; **17** (4): 046011.
- [Tuncer *et al.*, 2021] **Tuncer, T.; Dogan, S. & Acharya, U. R.**: , 2021; Automated EEG signal classification using chaotic local binary pattern; *Expert Systems with Applications*; **182**: 115175.
- [Uribe *et al.*, 2019] **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. & Attux, R.**: , 2019; A correntropy-based classifier for motor imagery brain-computer interfaces; *Biomedical Physics & Engineering Express*; **5** (6): 065026.
- [Van der Lubbe *et al.*, 2021] **Van der Lubbe, R. H.; Sobierajewicz, J.; Jongsma, M. L.; Verwey, W. B. & 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.
- [Van der Maaten & Hinton, 2008] **Van der Maaten, L. & Hinton, G.**: , 2008; Visualizing data using t-sne.; *Journal of machine learning research*; **9** (11).
- [Van Der Maaten *et al.*, 2009] **Van Der Maaten, L.; Postma, E. & Van den Herik, J.**: , 2009; Dimensionality reduction: a comparative; *J Mach Learn Res*; **10** (66-71): 13.
- [Van Diessen *et al.*, 2015] **Van Diessen, E.; Numan, T.; Van Dellen, E.; Van Der Kooi, A.; Boersma, M.; Hofman, D.; Van Luterveld, R.; Van Dijk, B.; Van Straaten, E.; Hillebrand, A. et al.**: , 2015; Opportunities and methodological challenges in eeg and meg resting state functional brain network research; *Clinical Neurophysiology*; **126** (8): 1468–1481.
- [Van Mierlo *et al.*, 2014] **Van Mierlo, P.; Papadopoulou, M.; Carrette, E.; Boon, P.; Vandenberghe, S.; Vonck, K. & Marinazzo, D.**: , 2014; Functional brain connectivity from eeg in epilepsy: Seizure prediction and epileptogenic focus localization; *Progress in neurobiology*; **121**: 19–35.
- [Vaquerizo-Villar *et al.*, 2023] **Vaquerizo-Villar, F.; Gutiérrez-Tobal, G. C.; Calvo, E.; Álvarez, D.; Kheirandish-Gozal, L.; Del Campo, F.; Gozal, D. & Hornero, R.**: , 2023; An explainable deep-learning model to stage sleep states in children and propose novel eeg-related patterns in sleep apnea; *Computers in Biology and Medicine*; **165**: 107419.
- [Värbu *et al.*, 2022] **Värbu, K.; Muhammad, N. & Muhammad, Y.**: , 2022; Past, present, and future of eeg-based bci applications; *Sensors*; **22** (9): 3331.
- [Varone *et al.*, 2021] **Varone, G.; Bouliba, W.; Lo Giudice, M.; Benjdira, B.; Mammone, N.; Ieracitano, C.; Dashtipour, K.; Neri, S.; Gasparini, S.; Morabito, F. C. et al.**: , 2021; A machine learning approach involving functional connectivity features to classify rest-eeg psychogenic non-epileptic seizures from healthy controls; *Sensors*; **22** (1): 129.

## REFERENCES

## REFERENCES

- [Varoquaux, 2018] **Varoquaux, G.**: , 2018; Cross-validation failure: small sample sizes lead to large error bars; *Neuroimage*; **180**: 68–77.
- [Varsehi & Firoozabadi, 2021] **Varsehi, H. & Firoozabadi, S. M. P.**: , 2021; An EEG channel selection method for motor imagery based brain-computer interface and neurofeedback using granger causality; *Neural Networks*; **133**: 193–206.
- [Veena & Anitha, 2020] **Veena, N. & Anitha, N.**: , 2020; A review of non-invasive bci devices; *Int. J. Biomed. Eng. Technol.*; **34** (3): 205–233.
- [Velasquez-Martinez et al., 2020a] **Velasquez-Martinez, L.; Caicedo-Acosta, J. & Castellanos-Dominguez, G.**: , 2020a; Entropy-based estimation of event-related de/synchronization in motor imagery using vector-quantized patterns; *Entropy*; **22** (6): 703.
- [Velasquez-Martinez et al., 2020b] **Velasquez-Martinez, L.; Zapata-Castano, F. & Castellanos-Dominguez, G.**: , 2020b; Dynamic modeling of common brain neural activity in motor imagery tasks; *Frontiers in Neuroscience*; **14**: 714.
