

# Machine Learning Approaches for Neuro-engineering Applications

Andrés Marino Álvarez-Meza, Ph.D.

Signal Processing and Recognition Group - (SPRG)  
Universidad Nacional de Colombia  
Manizales, Colombia  
ACHIRP - January, 2021

# Outline

1 Signal Processing and Recognition Group - UNAL

2 Brain computer interfaces

3 Challenges

4 SPRG approaches

- Kernel-based connectivity analysis for BCI
- Deep&Wide Learning for BCI

5 Conclusions

6 Ongoing work

# Contents

- 1 Signal Processing and Recognition Group - UNAL
- 2 Brain computer interfaces
- 3 Challenges
- 4 SPRG approaches
  - Kernel-based connectivity analysis for BCI
  - Deep&Wide Learning for BCI
- 5 Conclusions
- 6 Ongoing work

# Universidad Nacional de Colombia sede-Manizales (UNAL)



# Universidad Nacional de Colombia sede-Manizales (UNAL)



# Signal Processing and Recognition Group

Since 1998 - Dir.: Prof. Germán Castellanos



# Signal Processing and Recognition Group

## Research and academic interests

### Current Courses:

- Signals and Systems (Electrical and Electronic Eng.).
- Signal theory (Electrical and Electronic Eng.).
- Digital signal processing (Electrical and Electronic Eng.).
- Data analysis (Electrical and Electronic Eng.).
- Image processing (Electrical and Electronic Eng.).
- Stochastic process (M.Sc. and Ph.D. in automatics).
- Machine learning (M.Sc. and Ph.D. in automatics).
- Advanced Machine Learning (M.Sc. and Ph.D. in automatics).

# Signal Processing and Recognition Group

## Research and academic interests

### Research interests:

- Data analysis.
- Smart agriculture.
- Computer vision.
- *Computer-aided systems from biosignal data.*
- *Neuro-engineering.*

# Signal Processing and Recognition Group

## Cloud computing resources



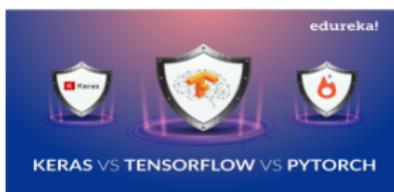
# kaggle™



Free! Cloud Server  
TPU & GPU

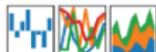


Google  
colab



## pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



	ScikitLearn	TensorFlow	PyTorch
0	Classification	450	400
1	Reg.	510	400
2	Open	590	500
3	Rec.	670	540
4	NeuralNet	660	360
5	ShallowNet	1050	1170
6	ShallowNet_U	2200	2100
7	Other	1000	1000

## scikit learn

# Signal Processing and Recognition Group

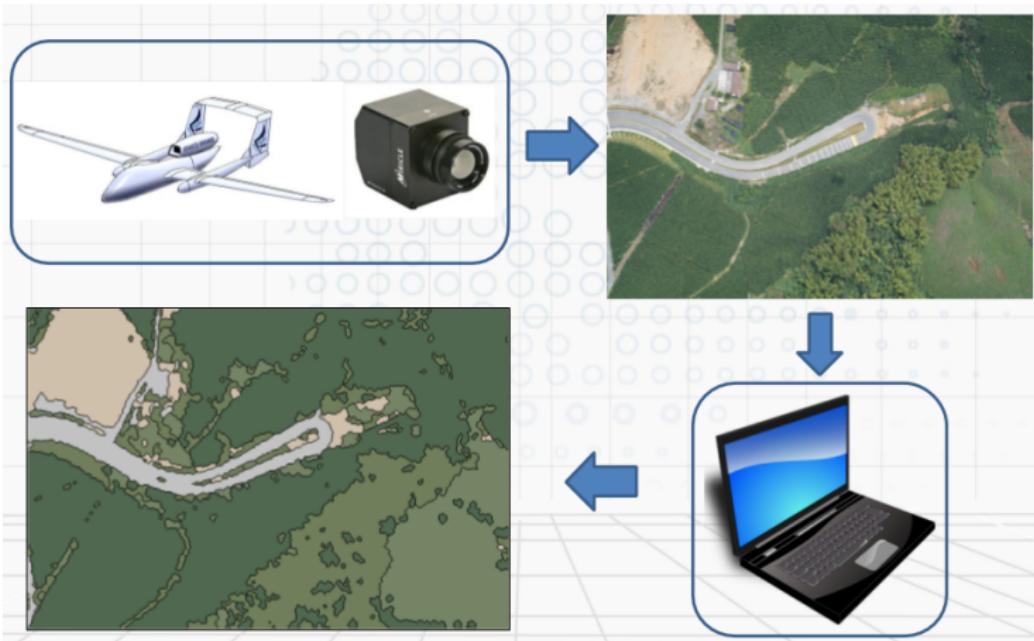
## Minciencias and UNAL projects



**Video-based activity recognition  
(Surveillance systems)**

# Signal Processing and Recognition Group

## Minciencias and UNAL projects

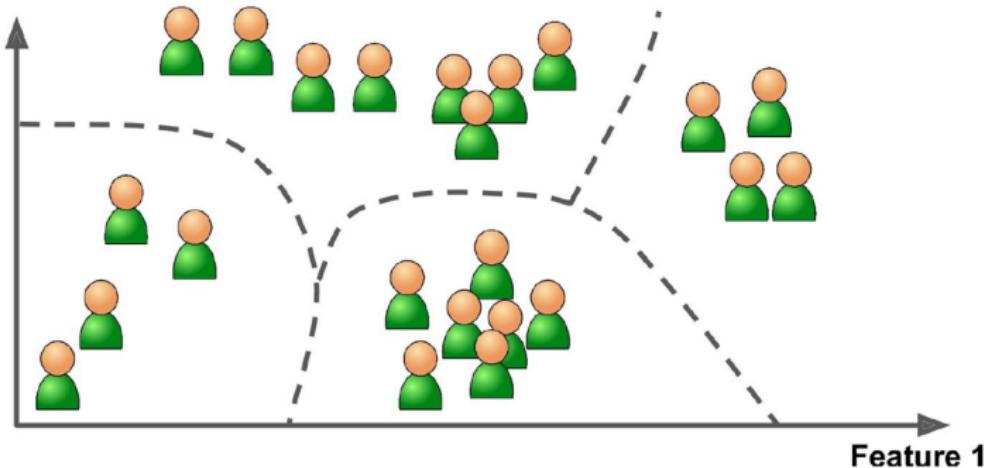


**Smart agriculture  
(Aerial image segmentation)**

# Signal Processing and Recognition Group

## Minciencias and UNAL projects

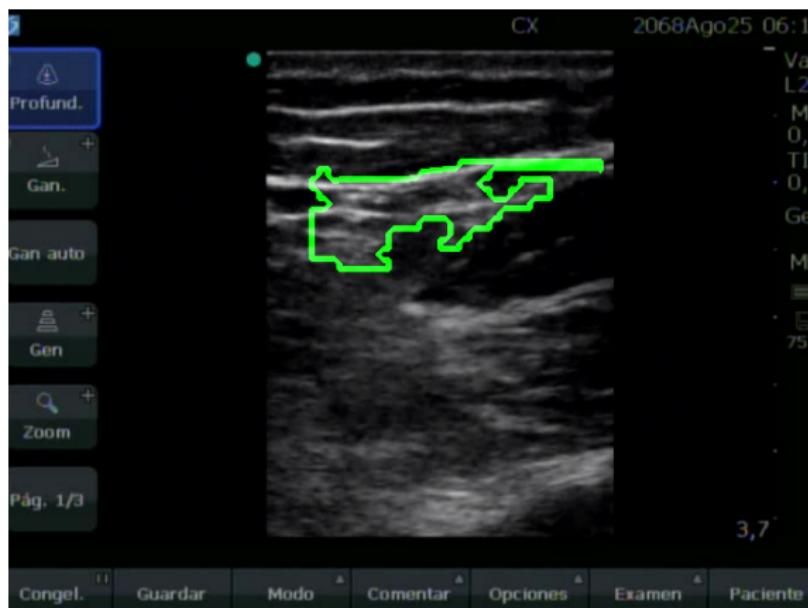
Feature 2



Smart agriculture  
(Data analysis)

