# Atlas-based classification algorithms for identification of informative brain regions in fMRI data

**Juan E. Arco[1], Paloma Díaz-Gutiérrez[1], Javier Ramírez[2], María Ruz[1*]**

[1]Mind, Brain and Behavior Research Centre (CIMCYC), University of Granada, Granada, Spain

[2]Department of Signal Theory, Networking and Communications, University of Granada, Granada, Spain

*** Correspondence:**
María Ruz
mruz@ugr.es

**Abstract**

Multi-voxel pattern analysis (MVPA) has been successfully applied to neuroimaging data due to its larger sensitivity compared to univariate traditional techniques. Although a Searchlight strategy that locally sweeps all voxels in the brain is the most extended approach to assign functional value to different regions of the brain, this method does not offer information about the directionality of the results and it does not allow studying the combined patterns of more distant voxels.

In the current study, we examined two different alternatives to searchlight. First, an atlas-based local averaging (ABLA, Schrouff et al., 2013a) method, which computes the relevance of each region of an atlas from the weights obtained by a whole-brain analysis. Second, a Multiple-Kernel Learning (MKL, Rakotomamonjy et al., 2008) approach, which combines different brain regions from an atlas to build a classification model. We evaluated their performance in two different scenarios where differential neural activity between conditions was large vs. small, and employed nine different atlases to assess the influence of diverse brain parcellations.

Results show that all methods are able to localize informative regions when differences were large, demonstrating stability in the identification of regions across atlases. Moreover, the sign of the weights reported by these methods provides the sensitivity of multivariate approaches and the directionality of univariate methods. However, in the second context only ABLA localizes informative regions, which indicates that MKL leads to a lower performance when differences between conditions are small. Future studies could improve their results by employing machine learning algorithms to compute individual atlases fit to the brain organization of each participant.

**Keywords: Multi-voxel pattern analysis, Multiple-Kernel Learning, Searchlight, Atlas-based local averaging, fMRI, permutation testing.**

## 1 Introduction

The use of machine learning in neuroscience has increased exponentially in the last years. Several studies have employed these methods in clinical contexts, providing tools for computer-aided diagnosis of different neurological disorders, such as Alzheimer's (Arco et al., 2015), Parkinson's (Choi et al., 2017), epilepsy (Del Gaizo et al., 2017) or brain computer interfaces in quadriplegic patients (Blankertz et al., 2007; Nurse et al., 2015). Here, obtaining the maximum decoding performance is the main aim, whereas the source of information is not of interest. On the other hand, machine learning methods are also used to study the brain regions involved in different cognitive operations (Haxby et al., 2014), and here the main goal is not prediction itself. Hebart and Baker (2017) remarked the importance of differentiating multivariate decoding for prediction and for interpretation as two independent frameworks. In the interpretation context, Multi-voxel pattern analysis (MVPA) provides larger sensitivity than classic univariate approaches (Haynes and Rees, 2006; Norman et al., 2006), as it localizes where information is contained based on the distribution of spatial patterns. However, MVPA brings some crucial points to be considered: is it possible to use these techniques in a different context from which they were developed for? If so, would it be necessary to modify the existing algorithms to accomplish the new goals? Finding the most adequate method for each specific context is of vital importance, and in the current investigation we aimed to compare the sensitivity of different approaches and to propose some variations to assess the suitability of these methods in the field of Cognitive Neuroscience.

From the identification perspective, classification is simplest when performed in a region of interest (ROI) based on *a priori* knowledge. The accuracy of the algorithm highly depends on how well the regional hypothesis fits the observed data. Haxby et al. (2001) demonstrated that the representations of faces and objects were differentially distributed in the ventral temporal cortex, whereas Haynes and Rees (2005) showed that there is an orientation-selective processing in the primary visual cortex (V1). Other studies detected distributed patterns of activity in the visual cortex (Cox and Savoy, 2003; Kamitani and Tong, 2005), whereas Poldrack (2007) highlighted the Type 1 error reduction when a statistical test is applied to each ROI. However, when there is not a straightforward hypothesis regarding the regions involved in specific computations, the whole brain may need to be explored. The main drawback of whole-brain analyses is related to the curse of dimensionality: in fMRI studies, there are usually many more features (e.g. voxels) than samples (e.g. images or volumes), which complicates the definition of a classification model to separate the two classes (Fort and Lambert-Lacroix, 2005). Alternatively, feature-selection methods select a subset of informative features (e.g. voxels in fMRI) that will be the input of the subsequent classification. As an example, *t*-tests can be used to restrict the voxels fed to the classifier to those that differ between the classes. Previous studies have employed this method to localize the regions associated with different pathologies and psychological contexts (Arco et al.,

2015; Balci et al., 2008; Haynes and Rees, 2005; De Martino et al., 2008; see Mwangi et al., 2014, for a detailed review).

One of the most appealing approaches for identification of cognitive informative regions is the Searchlight technique (Kriegeskorte et al., 2006), a method that offers results potentially easier to interpret due to its larger spatial precision without the need to define specific ROIs. Searchlight produces maps of accuracies from small spherical regions centered on each voxel of the brain. For each sphere, a classification analysis is performed, and the decoding performance is assigned to the central voxel. Many studies have successfully used this technique (e.g. Chen et al., 2017; Cichy et al., 2016; Coutanche et al., 2011; González-García et al., 2017; Loose et al., 2017; Qiao et al., 2017). However, it also has some disadvantages and limitations to consider. Searchlight performance depends on the size of the sphere; the larger its radius, the larger the number of significant voxels, even when the size of the informative regions stayed fixed (Etzel et al., 2013; although see Arco et al. 2016). Another drawback is that the accuracy of the classification within a certain sphere is associated with the central voxel, which obviates the possibility that only a few voxels of the total in the sphere truly contain information. As a result, some voxels may be marked as significant only because they are at the center of a sphere that contains informative voxels, leading to somewhat distorted results (see Figure 3 in Etzel et al., 2013 for an extreme example). Another problem is its high computational cost. Each Searchlight analysis entails a massive number of classifications (one for each voxel of the brain), increasing the computational time compared to other simpler approaches. This time cost increases exponentially when different values of the parameters associated with the classifier are evaluated to find the one with the largest performance (grid search) and also when permutation tests are used to evaluate the statistical significance of the results.

There are other alternatives based on atlas that do not suffer from this large computational cost. This is the case of Multiple Kernel Learning (MKL, Lanckriet et al., 2004), a method that uses *a priori* templates of brain organization to guide the decision of the classifier. Specifically, this approach extracts information from brain parcellations provided by an atlas to maximize the performance of the classification algorithm, and ranks the regions according to their importance in the decision. A crucial aspect is the two-level hierarchical model that this approach entails. The regions used for classification have an associated weight, which indicates their contribution to the model. Voxels within each region have a similar weight value. Thus, MKL offers information both at the region and at the voxel level. Previous studies have used this method in the context of neuroimaging, e.g. discrimination between Parkinson's neurological disorders (Adeli et al., 2017; Filippone et al., 2013), identification of attention deficit hyperactivity disorder (ADHD) patients (Dai et al., 2017; Qureshi et al., 2017) and localization of informative regions (Schrouff et al., 2018). This approach leads to a sparse solution, which means that only a subset of regions is selected to contribute to the decision function (similarly to feature-selection methods). However, this decreases its ability to detect informative regions, which is not recommended when identification of informative areas is the main aim. Schrouff et al. (2013a) proposed another decoding-based method based on local averages of the weights from each region defined in an atlas. This is known as Atlas-

based local averaging (ABLA). First, a whole-brain classification is performed, leading to a weight map summarizing the contribution of each voxel. Then, the weights defined in each region of the atlas are averaged and normalized by the size of the region. This yields a score of the informativeness of each region. Hence, this approach builds only one classification model since the summary of the weights is done *a posteriori*. In contrast, MKL combines the different regions of the atlas as part of the learning process, so that using a different atlas will result in a different classification, with the subsequent increase in computational cost compared with ABLA.

Previous research has usually employed atlas-based methods in classification contexts, where the main aim is to obtain the largest accuracy possible. However, the validity of these approaches in an identification scenario, where the goal is to find the informative brain regions during a certain mental operation, is yet unknown. Therefore, in this study, we evaluated the performance of different atlas-based approaches in an fMRI experiment, in two contexts with differential changes in neural activity. To do so, we modified the MKL and ABLA methods to better fit the requirements of an identification context instead of a classification one. Specifically, we included an L2-version of MKL, which avoids sparsity by allowing all regions of the corresponding atlas to contribute to the model. We compared the results obtained by MKL and ABLA methods to those obtained by Searchlight, as this approach is mainstream in recent neuroimaging research. In our study, we employed nine different atlases to examine how different brain parcellations influenced the identification of informative regions of MKL and ABLA. We predicted that L2-MKL and ABLA would be more sensitive than L1-MKL, and that they would show a larger overlap with Searchlight results. For a contrast with large differences in the neural activity, we expected overlap between the significant regions obtained by all the approaches. However, this overlap would decrease for the contrast testing more subtle difference in neural activity. In this case, we hypothesized that the specific organization of the brain proposed by each atlas would highly affect the identification of significant regions.

## 2 Material

### 2.1 Participants

Twenty-four students from the University of Granada (M = 21.08, SD = 2.92, 12 men) took part in the experiment and received an economic remuneration (20-25 euros, depending on performance). All of them were right-handed with normal to corrected-to-normal vision, no history of neurological disorders, and signed a consent form approved by the local Ethics Committee. The sample size was chosen according to previous studies that focused on a very similar paradigm (see Gaertig et al., 2012; Chang and Sanfey, 2013 and Grecucci et al., 2013).

### 2.2 Image acquisition

fMRI data were acquired using a 3T Siemens Trio scanner at the Mind, Brain and Behavior Research Centre (CIMCYC) in Granada (Spain). Functional images were obtained with a T2*-weighted echo planar imaging (EPI) sequence, with a TR of 2000 ms. Thirty-two descendent slices with a thickness of 3.5 mm (20% gap) were obtained (TE = 30 ms, flip

angle = 80°, voxel size of 3.5x3.5x3.5 $mm^3$). The sequence was divided in 8 runs, consisting of 166 volumes each. After the functional sessions, a structural image of each participant with a high-resolution T1-weighted sequence (TR = 1900 ms; TE = 2.52 ms; flip angle = 9°, voxel size of 1 $mm^3$) was acquired.

We used SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12) to preprocess and analyse the neuroimaging data. The first 3 volumes were discarded to allow for saturation of the signal. Images were realigned and unwarped to correct for head motion. Afterwards, T1 images were coregistered with the realigned functional images. Then we used slice timing correction to account for differences in slice acquisitions.

