A hybrid method to select morphometric features using tensor completion and F-score rank for gifted children identification

ZHANG Jin†, FENG Fan‡, HAN TianYi, DUAN Feng*, SUN Zhe†, CAIIFA Cesar F., SOLÉ-CASALS Jordi‡,5,6*

1 College of Computer Science, Nankai University, Tianjin 300350, China; 2 Department of Artificial Intelligence, Nankai University, Tianjin 300350, China; 3 Computational Engineering Applications Unit, Head Office for Information Systems and Cybersecurity, RIKEN, 351-0198 Saitama, Japan; 4 Instituto Argentino de Radioastronomía-CCT La Plata, CONICET/CIC-PBA/UNLP, 1894 V. Elisa, Argentina; 5 Department of Psychiatry, University of Cambridge, Cambridge CB2 0SZ, United Kingdom; 6 Data and Signal Processing Group, University of Vic-Central University of Catalonia, 08500 Vic, Catalonia, Spain

Received November 30, 2020; accepted June 18, 2021; published online August 10, 2021

Gifted children are able to learn in a more advanced way than others, probably due to neurophysiological differences in the communication efficiency in neural pathways. Topological features contribute to understanding the correlation between the brain structure and intelligence. Despite decades of neuroscience research using MRI, methods based on brain region connectivity patterns are limited by MRI artifacts, which therefore leads to revisiting MRI morphometric features, with the aim of using them to directly identify gifted children instead of using brain connectivity. However, the small, high-dimensional morphometric feature dataset with outliers makes the task of finding good classification models challenging. To this end, a hybrid method is proposed that combines tensor completion and feature selection methods to handle outliers and then select the discriminative features. The proposed method can achieve a classification accuracy of 93.1%, higher than other existing algorithms, which is thus suitable for the small MRI datasets with outliers in supervised classification scenarios.

Citation: Zhang J, Feng F, Han T Y, et al. A hybrid method to select morphometric features using tensor completion and F-score rank for gifted children identification. Sci China Tech Sci, 2021, 64: 1863–1871, https://doi.org/10.1007/s11431-020-1876-3

1 Introduction

Intelligence can be defined as the ability to perceive, comprehend and infer information within a context. Gifted individuals are those with outstanding cognitive abilities and creativity and are said to possess giftedness [1]. They not only have a higher intellectual ability, but also learn faster than the average individual in a quantitatively different way, presumably due to neurophysiological differences [2]. The neurological differences mean that gifted individuals may experience different neural development trajectories than neurotypical individuals during their childhood, leading to greater inter-connectivity between neural pathways [3].

To map the network of anatomically connected regions in an individual human brain, several tools for magnetic resonance imaging (MRI) were developed during decades of neuroscience research, which provided the opportunity to identify gifted children more qualitatively from the perspective of neural network connectivity instead of simply using intelligence quotient (IQ) test scales for quantitative testing. There are two approaches that can test anatomical connectivity: tractography extracted from diffusion-weighted imaging (DWI) [4–6] and morphology patterns calcu-
lated from structural covariance network (SCN) [7–9].

Diffusion-weighted tractography can visually reconstruct the trajectory of white matter by tracking the main diffusion directions of water molecules. However, these are disturbed by head movements [10] and the method generates a large number of false-positive connections [11].

Structural covariance analysis can construct whole-brain networks in two steps. Firstly, for each brain region in multi-model imaging data, one single morphometric feature is measured. After that, the covariance between each region is estimated, resulting in a single SCN [12,13]. However, the performance of this method heavily depends on the size of the collected MRI data [14].

Due to the artifacts in the MRI acquisition process and the small size of available samples in the dataset, these two methods are still insufficient to establish the human cortical connectivity network, which is not suitable for small MRI dataset classification.

Thus, we turn our attention to the MRI data itself, and use the morphometric features instead of the brain cortical connectivity to identify gifted children.