# Signal Processing and Recognition Group

## Minciencias and UNAL projects



**Ultra-sound image processing for nerve segmentation  
(Anaesthetic effects)**

# Signal Processing and Recognition Group

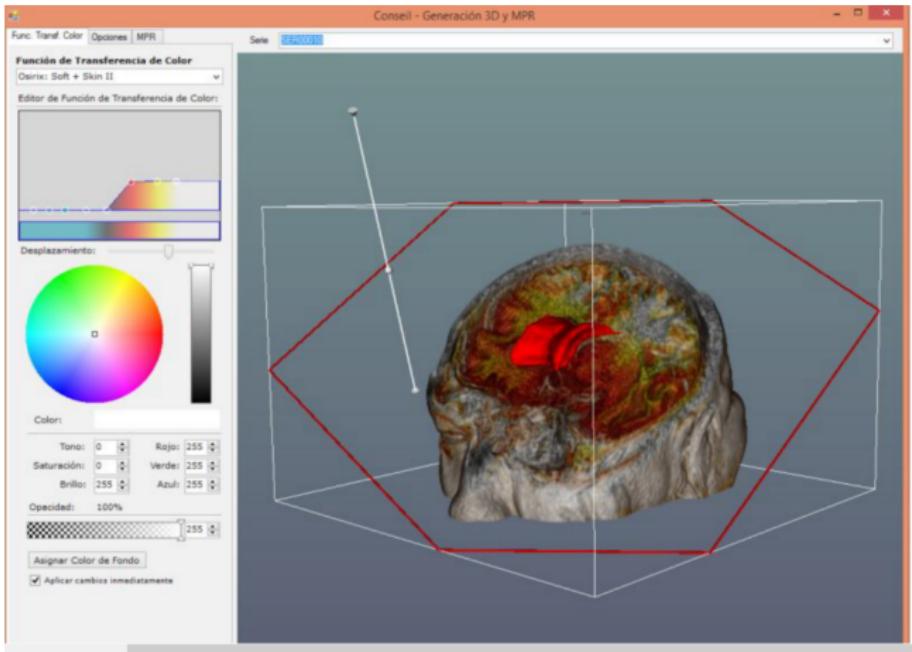
## Minciencias and UNAL projects



Infrared thermography image processing  
(Anaesthetic effects)

# Signal Processing and Recognition Group

## Minciencias and UNAL projects



**MRI processing to support  
Deep Brain Stimulation (Parkinson's disease)**

# Contents

1 Signal Processing and Recognition Group - UNAL

2 Brain computer interfaces

3 Challenges

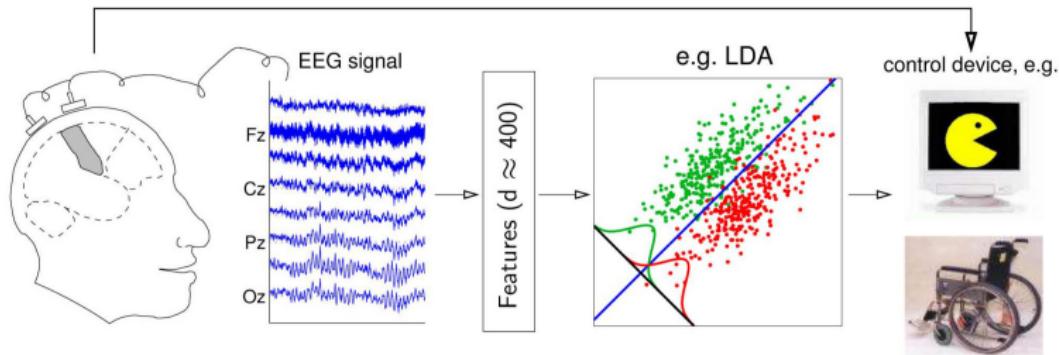
4 SPRG approaches

- Kernel-based connectivity analysis for BCI
- Deep&Wide Learning for BCI

5 Conclusions

6 Ongoing work

# Brain Computer Interfaces (BCI)



source: TU Charite Campus Benjamin Franklin - Machine learning for BCI

BCI: translation of human intentions into a technical control signal without using activity of muscles or peripheral nerves

# Brain Computer Interfaces (BCI)



- Assistive care.
- Gaming and entertainment.
- Cognitive improvement.
- Restoring neural and/or behavioral function.

BCI has become one of the most interesting alternatives to support automatic systems able to interpret brain functions.

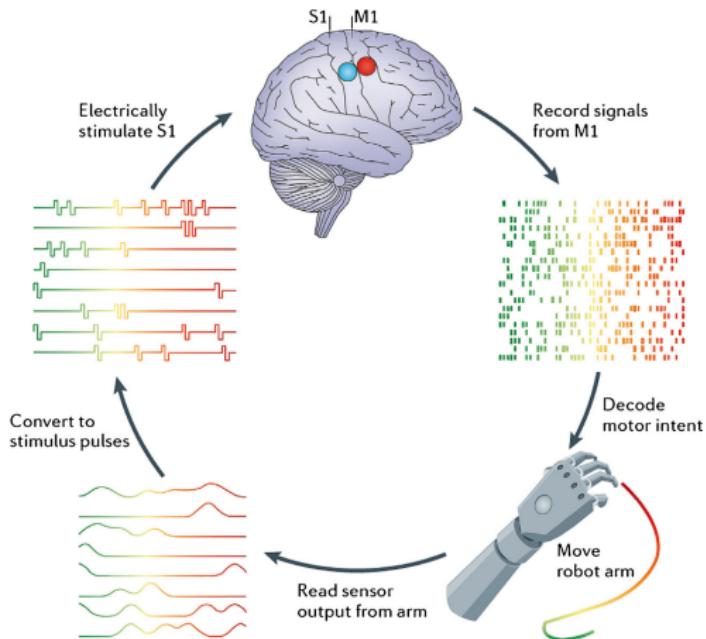
# BCI interesting remarks

- Grey Walter demonstrated use of **non-invasively brain recordings** from a human subject **to control a slide projector (1964)** [Graimann et al., 2010].
- Fetz demonstrated **increase of the firing rates** of neurons in the **motor cortex** in non-human primates along with **auditory or visual feedback (1969)** [Fetz, 1969].
- BCI was coined by Jaceques Vidal in **1971** aiming to interface human brain with computers (University of California) [Vidal, 1977].

# BCI interesting remarks

- The field of BCI has expanded for both **invasive** and **non-invasive** neural recordings in humans and animals.
- Nowadays, both **sensorimotor** and **cognitive functions** are studied, incorporating **feedback** mechanisms in **closed loop systems** [Miranda et al., 2015].

# Idealized bidirectional BCI



*Decoding:* extracting information from neural responses

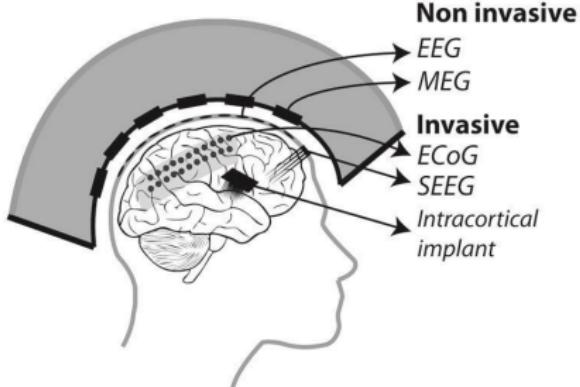
*Encoding:* representing the information of neural responses

# Contents

- 1 Signal Processing and Recognition Group - UNAL
- 2 Brain computer interfaces
- 3 Challenges
- 4 SPRG approaches
  - Kernel-based connectivity analysis for BCI
  - Deep&Wide Learning for BCI
- 5 Conclusions
- 6 Ongoing work

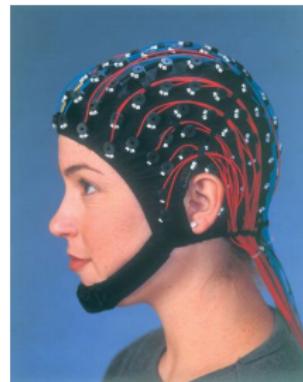
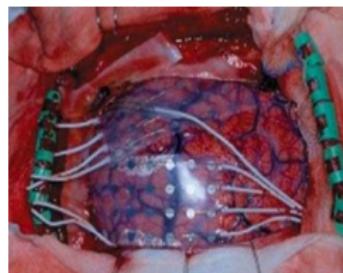
# BCI technologies

- **Invasive:** detection of *single neuron activity* by intra-cortical electrodes into the gray matter (rats and monkeys).
- **Partially invasive:** recording electrocorticographic (ECoG) implants inside the skull but outside the gray matter.
- **Noninvasive:** recording of the brain activity at a *macro level*.

