

Article

Hypergraph-Based Multitask Feature Selection with Temporally Constrained Group Sparsity Learning on fMRI

Youzhi Qu ¹, Kai Fu ¹, Linjing Wang ², Yu Zhang ², Haiyan Wu ³ and Quanying Liu ^{1,*}

¹ Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen 518055, China; quyz@mail.sustech.edu.cn (Y.Q.); kaifu1072@gmail.com (K.F.)

² Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China; linjingwang@outlook.com (L.W.); yu.zhang.ust@gmail.com (Y.Z.)

³ Centre for Cognitive and Brain Sciences and Department of Psychology, University of Macau, Macau 999078, China; haiyanwu@um.edu.mo

* Correspondence: liuqy@sustech.edu.cn

Abstract: Localizing the brain regions affected by tasks is crucial to understanding the mechanisms of brain function. However, traditional statistical analysis does not accurately identify the brain regions of interest due to factors such as sample size, task design, and statistical effects. Here, we propose a hypergraph-based multitask feature selection framework, referred to as HMTFS, which we apply to a functional magnetic resonance imaging (fMRI) dataset to extract task-related brain regions. HMTFS is characterized by its ability to construct a hypergraph through correlations between subjects, treating each subject as a node to preserve high-order information of time-varying signals. Additionally, it manages feature selection across different time windows in fMRI data as multiple tasks, facilitating time-constrained group sparse learning with a smoothness constraint. We utilize a large fMRI dataset from the Human Connectome Project (HCP) to validate the performance of HMTFS in feature selection. Experimental results demonstrate that brain regions selected by HMTFS can provide higher accuracy for downstream classification tasks compared to other competing feature selection methods and align with findings from previous neuroscience studies.



Citation: Qu, Y.; Fu, K.; Wang, L.; Zhang, Y.; Wu, H.; Liu, Q. Hypergraph-Based Multitask Feature Selection with Temporally Constrained Group Sparsity Learning on fMRI. *Mathematics* **2024**, *12*, 1733. <https://doi.org/10.3390/math12111733>

Academic Editors: Chengsheng Mao and Jie Xu

Received: 30 April 2024

Revised: 27 May 2024

Accepted: 30 May 2024

Published: 2 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: hypergraph; multitask learning; feature selection; fMRI

MSC: 92C50

1. Introduction

Neuroscientists have long been interested in uncovering the relationship between external stimuli (e.g., task stimuli, neuromodulation) and neural activity. Localizing stimulus-related brain regions is critical for understanding the functional mechanisms of the brain. Feature selection helps in this process by identifying relevant features and eliminating irrelevant ones from the original feature space, which is widely used in multitask learning (MTL) [1]. Since tasks in most MTL applications are related, it is natural to consider applying specific regularization terms to find subsets of tasks that share the same features, for instance, adding group lasso regularization [2] with an $\ell_{2,1}$ norm penalty to predict disease progression [3,4]. Therefore, designing appropriate regularization is crucial to model performance, which often relies on prior knowledge of the nature of the data [5–8]. To accurately localize stimulus-related brain regions, it is essential to develop feature selection methodologies that consider the spatiotemporal characteristics of the neural data.

Functional magnetic resonance imaging (fMRI) is a noninvasive imaging technique to measure time-varying blood oxygenation level-dependent (BOLD) dynamics in 3D brain voxels [9]. fMRI studies often segment brain regions according to atlases (e.g., automatic anatomical labeling (AAL) atlas [10] or multimodal parcellation (MMP) atlas [11]), then average the BOLD signals over voxels in different regions, calculate the correlation of

neural signals in regions of interest (ROIs), and obtain an adjacency matrix, in other words, the brain functional connectivity [12–14]. To better leverage the temporal information in neural dynamics, we integrate multitask learning into fMRI feature selection. Specifically, we segment the entire fMRI time series into multiple parts, each treated as a distinct task, thereby enabling a multitask feature selection approach. Consistent with the group lasso regularized multitask feature selection method, we employ the $\ell_{2,1}$ norm regularization to enhance the group sparsity. Additionally, we introduce smoothness regularization to accommodate the temporal correlations in fMRI data, addressing the higher similarity between close time points.

Data samples in neuroimaging datasets are limited compared to the large datasets required for deep learning. Therefore, preserving the locality and similarity between instances is crucial to make better use of the data. The hypergraph structure utilizes higher-order relationships, overcoming the limitation of the traditional graph structure that expresses data dependencies only in a pairwise manner [15–17]. A method is proposed to incorporate a hypergraph into multitask learning, which treats multiple data modalities (T1 and PET) as multiple tasks for finding biomarkers for Alzheimer’s disease [15]. Instead of using multimodal neural images, we treat different time periods in unimodal time series data (e.g., fMRI) as multiple tasks to capture time-varying information in the brain. We constructed a hypergraph structure to represent the high-order relationships among subjects, allowing us to fully utilize the limited data samples in neuroimaging. In this study, we introduced a hypergraph regularization term to enhance consistency among subjects connected by hyperedges. By integrating hypergraph regularization with multitask learning, we enhanced the effectiveness of feature selection.

To fully exploit the temporal information and high-order correlations of data samples, we propose a hypergraph-based multitask feature selection (HMTFS) method for fMRI data and apply it to reveal task-related brain regions, as illustrated in Figure 1. To validate the effectiveness of feature selection using HMTFS, we conduct classification and t-SNE visualization experiments with comprehensive ablation tests using a large HCP fMRI dataset. Additionally, we employ the HMTFS method to visualize task-related brain regions across 23 cognitive tasks in the HCP dataset and predicted the difficulty of language tasks. Our main contributions are summarized as follows: (1) by integrating the hypergraph structure and multitask learning, HMTFS effectively utilizes the temporal information between tasks and the high-order correlation between subjects to select features that enhance the downstream classification tasks; (2) a threshold-based method is proposed to construct a hypergraph, and comparative experimental results show that the threshold-based hypergraph construction method is superior to the k-nearest neighbors (kNN) method; (3) by incorporating the temporally constrained group sparsity and smoothness regularizations, HMTFS effectively captures the time-varying information in the brain; (4) as a feature selection tool, HMTFS not only efficiently identifies task-related brain regions but also performs well in challenging scenarios such as predicting task difficulty.

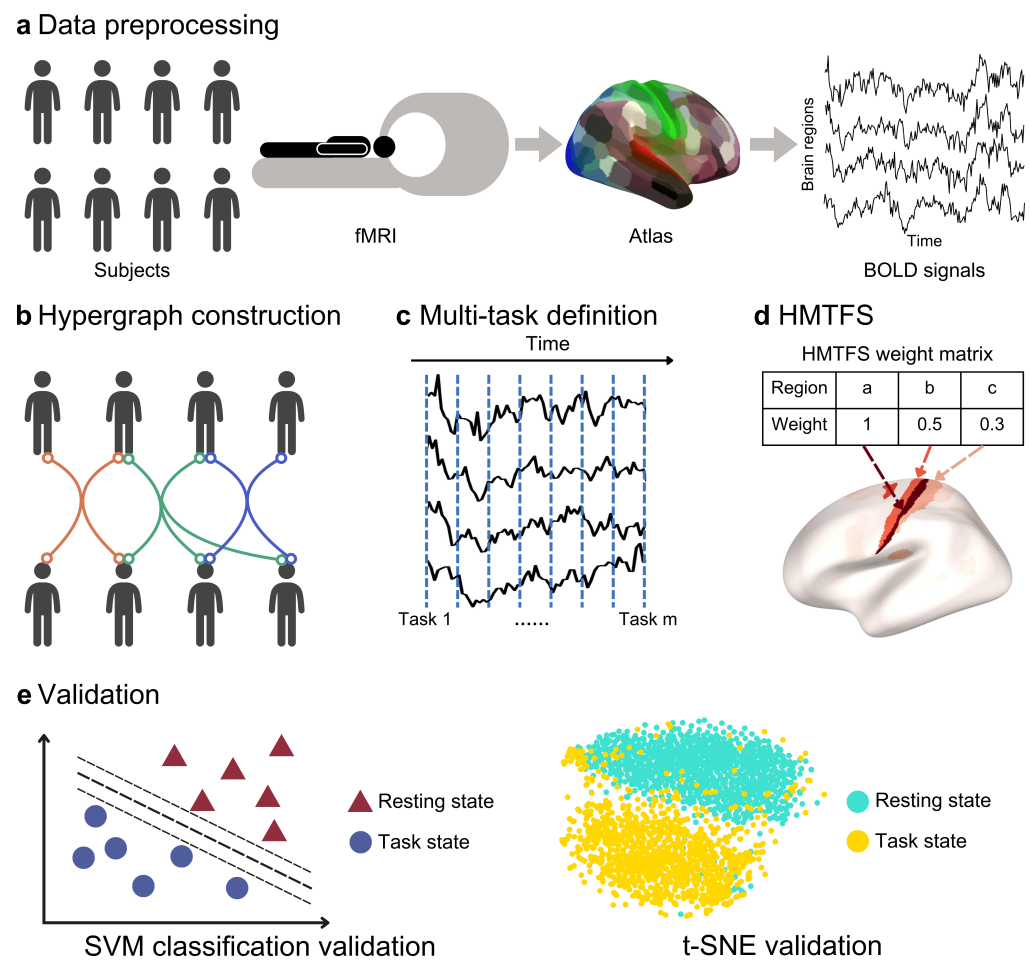


Figure 1. The Hypergraph-based multitask feature selection (HMTFS) framework for fMRI data analysis. (a) Data preprocessing: preprocessing fMRI data and parcellating brain with atlas. (b) hypergraph construction: employing the threshold-based method to construct a hypergraph based on intersubject distances. (c) Multitask definition: dividing the entire fMRI into multiple tasks by time slicing. (d) HMTFS: applying HMTFS to locate task-related brain regions. (e) Validation: using support vector machine (SVM)-based classification and t-SNE visualization to test the effectiveness of HMTFS on feature selection.

