

CNN-based Framework using Spatial Dropping for Enhanced Interpretation of Neural Activity in Motor Imagery Classification

Diego Fabian Collazos Huertas (✉ dfcollazosh@unal.edu.co)

Universidad Nacional de Colombia Sede Manizales <https://orcid.org/0000-0002-0434-3444>

Andres Marino Alvarez Meza

Universidad Nacional de Colombia Sede Manizales

German Castellanos Dominguez

Universidad Nacional de Colombia Sede Manizales

Research

Keywords: Motor imagery, convolutional neural networks, spatial dropping

Posted Date: May 21st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-27669/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on September 3rd, 2020. See the published version at <https://doi.org/10.1186/s40708-020-00110-4>.

CNN-based Framework using Spatial Dropping for Enhanced Interpretation of Neural Activity in Motor Imagery Classification

D.F. Collazos-Huertas · A.M. Álvarez-Meza ·
G. Castellanos-Dominguez
Signal Processing and Recognition Group,
Universidad Nacional de Colombia, Manizales
{dfcollazosh,amalvarezme,cgcastellanosd}@unal.edu.co

Received: date / Accepted: date

Abstract Interpretation of brain activity responses using Motor Imagery (MI) paradigms is vital for medical diagnosis and monitoring. Assessed by machine learning techniques, identification of imagined actions is hindered by substantial intra and inter subject variability. Here, we develop an architecture of Convolutional Neural Networks (CNN) with enhanced interpretation of the spatial brain neural patterns that mainly contribute to the classification of MI tasks. Two methods of 2D-feature extraction from EEG data are contrasted: Power Spectral Density and Continuous Wavelet Transform. For preserving the spatial interpretation of extracting EEG patterns, we project the multi-channel data using a topographic interpolation. Besides, we include a spatial dropping algorithm to remove the learned weights that reflect the localities not engaged with the elicited brain response. Obtained results in a bi-task MI database show that the thresholding strategy in combination with Continuous Wavelet Transform improves the accuracy and enhances the interpretability of CNN architecture, showing that the highest contribution clusters over the sensorimotor cortex with differentiated behavior between μ and β rhythms.

Keywords

Motor imagery, convolutional neural networks, spatial dropping.

1 Introduction

1 The Motor Imagery (MI) paradigm is a form of Brain-Computer Interface (BCI)
2 that performs the imagination of a motor action without any real execution, relying
3 on the similarities between imagined and executed actions at the neural level. MI
4 is usually measured with Electroencephalography (EEG) to register brain activity
5 on the scalp surface. Thus, assessment and interpretation of MI brain dynamics in

D.F Collazos-Huertas
Universidad Nacional de Colombia
E-mail: dfcollazosh@unal.edu.co

6 the sensorimotor cortex may contribute to applications ranging from evaluation of
7 pathological conditions and rehabilitation of motor functions [1,2], motor learning
8 and performance [3], improving the learning of different abilities [4], among others.
9 One of the main challenges in implementing MI practice is to recognize and identify
10 the imagined actions since EEG signals have substantial intra and inter subject
11 variability [5].

12 Currently, there is an increasing interest in deep learning models that are com-
13 posed of multiple processing layers of inference using data representations with
14 multiple levels of abstraction. In discriminating physiological signals, Convolu-
15 tional Neural Networks (CNN) become the leading deep learning architectures due
16 to their regularization structure and degree of translation invariance [6], yielding an
17 outstanding ability in transferring knowledge between apparently different tasks
18 of classification [7,8]. Thus, CNN models are useful in learning features related
19 to brain imaging and neuroscience discovery [9]. Nevertheless, for applications in
20 MI tasks, designing an available end-to-end CNN architecture remains a challenge
21 due to several restrictions: their large number of hyperparameters to be learned in-
22 crease the computational burden (being unsuitable for online processing [10]), and
23 complicated multilayer integration to encode relevant features at every abstraction
24 level of the input EEG data [11].

25 Another unresolved issue is the interpretability of results provided by CNN
26 models [12]. That is, along with the improved accuracy, the learned features can
27 be hard to understand within the context of the original MI paradigm. The value
28 of neural activity interpretation becomes evident in purposes like a medical diag-
29 nosis, monitoring, and computer-aided learning [13]. As a tool in image processing,
30 CNN architecture has been discussed for enhancing the physiological explanation
31 of MI paradigms represented by multiple time series (time dimension), which re-
32 flect the brain responses across the sensorimotor cortex (spatial dimension), and
33 commonly related to μ and β rhythms (spectral dimension). For representing lo-
34 cal and global structures in CNN models, therefore, the extraction of time-series
35 features is increasingly realized as a multi-dimensional tensor that retains the
36 EEG data structure throughout the learning process, by adequately encoding the
37 spatio/spectro-temporal relationships of the measured MI responses [14]. Never-
38 theless, CNN models should extract the structure of multi-dimensional images
39 properly to preserve the domain information of interest. Intending to make the
40 learned features more interpretable in MI tasks, two main aspects are to be con-
41 sidered to retain the spatial locality of CNN models: *i*) Improving the 2D-feature
42 extraction from EEG data for feeding the CNN models, and *ii*) Enhancing the
43 image-based EEG representation to integrate spatial domain knowledge with the
44 extracted 2D spectro-temporal features.

45 For building 2D-maps in discrimination of MI tasks, several algorithms of fea-
46 ture extraction are employed in CNN models, including the following: Common
47 Spatial Patterns due to the high recognition rate and computational simplicity [15];
48 Event-Related Synchronization to capture the channel-wise temporal dynamics of
49 the power signal [16]; Empirical Mode Decomposition to deal with EEG non-
50 stationarity [17,18]; and time-frequency planes using the Fourier and Wavelet
51 Transforms are frequently extracted because they allow a more straightforward
52 interpretation [19,20,21,22], being the latter decomposition better suited to deal
53 with sudden changes in EEG signals. Nonetheless, the extracted 2D images tend to
54 have substantial variability in patterns across trials due to inherent nonstationar-

55 ity, artifacts, a poor signal-to-noise ratio of EEG signals, individual differences in
56 cortical functioning (like subjects exhibiting activity in different frequency bands).

57 Concerning the integration of electrode montages with the extracted 2D fea-
58 tures, topographical representations are applied, involving either local or global
59 spline techniques to interpolate the spatial distribution of the potential field on
60 the scalp from distributed electrode arrays. For low electrodes distributions, ad-
61 equate mapping is the spherical spline interpolation [23]. For handling the lack
62 of significant samples in smaller datasets, one strategy of integration is to incor-
63 porate prior knowledge to optimize the neural network structure. For instance,
64 pre-trained networks are used, but assuming a substantial similarity between pre-
65 training and target sets [24,25,26]. Otherwise, some ambiguity may remain in the
66 foolproof nature of the pre-trained network methodology [27]. In the case of MI
67 tasks, there are very few accessible datasets having some differences in implement-
68 ing the paradigm. Another integration approach is to have some form of spatial
69 dropping algorithm to remove candidate localities, which are known to be not en-
70 gaged with the elicited brain response. Relying on the fact that motor imagery
71 responses are directly related to electrocortical activity over the sensorimotor area,
72 the spatial dropping can be performed either subject-independent by excluding all
73 electrodes out of the motor cortex before training and validation [28,29,30], or
74 by thresholding the electrode contribution after training and validation for each
75 subject.

76 Here, we develop a CNN architecture with an enhanced interpretation of the
77 spatial activity of brain neural patterns that mainly contribute to the classifica-
78 tion of MI tasks (left and right hand). Following the approach developed by [31],
79 the CNN framework is designed, for which we validate two commonly used tech-
80 niques of feature extraction from EEG data: Power spectral density and Contin-
81 uous Wavelet Transform. For preserving the spatial interpretation of extracting
82 EEG patterns, we project the multi-channel data using a topographic interpo-
83 lation. Besides, we include a spatial dropping algorithm to remove the learned
84 weights that reflect the localities not engaged with the elicited brain response.
85 Obtained results in a bi-task MI database show that the thresholding strategy is
86 desirable since the highest contribution clusters over the sensorimotor area with
87 differentiated behavior between μ and β bands. The agenda of the present paper
88 is as follows: Section 2 describes the collection of MI data used for validation.
89 Besides, it presents the fundamentals of feature extraction of time-frequency (t - f)
90 EEG Patterns and describes the design of Convolutional Neural Networks, in-
91 cluding the spatial dropping strategies for Motor Imagery classification. Further,
92 Section 3 provides a summary of the classifier accuracy performed by the extracted
93 t - f vectors and evaluates the interpretability of learning weights for distinguish-
94 ing between MI tasks. Lastly, Section 5 gives critical insights into the performed
95 interpretation and accuracy, and address some limitations and possibilities of the
96 presented CNN-based framework.

97 2 Materials and Methods

98 *Description of Bi-task MI Database and Preprocessing:* We perform experimental
 99 validation with nine subjects ($N_S=9$) of Dataset 2a¹, holding EEG signals ac-
 100 quired from the scalp by a C -channels montage ($C=22$). Each raw EEG channel
 101 $\mathbf{x}^c \in \mathbb{R}^T$ was sampled at 250 Hz (i.e., at sample rate $\Delta t=0.004$ s) and passed through
 102 a five-order bandpass Butterworth filter within $\Omega=[8, 30]$ Hz. Since earlier works
 103 have shown that electrical brain activities prompted by motor tasks are frequently
 104 related to μ and β rhythms [32], the spectral range is split into the following band-
 105 widths of interest: $\Delta f \in \{\mu \in [8-12], \beta_{\text{low}} \in [16-20], \beta_{\text{med}} \in [20-24], \beta_{\text{high}} \in [24-28]\}$ Hz.

106 To perform each MI task (left and right hand), a short beep noticed the trial
 107 beginning followed by a fixation cross that appeared on the black screen within the
 108 first 2 s-interval. An arrow (cue) appeared during 1.25 s, pointing to the induced
 109 direction. Then, each subject performed a run of each MI task while the cross
 110 re-appeared within the time interval, starting from 3.25 to the recording end, T s.
 111 The recordings were collected in six runs separated by short breaks, performing
 112 $N_\lambda=72$ trials per class and each one lasting $T=7$ s. We only validated these labeled
 113 trials with the removed artifacts.