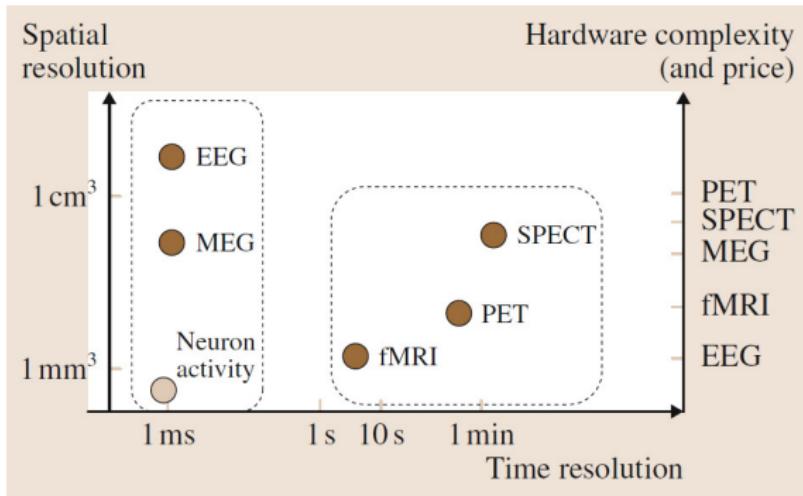


# BCI technologies

- **Invasive** techniques are more reliable and less noisy.  
**Drawbacks:** neurosurgery is required, long-term use or periodical replacement.
- **Noninvasive** are more safe and cheap. Subject preparation is fast being possible to perform real-time analysis.  
**Drawbacks:** noisy signal, time/spatial resolution issues.

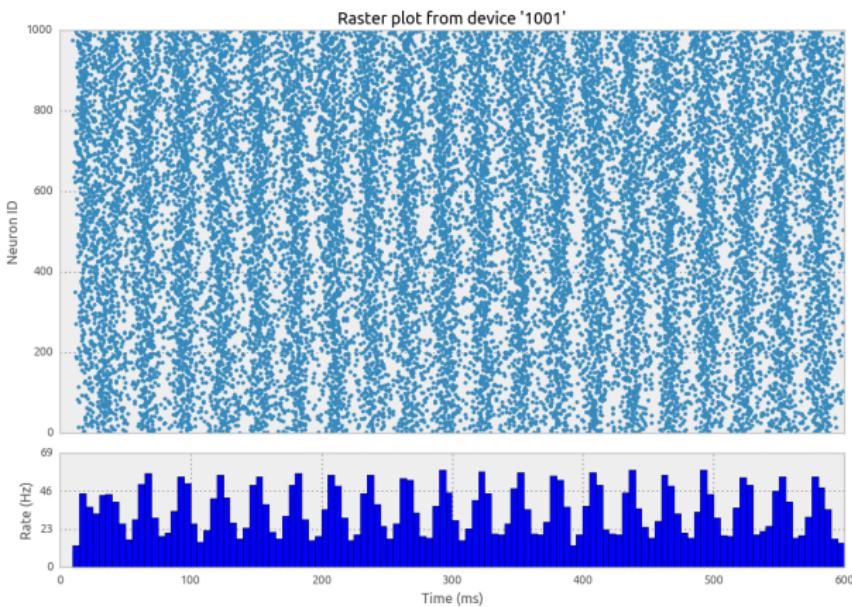


# BCI technologies



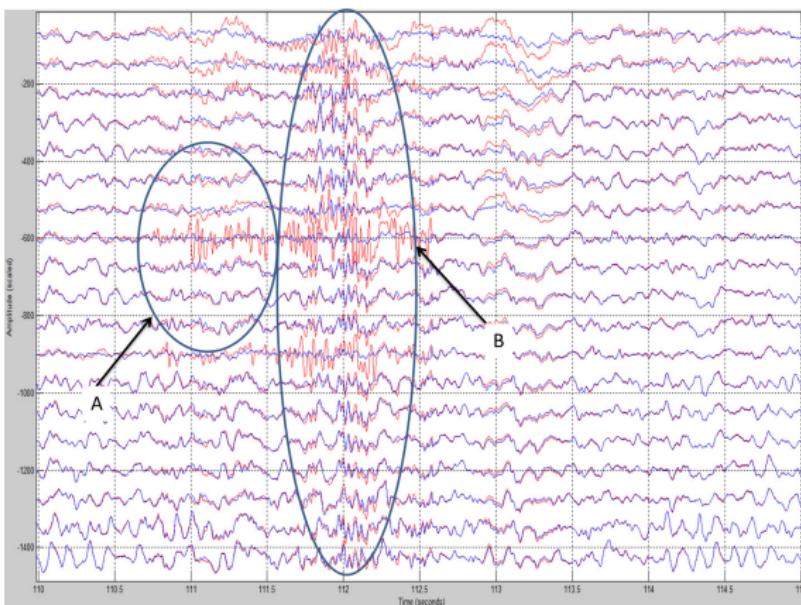
- Neuron activity - micro-array electrodes
- EEG: Electroencephalography
- MEG: Magnetoencephalography
- fNIR: functional near-infrared systems
- fMRI: functional magnetic resonance imaging
- SPECT: single photon emission computerized tomography
- PET: positron emission tomography

# BCI technologies



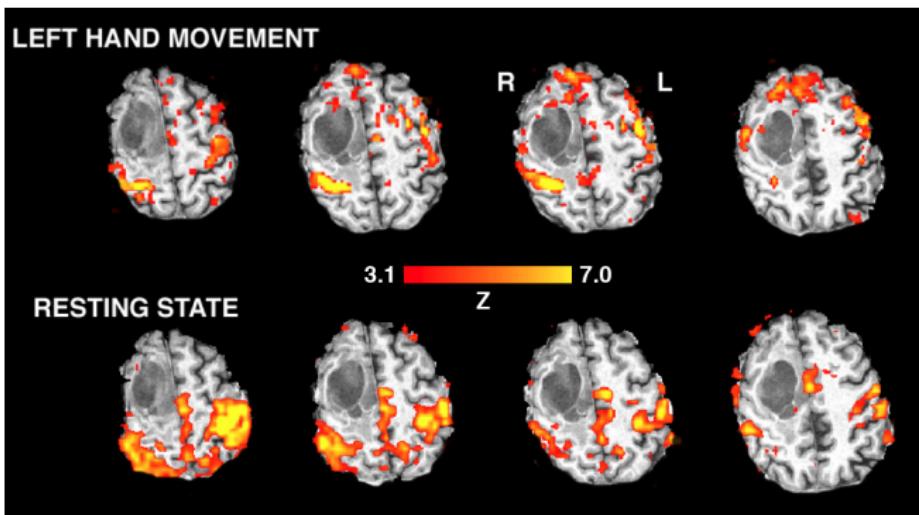
Neuron activity - spikes

# BCI technologies



EEG

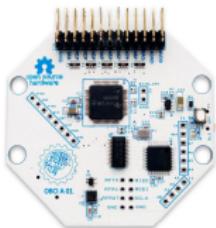
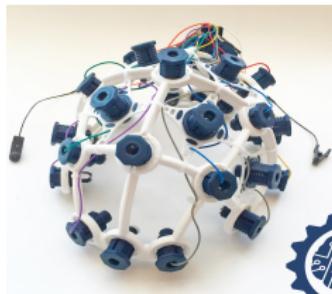
# BCI technologies



# BCI technologies

- Micro-array electrodes provide a good temporal and spatial resolution.  
**Drawback:** expensive, surgery is required.
- EEG has a good temporal resolution.  
**Drawback:** the worst spatial resolution.
- fMRI closest spatial resolution to the real neuron activity.  
**Drawback:** too expensive, poor temporal resolution.
- SPECT and PET do not provide resolution advantages and are expensive.
- fNIR is cheaper than fMRI, however, both of them are based on changes in cerebral blood flow (slow response).
- **EEG is the most practical technology for BCI.**

# BCI technologies



OPENBCI



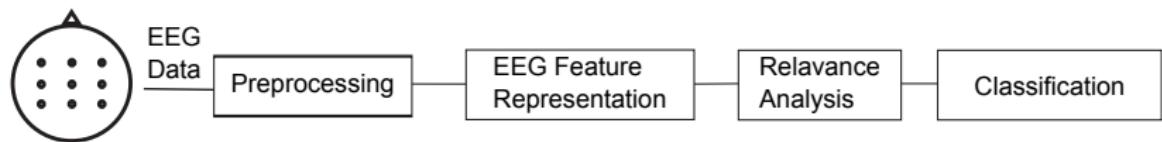
EMOTIV

Cheap EEG sensors

# BCI and Machine Learning - (ML)

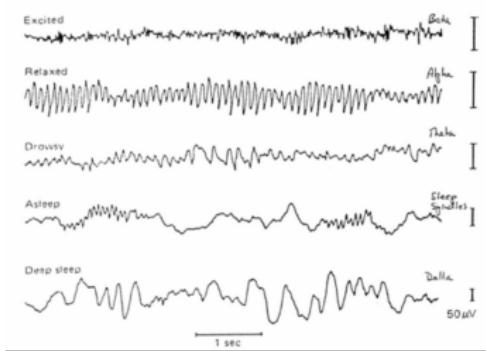
- **Supervised learning scenario:** to find a rule of **association** between pairs  $(x_n, y_n)$  of observed **inputs** and corresponding **targets** given an **expert**.
- **Unsupervised learning scenario:** the **only available** information is the observed **inputs**. The assumption is the presence of some **data regularities** since there is a process behind their generation.
- **Reinforcement learning scenario:** the system **interacts** with the **environment** by performing actions that **feedback** to the system in the form of **rewards** or **punishments**.