- [Velásquez-Martínez et al., 2013] **Velásquez-Martínez, L. F.; Álvarez-Meza, A. M. & Castellanos-Domínguez, C. G.**: , 2013; Motor imagery classification for BCI using common spatial patterns and feature relevance analysis; en *International Work-Conference on the Interplay Between Natural and Artificial Computation*; Springer; págs. 365–374.
- [Velasquez-Martinez et al., 2020c] **Velasquez-Martinez, L. F.; Zapata-Castano, F. & Castellanos-Dominguez, G.**: , 2020c; Dynamic modeling of common brain neural activity in motor imagery tasks; *Frontiers in Neuroscience*; **14**.
- [Verleysen & François, 2005] **Verleysen, M. & François, D.**: , 2005; The curse of dimensionality in data mining and time series prediction; en *International work-conference on artificial neural networks*; Springer; págs. 758–770.
- [Vidaurre et al., 2020] **Vidaurre, C.; Haufe, S.; Jorajuria, T.; Müller, K.-R. & Nikulin, V. V.**: , 2020; Sensorimotor functional connectivity: a neurophysiological factor related to BCI performance; *Frontiers in Neuroscience*; **14**: 1278.
- [Vidaurre et al., 2021] **Vidaurre, C.; Jorajuria, T.; Ramos-Murguialday, A.; Müller, K. R.; Gómez, M. & Nikulin, V. V.**: , 2021; Improving motor imagery classification during induced motor perturbations; *Journal of neural engineering*; **18** (4): 0460b1.
- [Vidaurre et al., 2023] **Vidaurre, C.; Gurunandan, K.; Idaji, M. J.; Nolte, G.; Gómez, M.; Villringer, A.; Müller, K.-R. & Nikulin, V. V.**: , 2023; Novel multivariate methods to track frequency shifts of neural oscillations in eeg/meg recordings; *NeuroImage*; **276**: 120178.
- [Vilela & Hochberg, 2020] **Vilela, M. & Hochberg, L. R.**: , 2020; Applications of brain-computer interfaces to the control of robotic and prosthetic arms; en *Handbook of clinical neurology*, tomo 168; Elsevier; págs. 87–99.
- [Volosyak et al., 2020] **Volosyak, I.; Rezeika, A.; Benda, M.; Gembler, F. & Stawicki, P.**: , 2020; Towards solving of the illiteracy phenomenon for vep-based brain-computer interfaces; *Biomedical Physics & Engineering Express*; **6** (3): 035034.
- [Wackernagel, 2003] **Wackernagel, H.**: , 2003; *Multivariate geostatistics: an introduction with applications*; Springer Science & Business Media.
- [Wackernagel, 2013] **Wackernagel, H.**: , 2013; *Multivariate geostatistics: an introduction with applications*; Springer Science & Business Media.
- [Wang et al., 2020a] **Wang, H.; Wang, Z.; Du, M.; Yang, F.; Zhang, Z.; Ding, S.; Mardziel, P. & Hu, X.**: , 2020a; Score-CAM: Score-weighted visual explanations for convolutional neural networks; en *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*; págs. 24–25.
- [Wang et al., 2020b] **Wang, H.; Xu, T.; Tang, C.; Yue, H.; Chen, C.; Xu, L.; Pei, Z.; Dong, J.; Bezerianos, A. & Li, J.**: , 2020b; Diverse feature blend based on filter-bank common spatial pattern and brain functional connectivity for multiple motor imagery detection; *IEEE Access*; **8**: 155590–155601.
- [Wang et al., 2021a] **Wang, J.-G.; Shao, H.-M.; Yao, Y.; Liu, J.-L. & Ma, S.-W.**: , 2021a; A personalized feature extraction and classification method for motor imagery recognition; *Mobile Networks and Applications*; 1–13.
- [Wang et al., 2019a] **Wang, K.; Zhai, D.-H. & Xia, Y.**: , 2019a; Motor imagination eeg recognition algorithm based on dwt, csp and extreme learning machine; en *2019 Chinese control conference (CCC)*; IEEE; págs. 4590–4595.
- [Wang et al., 2020c] **Wang, K.; Xu, M.; Wang, Y.; Zhang, S.; Chen, L. & Ming, D.**: , 2020c; Enhance decoding of pre-movement EEG patterns for brain-computer interfaces; *Journal of Neural Engineering*; **17** (1): 016033.