### 2.3 Design

Participants played the role of the responder in a modified Ultimatum Game (for theoretical background, which is not the focus of the present study, see Gaertig et al., 2012), deciding whether to accept or reject monetary offers made by different partners. If they accepted the offer, both parts earned their respective splits, whereas if they rejected it, neither of them earned money from that exchange. Offers consisted in splits of 10 Euros, which could be fair (5/5, 4/6) or unfair (3/7, 2/8, 1/9). The number on the left was always the amount of money given to the participant, and the one on the right was the one proposed by the partners for themselves.

Personal information about the partners was included as adjectives with different valence (16: half positive and half negative). For a third of the trials, the description were positive and another third negative, while the rest consisted on neutral trials with a text indicating absence of information ("no test"). Thus, the task contained two events in each trial, first a word (positive, negative or neutral in valence) and second two numbers (a monetary offer), to which participants had to respond. They performed a total of 192 trials, arranged in 8 runs (24 trials per run), in a counterbalanced order across participants. Each trial started with the word for 1000 ms, followed by a jittered interval lasting 5500 ms on average (4-7 s, +/0.25). Then, the numbers appeared for 500 ms followed by a second jittered interval (5500 ms on average, 4-7 s, +/0.25). Thus, participants read an adjective with a certain valence, and then they used this information to prepare to respond to the offer (second event). Thus, there is a preparatory process that leads to sustained activity along time. Because of this, the first event (words), was modelled as the duration of the word and the variable jittered interval, yielding a global duration ranging from 5 to 8 seconds. On the other hand, the second event was modelled as an impulse function (Dirac delta), i.e. with zero duration, as explained in Henson (2005), because this second event captures a different process. Once participants make a decision (accept the offer or not), the process ends. A large body of literature shows that preparatory processes extend in time (e.g. Bode and Haynes, 2009; González-García et al., 2017,2016; Sakai, 2008) whereas responding to a brief target does not (see the temporal duration of the potentials in Moser et al., 2014). This has been also measured by other neuroimaging methods, such as CNV ERP potential (Di Russo et al., 2017).

To test the reliability of the different approaches (sensitivity and overlap of the significant regions with those obtained by Searchlight), we focused on two different classification analyses. First, we aimed at discriminating between the observed pattern associated with accepting *vs.* rejecting offers (from now on, *decision* classification). The hand used to respond was counterbalanced across participants, which means that odd subjects used the right/left hand to accept/reject an offer, whereas in even subjects the order was the opposite. Second, we focused on distinguishing the positive *vs.* negative valence of the words (e.g. Lindquist et al., 2015; from now on, *valence* classification) that were equated in number of letters, frequency of use and arousal (Gaertig et al., 2012). Extensive previous research shows that motor responses associated with accepting *vs.* rejecting offers lead to large neural differences in motor cortex (e.g. Gabay et al., 2014; Scenario 1) whereas positive *vs.* negative information leads to very subtle activation differences (e.g. Lindquist et al., 2015, Scenario 2). We employed a Least-Squares Separate (LSS) model to obtain an accurate estimation of the neural activity (Turner et al., 2012). This method is based on iteratively fitting a new GLM for each event of every trial of the experiment. Thus, each model has two predicted BOLD time courses: one for the target event and a nuisance parameter estimate that represents the activation for the rest of the events of the same run. Previous studies have shown that this is the best approach for isolating the activity in contexts like this experiment (Abdulrahman and Henson, 2016; Arco et al., 2018), where overlap and collinearity among regressors are large.

### 2.4 Atlases

In this study, we used 9 atlases to assess the reliability of the informative regions obtained by the three atlas-based classification methods. They differ in three main aspects: the information that they use to cluster the brain regions (anatomical, functional or multimodal), the number of resulting regions (from 12 to 400) and the algorithms that implement the parcellation (a wide spectrum, from the *k*-means clustering to Bayesian models).

#### 2.4.1 BASC Cambridge

This atlas was computed from group brain parcellations generated by the BASC (Bootstrap Analysis of Stable Clusters) method, an algorithm based on *k*-means clustering to identify brain networks with coherent activity in resting-state fMRI (Bellec et al., 2010). These networks were generated from the Cambridge sample from the 1000 Functional Connectome Project (Liu et al., 2009). Based on this framework, different atlases were built depending on the number of networks defined (Urchs et al., 2015). In this study, we used four versions with 12, 20, 36 and 64 regions.

#### 2.4.2 AICHA

This atlas covers the whole cerebrum and is based on resting-state fMRI data acquired in 281 individuals (Joliot et al., 2015), and also relies on *k*-means clustering. One interesting feature is that it accounts for homotopy, relying on the assumption that a region in one hemisphere has a homologue in the other hemisphere. This leads to 192 homotopic region pairs (122 gyral, 50 sulcal and 20 gray nuclei).

### 2.4.3. Brainnetome

Fan et al (2016) introduced an atlas based on connectivity using in vivo diffusion MRI (dMRI) and fMRI data acquired in 40 subjects. It divides the human brain into 210 cortical and 36 subcortical regions, providing detailed information based on both anatomical and functional connectivity. The number of regions was computed with a cross-validation procedure to maximize consistency across subjects (Fan et al., 2014; Liu et al., 2013). All functional data, connections and brain parcellations are freely available at http://atlas.brainnetome.org.

### 2.4.4 Yeo2011

This atlas used a clustering algorithm to parcellate the cerebral cortex into networks of functionally coupled regions, employing fMRI data from 1000 subjects. The method employed assumes that each vertex of the cortex belongs to a single network (see Yeo et al., 2011). There are two versions available depending on the number of networks considered (7 or 17). We employed the latter for the subsequent analysis as it offers a more detailed parcellation of the brain. This atlas is preinstalled in Lead-dbs toolbox (http://www.lead-dbs.org).

### 2.4.5 Harvard-Oxford

Clustering in this atlas was performed with the automatic algorithm presented in Desikan et al. (2006), which subdivides structural magnetic resonance data of the human cerebral cortex into gyral based regions of interest (ROI). Its validity was evaluated by computing correlation coefficients and mean distances between these results and manually identified cortical ROIs. Forty-eight cortical regions were obtained from data of 37 subjects. The resulting atlas is freely distributed with FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/).

### 2.4.6 Schaefer

This atlas adds novel parcellations and a larger precision to the brain networks published in Yeo et al. (2011) by using a local gradient approach to detect abrupt transitions in functional connectivity patterns (Schaefer et al., 2017). These transitions are likely to reflect cortical areal boundaries defined by histology or visuotopic fMRI. The resulting parcellations were generated from resting-state fMRI based on 1489 participants (see original paper for further details). There are several versions of this atlas depending on the number of regions the brain is divided into (400, 600, 800 or 1000), but we selected the first one to maintain reasonable speed on computation analyses.

## 3 Methods

In this study, we considered four different algorithms based on linear classifiers. First, the atlas-based local averaging method (ABLA) presented in Schrouff et al. (2013a). Second, an

L1-MKL version of the algorithm introduced in Rakotomamonjy et al. (2008) and implemented in the PRoNTo toolbox (Schrouff et al., 2013b). Third, a modification of the L1-MKL to use an L2-norm instead of an L1 (from now on, L2-MKL) to avoid the sparsity that L1 leads to and the subsequent decrease in detecting informative regions. Finally, we used a Searchlight approach as a contrast common reference.

### 3.1 Atlas-based local averaging (ABLA)

This method is used after performing a whole-brain analysis in which all voxels of the brain are used as input to the classification algorithm. A linear classifier leads to a weight map where each value corresponds to the contribution of each voxel to the decision function. ABLA computes a normalization of the average weight for each region of an atlas that summarizes the importance of this region in a certain classification context. From a mathematical perspective, it is possible to specify a linear SVM (Bennett and Blue, 1998; Burges, 1998; Gaonkar et al., 2015) classification rule $f$ by a pair of $(\boldsymbol{w}, \boldsymbol{x})$, from equation:

$$f(\boldsymbol{x}_i) = \langle \boldsymbol{w}, \boldsymbol{x}_i \rangle + b \quad (1)$$

where $\boldsymbol{w}$ is the weight vector, $\boldsymbol{x}_i$ is the feature vector and $b$ is the error term. Thus, a point $x$ is classified as positive if $f(x) > 0$ or negative if $f(x) < 0$. The decision function is based on a linear rule that maximizes the geometrical margin between the two classes, which is obtained by solving the optimisation problem described in Boser et al. (1992):

$$\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_i \xi_i \quad \text{subject to} \quad y_i(\langle \boldsymbol{w}, \boldsymbol{x}_i \rangle + b) \geq 1 - \xi_i \quad \forall_i \xi_i \geq 0 \quad \forall_i \quad (2)$$

The solution to the optimization problem can be written as:

$$\boldsymbol{w} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} y_i \alpha_i \boldsymbol{x_i} \quad (3)$$

after applying the Lagrangian multipliers. Substituting the value of $\boldsymbol{w}$ in Equation 1, it is possible to rewrite the decision function in its dual form as

$$f(\boldsymbol{x}_i) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \alpha_i K(\boldsymbol{x}, \boldsymbol{x}_i) + b \quad (4)$$

where $\alpha_i$ and $b$ represent the coefficients to be learned from the examples and $K(\boldsymbol{x}, \boldsymbol{x}_i)$ is the kernel function characterizing the similarity between samples $\boldsymbol{x}$ and $\boldsymbol{x}_i$.

Once the significant model was obtained, we extracted the weight maps that guided the decision of the classifier. Then, we computed the normalized weight for each region in the atlas as the average of the absolute value of the weights contained in each region, as explained in Schrouff et al. (2013a). Equation 5 summarizes mathematically this computation:

$$NW_{ROI} = \frac{\sum_{v \in ROI} |W_v|}{m_{ROI}}$$

with $v$ representing the index of a voxel in the weight map, $W_v$ its weight and $m_{ROI}$, the number of voxels in region ROI.

### 3.2 Multiple Kernel Learning

This method combines the information from the different brain regions of an atlas to build the classification model, in contrast to the use in ABLA of the corresponding brain organization *a posteriori*. Specifically, MKL combines different kernels and optimizes their contribution to the model to obtain the highest performance. As a result, this approach offers information at two levels: regions and each voxel within them. Mathematically, the decision function is computed as a linear combination of all these basis kernels as stated in Lanckriet et al. (2004):

$$K(\boldsymbol{x}, \boldsymbol{x}') = \sum_{m=1}^{M} d_m K_m(\boldsymbol{x}, \boldsymbol{x}') \quad \text{with} \quad d_m \geq 0, \sum_{m=1}^{M} d_m = 1 \qquad (6)$$

where $M$ is the total number of kernels.