# BCI within a ML framework



A BCI system from a machine learning point of view

# BCI and ML issues



- Noisy signal
- Inter/Intra-subject variability
- Coding Time/Space/Freq. patterns
- Non-stationary data
- Redundant/Irrelevant information
- Lack of interpretability

# BCI and ML issues

- Estimating the **decoding/encoding models** required by BCIs is difficult on this diverse **high-dimensional** data.
- There are **multiple modalities** by which the signal characteristics differ between conditions.
- **Machine learning** tools are required to learn an appropriate neural **representation space**.

How can we represent brain activity using machine learning approaches to support BCI systems?

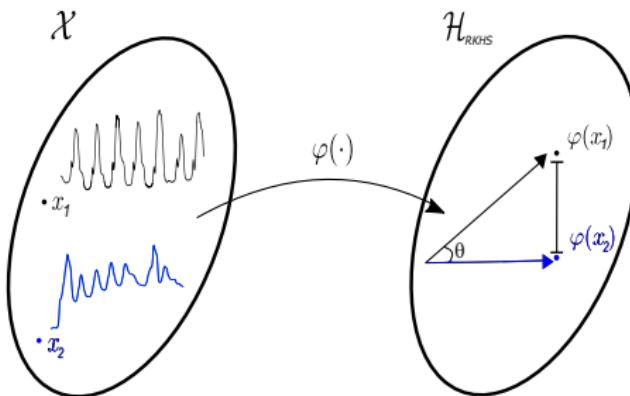
# Contents

- 1 Signal Processing and Recognition Group - UNAL
- 2 Brain computer interfaces
- 3 Challenges
- 4 SPRG approaches
  - Kernel-based connectivity analysis for BCI
  - Deep&Wide Learning for BCI
- 5 Conclusions
- 6 Ongoing work

# Kernel methods

- Kernel methods allow representing relevant data structures, even for non-stationary processes.  
[Principe, 2010, Muandet et al., 2017].
- Kernel methods handle non-linear operations on data by indirectly computing a mapping to a Reproducing Kernel Hilbert Space—(RKHS), where linear operations can be carried out [Nishiyama et al., 2020].

# Reproducing kernel Hilbert Space - (RKHS)



- Non-linear dependencies can be extracted in RKHS from input time-series [Muandet et al., 2017].
- Functional (non-directed) and effective (directed) connectivity are computed to code spatial patterns in BCI systems [Timme and Lapish, 2018].

# Effective connectivity based on Transfer Entropy - (TE)

- TE measures the deviation from the following generalized Markov condition [Vicente et al., 2011]:

$$p(y_{t+1}|\mathbf{y}_t^m, \mathbf{x}_t^n) - p(y_{t+1}|\mathbf{y}_t^m) \rightarrow 0 \quad (1)$$

$\mathbf{x}_t^n \in \mathbb{R}^n, \mathbf{y}_t^m \in \mathbb{R}^m$ : Markov processes.

$\mathbf{x} = \{x_t\}_{t=1}^l, \mathbf{y} = \{y_t\}_{t=1}^l$ : time-series, i.e., EEG channels.

- Based on the Shannon entropy estimator, TE yields:

$$\begin{aligned} TE(\mathbf{x} \rightarrow \mathbf{y}) &= H_S \left( \mathbf{y}_{t-1}^{dy}, \mathbf{x}_{t-u}^{dx} \right) - H_S \left( y_t, \mathbf{y}_{t-1}^{dy}, \mathbf{x}_{t-u}^{dx} \right) + \dots \\ &\dots + H_S \left( y_t, \mathbf{y}_{t-1}^{dy} \right) - H_S \left( \mathbf{y}_{t-1}^{dy} \right). \end{aligned} \quad (2)$$

# Kernel-based TE

- TE is extended using a matrix-based Renyi- $\alpha$  entropy estimator [Giraldo et al., 2015]:

$$H_\alpha(\mathbf{A}) = \frac{1}{1-\alpha} \log (\text{tr}(\mathbf{A}^\alpha)), \quad (3)$$

where  $\mathbf{A}$  is a Gramm (kernel) matrix (s.t.  $\text{tr}(\mathbf{A}) = 1$ ).

- The joint entropy is re-written as:

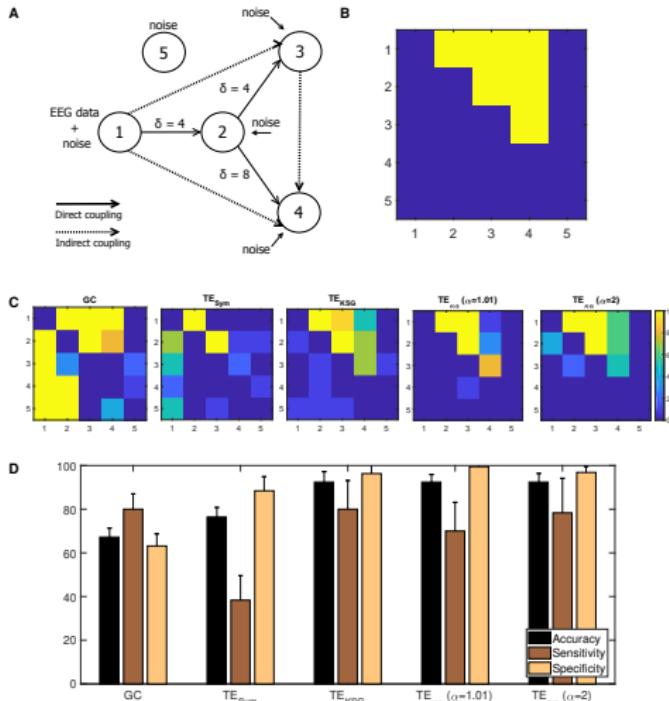
$$H_\alpha(\mathbf{A}, \mathbf{B}) = H_\alpha \left( \frac{\mathbf{A} \circ \mathbf{B}}{\text{tr}(\mathbf{A} \circ \mathbf{B})} \right). \quad (4)$$

- The conditional entropy is defined as:

$$H_\alpha(\mathbf{A}|\mathbf{B}) = H_\alpha(\mathbf{A}, \mathbf{B}) - H_\alpha(\mathbf{B}). \quad (5)$$

# Experiments: TE estimation

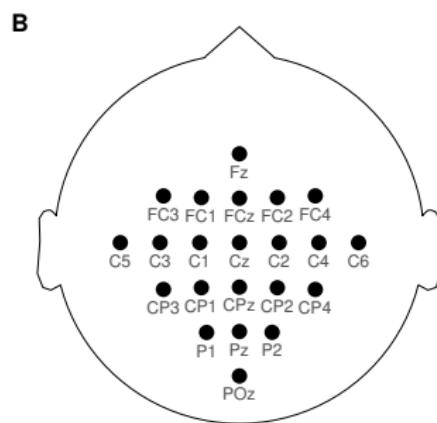
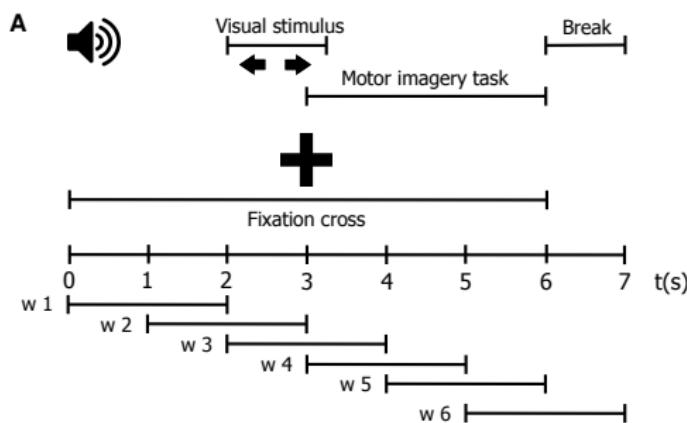
## Modified Kus model network



- $\delta \in \mathbb{N}$  stands for the time shift.
- Kernel-based TE determines the actual connections.
- TE-based can be useful to represent EEG data in BCI systems.
- For details see:  
[De La Pava Panche et al., 2019].