114 *Feature Extraction of t-f EEG Patterns:* In the first case, the feature set is ex-
 115 tracted from the Fourier decomposition method. So, provided the EEG sample
 116 frequency $F_s \in \mathbb{R}^+$, the Power Spectral Density (PSD) vector $\mathbf{s} = \{s_f \in \mathbb{R}^+ : f \in N_B\}$,
 117 with $N_B = \lfloor F_s/2 \rfloor$, is estimated through the nonparametric Welch's method that
 118 calculates the Fast Fourier Transform (FFT) algorithm on a set of $M \in \mathbb{N}$ overlap-
 119 ping segments, which are split from the preprocessed EEG data vector \mathbf{x}^c . Due to
 120 the non-stationary nature of EEG data, the piecewise stationary analysis is car-
 121 ried out over the set of the extracted overlapping segments that are windowed by
 122 a smooth-time weighting window $\alpha \in \mathbb{R}^\tau$ that lasts $\tau \in \mathbb{N}$ ($\tau < T$), yielding a set of the
 123 time segments $\{\mathbf{v}^m \in \mathbb{R}^\tau : m \in M\}$, where $v_t^m \in \mathbb{R}$ ($t \in \tau$) is t -th element of \mathbf{v}^m . So, the
 124 t -f patterns are extracted from EEG signals through the modified periodogram
 125 vector, $\mathbf{u} = \{u_f \in \mathbb{R}^+\}$, $\mathbf{u} \in \mathbb{R}^{N_B}$, computed as follows:

$$126 \quad u_f = \mathbb{E} \left\{ \left| \sum_{t \in \tau} v_t^m \exp(-j2\pi t f) \right|^2 : \forall m \in M \right\}. \quad (1)$$

127 Thus, the resulting PSD vector is computed with spectral components defined
 128 as $s_f = u_f / (M\nu)$, being $\nu = \mathbb{E} \{ |\alpha_t|^2 : \forall t \in \tau \}$, and $\mathbb{E} \{ \cdot \}$ – the expectation operator.

129 In the second case, the feature set is extracted from Continuous Wavelet Trans-
 130 form (CWT) that quantifies similarity between a given equally sampled time series
 131 at time spacing $\delta_t \in \mathbb{R}$ and a previously fixed base function $\psi(\eta)$, termed *mother*
 132 *wavelet* ruled by a dimensionless parameter vector $\eta \in \mathbb{R}$. Namely, each time element
 133 of the CWT vector $\boldsymbol{\zeta}^g \in \mathbb{C}^T$ is extracted from the preprocessed EEG time-series
 134 $\mathbf{z} \in \mathbb{R}^c$ at scale $g \in \mathbb{R}$ by accomplishing their convolution with the scaled and shifted
 135 mother wavelet in the form:

$$\boldsymbol{\zeta}_t^f = \sum_{\tau \in T} z_\tau \psi^* ((\tau - t)\delta_t, f), \quad (2)$$

136 where notation $(*)$ stands for the complex conjugate.

¹ BCI Competition IV, publicly available at www.bbc.de/competition/iv/

137 To build a picture showing amplitude variations through time in Eq. (2), both
 138 procedures of wavelet scaling g and translating through the localized time index
 139 $t \in T$ are used. As a result, the extracted wavelet coefficients provide a compact
 140 representation pinpointing energy distribution of EEG data in time and frequency
 141 domains. Therefore, the resulting CWT vector is computed with spectral compo-
 142 nents defined as $s_f = \mathbb{E} \left\{ \zeta_t^f : \forall t \in \tau \right\}$.

143 Having extracted the feature set, we further compute a real-valued representa-
 144 tive vector, $\rho^{r, \Delta f} \in \mathbb{R}^C$ for each trial $r \in R$, with electrode elements that accumulate
 145 the spectral contribution as follows:

$$146 \quad \rho^{r, c} = \sum_{\eta_{\min} \leq f \leq \eta_{\max}} s_f^{r, c}, f \in \Delta f \quad (3)$$

147 where the frequencies η_{\min} and η_{\max} determine each one of the bandwidths of
 148 interest $f \in \Delta f$, within the most discriminating MI information is assumed to be
 149 concentrated.

150 Then, we map the multi-channel data per patient on a 2D surface, aiming of
 151 preserving the spatial interpretation of extracted t - f patterns. In order to preserve
 152 the distance between electrodes in the 3D plane, we compute the topographic inter-
 153 polation matrix across all trials, $\{S(\rho^{r, \Delta f}) \in \mathbb{R}^{S \times S'} : \forall r \in R\}$, through the projecting
 154 matrix that maps each EEG trial field, $\rho^{r, \Delta f}$, as a 2-D circular view (looking down
 155 at the head top) using spherical splines that sizes $(S \times S')^2$, as detailed in [33].

156 *Motor Imagery Classification using Convolutional Neural Networks:* The proposed
 157 CNN architecture contains three learning stages: *i*) Convolutional layer that holds
 158 a set of kernel filters, $\{K_i \in \mathbb{R}^{K \times K} : i \in I\}$ (I is the number of used kernel filters),
 159 together with the corresponding bias vectors $\{b_i \in \mathbb{R}^{S \times S'}\}$, which are applied by a
 160 sliding window across each topographic map $S(\rho^{r, \Delta f})$, yielding the convolution
 161 feature map as below:

$$162 \quad \Xi^{r, i, \Delta f} = \gamma_1 \left(K_i \otimes S(\rho^{r, \Delta f}) + b_i \right), \Xi^{r, i, \Delta f} \in \mathbb{R}^{S \times S'} \quad (4)$$

163 where $\gamma_1(\cdot)$ is a non-linear activation function, and \otimes denotes the convolution
 164 operator. Of note, a zero-padding method is adopted to prevent losing the feature
 165 dimension, so that the output and input sizes of convolution mapping can be the
 166 same after the zero-padding procedure.

167 *ii*) Pooling layer that is a down-sampling stage to reduce the dimension of
 168 output neurons in $\Xi^{r, i, \Delta f}$ through a pool operator matrix $\bar{K} \in \mathbb{R}^{K' \times K'}$, with $K' \leq K$,
 169 aiming at decreasing the computational burden and the over-fitting issue. Then,
 170 each down-sampled map $\bar{\Xi}^{r, i, \Delta f}$ is rearranged into a vector form $\bar{\xi}^r \in \mathbb{R}^{GG'IN_f}$
 171 (with $G \leq S, G' \leq S'$) by concatenating all matrix rows across Δf and i domains.

172 *iii*) A fully-connected stage that includes a neural network with all neurons
 173 $\mathbf{h}^r(q) \in \mathbb{R}^{N_h(q)}$ connected directly to the outputs of preceding layer $q-1$ as follows:

$$174 \quad \mathbf{h}^r(q) = \gamma_2(\mathbf{W}(q)\mathbf{h}^r(q-1) + \beta(l)), q = \overline{2, Q} \quad (5)$$

175 where $\mathbf{h}^r(1) = \bar{\xi}^r$, \mathbf{W} , sizing $GG'IN_f \times N_h(q)$, is the weighting matrix that contains
 176 the connection weights between the preceding neurons and the hidden units N_h of
 177 layer q , $\beta(q) \in \mathbb{R}^{N_h(q)}$ is the bias neuron, and $\gamma_2(\cdot)$ is an activation function.

² function `topplot()` in *EGLAB* toolbox

As a result, we obtain the output vector set $\{\mathbf{y}^r = \mathbf{h}^r(Q)\}$, with $\mathbf{y}^r \in [0, 1]^{N_\lambda}$, representing N_λ mutually exclusive classes, so that the last layer is tied to the output dimension ($N_h = N_\lambda$).

Due to the CNN-model training back-propagates the discriminating information, through the tied weights, from the hidden spaces to the input data, we propose to assess the relevance of input feature mappings, employing the matrix $\mathbf{W}(q) \in \mathbb{R}^{D \times N_h}$ that holds the row vectors $\mathbf{w}_d^q \in \mathbb{R}^{N_h}$ with $D = GG'IN_f$. Based on the fact that each \mathbf{w}_d^q measures the contribution of input features to build the hidden space $\mathbf{h}^r(q)$, the relevance of d -th feature is assessed as the generalized mean of its corresponding reverse projection vector, that is, $\rho_d^q = \|\mathbf{w}_d^q\|_p$, yielding the vector $\boldsymbol{\rho}^q = \{\rho_d^q \in \mathbb{R}^+; \forall d \in D\}$, where notation $\|\cdot\|_p$ stands for l_p -norm. The obtained relevance vector $\boldsymbol{\rho}^q$ is reshaped into an estimated feature mapping matrix $\tilde{\Theta} \in \mathbb{R}^{S \times S'}$ that is computed for each Δf as follows:

$$\tilde{\Theta} = \phi \left(\mathbb{E} \left\{ \tilde{\Xi}_i; \forall i \in I \right\} \right) \quad (6)$$

where $\tilde{\Xi}_i \in \mathbb{R}^{G \times G'}$ is the reconstructed feature mapping for i -th kernel filter, and $\phi(\cdot)$ is an extrapolation operator that maps from $G \times G' \rightarrow S \times S'$. In this way, the obtained $\tilde{\Theta}$ highlights the spatial discriminative information projected from topographic maps.

3 Experiments

We validate the proposed CNN-based MI classification framework by appraising the following procedures: *i*) Preprocessing and extraction of t - f planes, evaluating the extraction methods of power spectral density and continuous wavelet transform, for which the corresponding parameter tuning is carried out, *ii*) Tuning of CNN architecture for MI discrimination, evaluating the spatial dropping algorithm proposed for preserving the interpretation of extracted 2-D features. Two dropping approaches are appraised: removing all electrodes out of the sensorimotor area before training and validation, and thresholding the electrode contribution after training and validation.

Extraction of t - f feature patterns: Each channel recording, $\mathbf{x}^c \in \mathbb{R}^T$, is split into $N_\tau = 5$ segments, $\{\mathbf{x}_\tau^c \in \mathbb{R}^\tau, \tau < T\}$, using a sliding window approach with a segment length $\tau = 2s$ with overlap $\delta\tau = 1s$. Within each segment \mathbf{x}_τ^c , PSD estimates are computed, fixing the following parameters: $\tau = 256, \delta\tau = 0.9\tau$. Likewise, we compute the CWT vector $\boldsymbol{\varsigma}^g$, selecting the Morlet wavelet as ψ that is frequently used in spectral analysis of EEG signals [34]. So, we extract the continuous wavelet coefficients within each time-segment using a Complex Morlet wavelet, adjusting the scaling value to $g = 16$ and the sampling period to $1/\Delta t$.