The decision function of the MKL problem is very similar to SVM (Equation 1) but adding the sum of the different kernels from the corresponding atlas:

$$f(\boldsymbol{x}_i) = \sum_m \langle \boldsymbol{w}_m, \boldsymbol{x}_i \rangle + b \qquad (7)$$

The MKL version considered in this study is based on the formulation presented in Rakotomamonjy et al. (2008), where a solution is obtained by solving the following optimization problem:

$$\text{minimize} \quad \frac{1}{2}\sum_m \frac{1}{d_m} \|\boldsymbol{w}_m\|^2 + C \sum_i \xi_i \quad \text{subject to}$$

$$y_i(\sum_m \boldsymbol{w}_m, \boldsymbol{x}_i + b) \geq 1 - \xi_i \quad \forall i \xi_i \geq 0 \quad \forall i \sum_m d_m = 1, \; d_m \geq 0 \; \forall m \qquad (8)$$

where $d_m$ is the contribution to the decision function of each region (see Rakotomamonjy et al., 2008 for a detailed explanation).

This MKL variation optimizes, in a simultaneous manner, the contribution to the decision function of every voxel within a region and the contribution of the region as a whole, in a two-level hierarchical model. In addition, the L1-norm (Tibshirani, 1996) constraint on $d_m$ enforces sparsity on some kernels, resulting in a zero-contribution of these regions: information from a region that is present in another is automatically discarded in one of them. Mathematically:

$$S = \sum_{i=1}^{n} |y_i - f(\boldsymbol{x}_i)| \qquad (9)$$

Thus, the L1-norm is based on minimizing the sum of the absolute differences between the target value ($y_i$) and the estimated values ($f(\boldsymbol{x}_i)$). This hierarchical model leads to two different weight maps: one that summarizes the contribution to the model of each region (region level), and another that provides the contribution of each voxel within its

corresponding region (voxel-level). The sparsity that this method entails can be very interesting in classification problems (Arco et al., 2015; Khedher et al., 2017; Plant et al., 2010), but it can also potentiate the instability of the selected regions and decrease the sensitivity in identification contexts (Baldassarre et al., 2017). For this reason, we applied a different version of MKL based on L2-norm instead of L1. In this case all regions defined by the atlas are used to build the model. Mathematically:

$$S = \sum_{i=1}^{n} (y_i - f(\boldsymbol{x}_i))^2 \quad (10)$$

Thus, the L2-norm relies on minimizing the sum of the square of the differences between the target value ($y_i$) and the estimated values ($f(\boldsymbol{x}_i)$).

In both versions of the MKL, we applied two preprocessing steps before classification: first, we applied a mean-centering to all kernels from each region of the atlas, a very common step in machine learning. This operation relies on subtracting the voxel-wise mean for each voxel across samples, which is computed on the training data to maintain the independence between the training and test subsets. Then, we normalized the kernel dividing each sample by its norm. Regions from which kernels are computed usually have different sizes, and larger regions would have a larger contribution to the model simply because of its larger size. This operation guarantees that all regions have an equal chance regardless of their sizes.

### 3.3 Searchlight

This method was introduced by Kriegeskorte et al. (2006) to identify the location of the neural activity that contains information about a given classification. It defines a sphere with a certain radius so that only the voxels inside this sphere are used to build the classification model. Performance is associated with the central voxel of the sphere. This procedure is repeated for all voxels in the brain, yielding a map of accuracies. Its main drawback is its local-multivariate nature: it extracts patterns of information from a reduced number of voxels, and this number is much smaller than the one obtained when the brain is evaluated as a whole (see Etzel et al., 2013 for additional considerations of this method).

In each sphere, we employed a support vector machine (SVM) classifier with a linear kernel due to its simplicity and the high performance reported by previous studies (Misaki et al., 2010; Pereira et al., 2009). A mathematical description of the SVM algorithm is provided in Section 3.1. We used a 12-mm radius sphere to strike a balance between sensitivity and spatial precision: smaller sizes may not detect some informative voxels whereas larger values can boost false-positives rates (Arco et al., 2016; Chen et al., 2011).

### 3.4 Performance and statistical significance

We performed a nested cross-validation to train the model and optimize the hyper-parameters of the classifier (soft-margin parameter, C), both in the ABLA and in the two MKL versions: L1-MKL and L2-MKL. In these situations, the C hyperparameter range was $[10^{-5}:10^{5}]$. Regarding Searchlight, we used a standard soft-margin parameter of C=1 for each SVM classifier due to the high performance that it provides according to previous studies

(e.g. Chanel et al., 2016; Dosenbach et al., 2010; Fan et al., 2008). The dataset comprised an fMRI experiment divided into 8 independent runs. To maintain the independence between training and testing, we used a *leave-one-run-out* cross-validation for the external loop (all methods) and the internal loop (MKL and SVM). This means that in the Searchlight approach, 7 runs were employed to train the classifier, using the remaining one for testing. In MKL and SVM, six runs were used for training, the seventh for validation and the last one for testing. We computed the balanced accuracy within participants to evaluate the performance of the model. For a binary classification, the balance accuracy is computed as the average of the accuracy obtained in the images belonging to each experimental condition individually, which increases the robustness of the performance evaluated when there is a different number of images of each class.

Statistical significance was assessed with the method proposed by Stelzer et al. (2013), with a slight difference when the procedure was applied to Searchlight or the atlas-based approaches. Unlike Searchlight, Atlas-based methods perform a whole-brain classification, obtaining a single global accuracy. The spatial information about the amount of information in each voxel is reflected on weights. Hence, whereas in Searchlight the significance was computed from accuracy maps, in the other methods weight maps were used instead (see also e.g. Haufe et al., 2014; Schrouff et al., 2018). First, the labels of the images were randomly shuffled. Then, the corresponding classification method (ABLA, MKL or Searchlight) was applied. This procedure was repeated 100 times in a within-subject classification, resulting in 100 permuted accuracy/weight maps per participant (accuracy for Searchlight and weight for the rest). A map from each individual was randomly picked following a Monte Carlo resampling with replacement (Forman et al., 1995), averaging the permuted maps and obtaining a permuted group map. This procedure was carried out 50000 times to build an empirical chance distribution. A voxel/region was considered significant if no more than 50 samples of the empirical distribution had a larger value than the one obtained without shuffling the labels, which corresponds to a cluster-defining primary-threshold of $p$=0.001 (50/50000). Once the image was thresholded, an empirical distribution of the cluster sizes of the 50000 permuted maps was built to compute the required family-wise error rate at the cluster level. After associating a $p$-value to each cluster, an FWE correction was applied ($p$=0.05) on all-cluster $p$-values to correct for multiple comparisons at the cluster level.

### 3.5 Comparison of different atlases

Following the procedure proposed by Schrouff et al. (2013a), we computed the Pearson correlation between the overlapping voxels of the weight maps obtained by the different atlases. Since ABLA organizes the weights *a posteriori* in regions from a whole-brain classification, it is only possible to compute this correlation for L1-MKL and L2-MKL. To do so, we calculated the overlap between the significant voxels obtained by each atlas, yielding a value ranging from 0 to 1. We employed permutation tests to assess the significance of the correlation coefficients using a similar framework as described in Section 3.4.

## 4 Results

In this section, we report the results obtained by the three approaches evaluated in this study: Atlas-based local averaging (ABLA), and the two versions of Multiple Kernel Learning (L1-MKL and L2-MKL). We compared the weight maps of these three methods with the accuracies map obtained by Searchlight by computing the overlap between significant voxels. Moreover, for L1 and L2-MKL we show the stability of the selected regions across atlases by computing a correlation between their overlapping-significant weight maps, using permutation tests to assess the significance of these correlations. We did not compute this correlation for the ABLA method because weights are exactly the same for all atlases. Additionally, we include the results obtained by these methods in two classification contexts (*decision* and *valence*) that lead to large or subtle differences between the conditions contrasted, to test the generalizability of the results of the different approaches.

### 4.1 Influence of the classification methods

We first focus on comparing the results obtained by ABLA, L1-MKL and L2-MKL in the *decision* classification. Table 1 summarizes these results in terms of accuracy and overlap between the significant regions identified by each method and those obtained by Searchlight (SL). The accuracies discussed in this section correspond to the ones obtained in the maximum overlap scenario, which does not mean that these accuracies were the absolute maximum itself. We further discuss the implications of this finding in Section 5. The first approach, ABLA, yielded a maximum overlap of 70.58%, and a corresponding accuracy of 81.51%. L1-MKL led to the same maximum overlap value, 70.58%, but a higher corresponding accuracy compared to ABLA: 85.02%. On the other hand, L2-MKL obtained a maximum overlap of 77.93%, whereas the accuracy was 70.65% after employing this approach.

We will now describe the results obtained in the *valence* classification. In this context, the ABLA method obtained a maximum overlap of 41.49%, with a corresponding accuracy of 51.77%. We assessed the significance of the accuracy by employing the non-parametric method described in Section 3.4. This last value is considerably lower than the one obtained in the *decision* classification and it likely reflects the subtle differences in the neural activity associated with the valence of a word. We observed that after applying the L1-MKL method, only one of the nine atlases employed led to a significant region that overlapped with Searchlight. However, the small size of this region (only 0.14% of the significant voxels obtained by L1-MKL overlapped with Searchlight results) highlights the inadequacy of this method to identify significant regions in a context like the *valence* one. With reference to L2-MKL, the maximum overlap slightly increased (3.81%), with a corresponding accuracy of 49.14%. Since the value of the overlap is considerably small as well, conclusions derived for L1-MKL results can be also applied to L2-MKL. All the results obtained by the three classification methods in the *valence* classification are summarized in Table 2.

### 4.2 Influence of the atlases

Previous section has depicted the influence of the different classification methods in the results. Moreover, this section will describe the effect of the different brain parcellations in the identification of informative brain regions for the three approaches used. In the first context (the *decision* classification), ABLA marked as informative similar regions regardless of the atlas employed. This similarity can be shown in Figure 1, where the spatial distribution of the significant regions for two different atlases is reported. Despite most of atlases led to the same results, we found a variability in terms of overlap between the different atlases. Specifically, the largest overlap score with Searchlight was obtained by the Harvard-Oxford atlas (70.58%), whereas the minimum value was derived from the Camb36 division of the brain (21.36%). Results for the rest of the atlases are summarized in Table 1.

When applying L1-MKL, the largest overlap value was obtained by the same atlas as with ABLA: the Harvard-Oxford, with a 70.58%. However, the minimum overlap corresponded to Schaefer atlas (14.5%). It seems that his method is more affected than ABLA by the different brain parcellations. As Figure 1 shows, the distribution of the significant regions is similar across atlases, but in this case, sensitivity is lower than ABLA for most atlases. Quantitative results for each brain parcellation are summarized in Table 1. For the last classification method used, L2-MKL, the parcellation derived from the Camb64 atlas yielded the largest accuracy and minimum overlap score (74.74% and 21.36%, respectively). This finding remarks that maximum overlap and accuracy is not usually simultaneously obtained. On the other hand, Harvard-Oxford led to a good accuracy value and the largest overlap (70.65% and 77.93%, respectively). Figure 1 shows how informative regions vary for two different atlases, whereas Table 1 includes numerical results about accuracy, overlap and number of significant voxels for all brain parcellations.