2. Related Work

2.1. Multitask Feature Selection Models

Multitask learning (MTL) is an inductive transfer mechanism aimed at enhancing generalization performance through the utilization of knowledge gained across various tasks [1,18]. MTL improves performance by leveraging the information shared among related tasks. In recent years, MTL has been widely applied across a diverse range of fields, such as medicine and computer vision [19–22].

Suppose we have m learning tasks $\{\mathcal{T}_i\}_{i=1}^m$; the training data of task i is $\mathcal{D}_i$ consisting of n_i training samples, i.e., $\mathcal{D}_i = \{\mathbf{x}_j^i, y_j^i\}_{j=1}^{n_i}$, where $\mathbf{x}_j^i \in \mathbb{R}^{d_i}$ is the j th training instance in $\mathcal{T}_i$ and y_j^i is its label. The training data matrix $\mathbf{X}^i = [\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_{n_i}^i]^T \in \mathbb{R}^{n_i \times d_i}$ for $\mathcal{T}_i$. The weight matrix $\mathbf{W} = [\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^m]$, where $\mathbf{w}^i \in \mathbb{R}^d$ denotes the feature weight vector in task i . In the homogeneous MTL, all tasks share the same feature space dimension, ensuring that the training data matrices for all tasks have consistent feature dimensions d ,

i.e., $d^i = d^j$ for $i \neq j$. The simplest approach to defining a loss function is to use the squared ℓ_2 norm. For homogeneous MTL objectives, this can be formulated as follows:

$$\min_W \frac{1}{2} \sum_{i=1}^m \|\mathbf{Y}^i - \mathbf{X}^i \mathbf{w}^i\|_2^2. \quad (1)$$

In multitask learning, regularization norms can be applied to the matrix $\mathbf{W}$ for feature selection. The pioneering research on multitask feature selection employed the $\ell_{2,1}$ norm within $\mathbf{W}$, promoting row sparsity to aid in identifying crucial features [23]. Further studies have expanded on this approach by employing various types of norms, such as $\ell_{\infty,1}$, for feature selection [5,24]. The $\ell_{p,q}$ norm can be utilized to address problems for multitask feature selection problems, with the objective function formulated as follows:

$$\min_W \frac{1}{2} \sum_{i=1}^m \|\mathbf{Y}^i - \mathbf{X}^i \mathbf{w}^i\|_2^2 + \lambda \|\mathbf{W}\|_{p,q}, \quad (2)$$

where the parameter λ balances the importance of the loss function and the regularization term, $\|\mathbf{W}\|_{p,q} \equiv \|(\|\mathbf{w}_1\|_p, \|\mathbf{w}_2\|_p, \dots, \|\mathbf{w}_d\|_p)\|_q$, and $\mathbf{w}_i$ represents the i th row of $\mathbf{W}$ and $\|\cdot\|_p$ denotes the ℓ_p norm applied to a vector.

2.2. Laplacian Sparse Coding

Neuroscience research focused on the visual cortex has demonstrated that the brain processes visual stimuli following the principles of sparse coding [25,26]. This approach utilizes a limited set of basis vectors in a linear combination to maximally retain information from the original high-dimensional data. Sparse coding effectively compresses data and has been widely adopted in fields such as image reconstruction, image denoising, and face recognition [27–30].

Given a set of signals $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n] \in \mathbb{R}^{d \times n}$ and codebook $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k] \in \mathbb{R}^{d \times k}$, the purpose of sparse coding is to find a sparse representation $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n] \in \mathbb{R}^{k \times n}$ that minimizes the reconstruction error, expressed as $\|\mathbf{S} - \mathbf{C}\mathbf{R}\|_F^2$. The simplest approach to enhancing the sparsity of $\mathbf{R}$ involves incorporating the ℓ_0 norm. However, the minimization problem becomes nonconvex due to the ℓ_0 norm, leading to NP-hard computational challenges. Recent research address the sparse coding challenge by adding the ℓ_1 norm to ensure the sparsity of $\mathbf{R}$ [31,32]. Thus, the problem can be formulated as follows:

$$\min \|\mathbf{S} - \mathbf{C}\mathbf{R}\|_F^2 + \lambda \|\mathbf{R}\|_{1,1}, \quad (3)$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix, the parameter λ governs the trade-off between sparsity and minimization error, and $\|\mathbf{R}\|_{1,1} = \sum_{i=1}^k \sum_{j=1}^n |r_{ij}|$.

Sparse coding independently encodes features, which may overlook similar features in the signal, leading to a loss of local information in the encoded features. To preserve the local relationships and information among features, introducing a graph structure can be an effective solution. In such a graph structure, similarity matrix $\mathbf{A}$ captures the similarities between vertices, where $\mathbf{A}_{ij}$ represents the similarity between vertex i and vertex j . The degree matrix $\mathbf{D}$ is a diagonal matrix whose diagonal entries $\mathbf{D}_{ii}$ sum up all the similarities connected to vertex i . The Laplacian matrix $\mathbf{L}$ is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. To effectively preserve the inherent structure and interrelationships within the data, the graph-based Laplacian sparse coding can be formulated as follows:

$$\min \|\mathbf{S} - \mathbf{C}\mathbf{R}\|_F^2 + \lambda \|\mathbf{R}\|_{1,1} + \gamma \text{Tr}(\mathbf{R}\mathbf{L}\mathbf{R}^T), \quad (4)$$

where $\text{Tr}(\cdot)$ denotes the trace of a matrix, and the parameter γ quantifies the strength of locality preservation among the features to be encoded.

In ordinary graph structures, each edge connects two vertices. In contrast, edges in a hypergraph can connect any number of vertices, allowing it to capture complex higher-order relationships that cannot be represented by simple pairwise connections. This feature

makes hypergraphs a powerful tool for modeling complex interactions and relationships. Hypergraphs have been widely applied in fields such as machine learning, bioinformatics, and social networks [16,33,34].

Let $G(V, E)$ denote a hypergraph, where V represents the set of vertices and E denotes the set of hyperedges. Each hyperedge e is considered a subset of V , and the weight of each hyperedge e is $a(e)$. A diagonal edge weight matrix $\mathbf{A} \in \mathbb{R}^{|E| \times |E|}$ contains these weights, with $\mathbf{A}_{ii} = a(e_i)$ for each hyperedge $a(e_i)$. The incident matrix $\mathbf{H} \in \mathbb{R}^{|V| \times |E|}$ is defined such that entry $h(v_i, e_j) = 1$ if $v_i \in e_j$, and $h(v_i, e_j) = 0$ otherwise. The vertex degree matrix, $\mathbf{D}_v \in \mathbb{R}^{|V| \times |V|}$, has diagonal entries corresponding to the degrees of the vertices. The degree of vertex v_i is calculated as $\sum_{e_j \in E} a(e_j) h(v_i, e_j)$. Similarly, the hyperedge degree matrix, $\mathbf{D}_e \in \mathbb{R}^{|E| \times |E|}$, contains diagonal entries that reflect the degrees of each hyperedge. The degree of hyperedge e_j is determined by $\sum_{v_i \in V} h(v_i, e_j)$. The unnormalized hypergraph Laplacian matrix [35,36] can be defined as $\mathbf{L}^h = \mathbf{D}_v - \mathbf{H} \mathbf{A} \mathbf{D}_e^{-1} \mathbf{H}^T$. The hypergraph-based Laplacian sparse coding can be formulated as follows [17]:

$$\min \|\mathbf{S} - \mathbf{C}\mathbf{R}\|_F^2 + \lambda \|\mathbf{R}\|_{1,1} + \gamma \text{Tr}(\mathbf{R} \mathbf{L}^h \mathbf{R}^T). \quad (5)$$

3. Materials and Methods

3.1. Hypergraph

As previously defined, the hypergraph $G(V, E)$ comprises a set of vertices V and a set of hyperedges E , with the weights of each hyperedge represented in the diagonal matrix $\mathbf{A}$. The incidence matrix $\mathbf{H}$ indicates the relationship between vertices and hyperedges, explicitly mapping which vertices are included in each hyperedge. Adopted from the method in [35], the normalized hypergraph Laplacian matrix is defined as follows:

$$\Delta = \mathbf{I} - \mathbf{D}_v^{-1/2} \mathbf{H} \mathbf{A} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-1/2}, \quad (6)$$

where $\mathbf{D}_v$ and $\mathbf{D}_e$ are diagonal matrices of vertex degrees and hyperedge degrees, respectively; $\mathbf{I}$ is an identity matrix.