- [Wang et al., 2019b] **Wang, M.; Hu, J. & Abbass, H. A.**: , 2019b; Stable eeg biometrics using convolutional neural networks and functional connectivity.; *Aust. J. Intell. Inf. Process. Syst.*; **15** (3): 19–26.
- [Wang et al., 2014] **Wang, X.; Wang, A.; Zheng, S.; Lin, Y. & Yu, M.**: , 2014; A multiple autocorrelation analysis method for motor imagery EEG feature extraction; en *The 26th Chinese Control and Decision Conference (2014 CCDC)*; IEEE; págs. 3000–3004.
- [Wang et al., 2021b] **Wang, Z.; Wang, F.; Liang, C. & Zhang, J.**: , 2021b; A time-varying method for brain effective connectivity analysis of emotional EEG data; en *2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*; IEEE; págs. 131–139.
- [Warren, 2015] **Warren, M. J.**: , 2015; Five ways to look at cohen's kappa; *Journal of Psychology & Psychotherapy*; **5** (4): 1.
- [Wei & Luo, 2010] **Wei, G. & Luo, J.**: , 2010; Sport expert's motor imagery: Functional imaging of professional motor skills and simple motor skills; *Brain research*; **1341**: 52–62.

- [Wei *et al.*, 2023] Wei, M.; Yang, R.; Huang, M.; Ni, J.; Wang, Z. & 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*.
- [Willett *et al.*, 2021] Willett, F. R.; Avansino, D. T.; Hochberg, L. R.; Henderson, J. M. & Shenoy, K. V.: , 2021; High-performance brain-to-text communication via handwriting; *Nature*; **593** (7858): 249–254.
- [Williams & Seeger, 2001] Williams, C. & Seeger, M.: , 2001; Using the nyström method to speed up kernel machines; en *Proceedings of the 14th annual conference on neural information processing systems*; CONF; págs. 682–688.
- [Wilroth *et al.*, 2023] Wilroth, J.; Bernhardsson, B.; Heskebeck, F.; Skoglund, M. A.; Bergeling, C. & Alickovic, E.: , 2023; Improving eeg-based decoding of the locus of auditory attention through domain adaptation; *Journal of Neural Engineering*; **20** (6): 066022.
- [Wriessnegger *et al.*, 2020] Wriessnegger, S. C.; Müller-Putz, G. R.; Brunner, C. & Sburlea, A. I.: , 2020; Inter-and intra-individual variability in brain oscillations during sports motor imagery; *Frontiers in human neuroscience*; **14**: 576241.
- [Wu & Mooney, 2018] Wu, J. & Mooney, R. J.: , 2018; Faithful multimodal explanation for visual question answering; *arXiv preprint arXiv:1809.02805*.
- [Wu *et al.*, 2020a] Wu, J.; Zhou, T. & Li, T.: , 2020a; Detecting epileptic seizures in EEG signals with complementary ensemble empirical mode decomposition and extreme gradient boosting; *Entropy*; **22** (2): 140.
- [Wu *et al.*, 2019] Wu, X.; Zheng, W.-L. & Lu, B.-L.: , 2019; Identifying functional brain connectivity patterns for EEG-based emotion recognition; en *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*; IEEE; págs. 235–238.
- [Wu *et al.*, 2020b] Wu, X.; Zheng, W.-L. & Lu, B.-L.: , 2020b; Investigating EEG-based functional connectivity patterns for multimodal emotion recognition; *arXiv preprint arXiv:2004.01973*.
- [Xiao *et al.*, 2018] Xiao, C.; Choi, E. & 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.
- [Xie *et al.*, 2019] Xie, J.; Liu, F.; Wang, K. & Huang, X.: , 2019; Deep kernel learning via random fourier features; *arXiv preprint arXiv:1910.02660*.
- [Xie *et al.*, 2018] Xie, X.; Yu, Z. L.; Gu, Z.; Zhang, J.; Cen, L. & Li, Y.: , 2018; Bilinear regularized locality preserving learning on riemannian graph for motor imagery BCI; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **26** (3): 698–708.
- [Xie & Oniga, 2020] Xie, Y. & Oniga, S.: , 2020; A review of processing methods and classification algorithm for EEG signal.; *Carpathian Journal of Electronic & Computer Engineering*; **12** (3).