On the one hand, due to the limitations of the MRI acquisition device or the scan parameters, the MRI artifact may affect the morphometric feature extraction, resulting in outliers which hinder the performance improvement of the classification model. Recently, research in tensor completion (TC), which is a higher-order extension of matrix completion, has achieved a consistent performance in a variety of real-world applications [15–18]. Given a tensor with incomplete entries, tensor completion methods such as recent low-rank based tensor completion (LRTC) [19] and simultaneous tensor decomposition and completion (STDC) [20] use low-rank factorization to model the available data entries, which are then used for completion to recover the missing values [21–23]. TC methods have already been used in electroencephalography (EEG) to recover missing samples, showing better performances than other simple imputation methods [24,25].

On the other hand, in machine learning scenarios, typically an enormous number of samples are required to ensure that there are enough samples to learn the rules of a high-dimensional space. The gifted children MRI dataset used in this paper contains 2156 features (7 morphometric features with 308 brain regions), which only consists of 29 samples. It is quite difficult to train classification models using such a small, high-dimensional datasets. Feature selection methods, such as principal components analysis (PCA) [26] and F-score [27] can reduce the feature dimension, which will be applied to enhance the classification models [28–30].

To address these two problems, here a hybrid method is proposed to select morphometric features and brain regions from the gifted children MRI dataset. Firstly, with a brain parcellation template, the morphometric feature matrix for each sample is extracted. Then, outlier masks for both the gifted and control groups are generated. These two group outlier masks and their morphometric features matrices are used as the input of the STDC algorithm. The whole outlier completed dataset is then split into a training set and a validation set. F-score rank algorithm, the proposed feature selection method, is then applied on the training set to obtain the features mask, which is applied on both the training and validation sets. Multiple machine learning techniques are then used to carry out the classification.

With this hybrid method, the well-trained classification models can achieve a 93.1% accuracy. Hence, the proposed method can be a good alternative to identify morphometric features and brain regions related to giftedness and can be applied to other small, high-dimensional datasets.

The rest of the paper is organized as follows. Details on the tensor completion method, the feature selection method and the experiment settings will be outlined in Section 2. Results on various feature selection methods and features of the giftedness will be discussed in Section 3. Finally, a conclusion will be drawn in Section 4.

2 Materials and methods

2.1 Gifted-children MRI dataset

The gifted children MRI dataset [12] used in this paper consists of 29 healthy right-handed male subjects with no history of neurological disorders. Table 1 shows the dataset details. The groups were not significantly different in terms of age. These gifted children not only have a superior IQ but can also achieve a high performance in various types of tasks, such as spatial, numerical, verbal reasoning, abstracting reasoning, and memory tasks [31].

With the same scanning procedures and parameters used in [12], all participants were examined on a 3T MRI scanner (Magnetom Trio Tim, Siemens Medical Systems). The raw (anonymised) MRI dataset are available in the Open-Neuro repository (https://openneuro.org/datasets/ds001988).

2.2 Brain region atlas and morphometric features

Each individual brain was parcellated into 308 cortical regions. The parcellation atlas was constructed based on the Desikan-Killiany atlas (68 regions) [32]. Each region defined in the Desikan-Killiany atlas was sub-parcellated into spa-

| Table 1 Class membership information of gifted children dataset |
|---------------------------------------------------------------|
| Group            | Control group (CG) | Gifted group (GG) |
| Subject count   | 14                | 15               |
| Average age     | 12.53             | 12.03            |
| Average IQ      | 122.71            | 148.80           |
tially contiguous regions by the backtrack algorithm [33] in FreeSurfer, so that the final parcels could be constrained by the original anatomical boundaries. All regions in the new parcellation were approximately equal in size (500 mm$^2$).

The original feature matrix for each sample consists of 7 morphometric features measured at each of the 308 brain regions. For each brain region, surface and volume-based morphometric features were estimated. Figure 1 shows the morphometric features (grey matter volume, cortical thickness, surface area, intrinsic curvature, mean curvature, curvature index, and folding index).

### 2.3 Outlier detection method

The raw morphometric feature matrix contains outliers due to the acquisition procedure and the applied pre-processing. The interquartile range (IQR) method based on box plots was used to detect outliers, as shown in Figure 2.