In this study, each vertex in the hypergraph represents a subject rather than a brain region. A crucial step in hypergraph construction is to assemble hyperedges according to the distance between vertices. The k -nearest neighbors (kNN) method can be applied to select a fixed-size subset of adjacent vertices, specifically k vertices, as hyperedges for each vertex [16,35]. We propose a threshold-based hypergraph construction method where a vertex is connected to a set of neighbors if their distance from the vertex is below a predefined threshold. Consequently, given N vertices in total, the corresponding incidence matrix $\mathbf{H} \in \mathbb{R}^{N \times N}$ has N hyperedges, each connecting a flexible-sized subset of vertices, in contrast to the fixed size of $k + 1$ vertices in the kNN-based method.

3.2. Hypergraph-Based Multitask Feature Selection Method

The proposed hypergraph-based multitask feature selection (HMTFS) method employs temporally constrained group sparsity regularization to encourage the intrinsic time correlation in fMRI and select a subset of shared features across multiple tasks. Additionally, hypergraph regularization in HMTFS captures the high-order information among subjects. After the preprocessing of the fMRI data, the BOLD signals are formatted into two-dimensional matrices representing brain regions across time. Each task represents a specific temporal segment, with brain regions serving as the feature dimensions. The m frames contained within the BOLD signals are defined as m learning tasks. For each task i , the training data $\mathcal{D}_i = \{\mathbf{X}^i, \mathbf{Y}^i\}_{i=1}^m$. The training data matrix $\mathbf{X}^i = [\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_N^i]^T \in \mathbb{R}^{N \times d}$ consists of training samples from N subjects for task i , where $\mathbf{x}_j^i \in \mathbb{R}^d$ represents the signals from d brain regions for subject j at time i . $\mathbf{Y}^i \in \mathbb{R}^N$ represents the class labels, which

indicate whether subjects are in a resting state or engaged in specific task states. The objective function of HMTFS can be formulated as follows:

$$\min_{\mathbf{W}} \frac{1}{2} \sum_{i=1}^m \|\mathbf{Y}^i - \mathbf{X}^i \mathbf{w}^i\|_2^2 + \lambda_1 \|\mathbf{W}\|_{2,1} + \lambda_2 \sum_{i=2}^m \|\mathbf{w}^i - \mathbf{w}^{i-1}\|_2 + \lambda_3 \Omega, \quad (7)$$

where $\mathbf{w}^i \in \mathbb{R}^d$ is the feature weight vector in the i th time period or equivalent task i , $\mathbf{W} = [\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^m] \in \mathbb{R}^{d \times m}$, and $\|\mathbf{W}\|_{2,1} = \sum_{i=1}^d \|\mathbf{w}_i\|_2$ denotes the $\ell_{2,1}$ norm of $\mathbf{W}$ with $\mathbf{w}_i$ denoting the i th row of $\mathbf{W}$. The hypergraph regularization term is defined as $\Omega = \sum_{i=1}^m \mathbf{w}^{iT} \mathbf{X}^{iT} \Delta \mathbf{X}^i \mathbf{w}^i$. Equation (7) consists of four terms: the first term indicates the averaged training loss across m tasks; the second term encourages the clustering of weights in $\mathbf{W}$ across multiple time periods, resulting an implicit feature selection for all tasks; the third term smoothes the difference between the weight vectors of successive time points; and the fourth term regularizes the smoothness among different vertices in the hypergraph based on hypergraph Laplacian Δ . Three hyperparameters λ_1 , λ_2 , and λ_3 balance the relative contribution of the corresponding regularization terms. The learnable parameter is $\mathbf{W}$, representing the importance of features corresponding to each brain region. The problem (7) is a convex optimization problem with respect to $\mathbf{W}$, ensuring the uniqueness of the solution. This learning problem can be easily solved by Adam optimizer. Once the problem (7) is solved, we can select the features by sorting the weight matrix $\mathbf{W}$. To be noted, feature selection with HMTFS is unattached to the downstream classification tasks.

3.3. Evaluation Datasets and Implementation Settings

In this study, we utilize fMRI data from 1075 subjects in the public large-scale Human Connectome Project (HCP) S1200 dataset [37]. For HCP dataset, the resting-state fMRI (rs-fMRI) and task fMRI are preprocessed using the standard HCP pipeline [38], followed by brain parcellation into 360 regions using the MMP atlas [11]. To ensure robust validation, we employ a 5-fold cross-validation strategy to evaluate the proposed HMTFS method. For the training data, we first calculate the Euclidean distance between BOLD signals as intersubject distances, then apply the threshold-based method to construct the hypergraph, with each hyperedge weight set to 1. Figure 2f,g illustrate the classification accuracy under identical conditions using two different hypergraph construction methods. The threshold-based method has a clear peak at $k = 45$, whereas the kNN-based method plateaus. We test the sensitivity of three hyperparameters for regularization (λ_1 , λ_2 , and λ_3 in Equation (7)) on the classification performance, with λ_1 , λ_2 , and $\lambda_3 \in \{10^2, 10^1, 10^0, 10^{-1}, 10^{-2}\}$. The performance peaks at $\lambda_1 = 10$, $\lambda_2 = 1$, and $\lambda_3 = 10$, as shown in Figure 2h. Thus, parameters are set with a threshold $k = 45$, and $\lambda_1 = 10$, $\lambda_2 = 10$, $\lambda_3 = 1$ in Equation (7). The HMTFS model is trained using an Adam optimizer [39] with a fixed learning rate of 0.001 and parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$ for 50 epochs. To further investigate the effectiveness of the selected brain regions, a support vector machine (SVM) classifier is employed to differentiate between resting and task states using the BOLD signals of the top 10 selected brain regions. The effectiveness of the selected features is quantified by classification accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under receiver operating characteristic curve (AUC).