Implementing either method of feature extraction, we perform validation in four different scenarios for spectral bandwidths of interest $f \in \Delta_f$: A) μ , B) β , C) $\mu \cup \beta$, and D) $\mu \cup \beta_{\text{low}} \cup \beta_{\text{med}} \cup \beta_{\text{high}}$.

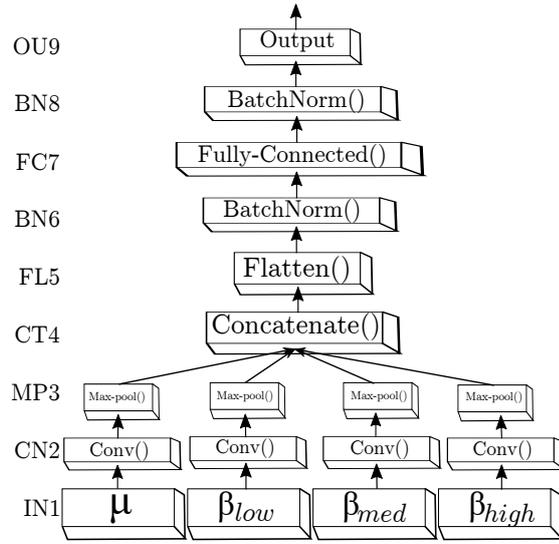


Fig. 1 The proposed CNN structure that is based on *Wide&Deep* neural network handling multiple inputs. The first layer (IN1) is the input, second (CN2) and third layers (MP3) are hidden and accomplish the feature mapping generation, while the block, ranging from the output of layer CT4 to the OU9 layer, comprises the classification stage.

217 *Proposed CNN Architecture for MI Discrimination.* The adopted multiple input
 218 CNN model is based on the non-sequential *Wide&Deep* neural network [35] that
 219 performs learning of deep patterns (using the deep path) under simple rules
 220 (through the short path), having the following units:

- 221 – **IN1**: Input layer that holds an image set sizing 42×56 .
- 222 – **CN2**: Convolutional layer (first hidden layer). We use two spatial filters that
 223 perform two resulting feature maps, sizing 42×56 . Each convolution kernel has
 224 a size of 3×3 , using a stride of one sample. In addition, this layer incorporates
 225 a rectified linear unit *ReLU* [36] through the activation function $\gamma_1(\cdot)$.
- 226 – **MP3**: Max-pooling layer (second hidden layer). This layer sub-samples the
 227 resulting mapping that picks up the maximum value of each feature map to
 228 reduce the number of output neurons, also using a stride of one sample. Thus,
 229 each feature mapping in **CN2** is down-sampled using a pool size of 2×2 , re-
 230 sulting in a matrix of size 21×28 .
- 231 – **CT4**: Concatenate layer, linking together of all resulting MP3 feature maps
 232 into a single block.
- 233 – **FL5**: Flatten layer that arranges the set of concatenated feature maps from
 234 CT4 into a single 1D array. So, the map is vectorized into a one-dimensional
 235 array of size $(21)(28)(2)(4) = 1176$ points, resulting from 2 spatial filters, and 4
 236 bandwidths of interest.
- 237 – Batch normalization (BN) layers (**BN6** and **BN8**) that address the vanishing
 238 and exploding gradient problems presented in fully-connected networks. To
 239 cope with this issue, all inputs of the previous layer at each batch are zero-
 240 scored, holding the mean activation close to 0 and the activation standard
 241 deviation close to 1.

- 242 – **FC7**: Fully-connected layer (third hidden layer) that is linked to each neuron of
 243 **OU9**, holding h_u neurons for which the weight values are regularized through
 244 the parameters (l_1, l_2) using the Elastic Net regularization. According to [37],
 245 Elastic Net is used for preventing over-fitting by penalizing a model having
 246 large weights, and can be used more naively, e.g., when little prior knowledge
 247 is available about the dataset. This layer uses a rectified linear unit *ReLU* as
 248 the activation function $\gamma_2(\cdot)$. The following parameter setting of **FC7** is fixed:
 249 . Number of neurons are fixed through an exhaustive grid search within
 250 $h_u=[50, 100, \dots, 550]$.
 251 . Activation functions of **FC7** and **OU9** are *ReLU* and *softmax* function,
 252 respectively.
 253 . The learning rate is fixed at $lr=\exp(-3)$.
 254 . The optimizer used is the *Adam algorithm* and the *loss function* used is the
 255 mean squared error (MSE).
 256 . The regularization parameters l_1 and l_2 are tuned by a grid search around
 257 $[0.001, 0.01, 0.1]$.
 258 – **OU9**: Output layer having two neurons, each one representing either task label
 259 (left hand or right hand). This layer is fully connected to **FC7** and uses the
 260 *softmax* procedure as the activation function $\gamma_2(\cdot)$.

261 *Evaluating Metrics of Classifier Performance.* As a measure of performance, the
 262 classifier accuracy $a_c \in \mathbb{R}[0, 1]$ is computed as follows:

$$263 \quad a_c = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (7)$$

264 where T_P , T_N , F_P and F_N are true-positives, true-negatives, false-positives, and
 265 false-negatives, respectively.

266 Besides, the kappa value, $\kappa \in \mathbb{R}[0, 1]$, is computed to evaluate the accuracy per-
 267 formance when removing the impact of random classification as follows [38]:

$$268 \quad \kappa = (a_c - p_e)/(1 - p_e), \quad (8)$$

269 where $p_e=0.5$ for bi-label problems.

270 A cross-validation scheme is used to evaluate the CNN-based classifier perfor-
 271 mance. Thus, the set of training trials per subject is randomly partitioned using
 272 a stratified 10-fold cross-validation to generate the set of validation trials. This
 273 procedure is repeated ten times by shifting the test and training dataset.

274 4 Results

275 *Performed Accuracy of Extracted t-f Planes.* Initially, we discuss the classifier per-
 276 formance of the computed PSD vectors of contribution, ρ^{r, Δ_f} . In each one of the
 277 tested scenarios for spectral bandwidths of interest, parameter tuning is carried
 278 out to achieve the maximum accuracy within the MI interval [3-5], s . As seen
 279 in Table 1, the use of only one rhythm (μ or β) is not enough to reach the best
 280 values of accuracy. Moreover, β waveform drops to 80%. Their combination $\mu \cup \beta$
 281 barely helps the classifier rule. Thus, the last validating scenario (i.e., D) reaches
 282 the best performance on average across all subjects, meaning that the inclusion of

283 more detailed information of β subbands allows improving the accuracy of PSD
 284 vectors. Concerning the individual performance, values achieved by A02T, A01T,
 285 A04T, and A05T are the lowest, while A08T, A09T, and A03T accomplish the best
 286 results. Regarding the CWT-based contribution vectors, the bottom part of Ta-
 287 ble 1 shows that the use of every spectral bandwidth scenario allows enhancing
 288 the performed results, but without statistical difference between them when aver-
 289 aging across the subject set. Furthermore, the accuracy of CWT-based vectors is
 290 comparable to the one obtained by the best case of PSD-based extraction vectors,
 291 having a very similar ranking of individual performance.

<i>PSD</i>				
<i>Subject</i>	μ	β	$\mu \cup \beta$	$\mu \cup 3\beta$
A08T	97.1±3.5	84.9±6.9	93.3±3.9	96.8±3.9
A09T	92.1±4.8	89.0±6.1	96.5±4.3	96.6±4.1
A03T	91.1±5.6	77.5±6.4	91.2±6.5	91.3±4.2
A06T	80.0±6.9	81.6±5.5	82.5±6.1	86.7±5.8
A07T	83.8±8.7	78.9±4.3	82.0±4.9	85.1±8.6
A05T	76.2±5.2	79.0±6.9	80.8±6.6	84.0±4.5
A04T	84.7±6.2	79.3±7.6	80.7±7.5	83.7±5.4
A01T	82.6±3.4	80.5±5.4	82.6±5.6	83.5±7.0
A02T	80.4±6.8	77.8±6.2	81.1±5.3	82.5±6.4
Average h_u	200	350	350	300
Average a_c	85.3±5.7	80.9±6.1	85.6±5.6	87.8±5.5
<i>CWT</i>				
A03T	96.4±3.6	95.0±4.6	95.0±4.6	94.2±2.9
A08T	96.3±4.9	95.4±5.1	94.0±4.6	94.0±5.6
A09T	94.0±5.4	94.8±5.6	94.8±4.2	92.3±5.9
A07T	85.7±7.9	83.4±6.8	86.4±6.7	87.4±5.4
A06T	83.9±5.6	84.1±5.1	86.7±7.2	86.7±7.2
A04T	86.9±7.7	84.6±8.4	85.4±7.3	85.3±7.2
A01T	83.2±6.7	82.5±5.1	83.4±5.5	82.5±3.9
A05T	79.8±6.3	76.8±9.0	79.1±4.8	82.2±4.9
A02T	81.8±7.2	84.0±8.2	83.8±6.5	81.7±7.2
Average h_u	350	400	250	250
Average a_c	87.5±6.1	86.7±6.4	87.6±5.7	87.4±5.6

Table 1 Performed accuracy within the MI segment using the whole electrode montage ($C=22$). The best figure achieved by each individual is marked in black.

292 In terms of the tuned CNN parameters, their values averaged across the subject
 293 set show that the training scenario achieving the best accuracy ($\mu \cup \beta_{\text{low}} \cup \beta_{\text{med}} \cup \beta_{\text{high}}$)
 294 demands from the PSD-based vectors more hidden units h_u than in the case of
 295 CWT planes. A similar situation holds in the scenario $\mu \cup \beta$ that also performs
 296 high accuracy. When extracting the t - f vectors from a single rhythm (μ or β), the
 297 PDS-based representation demands less hidden units but achieves lower accuracy.