The difference between $Q_1$ and $Q_3$ is called the IQR:

$$
IQR = Q_3 - Q_1.
$$

(1)

The decision boundary is set to 1.5 times the IQR. Any morphometric feature value which is smaller than the lower boundary (LB) or larger than the upper boundary (UB) should be considered as an outlier.

$$
LB = Q_1 - 1.5 \times IQR,
UB = Q_3 + 1.5 \times IQR.
$$

(2)

![Figure 1](image)

**Figure 1** (Color online) Pipeline for morphometric feature extraction. Each brain was parcellated into 308 regions. Volume and surface parcellation templates were used. All MRI data were mapped to the same cortical parcellation templates to produce a 7 × 308 feature matrix for each subject.

2.4 Outlier completion method: STDC

Tensor representation has been widely studied for multiple scenarios where only a subset of entries is missing. Considering an $n^{th}$ order tensor $X$, the Canonical-Polyadic (CP) tensor decomposition [34] can be described as follows:

$$
X = \sum_{r=1}^{R} \alpha_r u_r \otimes u_r \otimes \cdots \otimes u_r,
$$

(3)

where $u_r$ is the decomposed factor along the $k^{th}$ mode, which can be referred to as the $k^{th}$ mode factor.

Given an $n^{th}$ order tensor $X_0 \in \mathbb{R}^{l_1 \times l_2 \times \cdots \times l_n}$ and a subset of missing entries $Q$, the tensor completion methods try to find a tensor $X'$ with its components $Z, V_1, \cdots, V_n$ such that $X'_r$ and $X'$ have the same observed entries, which can be described as follows:

$$
X = Z \times V_1^T \times \cdots \times V_n^T
$$

and

$$
\Omega(X) = \Omega(X_0),
$$

(4)

where $Z$ is an $n^{th}$ order tensor of the same size as $X_0$, and each $V_i$ denotes a $l_i \times l_i$ matrix. If $X'$ is a low-rank tensor, then the core tensor $Z$ is of low rank, or $V_1, \cdots, V_n$ are a set of low-rank matrices. Such a low-rank property is usually regarded as a global prior in tensor completion.

To find the proper $X', Z, V_1, \cdots, V_n$, the STDC method defines an augmented Lagrange function as follows:

$$
L(X', Z, V_1, \cdots, V_n, \lambda, \mu) = \sum_{k=1}^{n} \alpha_k \|V_k\|_F^2 + \mu \text{tr}(V_k \otimes \cdots \otimes V_k) L(V_k \otimes \cdots \otimes V_k) + \gamma \|Z - Z \times V_1^T \times \cdots \times V_n^T\|_F^2 + \frac{\mu}{2} \|X' - Z \times V_1^T \times \cdots \times V_n^T\|_F^2.
$$

(5)

The STDC [20] method estimates the decomposed latent factors using partially observed data. It is possible to compute the missing entries from the estimated latent factors. The factor prior, or regularization, is often applied on decomposed components based on low-rank structure [19] in the STDC framework. Patterns can then be obtained in latent matrices, yielding the minimum number of rank-one tensors. The STDC [20] can be described as Algorithm 1.

The raw morphometric feature matrix is treated as the incomplete $n^{th}$ tensor $X_0$. For both gifted and control groups, their outlier mask matrix is treated as matrix $V_k$.

2.5 Feature selection methods

After tensor completion, the outliers in the feature matrix for both the gifted and control groups will be completed. After splitting the dataset into training sets and validation sets, several feature selection methods, as well as the proposed F-score rank method, are applied only on the training sets to explore which morphometric features and brain regions can
Algorithm 1: STDC

**Input:** an incomplete tensor $X \in \mathbb{R}^{n \times l \times r \times s}$, the parameters $\alpha, \beta, \gamma, \mu, \nu$

1. Initialize $X, Y, Z, V_0, \ldots, V_n$ by
   
   $V_i$: identify matrix
   $X, Z: X_0$
   $Y$: a tensor with all zero elements
2. Let $t = 1$ and $\mu = \mu$
3. Do
   
   For $t = 1$ to $n$
   
   Optimize $V_i, X, Z$
   
   Update $Y = Y + \mu (X - Z) \times V_i^T \times \cdots \times V_n^T$
   
   $\mu = \rho \mu, \rho \in [1.1, 1.2]$
   
   While $\|X - Z \times V_1^T \times \cdots \times V_n^T\|_F > 10^9 \|V_n\|_F$

**Output:** the submanifolds $V_1, \ldots, V_n$

the complete tensor $Z$

achieve a better performance on gifted children identification. A features mask is then generated to indicate the selected morphometric features and brain region indices.