- [Xiong *et al.*, 2020] Xiong, X.; Yu, Z.; Ma, T.; Luo, N.; Wang, H.; Lu, X. & Fan, H.: , 2020; Weighted brain network metrics for decoding action intention understanding based on EEG; *Frontiers in Human Neuroscience*; **14**: 232.
- [Xu *et al.*, 2018] Xu, B.; Zhang, L.; Song, A.; Wu, C.; Li, W.; Zhang, D.; Xu, G.; Li, H. & Zeng, H.: , 2018; Wavelet transform time-frequency image and convolutional network-based motor imagery eeg classification; *Ieee Access*; **7**: 6084–6093.
- [Xu *et al.*, 2020a] Xu, J.; Grosse-Wentrup, M. & Jayaram, V.: , 2020a; Tangent space spatial filters for interpretable and efficient riemannian classification; *Journal of Neural Engineering*; **17** (2): 026043.
- [Xu *et al.*, 2020b] Xu, J.; Zheng, H.; Wang, J.; Li, D. & Fang, X.: , 2020b; Recognition of eeg signal motor imagery intention based on deep multi-view feature learning; *Sensors*; **20** (12): 3496.
- [Xu *et al.*, 2020c] Xu, M.; Yao, J.; Zhang, Z.; Li, R.; Yang, B.; Li, C.; Li, J. & Zhang, J.: , 2020c; Learning eeg topographical representation for classification via convolutional neural network; *Pattern Recognition*; **105**: 107390.
- [Yang *et al.*, 2021a] Yang, J.; Ma, Z. & Shen, T.: , 2021a; Multi-time and multi-band csp motor imagery eeg feature classification algorithm; *Applied Sciences*; **11** (21): 10294.
- [Yang *et al.*, 2021b] Yang, J.; Yu, H.; Shen, T.; Song, Y. & Chen, Z.: , 2021b; 4-class mi-eeg signal generation and recognition with cvae-gan; *Applied Sciences*; **11** (4): 1798.
- [Yang *et al.*, 2021c] Yang, L.; Song, Y.; Ma, K.; Su, E. & Xie, L.: , 2021c; A novel motor imagery eeg decoding method based on feature separation; *Journal of Neural Engineering*; **18** (3): 036022.
- [Yao *et al.*, 2020] Yao, Y.; Ding, Y.; Zhong, S. & Cui, Z.: , 2020; EEG-based epilepsy recognition via multiple kernel learning; *Computational and Mathematical Methods in Medicine*; **2020**.
- [Yeh *et al.*, 2021] Yeh, C.-H.; Jones, D. K.; Liang, X.; Descoteaux, M. & Connelly, A.: , 2021; Mapping structural connectivity using diffusion MRI: Challenges and opportunities; *Journal of Magnetic Resonance Imaging*; **53** (6): 1666–1682.
- [Yeh *et al.*, 2012] Yeh, Y.-R.; Lin, T.-C.; Chung, Y.-Y. & Wang, Y.-C. F.: , 2012; A novel multiple kernel learning framework for heterogeneous feature fusion and variable selection; *IEEE Transactions on multimedia*; **14** (3): 563–574.
- [Yen *et al.*, 2023] Yen, C.; Lin, C.-L. & Chiang, M.-C.: , 2023; Exploring the frontiers of neuroimaging: a review of recent advances in understanding brain functioning and disorders; *Life*; **13** (7): 1472.
- [Yilmaz *et al.*, 2018] Yilmaz, C. M.; Kose, C. & 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.

- [Yu & Yu, 2021] **Yu, J. & Yu, Z. L.**: , 2021; Cross-correlation based discriminant criterion for channel selection in motor imagery BCI systems; *Journal of Neural Engineering*.
- [Yu et al., 2020] **Yu, Z.; Ma, T.; Fang, N.; Wang, H.; Li, Z. & Fan, H.**: , 2020; Local temporal common spatial patterns modulated with phase locking value; *Biomedical Signal Processing and Control*; **59**: 101882.
- [Yuan et al., 2021] **Yuan, K.; Chen, C.; Wang, X.; Chu, W. C.-w. & Tong, R. K.-y.**: , 2021; Bci training effects on chronic stroke correlate with functional reorganization in motor-related regions: A concurrent eeg and fmri study; *Brain sciences*; **11** (1): 56.