298 Figure 2 displays the dependency of CNN hidden units on the obtained accu-
 299 racy. Compared to the best score that is achieved by the individually tuned value
 300 of h_u , the deterioration in performance is noticeable (nearly 5%) when decreasing
 301 the number of units in every trained CNN model. At the same time, the computa-
 302 tional burden can reduce, on average, about quarter time. Moreover, the variations
 303 of accuracy by changing the amount of h_u indicate a similar complexity of both
 304 measured extraction approaches.

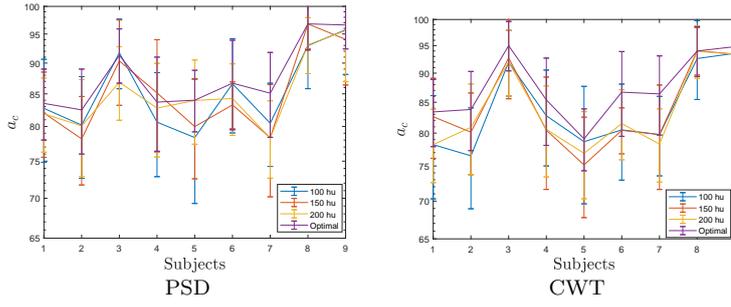


Fig. 2 Dependence of CNN hidden units and the individual accuracy. Label "optimal" is the case of individually tuned CNN model.

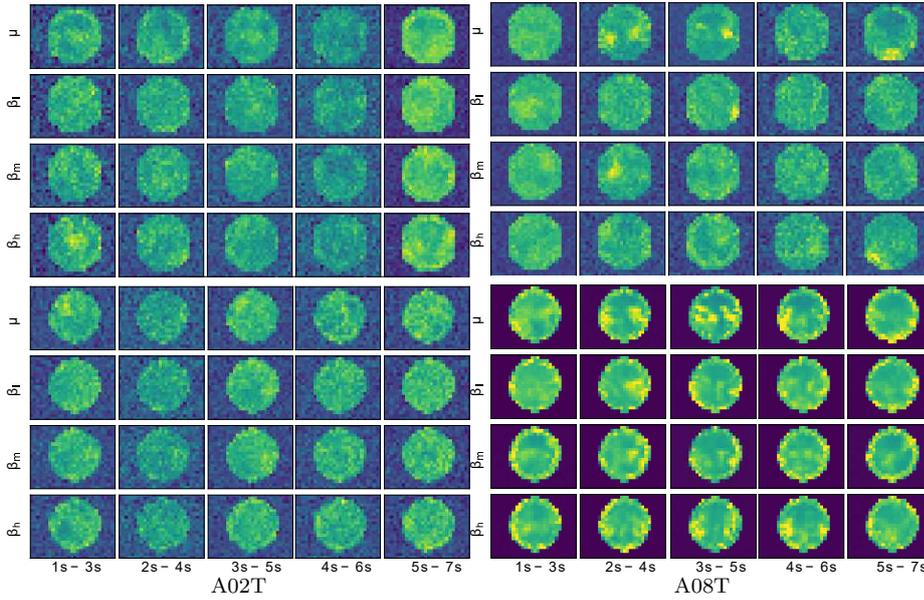


Fig. 3 Reconstructed topographic maps for experiment *D* including all bandwidths across the time domain in the best (A08T) and the worst (A02T) subjects, left and right column respectively.

305 *Interpretability of Brain Areas Activated by MI Tasks.* Intending to give the inter-
 306 pretability of the extracted input t - f vectors, we represent the feature mapping
 307 graphically (topoplot) to highlight the spatial distribution of the assessed discrimi-
 308 native ability. Each topoplot depicts the proposed assessment $\hat{\Theta}(\rho^r, \Delta^f)$ computed
 309 in Eq. (6) by which we reconstruct the input feature image from the trained
 310 CNN weights to estimate the contribution of the electrodes, under the assumption
 311 that the higher the reconstructed weight, the more important the discriminating
 312 strength of electrodes. Of note, those interpolated values falling out the electrode
 313 space are assumed as meaningless. This situation may arise since the network ini-
 314 tializes the weight set with random values, including the background pixels. So, the

315 variability and poor signal-to-noise rate result in false augmentation of background
 316 localities, as it is the case of subjects reaching low discrimination ability.

317 The top row of Fig. 3 displays the PSD-based spatial distribution reconstructed
 318 for the best training scenario (D) within each time segment. As seen, the topoplots
 319 of A02T (the worst individual) present the spectral bandwidths contributing much
 320 alike with values mainly spread all over space, including places out of the electrode
 321 space. Besides, the contribution estimates are low and tend to be noisy. Another
 322 fact to mention is that brain activity notably increases within the last time seg-
 323 ment, for which the MI activity is assumed to have already vanished. By con-
 324 trast, the best-achieving subject A08T has some relevant localities, which gather
 325 in places of either brain hemisphere and within the MI interval, fading at the time
 326 window [4-6] s.

327 In turn, the bottom row depicts the CWT-based topoplots assessed by the same
 328 training scenario (i.e., D), showing that the obtained spatial distribution of A02T
 329 still presents the spectral bandwidths that contribute similarly. However, several
 330 spatial clusters appear, and the amount of meaningless estimates decreases. Never-
 331 theless, a notably enhanced topographic representation is performed by A08T, for
 332 which the CWT-based vectors result in values adequately accommodated within
 333 the electrode space, regardless of the window time. Furthermore, the contribution
 334 concentrates on electrode neighborhoods clearly defined and changing over time.
 335 Thus, the μ rhythm shows the sensorimotor electrodes contribute the most, being
 336 more evident their importance at the window [3-5]s, right at the MI period.

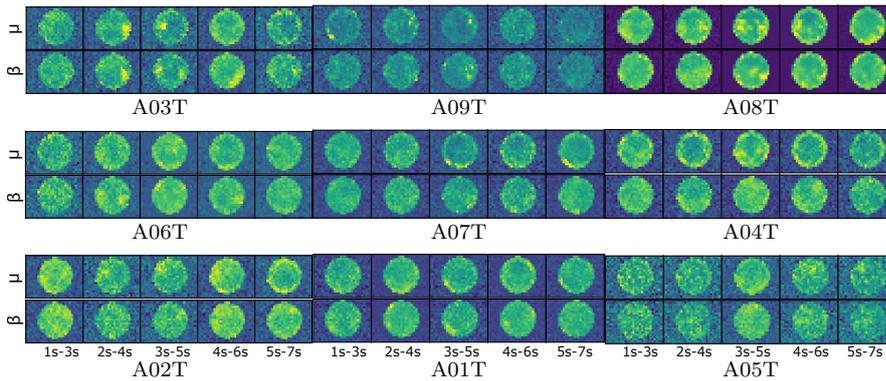


Fig. 4 Topographic representation of scenario C computed for the individually tuned CWT feature extraction.

337 Figure 4 displays the topoplots individually computed for the CWT-based
 338 feature extraction under the scenario C (i.e., $\mu \cup \beta$), showing that the brain activity
 339 tends to gather over some electrodes in most of the subjects. Also, the brain
 340 activity between neighboring time windows changes smoothly, at least, in the
 341 subjects performing high accuracy. As the discrimination ability of individuals
 342 decreases, the topographic representations become more blurred, meaning that
 343 the learned weights are still severely affected by the variability of captured EEG
 344 data. This situation is more visible in A05T (performing the worst) that has a lot

345 of learning weights out of the scalp area, evidencing that the CNN model is likely
 346 to be overtrained.

347 *Performance of Spatial Dropping Strategies.* Two approaches are evaluated: *i*) Re-
 348 moving all electrodes out of the sensorimotor area before training and validation,
 349 and *ii*) Thresholding the electrode contribution after training and validation.

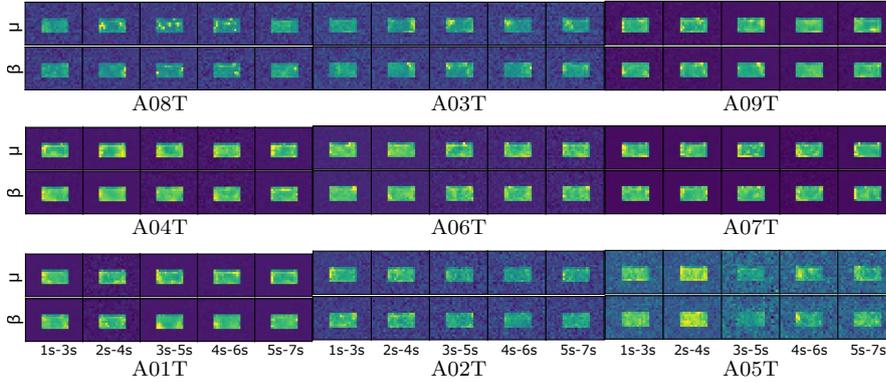


Fig. 5 Assessed weights of relevance by removing all electrodes out of the sensorimotor area computed for the case of CWT-based feature extraction and $\mu \cup \beta$ scenario.

350 The first spatial dropping strategy is implemented by just including all elec-
 351 trodes belonging to the motor cortex region (that is, C3,9,10,11,C4,14,15,16,17,18),
 352 following the spatial electrode distribution reported by [39]. Figure 5 depicts the
 353 estimated topoplots of the individuals ranked by decreasing accuracy and shows
 354 that the brain activity gathers more prominently over some lateral sensorimotor
 355 electrodes in most of the subjects. Moreover, the brain activity between neigh-
 356 boring time-windows changes smoothly, holding the highest contribution within the
 357 segments of MI (2-4 and 3-5 s). In the first subject triad (A08T, A03T, and A09T),
 358 the contribution of either rhythm (μ or β) differs. Besides, the amount of learning
 359 values out the scalp is considerably smaller than in the previous case. Nonetheless,
 360 the topographic representations of the subjects having the worst accuracy (A02T
 361 and A05T) remain still blurred.

362 For the second dropping strategy, Fig. 6 represents the thresholded values,
 363 showing the presence of several electrodes with a relevant contribution. So, the
 364 top subject triad holds the learned weights located on the lateral zones, having
 365 the highest contribution near the sensorimotor area with differentiated behavior
 366 between μ and β rhythms. As expected, the central localities near the longitudinal
 367 fissure have zero-valued weights. However, as the individual performance decreases,
 368 the number of relevant electrodes rises due to the increased variability. Moreover,
 369 the variance of the captured EEG data for the worst-performing subjects is so
 370 strong that they have a distorted topoplot with values out of the scalp. Still, these
 371 subjects present relevant electrodes, unlike the previous approaches achieve.