Regarding the *valence* contrast, results were highly affected by the atlas used. We found a large consistency in the significant regions obtained by ABLA and Searchlight when the Cambridge12 atlas was employed. However, this atlas was not the only one that led to a good performance. Specifically, the brain parcellations provided by AICHA and Harvard-Oxford also identified informative regions similar to those obtained by Searchlight. Most importantly, these regions contain areas that have been reported by previous research (e.g. ventromedial prefrontal cortex, Lindquist and Mejia, 2015), which supports the reliability of the results. Figure 2 includes the significant results associated with different atlases and the ABLA approach in this *valence* classification, whereas Table 2 provides additional information to these results.

Unlike ABLA, the two method based on MKL hardly detected reliable information for all the atlases employed. With reference to L1-MKL, each brain parcellation led to a completely different distribution of the informative voxels. However, none of the nine atlases that we employed yielded an accuracy that surpassed the chance level, so that the subsequent model did not provide useful information about where the information regions were located. Results were very similar when L2-MKL was employed. In this case, models derived from some atlases surpassed the chance level, but they were not able to identify the regions that

truly contained information. Figure 2 shows the informative voxels that both L1-MKL and L2-MKL yielded, respectively, whereas Table 2 summarizes the results obtained by these two approaches.

### 4.3 Stability of the weights across atlases

We compared the weight maps across the different atlases for L1-MKL and L2-MKL, in the two classification contexts. Due to the nature of the ABLA method, weights are the same for all the atlases since different brain parcellations are used once the model is built. For this reason, we only compute the stability of the subsequent weights for the two approaches based on MKL. In the *decision* classification, the correlation values obtained by the first 6 atlases (Camb12, Camb20, Camb36, Camb64, AICHA and Yeo2011) ranging from 0.882 to 0.975 when the L1-MKL was used. The weight maps derived from the Harvard-Oxford atlas also yielded a large similarity to these 6 atlases, but this correlation decreased when the Brainnetome atlas was employed. By contrast, the Schaefer atlas led to very different weights compared to any of the other atlases. These results suggest that, for this contrast, the decision function derived from L1-MKL is based on the same voxels. Moreover, the contribution of these voxels to the classifier decision is stable for all brain parcellations proposed by each atlas. Table 3 summarizes the correlation values obtained between each pair of atlases.

Regarding L2-MKL, it yielded very similar weight maps regardless of the atlases used. It is worth noting the large correlation between each pair of atlases (see Table 4), even with the Schaefer atlas that yielded very different weights when L1-MKL was applied. Only maps provided by Yeo2011 and Brainnetome are slightly less similar to those obtained by the four atlases, whereas both show a large correlation with the others. The rest of the atlases present correlation values close to 1. Different atlases led to very similar results, highlighting the robustness of L2-MKL in the identification of informative regions. Moreover, this finding shows the low influence that the brain parcellation has in the results, which validates the use of these atlas-based methods even without a prior hypothesis about the brain organization in a specific process.

For the *valence* classification, the localization of the informative regions was so variable that results derived from most atlases did not overlap the ones obtained by the others. For this reason, we could compute the correlation between AICHA, Harvard- Oxford and Brainnetome for L1-MKL because they were the only atlases that shared some informative voxels, yielding a maximum overlap of 0.428 (see Table 5). Results obtained by L2-MKL also showed a reduced overlap between the weight maps and we could only correlate the significant results of AICHA, Yeo2011 and Schaefer atlases. In this case, the maximum correlation was obtained by Yeo2011 and Schaefer, yielding a value of 0.99 (see Table 6). Nevertheless, this value was obtained from a small region since significant results provided by these two atlases were considerably different.

### 4.4 Directionality of the weights

In the *decision* classification, we evaluated not only the localization of the informative weights but their sign. Due to the nature of the contrast, it was expected that weights were

organized according to their sign in a specific hemisphere for each group of subjects. Figure 3 shows the distribution of the significant voxels depending on the sign of their weights for the ABLA method, comparing them with results obtained by univariate analysis. It is remarkable that participants who accepted the offer with the right hand and rejected it with the left hand (odd group) show a cluster of positive weights in the left hemisphere and a cluster of negative weights in the right hemisphere. On the other hand, these results are shifted when results from even participants were evaluated: the right hemisphere contains weights associated with accepting an offer, whereas the left hemisphere shows the negative weights. These results are consistent with those obtained by univariate analysis. Specifically, results from the odd group are quite similar from the Acceptance>Reject contrast, and results from the even group have a lot of similarities with the Acceptance<Reject contrast.

Results for both MKL methods are very similar to the ones obtained by ABLA. Regardless of the differences in the spatial localization of the information (already commented in previous sections), the weights follow the same distribution as ABLA. For those participants that accepted the offer by employing their right hand, the weights in the right hemisphere are positive, whereas the same hemisphere in the group of people that used their left hand to accept the offer contains negative weights. The signs of the significant voxels for the L1-MKL and L2-MKL methods are exhibited in Figures 3.

## 5 Discussion

In this study, we evaluated atlas-based methods, alternative to Searchlight, to localize the informative regions involved in cognitive functions. We extracted the weight maps from three atlas-based classification approaches (ABLA, L1-MKL and L2-MKL) and evaluated the statistical significance of each region. We used these methods in two different contexts. In the first one, where the two classes generated large differences in the observed pattern, L2-regularization resulted the best option for identification purposes. Moreover, atlas-based approaches showed a large stability in the informative regions found regardless of the atlas employed, which highlights the adequacy of these methods. In contrast, when the differences in the activity associated with each class were much subtler, only the ABLA approach showed certain stability in the informative regions across the atlases. However, both L1-MKL and L2-MKL were highly affected by the specific brain organization reflected in the atlases.

### 5.1 Influence of the classification methods

We have found that maximum accuracy and overlap do not usually concur, especially when detecting subtle differences in the observed pattern. In the *decision* classification, we found differences across the methods in terms of overlap and accuracy. L1-MKL usually obtained a larger accuracy than ABLA and L2-MKL for the different atlases, but a lower spatial overlap with Searchlight results. According to these results, we can separate the different approaches in two groups: on the one hand, ABLA and L2-MKL; on the other, L1-MKL. The reason for this difference is the regularization used by each method: while ABLA and L2-MKL use an L2-norm regularization, L1-MKL employs an L1. L1-norm provides sparse solutions since it only selects a subset of regions that contain predictive information,

while the rest are automatically driven to zero. This can be helpful from the classification standpoint since it leads to larger accuracies: when a lower number of features are considered, the dimensionality of the data is reduced, which facilitates finding the optimal solution to the classification problem. However, our results show that the model with the largest overlap is not usually the most accurate. This is consistent with previous studies, e.g. the extreme case reported by Sona et al. (2007). They proposed a model for decoding subjective perception of participants from their neural data while viewing movie segments. They found a framework that yielded a large performance, but the regions that guided the classifier were partially contained in the ventricles and other regions with large physiological noise. This means that their algorithm performed consistently well in the classification task, but it did not provide any useful information for a better understanding of the human brain. Our results support the need of clearly separating the use of multivariate decoding for prediction and for identification (Hebart and Baker, 2017) in addition to highlight the importance of selecting the methods that best fit the desired aim.

In the *valence* classification, we also found differences across the methods in terms of overlap and accuracy, but in this case these differences were even larger. ABLA was the only method that obtained some overlap with the voxels marked as informative by Searchlight, whereas L1-MKL and L2-MKL hardly detected those significant regions. The key of this finding is the classification problem itself. Evaluating whether a participant responded with the right/left hand to a stimulus generates large differences in the observed pattern and it is easy for a classifier to find a hyperplane that maximizes the separation of the two classes. On the other hand, isolating regions with a differential involvement in valence processing is much harder, as shown by recent metanalytic approaches (Lindquist et al., 2015). Moreover, previous research has suggested that decoding accuracies from the prefrontal cortex (PFC) are lower than in other regions like the visual sensory cortex. This could reflect an important role of neural coding in the PFC. In fact, most studies evaluated in Bhandari et al., (2018) yielded a just above-chance accuracy (see Figure 1 in Bhandari et al., 2018), similar to the accuracy obtained by ABLA in the *valence* classification (51.77%). It is worth noting that permutation tests were employed to assess significance. We set a strict value for the cluster-defining primary-threshold ($p$=0.001) and then performed an FWE-correction to the $p$-values associated to each cluster to account for the multiple comparisons problem. In addition, the low value of the accuracy can also be influenced by the way information is represented in the brain. As an example, Dubois et al., (2015) could not retrieve information about face identity employing fMRI despite single-unit recordings demonstrated that this information was represented in the underlying neuronal populations. Hence, the level of accuracy of the classification in decoding fMRI data depends in part on how the underlying neural activity is organized. The highest accuracy possible given a specific contrast in a certain brain region can be far from maximal, and thus a low but statistically significant value can still be theoretically relevant (Hebart and Baker., 2017; Bhandari et al., 2018). Regarding the large difference in sensitivity obtained by the two approaches, the number of significant voxels depends on the brain regions involved in the cognitive function under study. In the *decision* classification, the number of regions (mainly motor-related), as well as their sizes, is large. On the other hand, the informative regions identified in the *valence* classification are related

to valence-dependent behaviour (vmPFC; Kuzmanovic et al., 2018), which has a considerably smaller size and can hardly be properly isolated in an atlas.

Our results show that ABLA provides a larger overlap than all the other methods in the two classification problems, especially in the *valence* one. In this case, most MKL results lead to a below-chance accuracy in addition to a poor or null overlap with Searchlight results. This discrepancy must be due to the different framework of ABLA. Both L1-MKL and L2-MKL consider the regions provided by the atlas to build the model as part of the learning process. Hence, if the parcellations derived from the atlas do not properly delimitate the different regions involved in the context under study, the resulting model would be suboptimal since it is based on non-valid assumptions about effective brain parcellations. On the other hand, ABLA builds the classification model from a whole-brain approach, which means that the atlas parcellations do not have any influence in the learning process. Instead, the brain organization is incorporated after building the model to summarize how informative each region is. For this reason, ABLA leads to a better performance in conditions of subtle or small differences between the conditions, although it is supposed to have a lower ability to detect informative regions compared with methods based on MKL when the atlas leads to a realistic approximation of the brain subdivisions.