1. **NONE feature selection**
   
   For each subject, a feature matrix is defined, which contains 7 morphometric features of 308 brain regions. For each feature matrix, a one-dimensional feature vector can be generated through vectorization. In this case, all the features in the raw feature matrix are used and are put into the classification models. In this paper, this is called the NONE feature selection method.

2. **VON feature selection**
   
   The von Economo atlas [35] subdivides the cortex into five different categories, according to the laminar structure of the brain cortex and its corresponding functional cortical specializations. Regions located in the primary motor cortex are classified as type 1. Association cortices are referred to as type 2 and type 3, while secondary and primary sensory areas are type 4 and type 5, respectively.

   For type 2 and type 3 brain cortex regions, there are significant differences in the neurobiological substrate of versatility between the gifted and control groups in terms of the cortical thickness (CT) morphometric features [12]. Thus, it is reasonable to select those types of brain regions as features for the classification system.

   In this case, all 7 morphometric features from these corresponding types 2 and 3 brain regions are treated as features, and in this paper, this is called the VON feature selection method.

3. **F-score feature selection**
   
   Given the training vectors $x_k, k = 1, \ldots, m$, the size of the gifted group $p$ and the size of the control group $q$, the F-score [27] of the $i^{th}$ feature can be defined as

$$F(i) = \frac{(p-1)(q-1)(\bar{x}_p^i - \bar{x}_q^i)^2 + (\bar{x}_q^i - \bar{x}_p^i)^2}{\sum_{k=1}^{p} (x_k^i - \bar{x}_p^i)^2 + \sum_{k=1}^{q} (x_k^i - \bar{x}_q^i)^2},$$

where $\bar{x}_p^i, \bar{x}_q^i$ are the average of the whole, gifted and control groups for the $i^{th}$ feature, respectively, and $x_k^p, x_k^q$ are the $i^{th}$ feature of the $k^{th}$ subject for each group.

   The numerator indicates the discrimination between the gifted and control sets, and the denominator indicates the dispersion within each of the two sets.

   In this case, a threshold is set to select the top highest features from the morphometric features and brain regions. In this paper, this method will be referred to as the F-score (FS) feature selection method.

4. **F-score rank feature selection**
   
   For the original F-score algorithm, the input will be a one-dimensional feature vector in which the morphometric features and brain regions are feature candidates which are all treated as independent items. However, the functions performed by adjacent brain regions may be similar, and the correlation between morphometric features and brain regions may be discarded by this method. Thus, to obtain more interpretable features and to further reduce the number of features, here the F-score rank feature selection method is proposed, which can cluster brain regions according to their correlation with morphometric features.

   Figure 3 shows the pipeline of the proposed F-score rank feature selection method. Firstly, the F-score value of 7 morphometric features for each brain region is calculated among the training set, yielding a two-dimensional F-score matrix. For each region, the F-score values are sorted by descending order, where the morphometric feature with the highest F-score value should be put in the first column. The morphometric feature and brain region indexes corresponding to the highest F-score value are considered to have the strongest correlation. Hence, only the first column is picked up and the number of brain regions for each group is counted. In this step, all 308 brain regions will be categorized into 7 groups (morphometric features). Finally, as with the F-score method, a threshold is set to select the morphometric features and brain regions from the statistical results, and a feature mask will be generated to define the feature index for the classification model input.

   In this case, a subset of both the morphometric features and the brain regions are selected by categorizing brain regions into morphometric feature groups. In this paper, this method will be referred to as the F-score rank (FSR).