- [Yuksel & Olmez, 2015] **Yuksel, A. & Olmez, T.**: , 2015; A neural network-based optimal spatial filter design method for motor imagery classification; *PloS one*; **10** (5): e0125039.
- [Zapała et al., 2020] **Zapała, D.; Zabielska-Mendyk, E.; Augustynowicz, P.; Cudo, A.; Jaśkiewicz, M.; Szewczyk, M.; Kopiś, N. & Francuz, P.**: , 2020; The effects of handedness on sensorimotor rhythm desynchronization and motor-imagery bci control; *Scientific reports*; **10** (1): 2087.
- [Zeiler & Fergus, 2014] **Zeiler, M. D. & Fergus, R.**: , 2014; Visualizing and understanding convolutional networks; en *Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I* 13; Springer; págs. 818–833.
- [Zhang et al., 2021a] **Zhang, A.; Lipton, Z. C.; Li, M. & Smola, A. J.**: , 2021a; Dive into deep learning; *arXiv preprint arXiv:2106.11342*.
- [Zhang et al., 2020a] **Zhang, B.; Yan, G.; Yang, Z.; Su, Y.; Wang, J. & Lei, T.**: , 2020a; Brain functional networks based on resting-state EEG data for major depressive disorder analysis and classification; *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **29**: 215–229.
- [Zhang et al., 2021b] **Zhang, C.; Kim, Y.-K. & Eskandarian, A.**: , 2021b; Eeg-inception: an accurate and robust end-to-end neural network for eeg-based motor imagery classification; *Journal of Neural Engineering*; **18** (4): 046014.
- [Zhang et al., 2020b] **Zhang, D.; Chen, K.; Jian, D. & Yao, L.**: , 2020b; Motor imagery classification via temporal attention cues of graph embedded eeg signals; *IEEE journal of biomedical and health informatics*; **24** (9): 2570–2579.
- [Zhang et al., 2020c] **Zhang, K.; Xu, G.; Han, Z.; Ma, K.; Zheng, X.; Chen, L.; Duan, N. & Zhang, S.**: , 2020c; Data augmentation for motor imagery signal classification based on a hybrid neural network; *Sensors*; **20** (16): 4485.
- [Zhang et al., 2021c] **Zhang, K.; Robinson, N.; Lee, S.-W. & Guan, C.**: , 2021c; Adaptive transfer learning for eeg motor imagery classification with deep convolutional neural network; *Neural Networks*; **136**: 1–10.
- [Zhang et al., 2017] **Zhang, L.; Chen, Y.; Tan, X.; He, C. & Zhang, L.**: , 2017; An improved self-training algorithm for classifying motor imagery electroencephalography in brain-computer interface; *Journal of Medical Imaging and Health Informatics*; **7** (2): 330–337.
- [Zhang et al., 2018a] **Zhang, R.; Xiao, X.; Liu, Z.; Jiang, W.; Li, J.; Cao, Y.; Ren, J.; Jiang, D. & Cui, L.**: , 2018a; A new motor imagery EEG classification method FB-TRCSP+ RF based on CSP and random forest; *IEEE Access*; **6**: 44944–44950.
- [Zhang et al., 2019] **Zhang, R.; Li, X.; Wang, Y.; Liu, B.; Shi, L.; Chen, M.; Zhang, L. & Hu, Y.**: , 2019; Using brain network features to increase the classification accuracy of MI-BCI inefficiency subject; *IEEE Access*; **7**: 74490–74499.
- [Zhang et al., 2021d] **Zhang, R.; Zong, Q.; Dou, L.; Zhao, X.; Tang, Y. & Li, Z.**: , 2021d; Hybrid deep neural network using transfer learning for eeg motor imagery decoding; *Biomedical Signal Processing and Control*; **63**: 102144.
- [Zhang et al., 2021e] **Zhang, T.; Han, Z.; Chen, X. & Chen, W.**: , 2021e; Subbands and cumulative sum of subbands based nonlinear features enhance the performance of epileptic seizure detection; *Biomedical Signal Processing and Control*; **69**: 102827.
- [Zhang et al., 2018b] **Zhang, W.; Tan, C.; Sun, F.; Wu, H. & Zhang, B.**: , 2018b; A review of EEG-based brain-computer interface systems design; *Brain Science Advances*; **4** (2): 156–167.