### 5.2 Influence of the atlases

Results show that the specific brain parcellations of each atlas impact the spatial accuracy of the different methods only when differences in the observed pattern are small, but not when these are large. In the *decision* classification, there was a large consistency among the significant regions obtained by all methods across the different atlases. These results carry important implications. Atlas-based approaches are assumed to have a large dependency on the way brain parcellations are computed. Our results evidence that atlases can be of help to identify informative regions even when the brain parcellations provided by the atlases do not perfectly delimitate the regions involved in the cognitive function under study. However, according to the results obtained in the *valence* classification, there are other contexts where these atlases are not accurate enough to guarantee a good performance in the identification of the sources of information. This is probably related to the size and the specific shape of the region involved in a certain cognitive function, such as the aforementioned ventromedial prefrontal cortex (vmPFC) associated with the *valence* classification. The only region that ABLA marked as significant in the Camb12 parcellation is the one that contains the vmPFC, which shows that this method was able to identify where the information was located. Nevertheless, this region has a massive size in this atlas, and these atlas-based methods consider each region as a whole, and thus a large number of voxels are marked as significant only because they are in the same region as the one that is really informative. However, using atlases with more subdivisions implies that regions are much smaller. Here, the parcellation proposed by the atlas is a good match to the spatial organization of a small structure such as the vmPFC, leading to a higher sensitivity and spatial accuracy.

The number of subdivisions of the atlases also influenced the performance of the three algorithms employed. In the *decision* classification, the optimum value in terms of overlap

was obtained by the 36 regions that the Camb36 atlas is divided into. We hypothesize that the number of regions is also important to obtain these results. Using an atlas with few subdivisions means that it is more likely to find an informative region, despite the small ratio between the voxels that are really significant and the total number of voxels that comprise the region. Instead, a large number of parcellations means that the classifier has to be much subtler in the identification of informative regions. The parcellations derived from Schaeffer add larger precision and subdivisions to the brain networks published by the Yeo2011 atlas. However, results show a better performance in terms of sensitivity when the simplest approach was used. This strongly indicates that using atlases that do not reflect an accurate representation of the brain organization is similar to choosing a large Searchlight sphere where only a few voxels within this sphere are informative (Etzel et al., 2013). Using a large radius increases the probability of marking as significant voxels that are not, increasing the false-positives rate.

### 5.3 Stability of the weights across atlases

We have found a large correlation between the significant weight maps obtained by different atlases in the *decision* classification. The magnitude of the weight of a specific voxel quantifies the contribution of this voxel to the model, and its sign informs us about its relationship to the *Accept* or to the *Reject* class. For the L1-MKL approach, we found large correlation values for all atlases except for the Schaefer one. Hence, the resulting weights associated with each model are very similar, which highlights the stability of the classification methods regardless of the atlas used. Interestingly, we found the largest correlations in the weight maps obtained by the four Cambridge atlases, which are all derived from the same clustering algorithm (BASC). This result supports the idea that the mathematical framework employed to delimitate the different brain regions is important since it can influence the success of the subsequent analyses. On the other hand, the poor performance of L1-MKL when the Schaefer atlas is used can be due to the conjunction of a sparse method and an atlas with a large number of regions, as mentioned in the previous section. It is important to note that our results do not invalidate the use of ambitious atlases aiming at obtaining a detailed parcellation of each cortical region. However, if these parcellations are not informative enough about how information is represented in the brain (unless an atlas is computed separately for each participant from his/her neural data: Gordon et al., 2017; Schaefer et al., 2018; Wang et al., 2015), sparse solutions are not recommended. Unlike L1-MKL, L2-MKL obtained a large correlation score between each pair of atlases (Table 4). Thus, L2-MKL adapts to different idiosyncrasies and leads to a common solution for different brain mappings. This means that the weight maps that guide the classification are essentially the same regardless of the atlas used so that it is possible to successfully employ this approach even without a clear hypothesis about the brain organization in a specific context.

Nevertheless, these conclusions are only valid when there are large differences in the observed pattern associated with the two classes to distinguish from. Our findings in the *valence* classification differ substantially from those obtained in the *decision* classification. L1-MKL results (summarized in Table 5) show that we could hardly compute the correlation

between two pairs of atlases: the first one, AICHA and Harvard-Oxford; the second, AICHA and BN. In addition, none of the significant results provided by these atlases share any voxel with the Searchlight results, which illustrates that weight maps are similar from a mathematical perspective, but make a null contribution to the neuroscience standpoint. Results obtained by L2-MKL are summarized in Table 6 and conclusions derived from them are essentially the same than L1-MKL. We could only compute the correlation between two pairs of atlases: Schaefer-AICHA and Schaefer-Yeo2011. From these three atlases, Schaefer is the one that leads to some overlap with the Searchlight approach: 3.81%. However, none of these significant voxels are shared by AICHA and Yeo2011. This reflects that the two versions of MKL are not able to identify small informative regions in contexts where differences in the observed pattern of the two conditions are minimum.

### 5.4 Directionality of the weights

One of the main advantages of using weights instead of accuracy is the directionality that they provide. We have evaluated the sign of each weight within the significant regions for each of the three atlas-based methods for the *decision* classification, where it is easy to evaluate whether the sign of the weight matches the psychological prediction. In this experiment, participants used one hand to accept an offer and the other one to reject it. The decoding analysis should mark as informative motor-related since this differentiates the conditions. However, it is worth remembering that the hand employed was counterbalanced across participants. We obtained exactly the expected results: the three approaches led to a map in which weights were organized according to their sign. For odd participants, regions associated with the acceptance of an offer (use of the right hand) were localized in the left hemisphere, with a positive sign. On the other hand, regions that contained information when the offer was rejected (left hand) were found in the right hemisphere, with a negative weight. More importantly, the informative regions for even participants shifted: positive weights were found in the right hemisphere, whereas weights with a negative sign were found in the left hemisphere. These results are very similar to those obtained by the univariate approach (see Figure 3): regions with a larger activation when participants accept/reject an offer match the sign of the weights of the different multivariate methods. However, atlas-based approaches use normalized data, which eliminates the differences in the global activation levels associated to each condition. Thus, these methods identify areas that show a different spatial distribution of the information, while the univariate approach purely relies on differences in the activation level. Despite this similarity, it is worth noting that multivariate approaches show a large sensitivity compared to a univariate one. We have employed a classification context where differences at the cluster level can also be picked up by univariate methods to highlight one of the main advantages of atlas-based approaches. Similarly to Searchlight, these techniques are able to extract information from fine-grained differential activation patterns, which results in a boost in sensitivity compared to univariate analysis. Figure 3 reveals the differences in sensitivity between multivariate and univariate methods in the *decision* context, whereas differences between these two approaches in the *valence* classification can be found in previous studies (see Figure 10 in Arco et al., 2018 and Figure 1 in Lindquist et al., 2015).

## 6 Conclusion

In this study, we compared three different atlas-based approaches to Searchlight to assess their ability to identify informative brain regions for cognitive contrasts that generate either large or small differences in the observed pattern. We have shown for the first time that these methods can be used as an alternative to Searchlight since they localize informative regions when there are large differences between the observed pattern associated with the two classes to distinguish from. In this case, results are consistent across atlases. Moreover, the use of weight maps provides additional information to accuracies, combining the sensitivity of decoding analyses and the directionality of univariate results. However, results change drastically when the differential observed pattern is much lower. Methods based on MKL are highly affected by the discrepancy of the shape of brain regions containing information and the one proposed by the atlases. On the other hand, ABLA is the only approach that identifies informative regions in accordance with previous research. Our results pave the way for finding a method that leads to a large spatial accuracy in the identification of subtle changes of the observed pattern. Future studies are needed to widen the findings of this study by evaluating the performance of these methods when the brain parcellations are specifically computed for each participant, which may substantially improve the neuroanatomical functional precision. This combination might boost their sensitivity and widen their adequacy in different contexts, especially when an accurate parcellation is crucial.

## 7 Funding

This work was supported by the Spanish Ministry of Science and Innovation through grant PSI2016-78236-P to M.R and the Spanish Ministry of Economy and Competitiveness through grant BES-2014-069609 to J.E.A. This research is part of J.E.A's activities for the PhD Program in Information and Communication Technologies of the University of Granada.

## 8 Acknowledgments

We are grateful to Janaina Mourão-Miranda for her kind help during the development of the algorithms employed in the current research.

## 9 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Table 1:** Summary of the results obtained for the different methods and atlases in the *decision* classification.

| Method | Atlas | Accuracy (%) | Significant voxels | SL voxels defined in atlas | Overlap with SL (%) | Regions | Significant regions |
|---|---|---|---|---|---|---|---|
| ABLA | Camb12 | 81.51 | 4704 | 4302 | 61.69 | 12 | 1 |
| L1-MKL | Camb12 | 86.2 | 4704 | 4302 | 61.69 | 12 | 1 |
| L2-MKL | Camb12 | 72.43 | 2654 | 4302 | 61.69 | 12 | 1 |
| ABLA | Camb20 | 81.51 | 3692 | 4302 | 58 | 20 | 1 |
| L1-MKL | Camb20 | 85.02 | 3692 | 4302 | 58 | 20 | 1 |
| L2-MKL | Camb20 | 65.21 | 2598 | 4302 | 58 | 20 | 1 |
| ABLA | Camb36 | 81.51 | 982 | 4302 | 21.36 | 36 | 1 |
| L1-MKL | Camb36 | 89.37 | 982 | 4302 | 21.36 | 36 | 1 |
| L2-MKL | Camb36 | 74.74 | 1000 | 4302 | 21.36 | 36 | 1 |
| ABLA | Camb64 | 81.51 | 3740 | 4302 | 63.34 | 64 | 1 |
| L1-MKL | Camb64 | 84.62 | 982 | 4302 | 21.36 | 64 | 1 |
| L2-MKL | Camb64 | 70.31 | 7613 | 4302 | 21.36 | 64 | 1 |
| ABLA | AICHA | 81.51 | 1802 | 3291 | 48 | 192 | 5 |
| L1-MKL | AICHA | 86.76 | 636 | 3291 | 19.23 | 192 | 1 |
| L2-MKL | AICHA | 69.53 | 2867 | 3291 | 60.13 | 192 | 11 |
| ABLA | Yeo2011 | 81.51 | 2731 | 3137 | 56.39 | 17 | 1 |
| L1-MKL | Yeo2011 | 87.34 | 2731 | 3137 | 56.07 | 17 | 1 |
| L2-MKL | Yeo2011 | 71.35 | 2731 | 3137 | 56.39 | 17 | 1 |
| ABLA | Harvard-Oxford | 81.51 | 4609 | 3389 | 70.58 | 48 | 2 |
| L1-MKL | Harvard-Oxford | 85.02 | 4609 | 3389 | 70.58 | 48 | 2 |
| L2-MKL | Harvard-Oxford | 70.65 | 6531 | 3389 | 77.93 | 48 | 5 |
| ABLA | Brainnetome | 81.51 | 2051 | 3057 | 47 | 246 | 9 |
| L1-MKL | Brainnetome | 78.77 | 904 | 3057 | 21.56 | 246 | 4 |
| L2-MKL | Brainnetome | 66.53 | 1129 | 3057 | 28.39 | 246 | 5 |
| ABLA | Schaefer | 81.51 | 1558 | 3137 | 42.9 | 400 | 23 |
| L1-MKL | Schaefer | 77.84 | 465 | 3137 | 14.5 | 400 | 6 |
| L2-MKL | Schaefer | 71.61 | 1926 | 3137 | 51.35 | 400 | 28 |