5. **Combined feature selection**
   
   Furthermore, the VON and F-score methods were combined to create a joint method. In this case, only type 2 and type 3 regions are considered when calculating the F-score value for morphometric features. In this paper, this method is referred to as the VFS feature selection method.
For all the feature selection methods, a feature mask will be generated, which can indicate the selected feature indices. Then, the feature mask is mapped to the training set for model training and to the validation set for evaluation.

2.6 Classification model and cross validation

Since the gifted MRI dataset only contains 29 subjects, traditional machine learning classification methods, such as the K-Nearest Neighbor (KNN) and the support vector machine (SVM), are considered in this paper instead of using deep learning network (DNN) methods.

Considering the size of the dataset, it is hard to execute k-fold validation. Thus, we applied the Leave-One-Out (LOO) cross validation, where the number of folds equals the number of instances in the dataset.

2.7 Evaluation metric

The accuracy (ACC) was calculated, which measures the proportion of correctly classified subjects for each group, to evaluate the performance of outlier completion methods and feature selection methods.

The proposed hybrid method for selecting discriminative features is evaluated in three steps. In the first step, the performance on outlier detection and completion is evaluated. The performance of the classification models before and after the outlier completion was tested without using any feature selection methods. In the second step, the classification performance of the several feature selection methods mentioned in this paper was evaluated. Furthermore, these feature selection methods were tested on both the original feature matrix and the completed feature matrix. In the last step, the number of selected features in the F-score rank method and the most discriminative features are discussed.

2.8 Hyperparameters and experimental settings

The number of features is the most important hyperparameter for each feature selection method. For the NONE feature selection method, all feature candidates are used for classification, so therefore the number of features is $7 \times 308$. The VON method only selected brain regions of type 2 (113 regions) and type 3 (71 regions). Based on the experimental results, the F-score method can achieve the best performance when the number of features is set to 200. For the F-score rank method, however, the maximum number of features is 308, when all morphometric features are selected. Choosing the top 3 morphometric features and brain regions within their categories can achieve the best performance, according to the experimental results.

For all the experiments reported in this paper, a Linux server running Ubuntu 16.04 was used. The server contains one Intel Core i7-6700 processor running at 2.6 GHz.

In summary, Figure 4 shows the whole pipeline for selecting discriminative features with our hybrid method.

Firstly, with the help of brain region parcellation, the morphometric features are estimated to form the original feature matrix, which contains many outliers. Using the IQR method, an outlier mask will be generated to indicate the
missing parts for the tensor completion method. Both the original feature matrix and its outlier mask will be used as the input of the tensor completion method. Then, the outlier completed matrix will be split into a training and a validation set, and feature selection methods will be executed during the LOO cross-validation. The selected feature indices will be mapped to the training set and the validation set, respectively. Finally, the training set feature vector will be used for training purposes and the validation set feature vector (only in one instance) will be used for evaluation purposes.

3 Results and discussion

3.1 Outlier detection and completion

(1) Outlier detection results

The results for outlier detection using IQR methods are presented in Figure 5 (for each morphometric feature) and Figure 6.

Figure 5 shows the number of outliers for each morphometric feature. The last three features have many more outliers than the first four features, which may be caused by the feature estimating method itself. Despite only 4% of the data

Figure 5 (Color online) Number of outliers for 7 morphometric features in the gifted and control groups. The last three features have the most outliers.
being outliers, these still have a huge impact on the accuracy of the classification.

The number of outliers for each subject was counted, as shown in Figure 6. It can be observed that the 7th subject in the gifted group (GG) and the 14th subject in the control group (CG) have the most outliers among all the subjects, which could make it hard for the classification model to correctly identify these two samples. In addition, all the subjects have outliers, which indicate that it is necessary to detect and fill them properly.

(2) Outlier completion with STDC

After handling the outliers, another outlier detection was conducted to evaluate the performance of the STDC method on outlier elimination, as presented in Figure 7.

In this step, the STDC was iteratively applied 3 times on the STDC output completed matrix. For each epoch, the previous completed matrix and its outlier mask served as the new input for the STDC method.

Compared to Figure 5, Figure 7 shows that the number of outliers is significantly decreased with the STDC method. This superior performance of the method can be seen as the number of outliers is dropped by an average of 3 times.