# Experiments: Motor Imagery classification

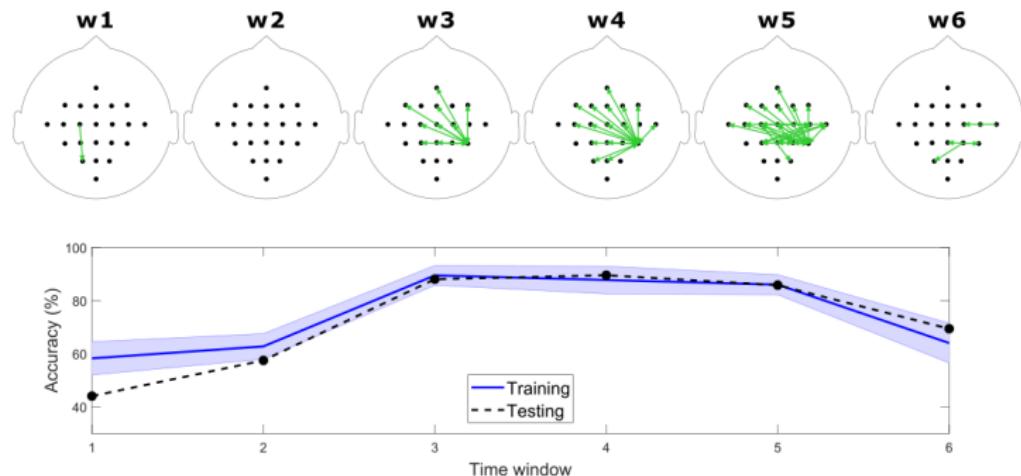
## BCI competition IV-2a dataset<sup>1</sup>



- Tested Motor Imagery-(MI) classes: Left hand vs. Right hand.
- Kernel-TE is used as feature representation.
- Variability-based feature selection and LDA classifier are used.

<sup>1</sup><http://www.bbci.de/competition/iv/desc2a.pdf>

# Experiments: Motor Imagery classification BCI competition IV-2a



Kernel-based TE favors the MI discrimination and the spatio-temporal data interpretability  
(see [De La Pava Panche et al., 2019]).

# Multi-output Gaussian Process with Spectral Mixture for connectivity analysis

- **Cramer's theorem:** a family of integrable functions  $\{\kappa_{ij}(\tau)\}_{i,j=1}^C$  holds the covariance functions of a weakly-stationary stochastic process:

$$\kappa_{ij}(\tau) = \int_{\mathbb{R}^n} e^{i\omega\tau} S_{ij}(\omega) d\omega, \quad (6)$$

$i$ : imaginary unit;  $S_{ij} : \mathbb{R} \rightarrow \mathbb{C}$  (positive definite);  
 $i, j$  time-series indexes.

- $S(\omega) = \mathbf{R}^H(\omega) \mathbf{R}(\omega)$ ; where  $\mathbf{R}(\omega) \in \mathbb{R}^{Q \times C}$ ;  
 $Q$ : decomposition rank.

# Multi-output Gaussian Process with Spectral Mixture for connectivity analysis

- $S = \{S_{ij}\}_{i,j=1}^C \in \mathbb{R}^{C \times C}$ : positive-definite complex-valued functions used as cross-spectral densities [Parra and Tobar, 2017]:
- The autocovariance function  $R_i(\omega)$  of the  $i$ -th channel is modeled as a complex-valued squared exponential:

$$R_i(\omega) = w_i \exp\left(-\frac{1}{4} \frac{(\omega - \mu_i)^2}{\sigma_i^2}\right) \exp(-\iota(\theta_i \omega + \phi_i)), \quad (7)$$

$w_i, \phi_i, \mu_i, \theta_i \in \mathbb{R}; \sigma_i \in \mathbb{R}^+$ .

- The cross-spectral density between channels  $i$  and  $j$  yields:

$$S_{ij}(\omega) = w_{ij} \exp\left(-\frac{1}{2} \frac{(\omega - \mu_{ij})^2}{\sigma_{ij}} + \iota(\theta_{ij} \omega + \phi_{ij})\right); \quad (8)$$

covariance:  $\sigma_{ij} \in \mathbb{R}^+$ ; mean:  $\mu_{ij} \in \mathbb{R}$ ; magnitude:  $w_{ij} \in \mathbb{R}$ ;  
 delay:  $\theta_{ij} \in \mathbb{R}$ ; phase:  $\phi_{ij} \in \mathbb{R}$ .

# Experiments: Motor Imagery classification BCI competition IV-2a

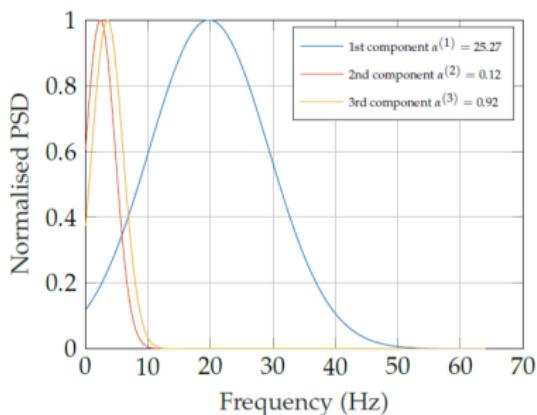
- The Fourier transform relates the resulting cross-spectral density and the kernel-based covariance function.
- A Multi-output Spectral Mixture Gaussian Process is used to extract a functional connectivity-based representation, solving:

$$-\log p(\mathbf{y}|\mathbf{t}, \Theta) = \frac{CT}{2} \log 2\pi + \frac{1}{2} \log |\mathbf{K}| + \frac{1}{2} \mathbf{y}^\top \mathbf{K}^{-1} \mathbf{y}, \quad (9)$$

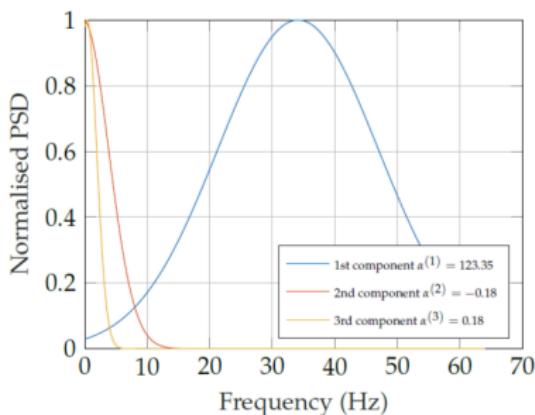
$$\Theta = \{w_i^{(q)}, \mu_i^{(q)}, \sigma_i^{(q)}, \theta_i^{(q)}, \phi_i^{(q)}, \sigma_i^2\}_{i=1,q=1}^{C,Q}.$$

- A spectral kernel is solved to quantify EEG channel relationships at automatically tuned frequency bands (for details see [Torres-Valencia et al., 2020]).

# Experiments: Motor Imagery classification BCI competition IV-2a



(a) Left hand MI



(b) Right hand MI

MOGP-spectral mixture codes relevant spatio-frequency patterns  
(alpha and beta rhythms- see [Torres-Valencia et al., 2020]).

# Experiments: Motor Imagery classification BCI competition IV-2a

**Table:** Classification results [%] for the MI dataset. The mean value is shown for comparison consistency against reported works.

Reference	Approach	Accuracy
[Qureshi et al., 2017]	ICA - ELM	94.29
[Li et al., 2018]	CSP - SVM	78.78
[Liang et al., 2016]	PDC - MEMD	70.22
[Elasuty and Eldawlatly, 2015]	DBN	73.44
[Gómez et al., 2019]	CSP - SVM	81.41
Ours [Torres-Valencia et al., 2020]	DMOSM-GP	82.11

- Competitive results are obtained holding spatio-frequency interpretability.
- More robust ML representations should be tested.