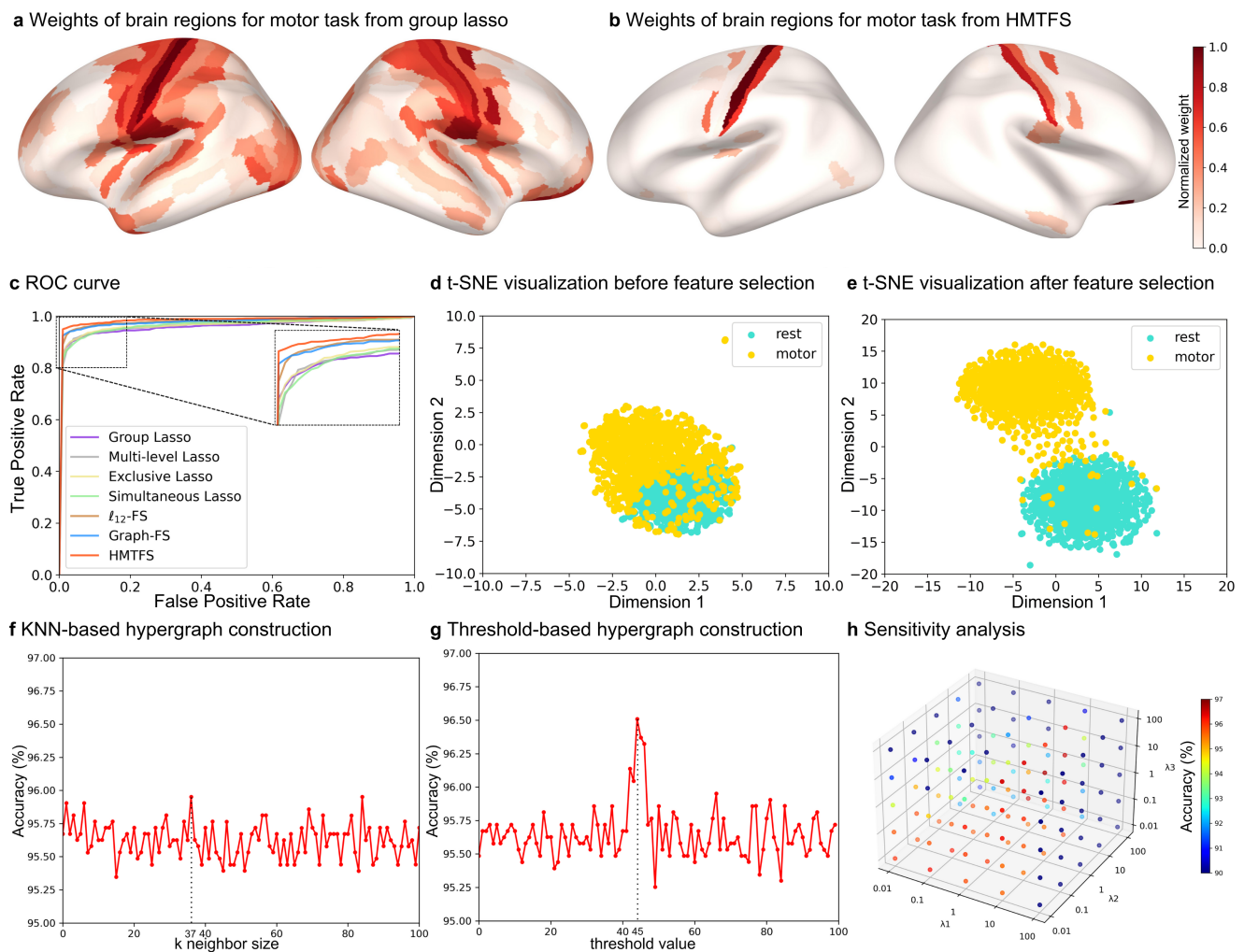


Figure 2. Validation of the hypergraph-based multitask feature selection method. (a) Visualization of brain regions weight obtained by group lasso method for motor tasks. (b) Visualization of brain region weights obtained by HMTFS method for motor tasks. (c) ROC curves comparing the classification effectiveness of HMTFS (shown in red) against other competing methods. (d) t-SNE visualization of the original data for rest and motor states. (e) t-SNE visualization of rest and motor states after feature selection. (f) Classification accuracy variation with k using the kNN-based hypergraph construction method. (g) Classification accuracy variation with k using the threshold-based hypergraph construction method. (h) Sensitivity test of three hyperparameters for regularization where λ_1, λ_2 , and $\lambda_3 \in \{10^{-2}, 10^{-1}, 10^0, 10^1, 10^2\}$. The color of the dots indicates the classification accuracy with different hyperparameters.

4. Results

4.1. Validation of the HMTFS Method

To validate the effectiveness of the HMTFS method in feature selection, we test its ability to identify task-related brain regions using resting-state and motor task fMRI data from the large-scale HCP dataset. We compare the HMTFS method with other multitask feature selection methods, including multitask group lasso (i.e., group lasso), multilevel lasso for MTL (i.e., multilevel lasso) [8], multitask exclusive lasso (i.e., exclusive lasso) [7], multitask simultaneous lasso (i.e., simultaneous lasso) [6], and ℓ_{1-2} norm multitask feature selection method (i.e., ℓ_{12} -FS) [40]. Additionally, we evaluate the impact of our hypergraph method against a traditional graph method, graph-based feature selection (Graph-FS), which utilizes graph structures to preserve locality and similarity among subjects. Motor task-related brain regions selected by group lasso and HMTFS methods are shown in Figure 2a,b, suggesting that the brain regions selected by HMTFS are more clustered in

the motor network. We utilize a linear SVM classifier to categorize the resting state versus motor state based on the top 10 brain regions most relevant to motor tasks, as identified by different feature selection methods. The ROC curves indicate that the HMTFS method can select the most informative features compared to other competing methods, as shown in Figure 2c. We visualize the fMRI data by t-SNE (Figure 2d,e), indicating that while whole-brain fMRI signals do not effectively distinguish between motor and resting states, clear differentiation is observed following feature selection using the HMTFS method. Table 1 presents a comparison of four metrics (accuracy, sensitivity, specificity, and AUC) between the proposed HMTFS method and other feature selection methods. The HMTFS method achieves a remarkable accuracy of 96.70%, clearly outperforming alternative methods in classifying between resting state and motor state. This superior accuracy highlights the effectiveness of the feature selection capabilities of the HMTFS method. We conduct an ablation study to validate the effectiveness of the regularization components within the HMTFS method, specifically the multitask smooth regularization and hypergraph regularization, which correspond to the third and fourth terms in Equation (7), respectively. For simplicity, we denote the hypergraph regularization term with the kNN-based method, the threshold-based hypergraph construction method, and the multitask smooth regularization term as K-H, T-H, and M, respectively. The results of the ablation study are shown in Table 2, confirming the necessity of threshold-based hypergraph construction and multitask smooth regularization.

Table 1. Classification performance (mean $\pm$ std) of different feature selection methods for distinguishing between resting state and motor state in the HCP dataset. The best results are displayed with bold font.

Method	Accuracy	Sensitivity	Specificity	AUC
Group lasso	93.25 $\pm$ 1.28	98.07 $\pm$ 0.78	88.34 $\pm$ 3.14	96.22 $\pm$ 1.46
Multilevel lasso	92.97 $\pm$ 1.41	96.91 $\pm$ 0.76	88.90 $\pm$ 3.17	96.38 $\pm$ 0.80
Exclusive lasso	93.53 $\pm$ 1.45	98.05 $\pm$ 0.75	89.05 $\pm$ 2.92	96.68 $\pm$ 0.57
Simultaneous lasso	92.32 $\pm$ 2.13	96.83 $\pm$ 2.67	87.86 $\pm$ 3.01	96.80 $\pm$ 0.63
ℓ_{12} -FS	95.30 $\pm$ 1.36	97.92 $\pm$ 1.08	92.71 $\pm$ 2.36	98.03 $\pm$ 0.74
Graph-FS	95.76 $\pm$ 1.00	98.88 $\pm$ 0.64	92.66 $\pm$ 1.76	98.12 $\pm$ 0.52
HMTFS	96.70 $\pm$ 0.17	99.34 $\pm$ 0.37	94.03 $\pm$ 0.50	98.75 $\pm$ 0.59

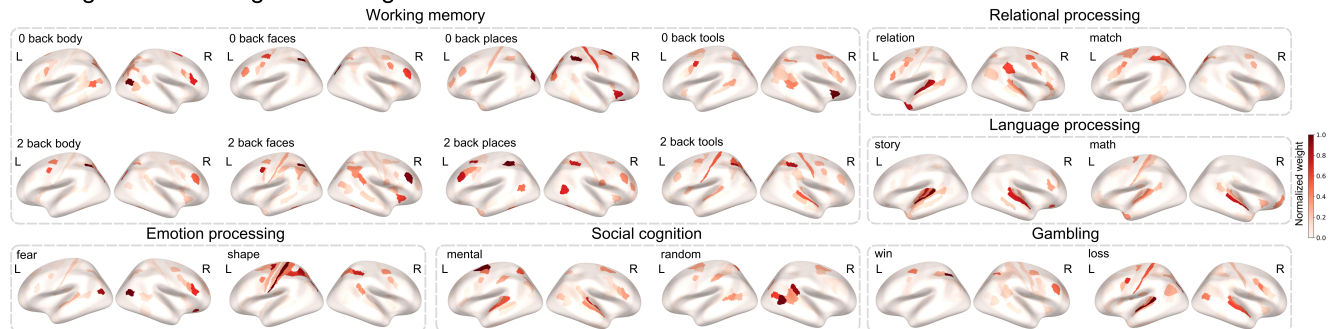
Table 2. Ablation study of hypergraph and multitask modules in HMTFS.

Method	Accuracy	Sensitivity	Specificity	AUC
Baseline (group lasso)	93.25 $\pm$ 1.28	98.07 $\pm$ 0.78	88.34 $\pm$ 3.14	96.22 $\pm$ 1.46
Baseline + K-H	95.95 $\pm$ 0.52	99.07 $\pm$ 0.28	92.80 $\pm$ 1.11	98.11 $\pm$ 0.55
Baseline + T-H	96.46 $\pm$ 0.93	99.45 $\pm$ 0.34	93.56 $\pm$ 1.69	98.71 $\pm$ 0.30
Baseline + T-H + M	96.70 $\pm$ 0.17	99.34 $\pm$ 0.37	94.03 $\pm$ 0.50	98.75 $\pm$ 0.59

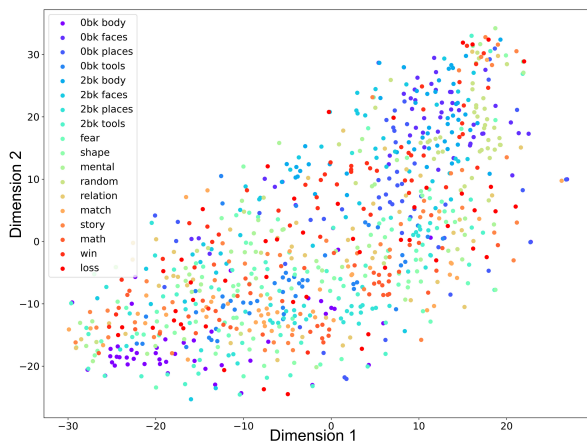
4.2. The Application of HMTFS Method for Cognitive Tasks