- [Zhang et al., 2021f] **Zhang, W.; Liang, Z. & Liu, Z.**: , 2021f; Combination of variational mode decomposition for feature extraction and deep belief network for feature classification in motor imagery electroencephalogram recognition.; *Sensors & Materials*; **33**.
- [Zhang et al., 2020d] **Zhang, X.; Lu, D.; Shen, J.; Gao, J.; Huang, X. & Wu, M.**: , 2020d; Spatial-temporal joint optimization network on covariance manifolds of electroencephalography for fatigue detection; en *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; IEEE; págs. 893–900.
- [Zhang et al., 2021g] **Zhang, X.; Yao, L.; Wang, X.; Monaghan, J.; Mcalpine, D. & Zhang, Y.**: , 2021g; A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers; *Journal of neural engineering*; **18** (3): 031002.
- [Zhang et al., 2018c] **Zhang, Y.; Nam, C.; Zhou, G.; Jin, J.; Wang, X. & Cichocki, A.**: , 2018c; Temporally constrained sparse group spatial patterns for motor imagery BCI; *IEEE transactions on cybernetics*; **49** (9): 3322–3332.
- [Zhao et al., 2021a] **Zhao, N.; Zhang, H.; Liu, T.; Liu, J.; Xiang, Y.; Shu, G.; Li, C.; Xie, J. & Chen, L.**: , 2021a; Neuromodulatory effect of sensorimotor network functional connectivity of temporal three-needle therapy for ischemic stroke patients with motor dysfunction: Study protocol for a randomized, patient-assessor blind, controlled, neuroimaging trial; *Evidence-Based Complementary and Alternative Medicine*; **2021**.
- [Zhao et al., 2019a] **Zhao, X.; Zhang, H.; Zhu, G.; You, F.; Kuang, S. & Sun, L.**: , 2019a; A multi-branch 3d convolutional neural network for eeg-based motor imagery classification; *IEEE transactions on neural systems and rehabilitation engineering*; **27** (10): 2164–2177.

## REFERENCES

## REFERENCES

- [Zhao et al., 2019b] **Zhao, Y.; Zhao, Y.; Durongbhan, P.; Chen, L.; Liu, J.; Billings, S.; Zis, P.; Unwin, Z. C.; De Marco, M.; Venneri, A.** et al.: , 2019b; Imaging of nonlinear and dynamic functional brain connectivity based on EEG recordings with the application on the diagnosis of alzheimer's disease; *IEEE transactions on medical imaging*; **39** (5): 1571–1581.
- [Zhao et al., 2020] **Zhao, Z.; Zhao, P. & Zhou, Q.**: , 2020; A hierarchical stacking extreme learning machine for multi-classification; en *2020 Chinese Automation Congress (CAC)*; IEEE; págs. 4176–4181.
- [Zhao et al., 2021b] **Zhao, Z.; Li, J.; Niu, Y.; Wang, C.; Zhao, J.; Yuan, Q.; Ren, Q.; Xu, Y. & Yu, Y.**: , 2021b; Classification of schizophrenia by combination of brain effective and functional connectivity; *Frontiers in Neuroscience*; **15**: 552.
- [Zhu et al., 2020] **Zhu, L.; Su, C.; Zhang, J.; Cui, G.; Cichocki, A.; Zhou, C. & Li, J.**: , 2020; EEG-based approach for recognizing human social emotion perception; *Advanced Engineering Informatics*; **46**: 101191.
- [Zhuang et al., 2020] **Zhuang, M.; Wu, Q.; Wan, F. & Hu, Y.**: , 2020; State-of-the-art non-invasive brain-computer interface for neural rehabilitation: A review; *Journal of Neurorestoratology*; **8** (1): 12–25.
- [Zoumpourlis & Patras, 2022] **Zoumpourlis, G. & Patras, I.**: , 2022; Covmix: Covariance mixing regularization for motor imagery decoding; en *2022 10th International Winter Conference on Brain-Computer Interface (BCI)*; IEEE; págs. 1–7.
- [Zuk et al., 2020] **Zuk, N. J.; Teoh, E. S. & Lalor, E. C.**: , 2020; Eeg-based classification of natural sounds reveals specialized responses to speech and music; *NeuroImage*; **210**: 116558.