**Table 2:** Summary of the results obtained for the different methods and atlases in the *valence* classification.

| Method | Atlas | Accuracy (%) | Significant voxels | SL voxels defined in atlas | Overlap with SL (%) | Regions | Significant regions |
|---|---|---|---|---|---|---|---|
| ABLA | Camb12 | 51.77 | 2095 | 911 | 41.49 | 12 | 1 |
| L1-MKL | Camb12 | 48.44 | 0 | 911 | 0 | 12 | 0 |
| L2-MKL | Camb12 | 50.71 | 0 | 911 | 0 | 12 | 0 |
| ABLA | Camb20 | 51.77 | 0 | 911 | 0 | 20 | 0 |
| L1-MKL | Camb20 | 48.18 | 0 | 911 | 0 | 20 | 0 |
| L2-MKL | Camb20 | 50.74 | 0 | 911 | 0 | 20 | 0 |
| ABLA | Camb36 | 51.77 | 0 | 911 | 0 | 36 | 0 |
| L1-MKL | Camb36 | 49.74 | 0 | 911 | 0 | 36 | 0 |
| L2-MKL | Camb36 | 49.22 | 406 | 911 | 0 | 36 | 1 |
| ABLA | Camb64 | 51.77 | 341 | 911 | 7.14 | 64 | 1 |
| L1-MKL | Camb64 | 46.88 | 549 | 911 | 0 | 64 | 1 |
| L2-MKL | Camb64 | 51.75 | 0 | 911 | 0 | 64 | 0 |
| ABLA | AICHA | 51.77 | 663 | 729 | 20.58 | 192 | 5 |
| L1-MKL | AICHA | 47.1 | 780 | 729 | 0 | 192 | 5 |
| L2-MKL | AICHA | 52.83 | 35 | 729 | 0 | 192 | 1 |
| ABLA | Yeo2011 | 51.77 | 0 | 709 | 0 | 17 | 0 |
| L1-MKL | Yeo2011 | 46.15 | 0 | 709 | 0 | 17 | 0 |
| L2-MKL | Yeo2011 | 50.97 | 1010 | 709 | 1.7 | 17 | 1 |
| ABLA | Harvard-Oxford | 51.77 | 439 | 715 | 21.4 | 48 | 1 |
| L1-MKL | Harvard-Oxford | 47.18 | 145 | 715 | 0 | 48 | 1 |
| L2-MKL | Harvard-Oxford | 48.33 | 0 | 715 | 0 | 48 | 0 |
| ABLA | Brainnetome | 51.77 | 438 | 738 | 7.99 | 246 | 3 |
| L1-MKL | Brainnetome | 43.34 | 349 | 738 | 0 | 246 | 3 |
| L2-MKL | Brainnetome | 51.19 | 137 | 738 | 0 | 246 | 1 |
| ABLA | Schaefer | 51.77 | 61 | 708 | 4.8 | 400 | 1 |
| L1-MKL | Schaefer | 46.24 | 123 | 708 | 0.14 | 400 | 2 |
| L2-MKL | Schaefer | 49.14 | 302 | 708 | 3.81 | 400 | 6 |

**Table 3:** Correlation between the significant weight maps across the different atlases after applying the L1-MKL method in the *decision* classification.

| L1-Multi Kernel Learning | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Atlas** | **Camb 12** | **Camb 20** | **Camb 36** | **Camb 64** | **AICHA** | **Yeo2011** | **Harvard-Oxford** | **Brainnetome** | **Schaefer** |
| **Camb12** | 1 | 0.974 | 0.906 | 0.936 | 0.889 | 0.937 | 0.539 | 0.383 | 0.144 |
| **Camb20** | 0.974 | 1 | 0.908 | 0.951 | 0.934 | 0.947 | 0.564 | 0.552 | 0.125 |
| **Camb36** | 0.906 | 0.908 | 1 | 0.975 | 0.933 | 0.911 | 0.497 | 0.568 | 0.1 |
| **Camb64** | 0.936 | 0.951 | 0.975 | 1 | 0.963 | 0.933 | 0.542 | 0.573 | 0.081 |
| **AICHA** | 0.889 | 0.934 | 0.933 | 0.963 | 1 | 0.882 | 0.61 | 0.549 | 0.088 |
| **Yeo2011** | 0.937 | 0.947 | 0.911 | 0.933 | 0.882 | 1 | 0.566 | 0.528 | 0.193 |
| **Harvard-Oxford** | 0.539 | 0.564 | 0.497 | 0.542 | 0.61 | 0.566 | 1 | 0.172 | 0.051 |
| **Brainnetome** | 0.383 | 0.552 | 0.568 | 0.573 | 0.549 | 0.528 | 0.172 | 1 | 0.109 |
| **Schaefer** | 0.144 | 0.125 | 0.1 | 0.081 | 0.088 | 0.193 | 0.051 | 0.109 | 1 |

**Table 4:** Correlation between the different atlases after applying the L2-MKL method in the *decision* classification.

| L2-Multi Kernel Learning | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Atlas** | **Camb 12** | **Camb 20** | **Camb 36** | **Camb 64** | **AICHA** | **Yeo2011** | **Harvard-Oxford** | **Brainnetome** | **Schaefer** |
| **Camb12** | 1 | 0.927 | 0.94 | 0.926 | 0.845 | 0.768 | 0.837 | 0.72 | 0.829 |
| **Camb20** | 0.927 | 1 | 0.958 | 0.973 | 0.776 | 0.71 | 0.795 | 0.64 | 0.762 |
| **Camb36** | 0.94 | 0.958 | 1 | 0.981 | 0.795 | 0.721 | 0.789 | 0.663 | 0.776 |
| **Camb64** | 0.926 | 0.973 | 0.981 | 1 | 0.805 | 0.738 | 0.798 | 0.67 | 0.783 |
| **AICHA** | 0.845 | 0.776 | 0.795 | 0.805 | 1 | 0.948 | 0.94 | 0.904 | 0.973 |
| **Yeo2011** | 0.768 | 0.71 | 0.721 | 0.738 | 0.948 | 1 | 0.897 | 0.857 | 0.945 |
| **Harvard-Oxford** | 0.837 | 0.795 | 0.789 | 0.798 | 0.94 | 0.897 | 1 | 0.849 | 0.945 |
| **Brainnetome** | 0.72 | 0.64 | 0.663 | 0.67 | 0.904 | 0.857 | 0.849 | 1 | 0.889 |
| **Schaefer** | 0.829 | 0.762 | 0.776 | 0.783 | 0.973 | 0.945 | 0.945 | 0.889 | 1 |

**Table 5:** Correlation between the significant weight maps across the different atlases after applying the L1-MKL method in the *valence* classification.

| L1-MKL | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Atlas** | **Camb 12** | **Camb 20** | **Camb 36** | **Camb 64** | **AICHA** | **Yeo2011** | **Harvard-Oxford** | **Brainnetome** | **Schaefer** |
| **Camb12** | 1 | - | - | - | - | - | - | - | - |
| **Camb20** | - | 1 | - | - | - | - | - | - | - |
| **Camb36** | - | - | 1 | - | - | - | - | - | - |
| **Camb64** | - | - | - | 1 | - | - | - | - | - |
| **AICHA** | - | - | - | - | 1 | - | 0.217 | 0.428 | - |
| **Yeo2011** | - | - | - | - | - | 1 | - | - | - |
| **Harvard-Oxford** | - | - | - | - | 0.217 | - | 1 | - | - |
| **Brainnetome** | - | - | - | - | 0.428 | - | - | 1 | - |
| **Schaefer** | - | - | - | - | - | - | - | - | 1 |

**Table 6:** Correlation between the significant weight maps across the different atlases after applying the L2-MKL method in the *valence* classification.

| L2-MKL | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Atlas** | **Camb 12** | **Camb 20** | **Camb 36** | **Camb 64** | **AICHA** | **Yeo2011** | **Harvard-Oxford** | **Brainnetome** | **Schaefer** |
| **Camb12** | 1 | - | - | - | - | - | - | - | - |
| **Camb20** | - | 1 | - | - | - | - | - | - | - |
| **Camb36** | - | - | 1 | - | - | - | - | - | - |
| **Camb64** | - | - | - | 1 | - | - | - | - | - |
| **AICHA** | - | - | - | - | 1 | - | - | - | 0.975 |
| **Yeo2011** | - | - | - | - | - | 1 | - | - | 0.99 |
| **Harvard-Oxford** | - | - | - | - | - | - | 1 | - | - |
| **Brainnetome** | - | - | - | - | - | - | - | 1 | - |
| **Schaefer** | - | - | - | - | 0.975 | 0.99 | - | - | 1 |

Figure captions:

**Fig 1** Significant voxels obtained by the Searchlight approach and the different classification methods for all the Cambridge12 and Harvard-Oxford atlases in the *decision* classification.

**Fig 2** Significant voxels obtained by the Searchlight approach and the different classification methods for all the Cambridge12 and Harvard-Oxford atlases in the *valence* classification.

**Fig 3** Summary of the results obtained for the *decision* classification by Searchlight, ABLA, L1-MKL, L2-MKL and univariate approaches. The three classification methods show large differences between the two groups considered (odd/even participants). Searchlight only provides information about the significance of each voxel itself, so that no separation between groups was considered.

**References:**


Abdulrahman, H., Henson, R.N., 2016. Effect of trial-to-trial variability on optimal event-related fMRI design: Implications for Beta-series correlation and multi-voxel pattern analysis. NeuroImage 125, 756-766.

Adeli, E., Guorong, W., Saghafi, B., An, L., Shi, F., Shen, D., 2017. Kernel-based joint feature selection and max-margin classification for early diagnosis of Parkinson's disease. Sci. Rep. 7, 41069.

Arco, J.E., González-García, C., Díaz-Gutiérrez, P., Ramírez, J., Ruz, M., (2018). Influence of activation pattern estimates and statistical significance tests in fMRI decoding analysis. Journal of Neuroscience Methods 308, 248-260.

Arco, J.E., González-García, C., Ramírez, J., Ruz, M., 2016. Comparison of different methods for brain decoding from fMRI beta maps. Poster presented at 22nd Annual Meeting of the Organization for Human Brain Mapping, Geneve, (Switzerland).

Arco, J.E., Ramírez, J., Puntonet, C.G., Górriz, J.M., Ruz, M., 2015. Short-term prediction of MCI to AD conversion based on longitudinal MRI analysis and neuropsychological tests. Innovation in Medicine Healthcare, 385-394.

Baldassarre L., Pontil, M., Mourão-Miranda, J., 2017. Combining Accuracy and stability for model selection in brain decoding. Frontier in Neuroscience 11:62.