372 Table 2 summarizes the bi-class performance achieved by each evaluated CNN-
 373 based framework, showing that every subject reaches a performance above $\sim 75\%$.

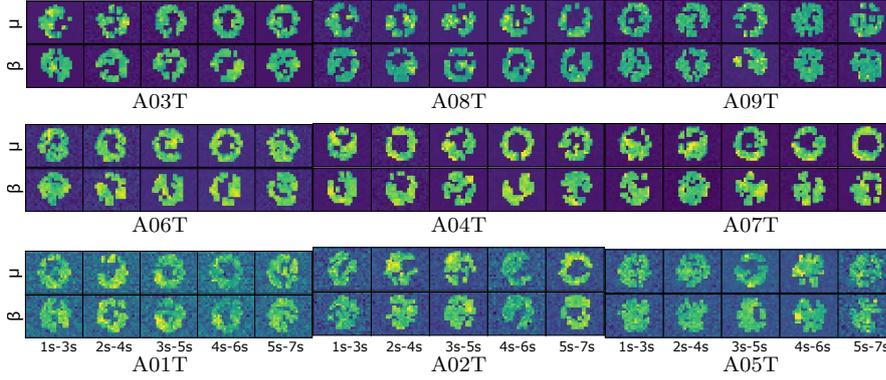


Fig. 6 CWT scenario C all subjects using relevant topoplot mask based relevance vector.

Subjects	[40]	[41]	CWT	κ	CWT^*	κ	CWT^{**}	κ
A03T	88.2	91.7	95.0±4.6	0.67	96.4±4.8	0.92	95.0±4.6	0.90
A09T	82.7	90.9	94.8±4.2	0.68	93.1±6.5	0.86	94.0±6.3	0.88
A08T	91.8	92.3	94.0±4.6	0.90	97.0±3.6	0.94	94.7±4.9	0.89
A06T	65.7	78.5	86.7±7.2	0.71	84.9±9.0	0.69	86.7±7.2	0.73
A07T	51.7	86.5	86.4±6.7	0.58	81.9±6.2	0.64	85.6±9.3	0.71
A04T	53.9	80.4	85.4±7.3	0.73	86.1±7.5	0.72	87.6±5.0	0.75
A02T	63.9	68.4	83.8±6.5	0.73	80.3±6.2	0.61	81.7±4.8	0.63
A01T	79.4	87.8	83.4±5.5	0.88	81.1±5.0	0.62	83.2±3.9	0.66
A05T	54.9	88.9	79.1±4.8	0.90	78.3±7.4	0.57	76.7±6.9	0.53
Average	70.2	85.0±7.4	87.6±5.7	0.75	86.6±6.2	0.73	87.4±5.7	0.74

Table 2 Classifier 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 subbands (μ and β) are included and the CNN parameters are tuned individually.

374 All achieved accuracy scores are competitive with other values performed by CNN-
 375 based approaches recently presented for motor imagery classification (left and right
 376 hand). It is worth noting that the use of either spatial dropping strategy results
 377 in small degradation of classifier accuracy or κ value.

378 5 Discussion and Concluding Remarks

379 We present an approach using CNN models to improve the interpretability of
 380 spatial contribution in terms of discriminating between MI tasks but preserving
 381 an adequate classification accuracy. Results obtained for on BCI Competition IV
 382 dataset 2a proves that the proposed deep learning framework allows improving
 383 accuracy along with revealing the electrodes with higher spatial relevance. Never-
 384 theless, the following aspects are to be regarded in the framework implementation:

385 *Feature Extraction of t - f Vectors.* For each estimated source, the t - f are extracted
 386 within each time window, generating an image containing temporal, frequential,
 387 and spatial information. With the aim of dealing with the non-stationary EEG
 388 nature, we evaluate the extraction of t - f patterns from the FFT-based periodogram

389 and Continuous Wavelet Transform. Then, all extracted t - f feature patterns are
 390 further interpolated to obtain the spatial distribution of activated brain areas
 391 through topographic maps. We obtain that both approaches are similar in terms
 392 of providing classifier performance and complexity of implementing CNN models.
 393 Besides, we evaluate four combining scenarios of μ and β rhythms, which differently
 394 influence the achieved accuracy. In the case of PSD estimates, only the inclusion
 395 of detailed information from three β subbands together with μ waveform provides
 396 the best system accuracy. By contrast, the CWT-based feature set gives high
 397 accuracy scores regardless of the evaluated subband combination. This result may
 398 be explained by the fact that CWT is more suitable for the decomposition of
 399 nonstationary data.

400 Nonetheless, the CWT-based vectors are preferable for interpretation purposes
 401 since the learned weights are less sensitive to the overtraining effect, clustering
 402 better on electrode neighborhoods clearly defined and smoothly changing over
 403 time following the MI paradigm.

404 *Spatial Interpretability of Activated MI Responses.* One aspect to remark is the dropout
 405 technique that CNN models include. Their high number of parameters makes them
 406 particularly prone to overheating, requiring regularization methods in practice and
 407 involving randomly modifying parameters during neural network training or infer-
 408 ence, or approximating this process [42]. Since adjacent pixels are highly correlated
 409 to the dropped pixels in Naive CNN dropout approaches, spatial dropout is devel-
 410 oped to withdraw an entire feature map across a channel [43,44]. Due to the
 411 interpretation of activated brain zones evoked by MI tasks can be performed by
 412 preserving spatial information in input multi-spectral images, we evaluate two
 413 strategies of spatial dropping to promote discarding of irrelevant image details:
 414 Including just the sensorimotor electrodes, and thresholding of the electrode con-
 415 tribution. Although the number learned values out the scalp decreases consider-
 416 ably in the former strategy, the topographic representations of subjects having low
 417 accuracy are still blurred, hindering the interpretation of analyzed brain activity.
 418 The use of full-set EEG electrodes has been already reported as difficult to achieve
 419 in practical MI applications, suggesting that the performance of CNN models can
 420 improve with fewer electrodes, which cover the motor cortex and sensorimotor
 421 cortex [45]. The obtained results show that the thresholding strategy is desirable
 422 since the highest contribution clusters over the sensorimotor area with differen-
 423 tiated behavior between μ and β bands. However, the high EEG data variability
 424 captured by the worst-performing subjects may still produce distorted topoplots
 425 with values out of the scalp, making difficult their understanding.

426 *Evaluated CNN Architecture for MI Discrimination.* The first design consideration
 427 is the number of convolutional layers, together with the type of end classifier. In
 428 MI tasks, 70% of CNN models use a rectified linear unit (ReLU) as the layers
 429 activation function, while the vast majority of classifier fully-connected layers em-
 430 ploy a softmax activation function [46]. The proposed network is fully-connected
 431 that relies on the *Wide&Deep* architecture to handle multiple inputs to learn deep
 432 patterns under simple rules. To improve the performance, the *Classifier Block* in-
 433 cludes batch normalization that is applied to the convolutional outputs before and
 434 after the fully-connected layer FC7. Also, we use the Elastic Net regularization

435 technique through the parameters (l_1, l_2) for preventing over-fitting by penalizing
436 a model having large weights.

437 However, some restrictions are to be mentioned: The first limitation to enhance
438 the performance of the evaluated CNN architecture is the small size of the exam-
439 ined dataset that holds just nine subjects with very different variability [47]. As a
440 result, the deterioration in performance is noticeable (nearly 5%) when decreasing
441 the number of units in each individual trained CNN model. Moreover, the small
442 data issue restricts the application of powerful approaches in deep learning like
443 augmentation or transfer learning, causing overfitting. Another concern is the ad-
444 equate sampling of the potential scalp field for topographic analysis that requires
445 a large number of electrodes [48].

446 As future work, to enhance the impact of tested Deep Learning models, we plan
447 to employ datasets that hold more labeled MI tasks, fusing CNNs with different
448 characteristics and architectures is also to be considered to learn more complex
449 relationships between spatial patterns and extracted t - f representations, making
450 the learned CNN weights be more accessible to interpret [49,50].

451 6 Declarations

452 *Data Availability Statement:* Publicly available datasets were analyzed in this study.
453 This data can be found here: <http://www.bbc.de/competition/iv/#download>.

454 *Funding:* This research was supported by Doctorados Nacionales 2017, conv.785
455 funded by COLCIENCIAS.

456 *Authors' contributions:* DF C-H, AM A-M, G C-D conceived of the presented idea.
457 DF C-H and AM A-M developed the theory based on EEG feature representation
458 and Convolutional neural networks and performed the computations. G C-D and
459 AM A-M verified the analytical methods. DF C-H and AM A-M investigate the
460 influence of the spatial dropout on the topographic map for improving the inter-
461 pretability of brain patterns and supervised the findings of this work. All authors
462 discussed the results and contributed to the final manuscript.

463 *Acknowledgments:*

464 References

- 465 1. C. Cannard, T. Brandmeyer, H. Wahbeh, and A. Delorme. *Self-health monitoring and*
466 *wearable neurotechnologies*, volume 168, pages 207–232. Elsevier, 01 2020.
- 467 2. M Xu, Z Wei, and D Ming. Research advancements of motor imagery for motor function
468 recovery after stroke. *Sheng wu yi xue Gong Cheng xue za zhi= Journal of Biomedical*
469 *Engineering= Shengwu Yixue Gongchengxue Zazhi*, 37(1):169–173, 2020.
- 470 3. A. Guillot and U. Debarnot. Benefits of motor imagery for human space flight: A brief
471 review of current knowledge and future applications. *Frontiers in Physiology*, 10:396, 2019.
- 472 4. L. Pillette, C. Jeunet, R. Nkambou, B. N’Kaoua, and F. Lotte. Towards artificial learning
473 companions for mental imagery-based brain-computer interfaces. *CoRR*, abs/1905.09658,
474 2019.
- 475 5. S. Marchesotti, M. Bassolino, A. Serino, H. Bleuler, and O. Blanke. Quantifying the role
476 of motor imagery in brain-machine interfaces. *Scientific reports*, 6:24076, 2016.