The cognitive tasks included in the HCP dataset are designed to target a broad range of brain functions, such as working memory, language processing, emotion processing, and social cognition. We applied the HMTFS method to select and visualize brain regions across 18 cognitive tasks in the HCP dataset. Figure 3a demonstrates that HMTFS effectively identified task-related brain regions with the results of brain region weights. Figure 3b,c further validate the effectiveness of HMTFS in selecting task-related brain regions by t-SNE visualization. The t-SNE visualization reveals that the original data from various cognitive tasks are intermingled, making it difficult to distinguish between them. However, the data corresponding to different tasks are clearly separable after feature selection using the HMTFS method.

a Weights of brain regions for cognitive tasks



b t-SNE visualization before feature selection



c t-SNE visualization after feature selection

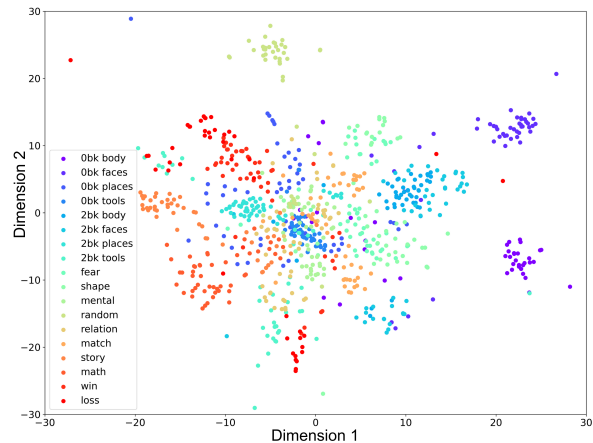


Figure 3. Results of the hypergraph-based multitask feature selection model applied to 18 cognitive tasks. (a) Visualization of brain region weights for the 18 cognitive tasks. (b) t-SNE visualization of the original data for the 18 cognitive tasks. (c) t-SNE visualization for the 18 cognitive tasks after feature selection.

In the validation process, we identified brain regions associated with motor tasks using the HMTFS method (Figure 2). To delve deeper into specific motor tasks within the HCP dataset, we applied the HMTFS method to identify brain regions associated with movements of the right hand, left hand, right foot, left foot, and tongue. The visualization of brain region weights is shown in Figure 4. The results clearly demonstrate the lateralization of brain function, with tasks involving the right side of the body activating the left hemisphere of the motor cortex and vice versa. Furthermore, the movements from the feet through the hands to the tongue correspond with the motor cortex's topographical layout from top to bottom. To further validate the effectiveness of selecting brain regions for specific motor tasks, we employed the hemodynamic response function (HRF) to generate BOLD signals corresponding to task stimuli, allowing us to compare the selected brain regions with visual regions. Figure 5 displays the HRF (shown in black), which outlines the designed structure of the specific motor tasks, with each task being performed twice during the entire motor task test. The results demonstrate that the motor-related brain regions selected by the HMTFS method (shown in red) achieve a better fit compared to those related to visual regions (shown in blue). The brain region weight matrix obtained from all cognitive tasks in the HCP dataset illustrates the overall mapping of brain regions associated with cognitive tasks, as shown in Figure 6. The visualization of the top 50 brain regions for each task reveals distinct differences between motor tasks and other cognitive tasks.

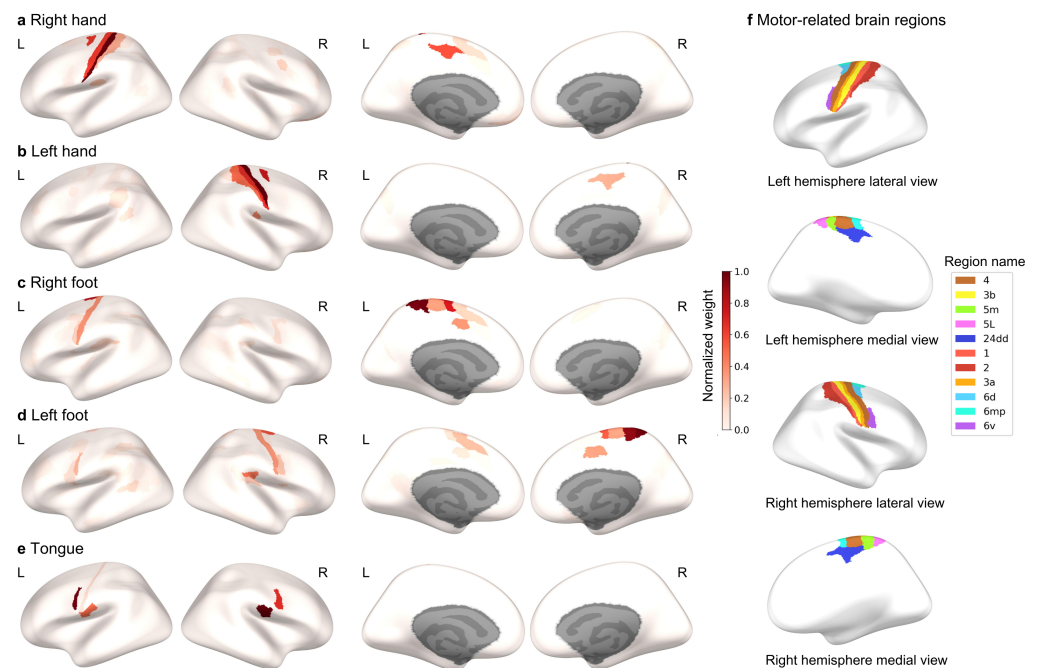


Figure 4. Visualization of brain region weights obtained by HMTFS method for motor tasks. (a) Visualization of brain region weights for the right-hand motor task. (b) Visualization of brain region weights for the left-hand motor task. (c) Visualization of brain region weights for the right-foot motor task. (d) Visualization of brain region weights for the left-foot motor task. (e) Visualization of brain region weights for the tongue motor task. (f) Visualization of motor-related brain regions according to the MMP atlas.

In addition to selecting brain regions related to different cognitive tasks based on task states, the HMTFS method also provides a significant advantage by not requiring prior knowledge of the experimental task structure. HMTFS can predict task performance of subjects that are not predefined. Within the HCP dataset, the language tasks are divided into story and math tasks. In the story task, participants listen to a brief story followed by responding to questions using a two-alternatives forced-choice format. Similarly, in the math task, participants solve addition or subtraction problems and select their answers from two provided options. We segmented the task trials into high and low difficulty levels for training the HMTFS method. The brain regions related to task difficulty selected by HMTFS are effective in predicting the level of task difficulty, as shown in Figure 7. Furthermore, as the tasks progress, both the story and math tasks show an increasing trend in predicted difficulty. This trend corresponds with the actual task design, where the specific question-answering component is positioned towards the end of the task, and the beginning of the task is not directly related to task difficulty.