# Deep&Wide Learning from EEG-based Topoplot maps

- Topoplot maps are built from the Power Spectral Density and the Continuous Wavelet Transform
- Convolutional Neural Network -(CNN) layers:

$$\Xi^{r,i,\Delta f} = \gamma_1 (\mathbf{K}_i \otimes \mathbf{S}(\rho^{r,\Delta f}) + \mathbf{b}_i), \quad (10)$$

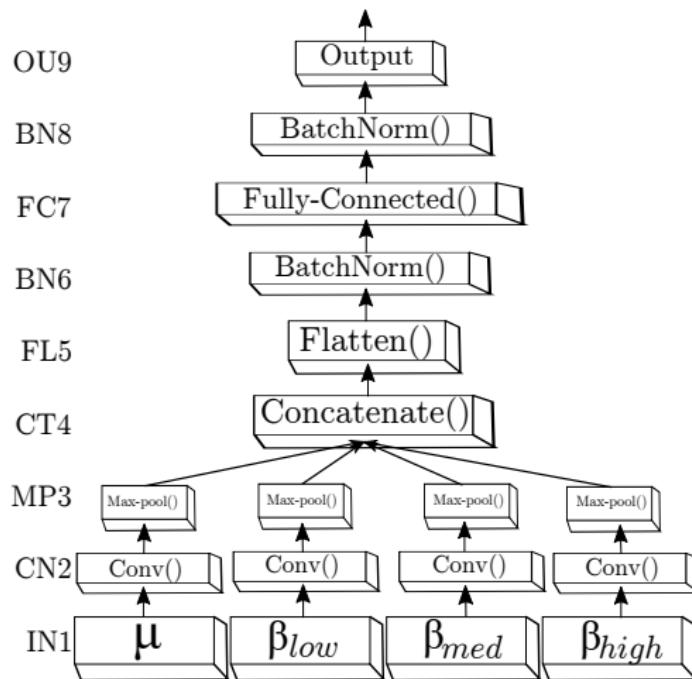
$\Xi^{r,i,\Delta f} \in \mathbb{R}^{S \times S'}$ : feature map;  $\{\mathbf{K}_i \in \mathbb{R}^{K \times K} : i \in I\}$ : kernel filters;  $\{\mathbf{b}_i \in \mathbb{R}^{SS'}\}$ : bias;  $\mathbf{S}(\rho^{r,\Delta f})$ : EEG-based topographic map.

- Fully connected layers:

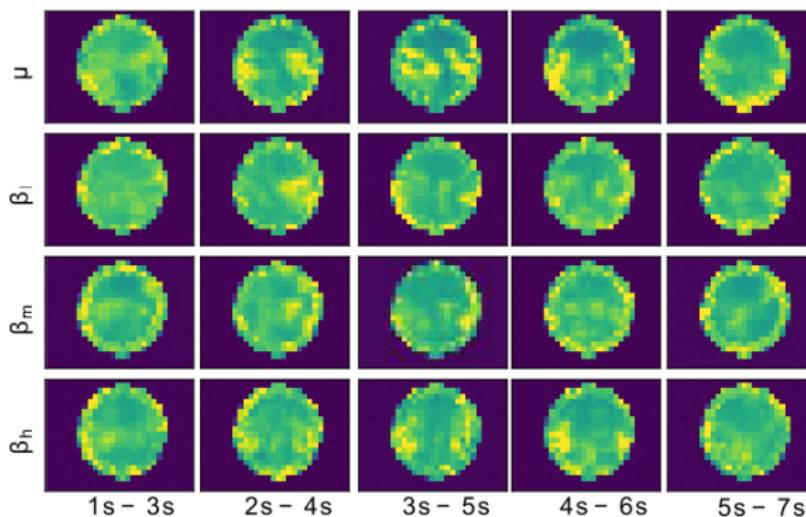
$$\mathbf{h}^r(q) = \gamma_2 (\mathbf{W}(q) \mathbf{h}^r(q-1) + \boldsymbol{\beta}(l)), \quad (11)$$

$\mathbf{h}^r(1) = \bar{\xi}^r$  (concatenation+flatten);  $\mathbf{W}(q) \in \mathbb{R}^{GG'IN_f \times N_h(q)}$ ;  $N_h$  : # neurons;  $\boldsymbol{\beta}(q) \in \mathbb{R}^{N_h(q)}$  : bias;  $\gamma_2(\cdot)$  : activation function.

# Deep&Wide Learning from EEG-based Topoplot maps



# Experiments: Motor Imagery classification BCI competition IV-2a



D&WL relevance analysis (spatio/temporal/frequency) from CWT-based topoplots (for details see [Collazos-Huertas et al., 2020]).

# Experiments: Motor Imagery classification

## BCI competition IV-2a



**Table 2 Bi-class accuracy of evaluated CNN training strategies, using the CWT-extracted vectors and either dropping strategy: CWT\* with sensorimotor electrodes and CWT\*\* with thresholding. In all compared cases, both sub-bands ( $\mu$  and  $\beta$ ) are included and the CNN parameters are tuned individually**

Subjects	[41]	[42]	CWT	$\kappa$	CWT*	$\kappa$	CWT**	$\kappa$
A03T	88.2	91.7	$95.0 \pm 4.6$	0.67	$96.4 \pm 4.8$	0.92	$95.0 \pm 4.6$	0.90
A09T	82.7	90.9	$94.8 \pm 4.2$	0.68	$93.1 \pm 6.5$	0.86	$94.0 \pm 6.3$	0.88
A08T	91.8	92.3	<u><math>94.0 \pm 4.6</math></u>	0.90	$97.0 \pm 3.6$	0.94	$94.7 \pm 4.9$	0.89
A06T	65.7	78.5	$86.7 \pm 7.2$	0.71	$84.9 \pm 9.0$	0.69	$86.7 \pm 7.2$	0.73
A07T	51.7	86.5	$86.4 \pm 6.7$	0.58	<u><math>81.9 \pm 6.2</math></u>	0.64	$85.6 \pm 9.3$	0.71
A04T	53.9	80.4	$85.4 \pm 7.3$	0.73	$86.1 \pm 7.5$	0.72	$87.6 \pm 5.0$	0.75
A02T	63.9	68.4	$83.8 \pm 6.5$	0.73	$80.3 \pm 6.2$	0.61	$81.7 \pm 4.8$	0.63
A01T	79.4	87.8	$83.4 \pm 5.5$	0.88	$81.1 \pm 5.0$	0.62	$83.2 \pm 3.9$	0.66
A05T	54.9	88.9	$79.1 \pm 4.8$	0.90	$78.3 \pm 7.4$	0.57	$76.7 \pm 6.9$	0.53
Average	70.2	$85.0 \pm 7.4$	$87.6 \pm 5.7$	0.75	$86.6 \pm 6.2$	0.73	$87.4 \pm 5.7$	0.74

D&WL achieves competitive MI discrimination results  
 (For details see [Collazos-Huertas et al., 2020]).

# Deep&Wide Learning as Electrophysiological Indicator in Mi Tasks

- Up to 30% of users may not develop enough coordination skills after MI training sessions because of inter and intra-subject variability.
- A D&WL approach is tuned to support pre-training neural desynchronization and initial training synchronization to predict the bi-class accuracy response:

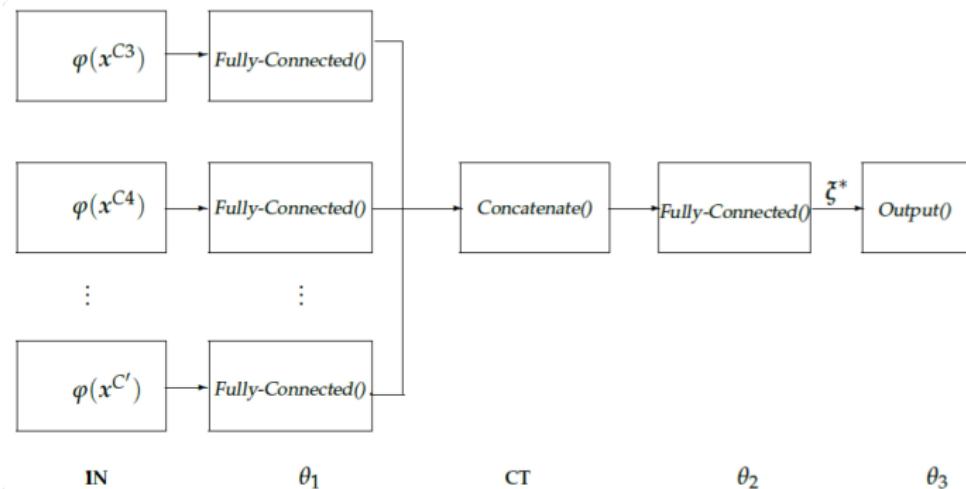
$$\min_{\pi} \mathbf{E}\{\|\psi(V_m) - (\theta_L \circ \dots \circ \theta_1)(\varphi(\mathbf{x}_m^c)|\pi)\|_2\}, \quad (12)$$

$\psi(V_m) \in \mathcal{Y}$ : assessed MI discrimination;

$\varphi(\mathbf{x}_m^c) \in \mathcal{X}$ : measured patterns from resting state (60s single trial before sessions) or initial training signals (2s before MI);

$\pi$  : D&WL parameters.