- 477 6. B. Rim, N. Sung, S. Min, and M. Hong. Deep learning in physiological signal data: A
478 survey. *Sensors (Basel, Switzerland)*, 20(4), February 2020.
- 479 7. S. Sakhavi, C. Guan, and S. Yan. Learning temporal information for brain-computer
480 interface using convolutional neural networks. *IEEE Transactions on Neural Networks
481 and Learning Systems*, 29(11):5619–5629, 2018.
- 482 8. R. Zemouri, N. Zerhouni, and D. Racocanu. Deep learning in the biomedical applications:
483 Recent and future status. *Appl. Sci.*, 9(8):1526, 2019.
- 484 9. S. Plis, D. Hjelm, R. Salakhutdinov, E. Allen, H. Bockholt, J. Long, H.. Johnson,
485 J. Paulsen, J. Turner, and V. Calhoun. Deep learning for neuroimaging: a validation
486 study. *Frontiers in Neuroscience*, 8:229, 2014.
- 487 10. H. Wu, Y. Niu, F. Li, Y. Li, B. Fu, G. Shi, and M. Dong. A parallel multiscale filter
488 bank convolutional neural networks for motor imagery eeg classification. *Frontiers in
489 Neuroscience*, 13:1275, 2019.
- 490 11. S. Amin, M. Alsulaiman, G. Muhammad, M. Bencherif, and M. Hossain. Multilevel
491 weighted feature fusion using convolutional neural networks for eeg motor imagery classi-
492 fication. *IEEE Access*, 7:18940–18950, 2019.
- 493 12. C. J Ortiz-Echeverri, S. Salazar-Colores, J. Rodriguez-Resendiz, and R. Gmez-Loenzo. A
494 new approach for motor imagery classification based on sorted blind source separation,
495 continuous wavelet transform, and convolutional neural network. *Sensors*, 19(20), October
496 2019.
- 497 13. C. Guan, L. Tih-Shih, G. Cuntai, S. Fung, D. Shuen, Y. Cheung, S. Teng, H. Zhang,
498 and K. Krishnan. Effectiveness of a brain-computer interface based programme for the
499 treatment of adhd: a pilot study. *Psychopharmacology bulletin*, 43(1):7382, 2010.
- 500 14. M. Doborjeh, N. Kasabov, and Z. Doborjeh. Evolving, dynamic clustering of
501 spatio/spectro-temporal data in 3d spiking neural network models and a case study on
502 eeg data. *Evolving Systems*, 9, 04 2017.
- 503 15. H. Yang, S. Sakhavi, K. K. Ang, and C. Guan. On the use of convolutional neural networks
504 and augmented csp features for multi-class motor imagery of eeg signals classification. In
505 *2015 37th Annual International Conference of the IEEE Engineering in Medicine and
506 Biology Society (EMBC)*, pages 2620–2623, 2015.
- 507 16. S. Sakhavi, C. Guan, and S. Yan. Parallel convolutional-linear neural network for motor
508 imagery classification. In *2015 23rd European Signal Processing Conference (EUSIPCO)*,
509 pages 2736–2740, 2015.
- 510 17. M. Taheri, S. and Ezoji and S.M. Sakhaei. Convolutional neural network based features
511 for motor imagery EEG signals classification in braincomputer interface system. *SN Appl.
512 Sci.*, 2(555), 2020.
- 513 18. X. Tang, W. Li, X. Li, W. Ma, and X. Dang. Motor imagery eeg recognition based
514 on conditional optimization empirical mode decomposition and multi-scale convolutional
515 neural network. *Expert Systems with Applications*, 149:113285, 2020.
- 516 19. J. Zhang, C. Yan, and X. Gong. Deep convolutional neural network for decoding motor
517 imagery based brain computer interface. In *2017 IEEE International Conference on Signal
518 Processing, Communications and Computing (ICSPCC)*, pages 1–5, 2017.
- 519 20. T. Uktveris and V. Jusas. Application of convolutional neural networks to four-class motor
520 imagery classification problem. *ITC*, 46:260–273, 2017.
- 521 21. H. K. Lee and Y. Choi. A convolution neural networks scheme for classification of motor
522 imagery eeg based on wavelet time-frequency image. In *2018 International Conference on
523 Information Networking (ICOIN)*, pages 906–909, 2018.
- 524 22. J. Yang, S. Yao, and J. Wang. Deep fusion feature learning network for mi-eeg classifica-
525 tion. *IEEE Access*, 6:79050–79059, 2018.
- 526 23. S Petrichella, L Vollere, F Ferreri, A Guerra, S Maatta, M Kononen, V Di Lazzaro, and
527 G Iannello. Channel interpolation in tms-eeg: a quantitative study towards an accurate
528 topographical representation. *Conference proceedings : ... Annual International Confer-
529 ence of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in
530 Medicine and Biology Society. Annual Conference*, 2016:989992, August 2016.
- 531 24. Y. Tabar and U. Halici. A novel deep learning approach for classification of eeg motor
532 imagery signals. *Journal of neural engineering*, 14 1:016003, 2017.
- 533 25. P. Thodoroff, J. Pineau, and A. Lim. Learning robust features using deep learning for
534 automatic seizure detection, 2016.
- 535 26. G. Xu, X. Shen, S. Chen, Y. Zong, C. Zhang, H. Yue, M. Liu, F. Chen, and W. Che. A
536 deep transfer convolutional neural network framework for eeg signal classification. *IEEE
537 Access*, 7:112767–112776, 2019.

- 538 27. R. D'Souza, P. Huang, and F. Yeh. Structural analysis and optimization of convolutional
539 neural networks with a small sample size. *Scientific Reports*, 10, 2020.
- 540 28. M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang. Eeg classification of motor imagery
541 using a novel deep learning framework. *Sensors (Basel, Switzerland)*, 19, 2019.
- 542 29. Y. Rong, X. Wu, and Y. Zhang. Classification of motor imagery electroencephalography
543 signals using continuous small convolutional neural network. *International Journal of*
544 *Imaging Systems and Technology*, pages 1–7, 2019.
- 545 30. X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, and L. Sun. A multi-branch 3d convolutional
546 neural network for eeg-based motor imagery classification. *IEEE Transactions on Neural*
547 *Systems and Rehabilitation Engineering*, 27:2164–2177, 2019.
- 548 31. P. Bashivan, I. Rish, M. Yeasin, and N. Codella. Learning representations from eeg with
549 deep recurrent-convolutional neural networks. *CoRR*, abs/1511.06448, 2016.
- 550 32. D. McFarland, L. Miner, T. Vaughan, and J. Wolpaw. Mu and beta rhythm topographies
551 during motor imagery and actual movements. *Brain Topography*, 12:177–186, 2004.
- 552 33. A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-
553 trial EEG dynamics including independent component analysis. *Journal of Neuroscience*
554 *Methods*, 134(1):9 – 21, 2004.
- 555 34. A.M. Alvarez-Meza, L.F. Velasquez-Martinez, and G. Castellanos-Dominguez. Time-series
556 discrimination using feature relevance analysis in motor imagery classification. *Neurocom-*
557 *puting*, 151:122–129, 2015.
- 558 35. H. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Cor-
559 rado, W. Chai, M. Ispir, R. Anil, Z. Haque, L. Hong, V. Jain, X. Liu, and H. Shah. Wide
560 & deep learning for recommender systems, 2016.
- 561 36. H. Ide and T. Kurita. Improvement of learning for CNN with ReLU activation by sparse
562 regularization. In *2017 International Joint Conference on Neural Networks (IJCNN)*,
563 pages 2684–2691, 2017.
- 564 37. V. Lawhern, A. Solon, N. Waytowich, S. Gordon, C Hung, and B. Lance. EEGNet: a
565 compact convolutional neural network for EEG-based brain–computer interfaces. *Journal*
566 *of Neural Engineering*, 15(5):056013, jul 2018.
- 567 38. F. Li, F. He, F. Wang, D. Zhang, Y. Xia, and X. Li. A novel simplified convolutional neural
568 network classification algorithm of motor imagery eeg signals based on deep learning.
569 *Applied Sciences*, 10(5):1605, 2020.
- 570 39. C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller. Bci competition
571 2008–graz data set a. *Institute for Knowledge Discovery (Laboratory of Brain-Computer*
572 *Interfaces)*, *Graz University of Technology*, 16, 2008.
- 573 40. S. Shahtalebi, A. Asif, and A. Mohammadi. Siamese neural networks for eeg-based brain-
574 computer interfaces, 02 2020.
- 575 41. B. Olivas-Padilla and M. Chacon-Murguia. Classification of multiple motor imagery using
576 deep convolutional neural networks and spatial filters. *Applied Soft Computing*, 75:461 –
577 472, 2019.
- 578 42. A. Labach, H. Salehinejad, and S. Valaee. Survey of dropout methods for deep neural
579 networks, 04 2019.
- 580 43. J. Tompson, R. Goroshin, A. Jain, Y. Lecun, and C. Bregler. Efficient object localization
581 using convolutional networks. In *The IEEE Conference on Computer Vision and Pattern*
582 *Recognition (CVPR)*, pages 648–656, 06 2015.
- 583 44. S. Park and N. Kwak. Analysis on the dropout effect in convolutional neural networks. In
584 *Computer Vision – ACCV 2016*, pages 189–204, 03 2017.
- 585 45. X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, and L. Sun. A multi-branch 3d convolutional
586 neural network for eeg-based motor imagery classification. *IEEE Transactions on Neural*
587 *Systems and Rehabilitation Engineering*, 27(10):2164–2177, 2019.
- 588 46. A. Craik, A. Kilicarslan, and J. L. Contreras-Vidal. Classification and transfer learning
589 of eeg during a kinesthetic motor imagery task using deep convolutional neural networks.
590 In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and*
591 *Biology Society (EMBC)*, pages 3046–3049, 2019.
- 592 47. D. Collazos-Huertas, J. Caicedo-Acosta, G. Castao-Duque, and C. Acosta-Medina. En-
593 hanced multiple instance representation using time-frequency atoms in motor imagery
594 classification. *Frontiers in Neuroscience*, 14:155, 2020.
- 595 48. C. Michel. Chapter 12 - high-resolution eeg. In Kerry H. Levin and Patrick Chauvel,
596 editors, *Clinical Neurophysiology: Basis and Technical Aspects*, volume 160 of *Handbook*
597 *of Clinical Neurology*, pages 185 – 201. Elsevier, 2019.