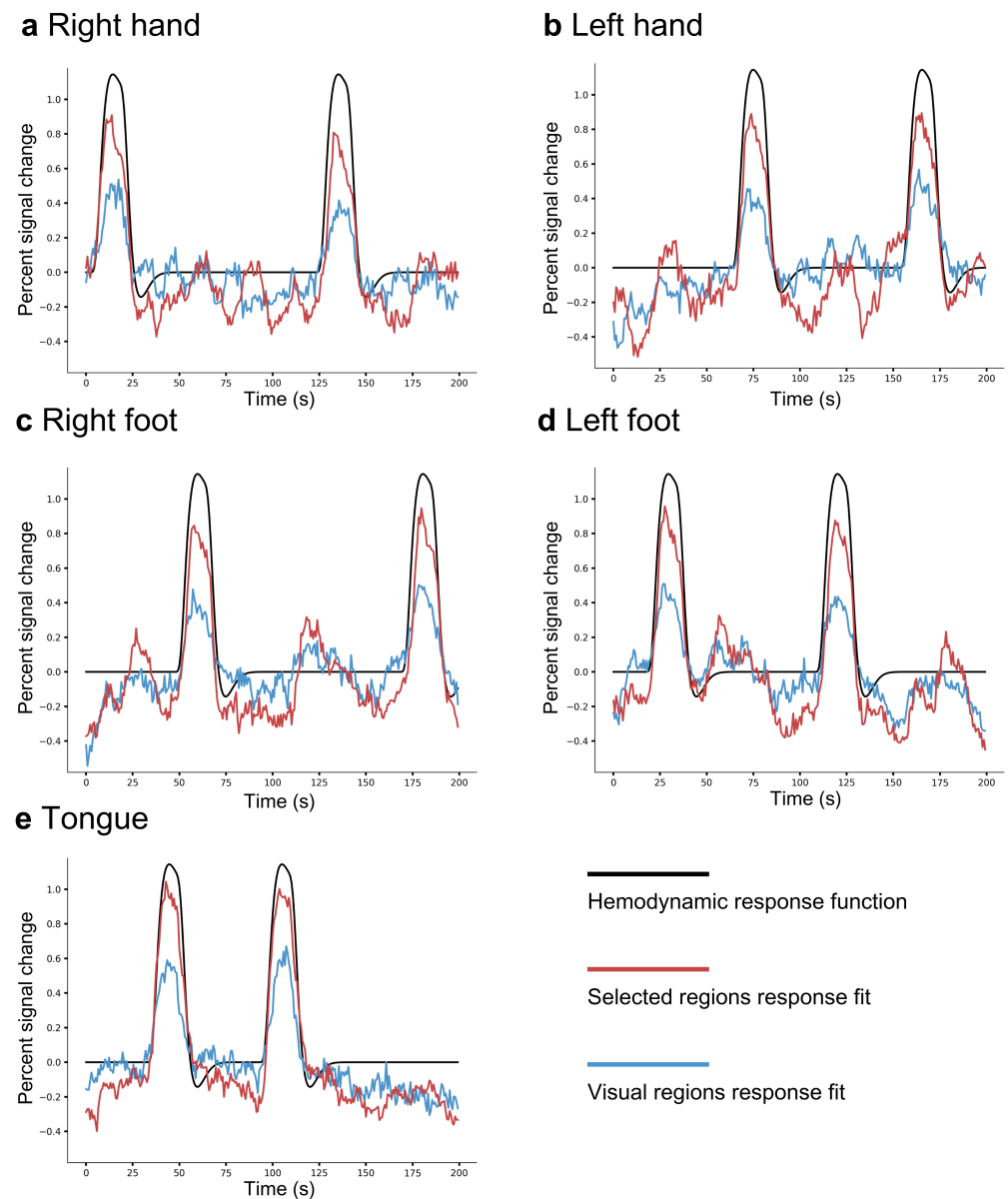


Figure 5. Hemodynamic response function (HRF) fitting results for motor tasks using brain regions selected by HMTFS method compared to visual-related brain regions. **(a)** Right-hand motor task HRF fitting results show that brain regions achieved an $R^2 = 0.6749$, compared to $R^2 = 0.4270$ for visual regions. **(b)** Left-hand motor task HRF fitting results show that brain regions achieved an $R^2 = 0.6554$, compared to $R^2 = 0.4069$ for visual regions. **(c)** Right-foot motor task HRF fitting results show that brain regions achieved an $R^2 = 0.6197$, compared to $R^2 = 0.3687$ for visual regions. **(d)** Left-foot motor task HRF fitting results show that brain regions achieved an $R^2 = 0.6133$, compared to $R^2 = 0.4104$ for visual regions. **(e)** Tongue motor task HRF fitting results show that brain regions achieved an $R^2 = 0.7356$, compared to $R^2 = 0.5566$ for visual regions.

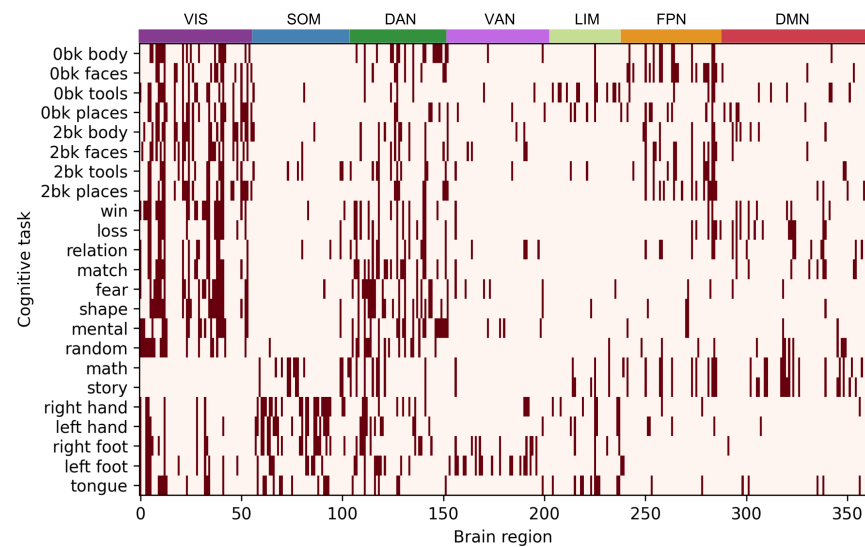
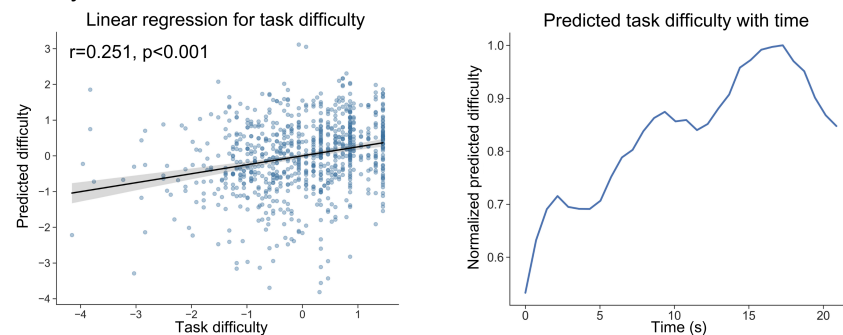


Figure 6. Cognitive task brain region mapping: Visualization of the top 50 brain regions for each task based on their weights determined by the HMTFS method. Each row represents the top 50 brain regions associated with a cognitive task, and colored cells indicate the brain regions involved in the task. The brain regions are organized according to the following brain network sequence: visual network (VIS), somatomotor network (SOM), dorsal attention network (DAN), ventral attention network (VAN), limbic network (LIM), frontoparietal network (FPN), and default mode network (DMN).

a Story task



b Math task

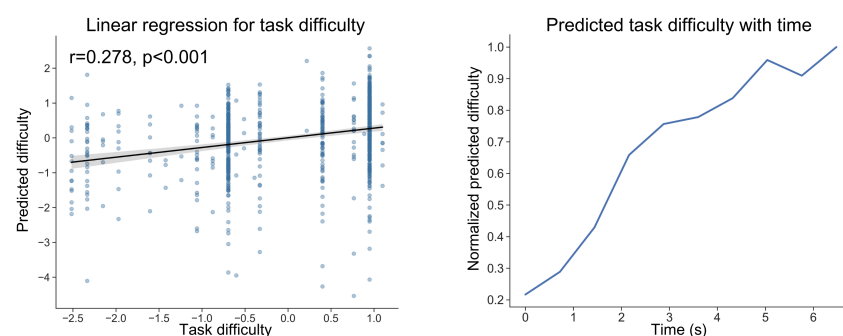


Figure 7. Task difficulty prediction results for language task, including story and math tasks. **(a)** Task difficulty prediction results for story task. The left panel illustrates the correlation between the task difficulty predicted by the HMTFS method and the actual task difficulty, showing a correlation with $r = 0.251$ and $p < 0.001$. The right panel shows the variations in predicted task difficulty at different times during the story task. **(b)** Task difficulty prediction results for math task. The left panel displays the correlation between the task difficulty predicted by the HMTFS method and the actual task difficulty, showing a correlation with $r = 0.278$ and $p < 0.001$. The right panel shows the variations in predicted task difficulty at different times during the math task.

5. Discussion

In this study, we proposed a hypergraph-based multitask feature selection framework with a threshold-based hypergraph construction method. By bridging the hypergraph structure for mining high-order correlation between subjects and multitask learning for preserving temporally constrained group sparsity, HMTFS allows us to select the most informative subset of fMRI features [35,41]. As demonstrated in Figure 2a, brain regions selected by HMTFS are more stable and clustered than by group lasso method. Additionally, HMTFS outperforms other multitask feature selection models and traditional graph-based methods in classifying resting state and motor state by effectively identifying task-related brain regions. These results demonstrate that the hypergraph structure effectively models the data, thereby further enhancing the performance of multitask feature selection. To further analyze the effectiveness of each component within HMTFS, the ablation study results presented in Table 2 highlight the importance of the multitask smooth regularization and hypergraph regularization components for feature selection. Hypergraph construction methods directly affect the quality of hypergraph learning [41–43]. As shown in Figure 2f,g, the threshold-based method achieves better feature selection results than the kNN-based method. Technically, kNN constructs a hypergraph by picking the nearest k neighbors; in contrast, the threshold-based method does not fix the number of neighbors, which provides flexibility to better capture higher-order correlations between subjects. In line with previous studies [15,44], the hypergraph exhibits a sweet-spot threshold, and the selected features of the hypergraph constructed at this sweet spot achieve the best classification performance. These findings highlight the benefits of using a threshold-based method for hypergraph construction and the importance of choosing an appropriate threshold, which both impact the quality of data modeling. Consequently, additional research is necessary to develop more effective hypergraph construction techniques for diverse datasets. Such advancements will improve the generalizability and robustness of hypergraph-based models in capturing the inherent complex relationships within the data.