# Deep&Wide Learning as Electrophysiological Indicator in Mi Tasks



Deep&Wide Learning arquitecture  
(For details see [Velasquez-Martinez et al., 2020])

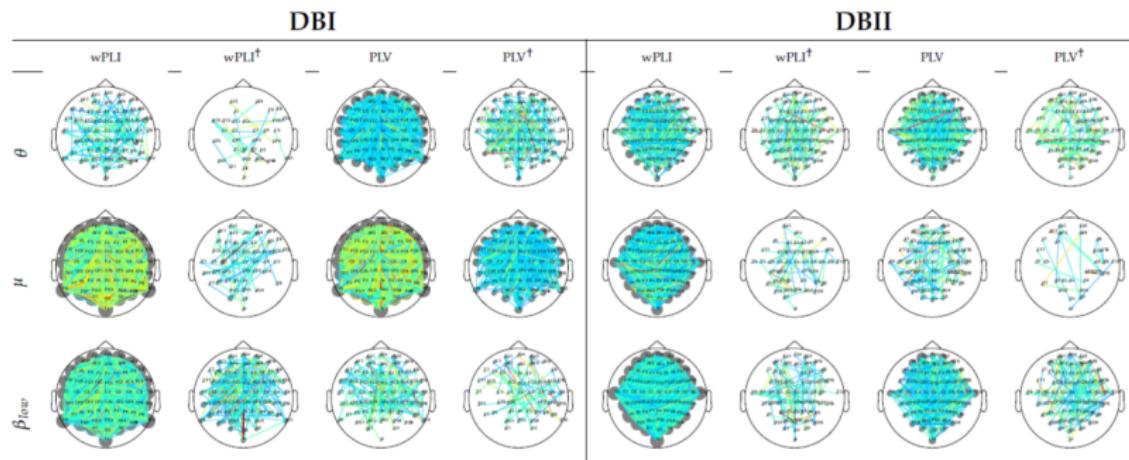
# Deep&Wide Learning as Electrophysiological Indicator in Mi Tasks

**Table 2.** Computed values of  $r$  for the indicator of initial training synchronization within the evaluated rhythm bandwidths:  $\mu$ ,  $\beta$ ,  $\mu + \beta$ . Notations LC, DRN, and LOO stand for Linear Correlation [40], Deep Regression Network, and leave-one-out-cross validation strategy, respectively. The best value per row is marked in bold.

<i>Rhythm</i>	<i>Electrode</i>	$\tau$ [s]				$\psi(\cdot)$	
		0.5	1.0	1.5	2.0	Mean	PCA <sub>1</sub>
<i>Subband</i>	<i>Configuration</i>						
$\mu$	2Ch(LC)	0.12	0.064	0.04	0.003	0.6	0.05
	6Ch(LC)	0.23	0.08	0.10	0.04	0.11	0.11
	2Ch(DRN $\xi^* = \xi_2$ )	0.13	0.064	0.13	<b>0.17</b>	0.06	0.17
	6Ch(DRN $\xi^* = \xi_2$ )	0.23	0.12	0.10	0.04	0.11	0.11
$\beta$	2Ch(LC)	0.11	0.06	0.08	0.02	0.07	0.06
	6Ch(LC)	0.14	0.04	0.006	0.016	0.11	0.07
	2Ch(DRN $\xi^* = \xi_2$ )	0.16	0.15	0.20	<b>0.23</b>	0.16	0.20
	6Ch(DRN $\xi^* = \xi_2$ )	0.19	0.05	0.23	<b>0.25</b>	0.21	0.20
$\mu + \beta$	2Ch(LC)	0.06	0.05	0.05	0.01	0.04	0.04
	6Ch(LC)	0.11	0.07	0.03	0.04	0.11	0.08
	2Ch(DRN $\xi^* = \xi_2$ )	0.08	0.06	0.10	<b>0.18</b>	0.11	0.09
	6Ch(DRN $\xi^* = \xi_2$ )	0.11	0.11	0.19	<b>0.21</b>	0.15	<b>0.21</b>
	2Ch(DRN $\xi^*$ )	0.84	0.80	0.94	0.91	0.78	0.83
	2Ch(DRN $\xi^*$ ) LOO	0.15	0.17	<b>0.24</b>	0.19	0.18	0.21
	6Ch(DRN $\xi^*$ )	0.87	0.77	0.93	<b>0.95</b>	0.82	0.82
	6Ch(DRN $\xi^*$ ) LOO	0.20	0.44	0.40	0.28	0.26	0.40

MI database (50 subjects): <http://gigadb.org/dataset/100295>

# Deep&Wide Learning as Electrophysiological Indicator in Mi Tasks



**Figure 1.** DNR weights mostly supporting the prediction performance, learned for wPLI and PLV predictors using two validation scenarios: leave-one-out cross-validation and leave-one-out cross-validation with Monte Carlo dropout (noted with <sup>†</sup>).

Relevance analysis from connectivity features  
(Monte Carlo Dropout enhancement).

# Contents

- 1 Signal Processing and Recognition Group - UNAL
- 2 Brain computer interfaces
- 3 Challenges
- 4 SPRG approaches
  - Kernel-based connectivity analysis for BCI
  - Deep&Wide Learning for BCI
- 5 Conclusions
- 6 Ongoing work

# Conclusions

- Current BCI technologies, i.e., EEG, coupled with ML approaches allow developing decoding/encoding systems to support assistive care, cognitive improvement, and entertainment tasks.
- The SPRG has been developing ML approaches based on kernel and deep learning methods to extract discriminant and interpretable patterns (spatial/temporal/frequency) in BCI.
- Still, the inter/intra subject variability issue and the inclusion of multi-view data, besides EEG, should be explored.

# Contents

- 1 Signal Processing and Recognition Group - UNAL
- 2 Brain computer interfaces
- 3 Challenges
- 4 SPRG approaches
  - Kernel-based connectivity analysis for BCI
  - Deep&Wide Learning for BCI
- 5 Conclusions
- 6 Ongoing work

# Ongoing work

## Localized Kernel alignment

- Localized kernel alignment to code nonstationary patterns:  
 $\bar{K} = \sum_{r=1}^R Q_r \bar{K}_r Q_r.$
- The diagonal matrices  $Q_r$  can be obtained solving a cosine-based cost (kernel alignment), where:

$$q_r(\mathbf{x}_n) = \begin{cases} \beta_0^r + \sum_{n'=1}^N \beta_{n'}^r \kappa_\beta(\mathbf{x}_n, \mathbf{x}_{n'}), & \text{if } y_n^r \neq 0 \\ 0, & \text{if } y_n^r = 0 \end{cases}. \quad (13)$$

For details see [Gil-Gonzalez et al., 2021].

# Ongoing work

## Correlated Chained Gaussian Processes

- CCGP to code non-stationary patterns from multi-input/multi-output data.
- CCGP relies on estimating the following joint distribution (from multiple independent GP priors):

$$p(\mathbf{Y}, \mathbf{f} | \mathbf{X}) = p(\mathbf{Y} | \boldsymbol{\theta}) p(\mathbf{f} | \mathbf{X}) = p(\mathbf{Y} | \boldsymbol{\theta}) \prod_{j=1}^J \mathcal{N}(f_j | \mathbf{0}, \mathbf{K}_{\mathbf{f}_j \mathbf{f}_j}), \quad (14)$$

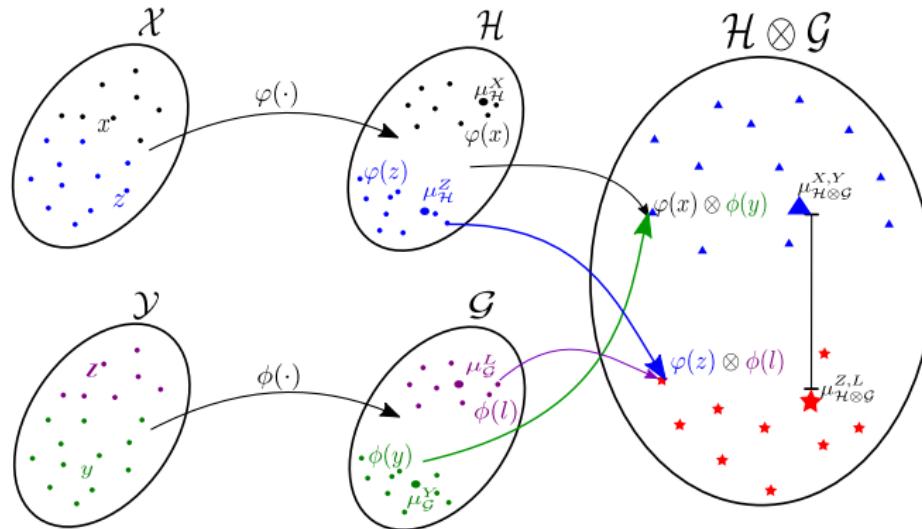
- The correlations between the GP latent functions are generated from a semi-parametric latent factor model-(SLFM):

$$f_j(\mathbf{x}_n) = \sum_{q=1}^Q w_{j,q} \mu_q(\mathbf{x}_n). \quad (15)$$

Work in progress...