-
- 598 49. S. Amin, M. Alsulaiman, G. Muhammad, M. Mekhtiche, and M. Hossain. Deep learning
599 for eeg motor imagery classification based on multi-layer cnns feature fusion. *Future*
600 *Generation Computer Systems*, 101:542 – 554, 2019.
- 601 50. Z. Wang, L. Cao, Z. Zhang, X. Gong, Y. Sun, and H. Wang. Short time fourier transfor-
602 mation and deep neural networks for motor imagery brain computer interface recognition.
603 *Concurrency and Computation: Practice and Experience*, 30(23):e4413, 2018.

Figures

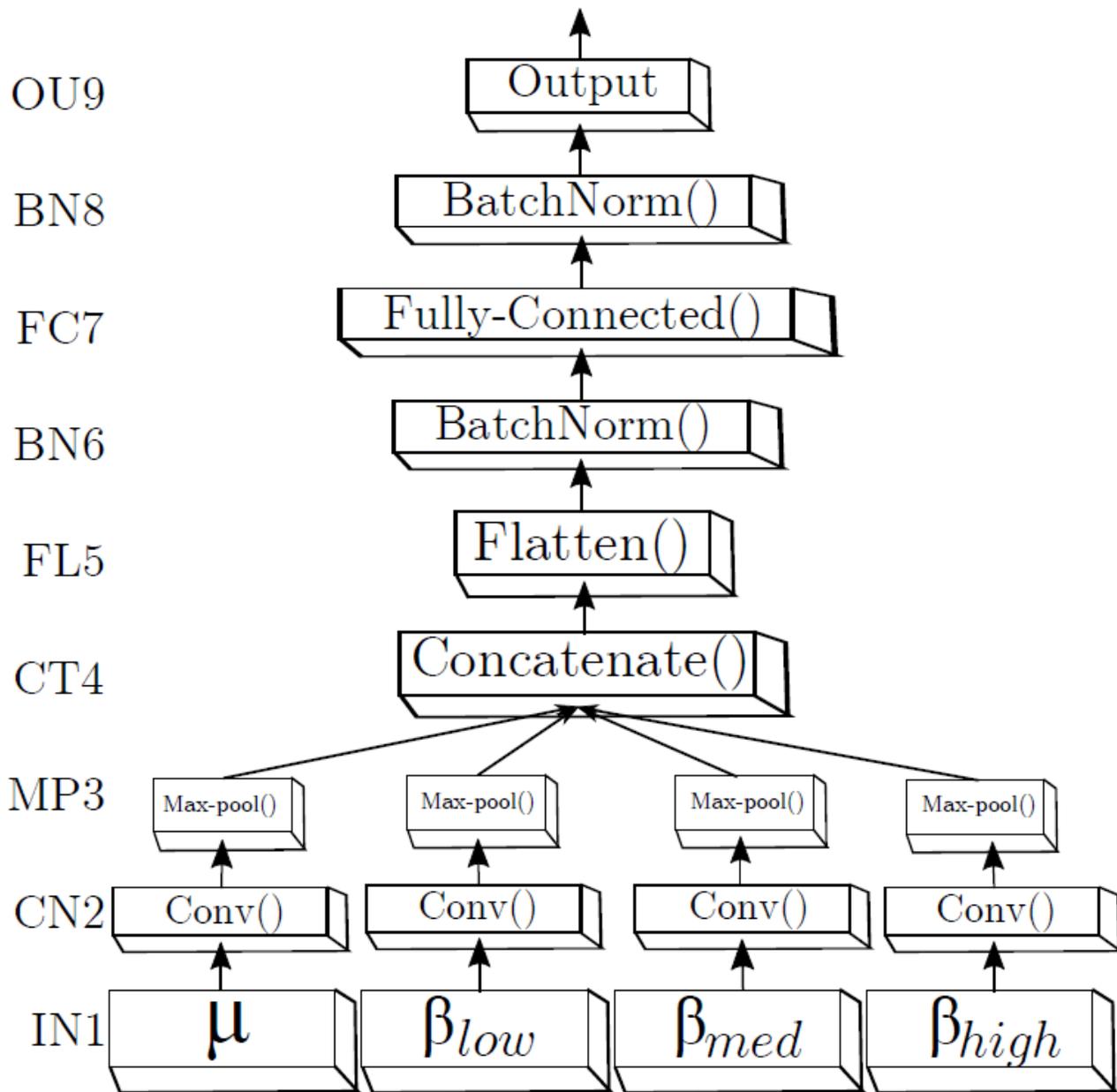


Figure 1

The proposed CNN structure that is based on Wide&Deep neural network handling multiple inputs. The first layer (IN1) is the input, second (CN2) and third layers (MP3) are hidden and accomplish the feature mapping generation, while the block, ranging from the output of layer CT4 to the OU9 layer, comprises the classification stage.

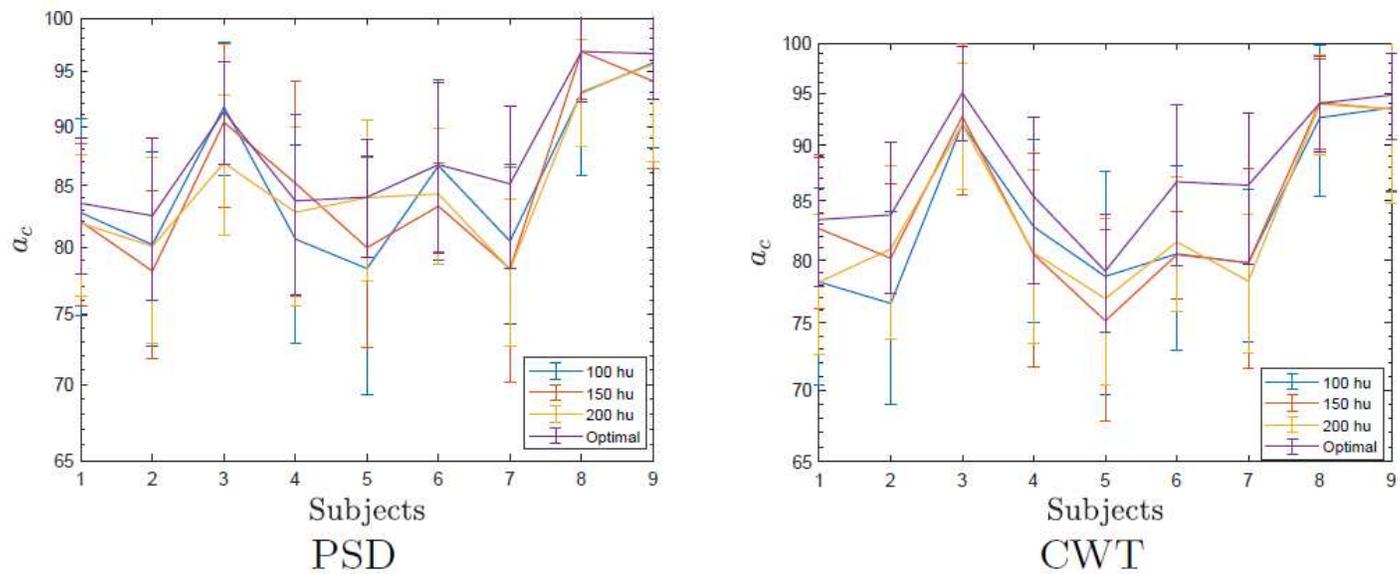


Figure 2

Dependence of CNN hidden units and the individual accuracy. Label "optimal" is the case of individually tuned CNN model.

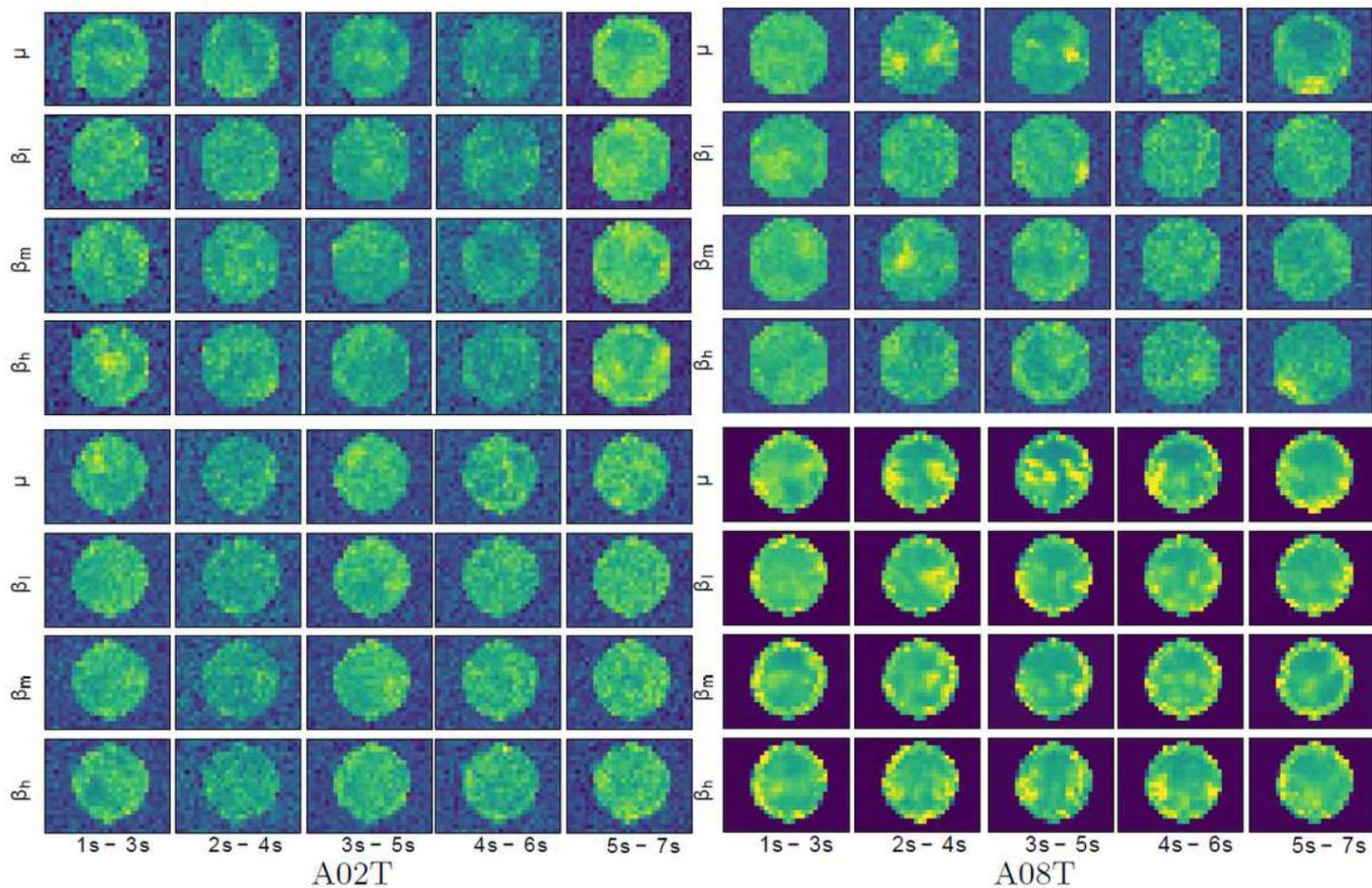


Figure 3

Reconstructed topographic maps for experiment D including all bandwidths across the time domain in the best (A08T) and the worst (A02T) subjects, left and right column respectively.