Moreover, the classification performance on the completed feature matrix was tested using the median value filling method, LRTC, and STDC by SVM and KNN.

The ACC results in Table 2 show that, after handling the outliers properly, the model can achieve a higher accuracy. It can be seen that the tensor completion algorithms such as LRTC and STDC are more suitable to handle the outliers.

3.2 Feature selection method performances

In this step, all the previously mentioned feature selection methods and our proposed F-score rank method were evaluated. These methods were also applied on the original morphometric feature matrix to remove the impact of STDC and thus depict a clearer evaluation of the methods.

Figure 8 shows the accuracy of SVM and KNN for each feature selection method applied to the original feature matrix and to the outlier completed feature matrix.

![Figure 6](image1.png)

**Figure 6** (Color online) Number of outliers for each subject in the gifted and control groups. All subjects have more or less outliers.

![Figure 7](image2.png)

**Figure 7** (Color online) Number of outliers for 7 morphometric features after iteratively applying the STDC method 3 times. Compared with Figure 5, the number of outliers significantly decreases.

![Figure 8](image3.png)

**Figure 8** (Color online) ACC for feature selection methods applied to the original morphometric feature matrix (RAW) and to the completed morphometric feature matrix (STDC), with SVM (a) and KNN (b).

| Outlier completion method | SVM  | KNN  |
|---------------------------|------|------|
| Fill median value         | 0.502| 0.622|
| LRTC                     | 0.766| 0.787|
| STDC                     | 0.793| 0.852|

Table 2 ACC for outlier completion methods with SVM and KNN. Notice that feature selection methods are not involved in this phase.
With the help of STDC, the outlier completed feature matrix can achieve a higher ACC than the original feature matrix, which indicates the outlier completion is also crucial for subsequent feature selection.

For all of feature selection methods, the proposed F-score rank method achieved the highest accuracy (93%) even without the STDC method being applied (86%).

To investigate the contributions of tensor completion and feature selection to the classification accuracy, experiments were run under three different scenarios: original feature matrix without feature selection methods; outlier completed feature matrix without feature selection methods; original feature matrix with the F-score rank method.

It can be observed that compared with using both STDC and F-score rank in Figure 8, using only one of these methods will result in a decrease in accuracy (as seen in Figure 9). In addition, compared with the baseline, the F-score rank feature selection method can achieve a slightly better accuracy than tensor completion, as shown in Figure 9. Thus, it is necessary to use both methods when facing a small dataset with outliers.

### 3.3 The discriminative features and brain regions

After outlier completion, the F-score rank method is applied on the completed feature matrix. During this phase, the statistical result vector (Figure 3) was collected, and the distribution of the number of brain regions among morphometric features (Figure 10) and the von Economo atlas (Figure 11) was examined.

Among these morphometric features, surface area, cortical thickness and gray matter volume are the most discriminatory. The number of regions that belongs to these three features account for 80% of all features.

Having selected the SA, CT and GM morphometric features, the distribution of the number of brain regions among the von Economo categories was also analysed. It was found that for all of the 226 selected brain regions, most of them belong to types 1, 2, 3 and 4.

### 4 Conclusion

Due to the uncertainty associated with MRI artifacts and heavy acquisition costs, it is not easy to identify gifted children from small, high-dimensional MRI datasets. In this paper, a hybrid method was proposed to identify gifted children. The novelty of the proposed method is that all the morphometric features and brain regions are explored in depth. The experimental results suggest that the combination of both methods can achieve a better performance than using only one method. Additionally, this hybrid method can be easily expanded to other similar scenarios.