# Ongoing work

## Hilbert Embedding-based distance - joint distributions



To include multi-view data from joint distributions in RKHS  
 Work in progress...

# Ongoing work

## D&WL holding kernel-based layers

- To develop D&WL-based representations from Random Fourier Features [Rahimi et al., 2007].
- Bochner's theorem [Rudin, 1962]:

$$\kappa(\mathbf{x} - \mathbf{y}) = \int p(\omega) e^{j\omega'(\mathbf{x}-\mathbf{y})} d\omega = \mathbf{E}\{\xi_\omega(\mathbf{x})\xi_\omega(\mathbf{y})\} \quad (16)$$

- The kernel-based connectivity analysis can be coupled within a D&WL framework- Work in progress...



# Ongoing work

## SPRG BCI-ML framework

The screenshot displays the SPRG BCI-ML framework interface. On the left, a code editor shows Python code for a stream consumer:

```

1 #!/usr/bin/python
2
3 # Loop consumer
4 def stream(self, data, topic: str, frame: int):
5
6     if topic == "eeg":
7
8         # eeg = self.buffer.eeg
9         eeg = self.resample(self.buffer.eeg, self.L)
10        eeg = self.centralize(eeg, normalize=True)
11
12        for i, line in enumerate(self.lines):
13            line.set_data(self.time, eeg[i] + 1 + i)
14
15        self.feed()
16
17    elif topic == "marker":
18        data = data.value
19        logging.warning(data)
20
21
22 if __name__ == '__main__':
23     Stream()

```

Below the code editor is a "Raw EEG" plot showing multiple channels (O2, O1, T4, C4, T3, Fz, Fp1) over time from -30 to 0 seconds. On the right, there's a circular diagram of electrode positions (Fp1, Fp2, F3, F4, C3, C4, T3, T4, O1, O2) and a table of saved montages.

- Open-Source Library for BCI-ML applications with stimuli synchronization and OpenBCI hardware compatibility (WiFi).
- MNE, Sklearn, Keras, etc - Python libraries compatibility.

# Thank you!

Andrés Marino Álvarez Meza, Ph.D.  
email: [amalvarezme@unal.edu.co](mailto:amalvarezme@unal.edu.co)

# References I

-  Collazos-Huertas, D., Álvarez-Meza, A., Acosta-Medina, C., Castaño-Duque, G., and Castellanos-Dominguez, G. (2020).  
Cnn-based framework using spatial dropping for enhanced interpretation of neural activity in motor imagery classification.  
*Brain Informatics*, 7(1):1–13.
-  De La Pava Panche, I., Alvarez-Meza, A. M., and Orozco-Gutierrez, A. (2019).  
A data-driven measure of effective connectivity based on renyi's  $\alpha$ -entropy.  
*Frontiers in neuroscience*, 13:1277.
-  Elasuty, B. and Eldawlatly, S. (2015).  
Dynamic bayesian networks for eeg motor imagery feature extraction.  
In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 170–173. IEEE.
-  Fetz, E. E. (1969).  
Operant conditioning of cortical unit activity.  
*Science*, 163(3870):955–958.
-  Gil-Gonzalez, J., Orozco-Gutierrez, A., and Alvarez-Meza, A. (2021).  
Learning from multiple inconsistent and dependent annotators to support classification tasks.  
*Neurocomputing*, 423:236–247.
-  Giraldo, L. G. S., Rao, M., and Principe, J. C. (2015).  
Measures of entropy from data using infinitely divisible kernels.  
*IEEE Transactions on Information Theory*, 61(1):535–548.

# References II

-  Gómez, V., Álvarez, A., Herrera, P., Castellanos, G., and Orozco, A. (2019).  
Short time eeg connectivity features to support interpretability of mi discrimination.  
In Vera-Rodriguez, R., Fierrez, J., and Morales, A., editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 699–706, Cham. Springer International Publishing.
-  Graimann, B., Allison, B., and Pfurtscheller, G. (2010).  
Brain–computer interfaces: A gentle introduction.  
In *Brain-Computer Interfaces*, pages 1–27. Springer.
-  Li, D., Zhang, H., Khan, M. S., and Mi, F. (2018).  
A self-adaptive frequency selection common spatial pattern and least squares twin support vector machine for motor imagery electroencephalography recognition.  
*Biomedical Signal Processing and Control*, 41:222–232.
-  Liang, S., Choi, K.-S., Qin, J., Wang, Q., Pang, W.-M., and Heng, P.-A. (2016).  
Discrimination of motor imagery tasks via information flow pattern of brain connectivity.  
*Technology and Health Care*, 24(s2):S795–S801.
-  Miranda, R. A., Casebeer, W. D., Hein, A. M., Judy, J. W., Krotkov, E. P., Laabs, T. L., Manzo, J. E., Pankratz, K. G., Pratt, G. A., Sanchez, J. C., et al. (2015).  
Darpa-funded efforts in the development of novel brain–computer interface technologies.  
*Journal of neuroscience methods*, 244:52–67.
-  Muandet, K., Fukumizu, K., Sriperumbudur, B., and Schölkopf, B. (2017).  
*Kernel Mean Embedding of Distributions: A Review and Beyond*.

# References III

-  **Nishiyama, Y., Kanagawa, M., Gretton, A., and Fukumizu, K. (2020).**  
Model-based kernel sum rule: kernel bayesian inference with probabilistic models.  
*Machine Learning*, 109(5):939–972.  
cited By 0.
-  **Parra, G. and Tobar, F. (2017).**  
Spectral mixture kernels for multi-output gaussian processes.  
In *Advances in Neural Information Processing Systems*, pages 6681–6690.
-  **Principe, J. C. (2010).**  
Information theoretic learning: Renyi's entropy and kernel perspectives.  
pages 1–45, New York, NY. Springer New York.
-  **Qureshi, M. N. I., Cho, D., and Lee, B. (2017).**  
Eeg classification for motor imagery bci using phase-only features extracted by independent component analysis.  
In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2097–2100.
-  **Rahimi, A., Recht, B., et al. (2007).**  
Random features for large-scale kernel machines.  
In *NIPS*, volume 3, page 5. Citeseer.
-  **Rudin, W. (1962).**  
*Fourier analysis on groups*, volume 121967.  
Wiley Online Library.

# References IV

-  Timme, N. M. and Lapish, C. (2018).  
A tutorial for information theory in neuroscience.  
*eNeuro*, 5(3).
-  Torres-Valencia, C., Orozco, Á., Cárdenas-Peña, D., Álvarez-Meza, A., and Álvarez, M. (2020).  
A discriminative multi-output gaussian processes scheme for brain electrical activity analysis.  
*Applied Sciences*, 10(19):6765.
-  Velasquez-Martinez, L., Caicedo-Acosta, J., Acosta-Medina, C., Alvarez-Meza, A., and Castellanos-Dominguez, G. (2020).  
Regression networks for neurophysiological indicator evaluation in practicing motor imagery tasks.  
*Brain Sciences*, 10(10):707.
-  Vicente, R., Wibral, M., Lindner, M., and Pipa, G. (2011).  
Transfer entropy—a model-free measure of effective connectivity for the neurosciences.  
*Journal of computational neuroscience*, 30(1):45–67.
-  Vidal, J. J. (1977).  
Real-time detection of brain events in eeg.  
*Proceedings of the IEEE*, 65(5):633–641.