Balci, S.K., Sabuncu, M.R., Yoo, J., Ghosh, S.S., Whitfield-Gabrieli, S., Gabrieli, J.D., Golland, P., 2008. Prediction of successful memory encoding from fMRI data. Med Image Comput Assist Inter. 11, 97-104.

Bellec, P., Rosa-Neto, P., Lyttelton, O.C., Benali, H., Evans, A.C., 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. NeuroImage 51, 1126-1139.

Bennett, K.P., Blue, J.A., 1998. A support vector machine approach to decision trees. 1998 IEEE International Joint Conference in Neural Networks Proceedings.

Bhandari, A., Gagne, C., Badre, D., 2018. Just above Chance: Is It Harder to Decode Information from Prefrontal Cortex Hemodynamic Activity Patterns? J Cogn Neurosci 30(10), 1473-1498.

Blankertz, B., Dornhege, G., Kraudelat, M., Müller, K.R., Curio, G., 2007. The non-invasive Berlin Brain-Computer Interface: fast acquisition of effective performance in untrained subjects. Neuroimage 37(2), 539-50.

Bode, S., Haynes, J.-D., 2009. Decoding sequential stages of task preparation in the human brain. Neuroimage 45 (2), 606-613.

Boser, B.E., Guyon, I., Vapnik, V., 1992. A training algorithm for optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 144-152.

Burges, C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121-167.

Chanel, G., Pichon, S., Conty, L., Berthoz, S., Chevallier, C., Grèzes, J., 2016. Classification of autistic individuals and controls using cross-task characterization of fMRI activity. Neuroimage: Clinical 10, 76-88.

Chang, L.J., Sanfey, A.G., 2013. Great expectations: neural computations underlying the use of social norms in decision-making. Social Cognitive and Affective Neuroscience 8(3), 277-284.

Chen, Y., Namburi, P., Elliott, L., Heinzle, J., Soon, C., Chee, M., Haynes, J., 2011. Cortical surface-based searchlight decoding. NeuroImage 56, 582-592.

Chen J., Leong, Y.C., Honey, C.J., Yong, C.H., Norman, K.A., Hasson, U., 2017. Shared memories reveal shared structure in neural activity across individuals. Nature neuroscience 20(1), 115-125.

Choi, H., Ha, S., Im, H. J., Paek, S. H., & Lee, D. S., (2017). Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. NeuroImage: Clinical 16, 586–594.

Cichy, R.M., Pantazis, D., Oliva, A., 2016. Similarity-based fusion of MEG and fMRI reveals spatio-temporal dynamics in human cortex during visual object recognition. Cerebral Cortex 26(8), 3563–3579.

Coutanche, M.N., Thompson-Schill, S.L., Schultz, R.T., 2011. Multi-voxel pattern analysis of MRI data predicts clinical symptom severity. NeuroImage 57(1), 113-123.

Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. NeuroImage 19, 261-270.

Craddock, R.C., James, G.A., Holtzheimer III, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp. 33(8), 1914-28.

Dai, D., Wang, J., Hua, J., He, H., 2012. Classification of ADHD children through multimodal magnetic resonance imaging. Front. Syst. Neurosci. 6(63).

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 43, 44-58.

Del Gaizo, J., Mofrad, N., Jensen, J. H., Clark, D., Glenn, R., Helpern, J., & Bonilha, L. (2017). Using machine learning to classify temporal lobe epilepsy based on diffusion MRI. Brain and Behavior, *7*(10), e00801.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968-980.

Di Russo, F., Berchicci, M., Bozzacchi, C., Perri, R.L., Pitzalis, S., Spinelli, D., 2017. Beyond the "Bereitschaftspotential": Action preparation behind cognitive functions. Neurosci Biobehav Rev. 78, 57-81.

Dosenbach, NU., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., et al., 2010. Prediction of individual brain maturity using fMRI. Science 329(5997), 1358-61.

Dubois, J., de Berker, A.O., Tsao, D.Y., 2015. Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. The Journal of Neuroscience 35(6), 2791-2802.

Etzel, J.A., Zacks, J.M., Braver, T.S., 2013. Searchlight analysis: promise, pitfalls, and potential. Neuroimage 78, 261-69.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., L, A.R., Fox, P.T., Eickhoff, S.B., Yu, C., Jiang, T., 2016. The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. Cerebral Cortex 26, 3508-3526.

Fan, L., Wang, J., Zhang, Y., Han, W., Yu, C., Jiang, T., 2014. Connectivity-based parcellation of the human temporal pole using diffusion tensor imaging. Cereb Cortex 24, 3365-3378.

Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., Lin, C.-J., 2008. LIBLINEAR: A library for large linear classification. J. March. Learn. Res. 9, 1871-1874.

Filippone, M., Marquand, A.F., Blain, C.R.V., Williams, S.C.R., Mourão-Miranda, J., Girolami, M., 2013. Probabilistic prediction of neurological disorders with a statistical assessment of neuroimaging data modalities. The annals of applied statistics 6(4), 1883-1905.

Forman, S., Cohen, J., Fitzgerald, M., Eddy, W., Mintum, M., Noll, D., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33, 636-647.

Fort, G., Lambert-Lacroix, S., 2005. Classification using partial least squares with penalized logistic regression. Bioinformatics 21, 1104-1111.

Gabay, A.S., Radua, J., Kempton, M.J., Mehta, M.A., 2014. The Ultimatum Game and the brain: A meta-analysis. Neuroscience and Biobehavioral Reviews 47, 549-558.

Gaertig, C., Moser, A., Alguacil, S., Ruz, M., 2012. Social information and economic decisión-making in the ultimatum game. Front Neurosci 6 (103).

Gaonkar, B., Shinohara, R., Davatzikos, C., Alzheimers Disease Neuroimaging Initiative, 2015. Interpreting support vector machine models for multivariate group analysis in neuroimaging. Med Image Anal. 24(1), 190-204

González-García, C., Arco, J.E., Palenciano, A.F., Ramírez, J., Ruz, M., 2017. Encoding, preparation and implementation of novel complex verbal instructions. NeuroImage 148, 264-273

González-García, C., Mas-Herrero, E., de Diego-Balaguer, R., Ruz, M., 2016. Task-specific preparatory neural activations in low-inference contexts. Brain Structure & Functions 8, 3997-4006.

Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J.M., Coalson, R.S., Nguyen, A.L., McDermott, K.B., Shimony, J.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Nelson, S.M., Dosenbach, N.U.F., 2017. Precision functional mapping of individual human brains. Neuron 95(4), 791-807.

Grecucci, A., Giorgetta, C., van't Wout, M., Bonini, N., Sanfey, A.G., 2013. Reappraising the Ultimatum: an fMRI study of emotion regulation and decision making. Cerebral Cortex 23(2), 399-410.

Haufe, S., Meinecke, F., Görgen, K, Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 87, 96-110.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293(5539), 2425-30.

Haxby, J.V., Connolly, A.C., Guntupalli, J.S., 2014. Decoding neural representational spaces using multivariate pattern analysis. Annu Rev. Neurosci. 37, 435-456.

Haynes, J.-D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.

Haynes, J.-D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat Neurosci. 8(5), 686-91.

Hebart, M.N., Baker, C.I., 2017. Deconstructing multivariate decoding for the study of brain function. NeuroImage, https://doi.org/10.1016/j.neuroimage.2017.08.005

Henson, R., 2005. Design efficiency in fMRI. URL http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency#VII._Should_I_treat_my_trials_as_events_or_epochs_.3F

Joliot, M., Jobard, G., Naveau, M., Delcroix, N., Petit, L., Zago, L., Crivello, F., Mellet, E., Mazoyer, B., Tzourio-Mazoyer, N., 2015. AICHA: An atlas of intrinsic connectivity of homotopic areas. Journal of Neuroscience Methods 254, 46-59.

Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat Neurosci. 8(5), 679-85.

Khedher, L., Illán, I.A., Górriz, J.M., Ramírez, J., Meyer-Baese, A., 2017. Independent component analysis-support vector machine-based computer aided diagnosis system for Alzheimer's disease with visual support. International Journal of Neural Systems 27(3), 8 1650050.

Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. PNAS 103, 3863-3868.

Kuzmanovic, B., Rigoux, L., Tittgemeyer, M., 2018. Influence of vmPFC on dmPFC Predicts Valence-Guided Belief Formation. The Journal of Neuroscience 38(37), 7996-8010.

Lanckriet, G.R.G., Cristianini, N., Bartlett, P., El Ghaoui, L., Jordan, M.I., 2004. Learning the Kernel Matrix with Semidefinite Programming. Journal of Machine Learning Research 5, 27-72.

Lindquist, K., Satpute, A., Wager, T., Weber, J., Barrett, L., 2015. The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. Cereb Cortex 26(5), 1910-1922.

Liu, H., Qin, W., Li, W., Fan, L., Wang, J., Jiang, T., Yu, C., 2013. Connectivity-based parcellation of the human frontal pole with diffusion tensor imaging. J Neurosci. 33, 6782-6790.

Liu, H., Stufflebeam, S.M., Sepulcre, J., Hedden, T., Buckner, R.L., 2009. Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. Proceedings of the National Academy of Sciences 106 (48).

Loose, L.S., Wisniewski, D., Rusconi, M., Goschke, T., Haynes, J.-D., 2017. Switch-independent task representations in frontal and parietal lobe. Journal of Neuroscience 37(33), 8033-8042.

Misaki, M., Kim, Y., Bandettini, P., Kriegeskorte, N., 2010. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. NeuroImage 53(1), 103-18.

Moser, A., Gaertig, C., Ruz, M., 2014. Social information and personal interests modulate neural activity during economic decision-making. Frontiers in Human Neuroscience 8, 31.

Mourão-Miranda, J., Bokde, A.L.W., Born, C., Hampel, H., Stetter, M., 2005. Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional fMRI data. NeuroImage 25, 980-995.

Mwangi, B., Tian, T.S., Soares, J.C., 2014. A review of feature reduction techniques in neuroimaging. Neuroinformatics 12(2), 229-244.

Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cognit. Sci. 10, 424–430.

Nurse, E.S., Karoly, P.J., Grayden, D.B., Freestone, D.R., 2015. A generalizable Brain-Computer-Interface (BCI) using machine learning for feature discovery. Plos One 10(6), 1-22.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: A tutorial overvie. NeuroImage 45, S199-209.

Plant, C., Teipel, S.J., Oswald, A., Böhm, C., Meindl, T., Mourão-Miranda, J., Bokde, A.W., Hampel, H., Ewers, M., 2010. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. NeuroImage 50(1), 162-74.

Poldrack, R.A., 2007. Region of interest analysis for fMRI. Soc Cogn Affect Neurosci. 2(1), 67-70.

Qiao, L., Zhang, L., Chen, A., Egner T, 2017. Dynamic trial-by trial recoding of task-set representations in the frontoparietal cortex mediates behavioral flexibility. Journal of Neuroscience 37(45), 11037-11050.