The HMTFS method has the potential to be a valuable interpretive tool in neuroscience. We utilized well-known motor tasks to assess the efficacy of the HMTFS method. During motor tasks, participants responded to visual cues by tapping their fingers, squeezing their toes, or moving their tongues [45]. HMTFS more effectively select motor-related brain regions compared to other feature selection models, as show in Figures 2 and 4. These regions include the premotor, primary motor, and primary somatosensory cortices, aligning well with the expected locations identified in prior neuroscience research [46,47]. The visualization of brain regions for motor tasks clearly illustrates the lateralization of brain function [48]. Specifically, the left motor cortex controls movements on the right side of the body and vice versa. Additionally, the motor cortex is organized into a continuous somatotopic homunculus, which controls body movements from the feet to the face [47,49]. The brain regions selected by the HMTFS method for each motor task correspond well with the somatotopic homunculus. The task-related brain regions identified by the HMTFS method effectively reflect and characterize the structure of tasks, as shown in Figure 5. The motor-related brain regions selected by HMTFS demonstrate an excellent fit to the corresponding task structures, with R^2 values consistently exceeding 0.6, surpassing those of visual-related regions. The visual-related brain regions also achieve R^2 values around 0.4, likely influenced by the visual stimuli presented during motor tasks. The visualization of brain region weights and t-SNE validation demonstrate the effectiveness of the HMTFS method in selecting task-related brain regions for various cognitive tasks in the HCP dataset, as shown in Figure 3. For instance, the working memory task involves sustained activity in the prefrontal cortex, which aligns with the brain regions selected by the HMTFS method [50]. Similarly, both the story and math tasks involve auditory processing of sounds and comprehension of speech, which is managed by the superior temporal gyrus [51,52]. The brain regions selected by the HMTFS method are consistent with this area. These results indicate that HMTFS effectively localizes task-related brain regions, which not only aids in understanding the functional mechanisms of the brain under various cognitive

tasks but also demonstrates the potential of HMTFS to identify brain regions associated with specific diseases.

Figure 6 demonstrates that different cognitive tasks engage distinct brain regions, exhibiting a distributed pattern. During complex cognitive tasks, multiple brain regions are activated simultaneously, rather than a single area working in isolation. This highlights the distributed processing of brain, particularly evident in functions related to working memory and semantic cognition [53,54]. Additionally, the mapping of cognitive tasks to brain regions illustrates the overlap among different tasks, which reflects the similarity between cognitive tasks. Tasks that perform similar functions often show overlapping activation in brain regions [55,56]. To further explore the applicability of the HMTFS method in neuroscience, we used the difficulty of language tasks as a test case. Figure 7 illustrates that the brain regions related to task difficulty identified by the HMTFS method not only predict the difficulty of the task but also reflect the progression of task difficulty over time. As the language tasks progress, both the story and math tasks exhibit an increase in difficulty, peaking during the question-answering phase at the end of the tasks. In summary, HMTFS can serve as a valuable interpretative tool in neuroimaging, capable of selecting task-related brain regions and applicable to more complex feature selections related to dynamic task challenges. The limitations of the HMTFS method include the necessity of constructing a hypergraph in advance, which introduces additional computational overhead as well. This could pose challenges when dealing with large-scale datasets. Furthermore, different settings in hypergraph construction affect its ability to model the data, thereby influencing the performance of HMTFS. This study primarily explored the efficacy of HMTFS in fMRI data. Future research should investigate its applicability in other neuroimaging modalities. In addition to localizing task-related brain regions, it is also crucial to explore the potential of HMTFS in identifying disease-related brain regions.

6. Conclusions

In conclusion, we proposed the HMTFS framework, which allows us to make full use of temporal information as well as high-order correlations between data samples. We verified that HMTFS can better capture task-related brain regions compared to other feature selection methods. In future, HMTFS has great potential as a feature selection tool in neuroscience to deepen our understanding of brain function.

Author Contributions: Conceptualization, Y.Q., Y.Z., H.W. and Q.L.; methodology, Y.Q., K.F. and L.W.; validation, Y.Q., K.F. and L.W.; writing—original draft preparation, Y.Q., K.F. and L.W.; writing—review and editing, Y.Z., H.W. and Q.L.; Y.Z., H.W. and Q.L. provided guidance and edits at various stages. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (2021YFF1200804), Shenzhen Science and Technology Innovation Committee (2022410129, KCXFZ20201221173400001, KJZD20230923115221044).

Data Availability Statement: In this study, we used a public large-scale fMRI dataset: Human Connectome Project (HCP) S1200 dataset. The HCP dataset is available at <https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release> (accessed on 1 March 2022).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, Y.; Yang, Q. A survey on multi-task learning. *IEEE Trans. Knowl. Data Eng.* **2021**, *34*, 5586–5609. [\[CrossRef\]](#)
2. Yuan, M.; Lin, Y. Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2006**, *68*, 49–67. [\[CrossRef\]](#)
3. Jie, B.; Liu, M.; Liu, J.; Zhang, D.; Shen, D. Temporally constrained group sparse learning for longitudinal data analysis in Alzheimer's disease. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 238–249. [\[CrossRef\]](#)
4. Zhou, J.; Liu, J.; Narayan, V.A.; Ye, J.; Modeling Disease Progression via Multi-Task Learning. Modeling disease progression via multi-task learning. *NeuroImage* **2013**, *78*, 233–248. [\[CrossRef\]](#)