This work was supported by the National Key R&D Program of China (Grant No. 2017YFE0129700), the National Natural Science Foundation of China (Key Program) (Grant No. 11932013), the National Natural Science Foundation of China (Grant No. 61673224), the Tianjin Natural Science Foundation for Distinguished Young Scholars (Grant No. 18JCJJQJC46100), and the Tianjin Science and Technology Plan Project (Grant No. 18ZXJMTG00250). J.S-C. work is also based upon work from COST Action CA18106, supported by COST (European Cooperation in...
1 Navas-Sánchez F J, Carmona S, Alemán-Gómez Y, et al. Cortical morphometry in frontoparietal and default mode networks in math-gifted adolescents. *Hum Brain Mapp*, 2016, 37: 1893–1902

2 Gross M U M. Exceptionally gifted children: Long-term outcomes of academic acceleration and nonacceleration. *J Education Gifted*, 2006, 29: 404–429

3 Navas-Sánchez F J, Alemán-Gómez Y, Sánchez-Gonzalez J, et al. White matter microstructure correlates of mathematical giftedness and intelligence quotient. *Hum Brain Mapp*, 2014, 35: 2619–2631

4 Hales P W, d’Arco F, Cooper J, et al. Arterial spin labelling and diffusion-weighted imaging in paediatric brain tumours. *NeuroImage-Clin*, 2019, 22: 101696

5 Raja R, Rosenberg G, Caprihan A. Review of diffusion MRI studies in chronic white matter diseases. *Neurosci Lett*, 2019, 694: 198–207

6 Assaf Y, Johansen-Berg H, Thiebaut de Schotten M. The role of diffusion MRI in neuroscience. *NMR Biomed*, 2019, 32: e3762

7 Yun J Y, Boedhoe P S W, Vriend C, et al. Brain structural covariance networks in obsessive-compulsive disorder: A graph analysis from the ENIGMA Consortium. *Brain*, 2020, 143: 684–700

8 Qi T, Schaadt G, Caferro R, et al. The emergence of long-range language network structural covariance and language abilities. *NeuroImage*, 2019, 191: 36–48

9 DuPre E, Spreng R N. Structural covariance networks across the life span, from 6 to 96 years of age. *Network Neurosci*, 2017, 1: 302–323

10 Walker L, Gozzi M, Lenroot R, et al. Diffusion tensor imaging in young children with autism: Biological effects and potential confounds. *Biol Psychiatry*, 2012, 72: 1043–1051

11 Maier-Hein K H, Neher P, Houde J C, et al. Tractography-based connectomes are dominated by false-positive connections. *bioRxiv*, 2019, doi: 10.1101/084137

12 Solé-Casals J, Serra-Grabulosa J M, Romero-Garcia R, et al. Structural brain network of gifted children has a more integrated and versatile topology. *Brain Struct Funct*, 2019, 224: 2373–2383

13 Bethlehem R A I, Romero-Garcia R, Mak E, et al. Structural covariance networks in children with autism or ADHD. *Cerebral Cortex*, 2017, 27: 4267–4276

14 Seiditz J, Vásá F, Shimm M, et al. Morphometric similarity networks detect microscale cortical organization and predict inter-individual cognitive variation. *Neuron*, 2018, 97: 231–247.e7

15 Yang J H, Zhao X L, Ji T Y, et al. Low-rank tensor train for tensor completion with total variation regularized tensor trains. *IEEE Trans Image Process*, 2020, 29: 6918–6931

16 Rouhi A, Nezamabadi-Pour H. Feature selection in high-dimensional data. In: Amini M, ed. Advances in Intelligent Systems and Computing. Vol 207. Berlin, Heidelberg: Springer, 2006

17 Tsagris M, Lagani V, Tsamardinos I. Feature selection for high-dimensional temporal data. *BMC Bioinf*, 2018, 19, 17

18 Ang J C, Mirzal A, Haron H, et al. Supervised, unsupervised, and semi-supervised feature selection: A review on gene selection. *IEEE ACM Trans Comput Biol Bioinf*, 2016, 13: 971–989

19 Rouhi A, Nezamabadi-Pour H. Feature selection in high-dimensional data. In: Amini M, ed. Advances in Intelligent Systems and Computing. Vol. 1123. Cham: Springer, 2020. 85–128

20 Romero-Garcia R, Atienza M, Clemmensen L H, et al. Effects of network resolution on topological properties of human neocortex. *NeuroImage*, 2012, 59: 3522–3532

21 van den Heuvel M P, Scholtens L H, Feldman Barrett L, et al. Bridging cytoarchitectonics and connectomics in human cerebral cortex. *J Neurosci*, 2015, 35: 13943–13948