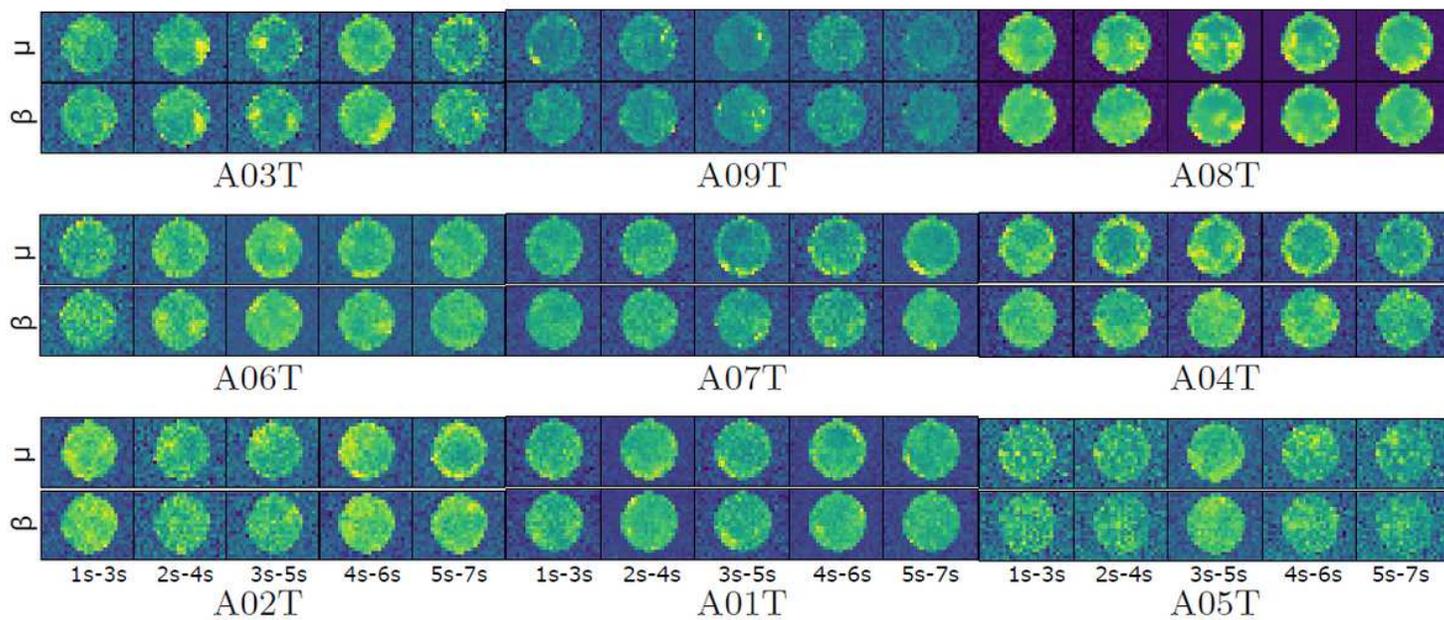


Figure 4

Topographic representation of scenario C computed for the individually tuned CWT feature extraction.

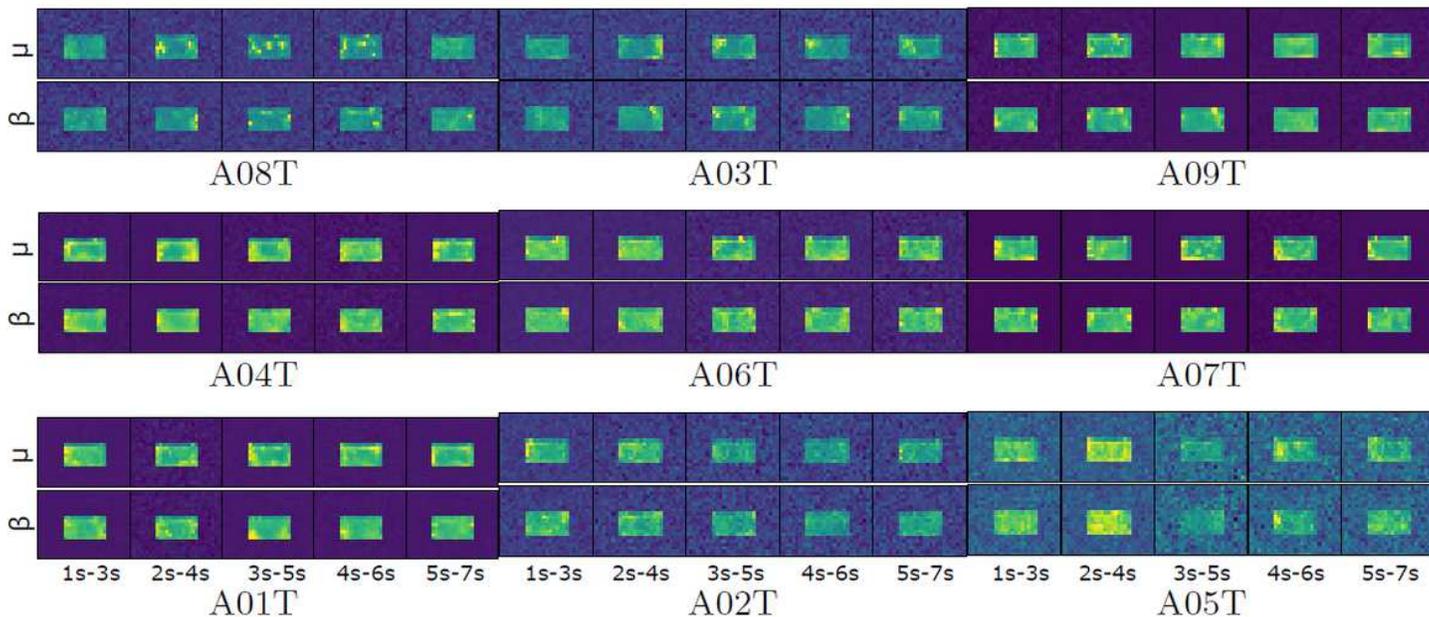


Figure 5

Assessed weights of relevance by removing all electrodes out of the sensorimotor area computed for the case of CWT-based feature extraction and μ U β scenario.

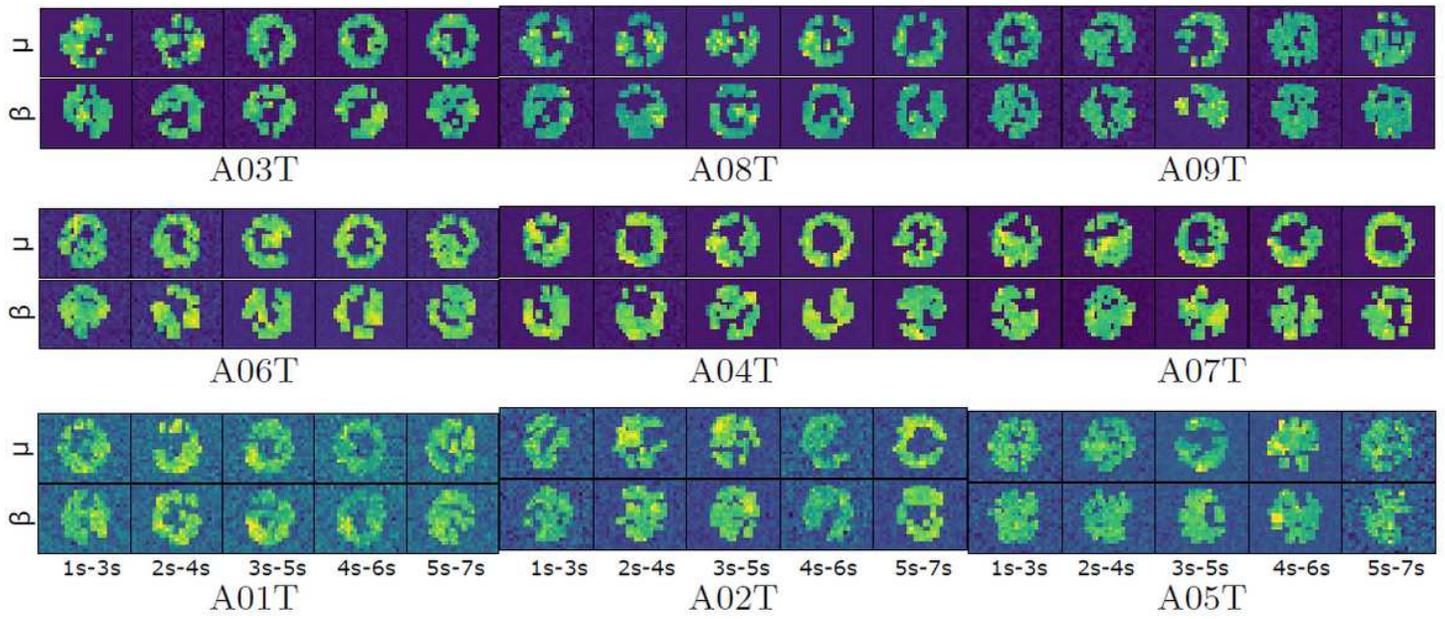


Figure 6

CWT scenario C all subjects using relevant topoplot mask based relevance vector.