5. Liu, H.; Palatucci, M.; Zhang, J. Blockwise coordinate descent procedures for the multi-task lasso, with applications to neural semantic basis discovery. In Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; pp. 649–656.
6. Turlach, B.A.; Venables, W.N.; Wright, S.J. Simultaneous variable selection. *Technometrics* **2005**, *47*, 349–363. [\[CrossRef\]](#)
7. Zhou, Y.; Jin, R.; Hoi, S.C.H. Exclusive lasso for multi-task feature selection. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; JMLR Workshop and Conference Proceedings; pp. 988–995.
8. Lozano, A.C.; Swirszcz, G. Multi-level lasso for sparse multi-task regression. In Proceedings of the 29th International Conference on Machine Learning, Edinburgh, UK, 26 June–1 July 2012; pp. 595–602.
9. Arthurs, O.J.; Boniface, S. How well do we understand the neural origins of the fMRI BOLD signal? *Trends Neurosci.* **2002**, *25*, 27–31. [\[CrossRef\]](#)
10. Tzourio-Mazoyer, N.; Landeau, B.; Papathanassiou, D.; Crivello, F.; Etard, O.; Delcroix, N.; Mazoyer, B.; Joliot, M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **2002**, *15*, 273–289. [\[CrossRef\]](#)
11. Glasser, M.F.; Coalson, T.S.; Robinson, E.C.; Hacker, C.D.; Harwell, J.; Yacoub, E.; Ugurbil, K.; Andersson, J.; Beckmann, C.F.; Jenkinson, M.; et al. A multi-modal parcellation of human cerebral cortex. *Nature* **2016**, *536*, 171–178. [\[CrossRef\]](#)
12. Logothetis, N.K. What we can do and what we cannot do with fMRI. *Nature* **2008**, *453*, 869–878. [\[CrossRef\]](#)
13. Buckner, R.L.; Krienen, F.M.; Yeo, B.T. Opportunities and limitations of intrinsic functional connectivity MRI. *Nat. Neurosci.* **2013**, *16*, 832–837. [\[CrossRef\]](#)
14. Zheng, S.; Punia, D.; Wu, H.; Liu, Q. Graph theoretic analysis reveals intranasal oxytocin induced network changes over frontal regions. *Neuroscience* **2021**, *459*, 153–165. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Shao, W.; Peng, Y.; Zu, C.; Wang, M.; Zhang, D.; Alzheimer’s Disease Neuroimaging Initiative. Hypergraph based multi-task feature selection for multimodal classification of Alzheimer’s disease. *Comput. Med. Imaging Graph.* **2020**, *80*, 101663. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Yin, W.; Qu, Y.; Ma, Z.; Liu, Q. Hyperntf: A hypergraph regularized nonnegative tensor factorization for dimensionality reduction. *Neurocomputing* **2022**, *512*, 190–202. [\[CrossRef\]](#)
17. Gao, S.; Tsang, I.W.H.; Chia, L.T. Laplacian sparse coding, hypergraph laplacian sparse coding, and applications. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 92–104. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Zhang, Y.; Yang, Q. An overview of multi-task learning. *Natl. Sci. Rev.* **2018**, *5*, 30–43. [\[CrossRef\]](#)
19. Graham, S.; Vu, Q.D.; Jahanifar, M.; Raza, S.E.A.; Minhas, F.; Snead, D.; Rajpoot, N. One model is all you need: Multi-task learning enables simultaneous histology image segmentation and classification. *Med. Image Anal.* **2023**, *83*, 102685. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Zhang, Y.; Wang, J.; Liu, Y.; Rong, L.; Zheng, Q.; Song, D.; Tiwari, P.; Qin, J. A multitask learning model for multimodal sarcasm, sentiment and emotion recognition in conversations. *Inf. Fusion* **2023**, *93*, 282–301. [\[CrossRef\]](#)
21. Zhang, D.; Shen, D.; Alzheimer’s Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer’s disease. *NeuroImage* **2012**, *59*, 895–907. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Yuan, X.T.; Liu, X.; Yan, S. Visual classification with multitask joint sparse representation. *IEEE Trans. Image Process.* **2012**, *21*, 4349–4360. [\[CrossRef\]](#)
23. Obozinski, G.; Taskar, B.; Jordan, M. Multi-task feature selection. *Stat. Dep. Berkeley Tech. Rep.* **2006**, *2*, 2.
24. Lee, S.; Zhu, J.; Xing, E. Adaptive multi-task lasso: With application to eQTL detection. In Proceedings of the 24th Annual Conference on Neural Information Processing Systems 23 Vancouver, BC, Canada, 6–9 December 2010.
25. Olshausen, B.A.; Field, D.J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **1996**, *381*, 607–609. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Tang, S.; Zhang, Y.; Li, Z.; Li, M.; Liu, F.; Jiang, H.; Lee, T.S. Large-scale two-photon imaging revealed super-sparse population codes in the V1 superficial layer of awake monkeys. *eLife* **2018**, *7*, e33370. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Duan, Y.; Wang, N.; Zhang, Y.; Song, C. Tensor-Based Sparse Representation for Hyperspectral Image Reconstruction Using RGB Inputs. *Mathematics* **2024**, *12*, 708. [\[CrossRef\]](#)
28. Li, Y.; Gui, G.; Cheng, X. From group sparse coding to rank minimization: A novel denoising model for low-level image restoration. *Signal Process.* **2020**, *176*, 107655. [\[CrossRef\]](#)
29. Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.S.; Ma, Y. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *31*, 210–227. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Nguyen, T.S.; Luong, M.; Kaaniche, M.; Ngo, L.H.; Beghdadi, A. A novel multi-branch wavelet neural network for sparse representation based object classification. *Pattern Recognit.* **2023**, *135*, 109155. [\[CrossRef\]](#)
31. Lee, H.; Battle, A.; Raina, R.; Ng, A. Efficient sparse coding algorithms. *Adv. Neural Inf. Process. Syst.* **2006**, *19*.
32. Donoho, D.L. For most large underdetermined systems of linear equations the minimal l_1 -norm solution is also the sparsest solution. *Commun. Pure Appl. Math. J. Issued Courant Inst. Math. Sci.* **2006**, *59*, 797–829. [\[CrossRef\]](#)
33. Murgas, K.A.; Saucan, E.; Sandhu, R. Hypergraph geometry reflects higher-order dynamics in protein interaction networks. *Sci. Rep.* **2022**, *12*, 20879. [\[CrossRef\]](#)
34. Wang, F.; Hu, F.; Chen, R.; Xiong, N. HLEGF: An Effective Hypernetwork Community Detection Algorithm Based on Local Expansion and Global Fusion. *Mathematics* **2023**, *11*, 3497. [\[CrossRef\]](#)

35. Zhou, D.; Huang, J.; Schölkopf, B. Learning with hypergraphs: Clustering, classification, and embedding. In Proceedings of the 20th Annual Conference on Neural Information Processing Systems 19, Vancouver, BC, Canada, 4–7 December 2006.
36. Agarwal, S.; Branson, K.; Belongie, S. Higher order learning with graphs. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA, 25–29 June 2006; pp. 17–24.
37. Van Essen, D.C.; Ugurbil, K.; Auerbach, E.; Barch, D.; Behrens, T.E.; Bucholz, R.; Chang, A.; Chen, L.; Corbetta, M.; Curtiss, S.W.; et al. The Human Connectome Project: A data acquisition perspective. *Neuroimage* **2012**, *62*, 2222–2231. [[CrossRef](#)] [[PubMed](#)]
38. Glasser, M.F.; Sotiropoulos, S.N.; Wilson, J.A.; Coalson, T.S.; Fischl, B.; Andersson, J.L.; Xu, J.; Jbabdi, S.; Webster, M.; Polimeni, J.R.; et al. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **2013**, *80*, 105–124. [[CrossRef](#)] [[PubMed](#)]
39. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
40. Zhang, J.; Miao, J.; Zhao, K.; Tian, Y. Multi-task feature selection with sparse regularization to extract common and task-specific features. *Neurocomputing* **2019**, *340*, 76–89. [[CrossRef](#)]
41. Gao, Y.; Zhang, Z.; Lin, H.; Zhao, X.; Du, S.; Zou, C. Hypergraph learning: Methods and practices. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *44*, 2548–2566. [[CrossRef](#)] [[PubMed](#)]
42. Bouhlel, N.; Feki, G.; Ben Ammar, A.; Ben Amar, C. Hypergraph learning with collaborative representation for image search reranking. *Int. J. Multimed. Inf. Retr.* **2020**, *9*, 205–214. [[CrossRef](#)]
43. Gao, Y.; Wang, M.; Tao, D.; Ji, R.; Dai, Q. 3-D object retrieval and recognition with hypergraph analysis. *IEEE Trans. Image Process.* **2012**, *21*, 4290–4303. [[CrossRef](#)] [[PubMed](#)]
44. Liu, M.; Zhang, J.; Guo, X.; Cao, L. Hypergraph regularized sparse feature learning. *Neurocomputing* **2017**, *237*, 185–192. [[CrossRef](#)]
45. Buckner, R.L.; Krienen, F.M.; Castellanos, A.; Diaz, J.C.; Yeo, B.T. The organization of the human cerebellum estimated by intrinsic functional connectivity. *J. Neurophysiol.* **2011**, *106*, 2322–2345. [[CrossRef](#)]
46. Barch, D.M.; Burgess, G.C.; Harms, M.P.; Petersen, S.E.; Schlaggar, B.L.; Corbetta, M.; Glasser, M.F.; Curtiss, S.; Dixit, S.; Feldt, C.; et al. Function in the human connectome: Task-fMRI and individual differences in behavior. *Neuroimage* **2013**, *80*, 169–189. [[CrossRef](#)] [[PubMed](#)]
47. Gordon, E.M.; Chauvin, R.J.; Van, A.N.; Rajesh, A.; Nielsen, A.; Newbold, D.J.; Lynch, C.J.; Seider, N.A.; Krimmel, S.R.; Scheidter, K.M.; et al. A somato-cognitive action network alternates with effector regions in motor cortex. *Nature* **2023**, *617*, 351–359. [[CrossRef](#)]
48. Güntürkün, O.; Ströckens, F.; Ocklenburg, S. Brain lateralization: A comparative perspective. *Physiol. Rev.* **2020**, *100*, 1019–1063. [[CrossRef](#)] [[PubMed](#)]
49. Penfield, W.; Boldrey, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain* **1937**, *60*, 389–443. [[CrossRef](#)]
50. Curtis, C.E.; D’Esposito, M. Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* **2003**, *7*, 415–423. [[CrossRef](#)] [[PubMed](#)]
51. Bhaya-Grossman, I.; Chang, E.F. Speech computations of the human superior temporal gyrus. *Annu. Rev. Psychol.* **2022**, *73*, 79–102. [[CrossRef](#)] [[PubMed](#)]
52. Yi, H.G.; Leonard, M.K.; Chang, E.F. The encoding of speech sounds in the superior temporal gyrus. *Neuron* **2019**, *102*, 1096–1110. [[CrossRef](#)] [[PubMed](#)]
53. Christophel, T.B.; Klink, P.C.; Spitzer, B.; Roelfsema, P.R.; Haynes, J.D. The distributed nature of working memory. *Trends Cogn. Sci.* **2017**, *21*, 111–124. [[CrossRef](#)] [[PubMed](#)]
54. McClelland, J.L.; Rogers, T.T. The parallel distributed processing approach to semantic cognition. *Nat. Rev. Neurosci.* **2003**, *4*, 310–322. [[CrossRef](#)] [[PubMed](#)]
55. Najafi, M.; McMenamin, B.W.; Simon, J.Z.; Pessoa, L. Overlapping communities reveal rich structure in large-scale brain networks during rest and task conditions. *Neuroimage* **2016**, *135*, 92–106. [[CrossRef](#)]
56. Qu, Y.; Jian, X.; Che, W.; Du, P.; Fu, K.; Liu, Q. Transfer learning to decode brain states reflecting the relationship between cognitive tasks. In Proceedings of the International Workshop on Human Brain and Artificial Intelligence, Vienna, Austria, 23–25 July 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 110–122.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.