Qureshi, M.N.I., Oh, J., Min, B., Jo, H.J., Lee, B., 2017. Multi-modal, multi-measure, and multi-class discrimination of ADHD with hierarchical feature extraction and extreme learning machine using structural and functional brain MRI. Front Hum Neurosci. 11(157).

Rakotomamonjy, A., Bach, F., Canu, S., Grandvalet, Y., 2008. SimpleMKL. Journal of Machine Learning 9, 2491-2521.

Rondina, J.M., Hahn, T., de Oliveira, L., Marquand, A.F., Dresler, T., Fallgatter, A.J., Shawe-Taylor, J., Mourão-Miranda, J., 2014. SCoRS—A method based on stability for feature selection and mapping in neuroimaging. IEEE Trans Med Imaging 33(1), 85-98.

Sakai, K., 2008. Task set and prefrontal cortex. Annu. Rev. Neurosci. 31, 219-245.

Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cerebral Cortex, 1-20.

Schrouff, J., Monteiro, J.M., Portugal, L., Rosa, M.J., Phillips, C., Mourão-Miranda, J., 2018. Embedding anatomical or functional knowledge in whole-brain multiple kernel learnig models. Neuroinformatics 16, 117-143.

Schrouff, J., Cremers, J., Garraux, G., Baldassarre, L., Mourão-Miranda, J., Phillips, C., 2013a. Localizing and comparing weight maps generated from linear kernel machine learning models. IEEE Explore, http://ieeexplore.ieee.org/document/6603572/.

Schrouff, J., Rosa, M.J., Rondina, J.M., Marquand, A.F., Chu, C., Ashburner, J., Phillips, C., Richiardi, J., Mourão-Miranda, J., 2013b. PRoNTo: Pattern Recognition for Neuroimaging Toolbox. Neuroinformatics 11(3), 319-337.

Sona, D., Veeramachaneni, S., Olivetti, E., Avesani, P., 2007. Inferring cognition from fMRI brain images. Int. Conf. Artif. Neural Netw. 869-878.

Stelzer, J., Chen, Y., Turner, R., 2013. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. NeuroImage 65, 69-82.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society 58(1), 267-288.

Turner, B., Mumford, J., Poldrack, R., Ashby, F., 2012. Spatiotemporal activity estimation for multivoxel pattern analysis with rapid event-related designs. NeuroImage 62(3), 1429-1438.

Urchs, S., Dansereau, C., Benhajali, Y., Bellec, P. Group multiscale functional template generated with BASC on the Cambridge sample. https://figshare.com/articles/Group_multiscale_functional_template_generated_with_BASC_on_the_Cambridge_sample/1285615

Wang, D., Buckner, R.L., Fox, M.D., Holt, D.J., Holmes, A.J., Stoecklein, S., Langs, G., Pan, R., Qian, T., Kuncheng, L., Baker, J.T., Stufflebeam, S.M., Wang, K., Wang, X., Hong, B., Liu, H., 2015. Parcellating cortical functional networks in individuals. Nature Neuroscience 18, 1853-1860.

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106, 1125-1165.

Figure 1

Click here to access/download;Figure;Figura 1.png

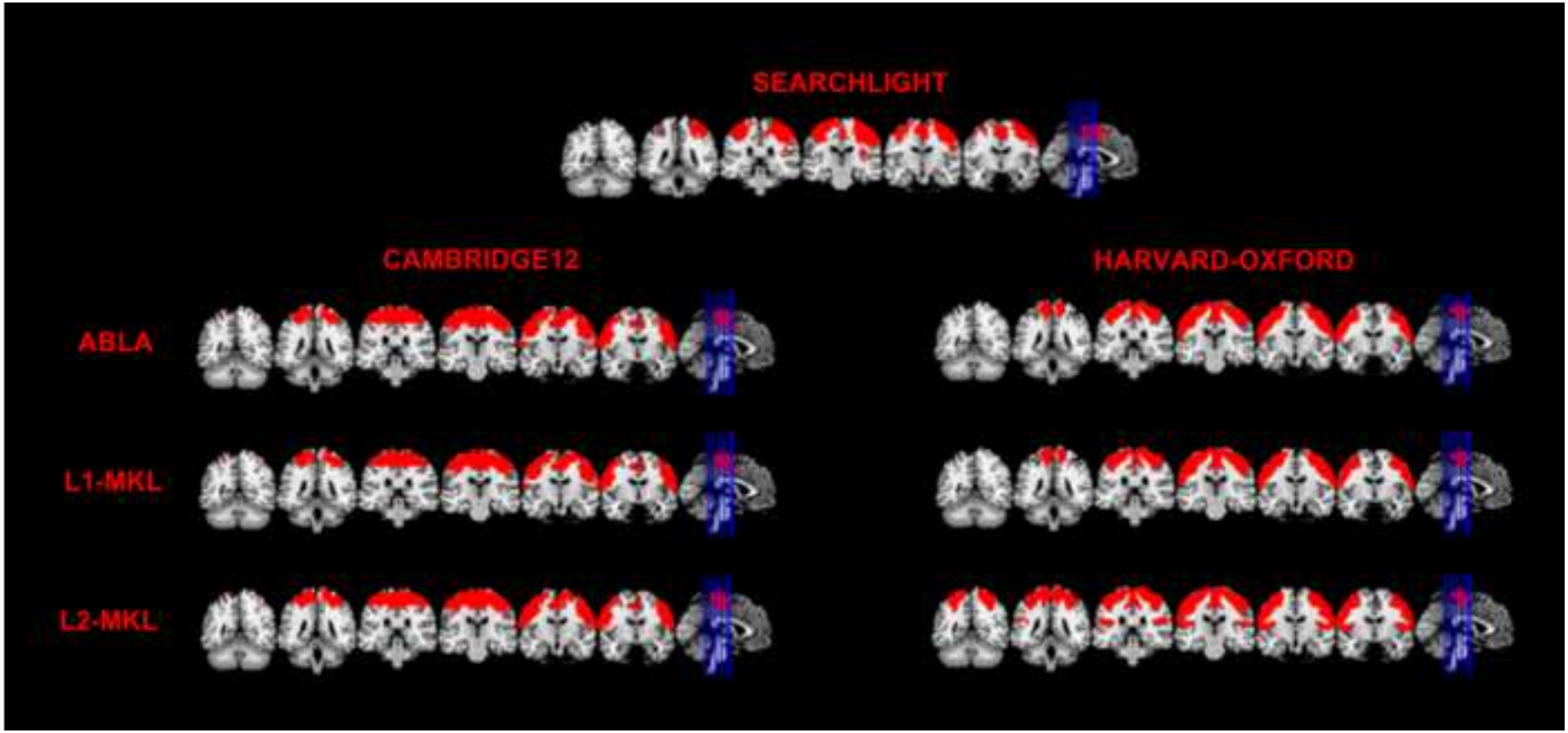

Figure 2

Click here to access/download;Figure;Figura 2.png

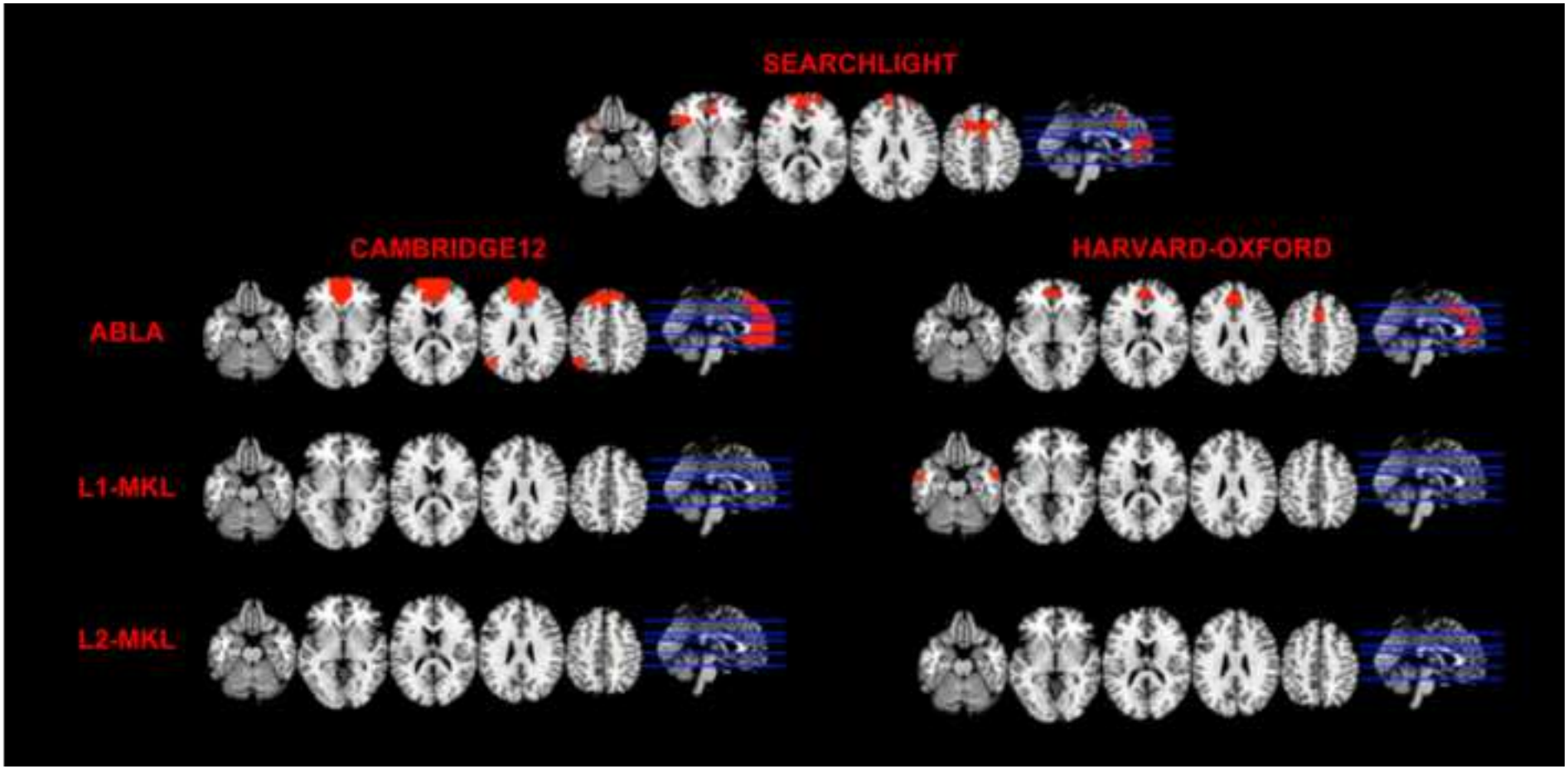

Figure 3

Click here to access/download;Figure;Figura 3.png