Deep Discriminative Learning for Autism Spectrum Disorder Classification

Mingli Zhang[1(✉)], Xin Zhao[2(✉)], Wenbin Zhang[3], Ahmad Chaddad[4], Alan Evans[1], and Jean Baptiste Poline[1]

[1] Montreal Neurological Institute, McGill University, Montreal, Canada
mingli.zhang@mcgill.ca
[2] College of Information Engineering, Dalian University, Dalian, China
zhaoxin@dlu.edu.cn
[3] University of Maryland, Baltimore County, MD, USA
[4] Guilin University of Electronic Technology, Guilin, China

Abstract. Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder characterized by deficiencies in social, communication and repetitive behaviors. We propose imaging-based ASD biomarkers to find the neural patterns related ASD as the primary goal of identifying ASD. The secondary goal is to investigate the impact of imaging-patterns for ASD. In this paper, we model and explore the identification of ASD by learning a representation of the T1 MRI and fMRI by fusioning a discriminative learning (DL) approach and deep convolutional neural network. Specifically, a class-wise analysis dictionary to generate non-negative low-rank encoding coefficients with the multi-model data, and an orthogonal synthesis dictionary to reconstruct the data. Then, we map the reconstructed data with the original multi-modal data as input of the deep learning model. Finally, the learned priors from both model are returned to the fusion framework to perform classification. The effectiveness of the proposed approach was tested on a world-wide cross-site (34) database of 1127 subjects, experiments show competitive results of the proposed approach. Furthermore, we were able to capture the status of brain neural patterns with the known input of the same modality.

1 Introduction

Autism spectrum disorder (ASD) is a structural and functional neurodevelopment disorder, it is also associated with weak communication skills, simple repetitive behavioral pattern and lowered concentration. The common way of diagnosis and treatment of ASD is based on symptoms, and thus, to identify a reliable biomarker is the main challenge [7]. Most diagnosis of ASD is confirmed at around 3 years old in the United States although, it is important to diagnose

A. Evans and J. P. Poline—co-last author.

S. Hartmann et al. (Eds.): DEXA 2020, LNCS 12391, pp. 435–443, 2020.
https://doi.org/10.1007/978-3-030-59003-1_29

ASD in the early stage of life for better treatment. Magnetic resonance imaging (MRI) based brain volumetric methods are commonly used to characterize ASD [13]. To better understand the origin of ASD for precise diagnosis, significant progress has been made using neural patterns of functional connectivity of functional magnetic resonance imaging (fMRI) data to caracterize brain changes related to ASD. Identification of Autism Spectrum Disorder from brain imaging provides biomarkers for the mechanisms of the pathology.

In recent years, many representation learning techniques such as discriminative dictionary learning (DDL) [11] and deep neural networks [1] are powerful algorithms to derive high-level latent features from high-dimensional [10] and multi-modal data [4]. DDL has been widely used in resting-state functional connectivity MRI analysis. Wang *et al.* [7] developed a low rank representation approach for multi-center ASD. Zhao *et al.* [12] presented an effective 3D convolutional neural network (CNN) based framework to derive discriminative overlap patterns of a spatial brain network that can characterize and identify ASD from healthy controls. However, considering the fact that ASD could be related to subtle feature changes in the brain, it would be difficult to train an end-to-end CNN directly without any pre-determined information, i.e., discriminative features. Most learning based method with extracted dependent or independent features (cortical thickness, cortical volume, connectome of fMRI) may result in a sub-optimal solution.

One of the challenge of ASD identification is to either estimate the corresponding cortical thickness of the subject under the same pre-processing pipeline or to find the correlation of these features for a given cortical area. The trained-rich matrix may be further processed to yield valuable informations that may be more clinically useful by the generation of gray matter thickness with computer-synthesized cortical volume, cortical surface area and thickness relationship.

In this study, we propose a novel multi-modal discriminative subspace learning approach named MMDL for identification of Autism Spectrum Disorder, by fusion of multi-modal brain imaging data. Different from the conventional modeling-based ASD identification methods, we use not only the priors learned by CNN-based learning, but also the priors from discriminative subspace learning. The fusion is performed in two aspects. First, training the dictionary pair learning (DPL) method. Then, the multi-modal features learned by DPL method and the original data as the input of the CNN. The first step can fully utilize the input data by improving the class-specific features of the original data. The CNN can boost the training performance. Capitalizing on the knowledge, the major contributions of this work are as follows:

– In this work, we propose a novel approach (MMDL), which fuses the classifier of discriminative dictionary learning and CNN to identify ASD. In this proposed MMDL method, instead of only using matrix factorization based discriminative dictionary learning, we also apply the CNN based learning to regularize the model. Specifically, during the CNN training, we initialize the reconstructed features from discriminative dictionary learning and the original data as the input of CNN, which boosts the input of CNN training.

Moreover, the trained dictionary pairs are also returned to the classifier fusion section to improve the identification of ASD performance.
- We demonstrate the classification performance of the proposed method on the functional connectivity matrix and gray matter (cortical surface area, cortical thickness and volume) of 1127 subjects from the challenging of predicting autism[1], the data is acquired from multiple sites with different protocols. The proposed model is much more accurate compared to the state-of-the-art. With one of the given features of gray matter, we can estimate the corresponding others of it.

2 Proposed MMDL Approach

The proposed MMDL approach incorporates deep CNN based training into the training framework, and guides the classify work with the learned priors. Figure 1 is an overview of the proposed framework. More details of each step are described as follows.

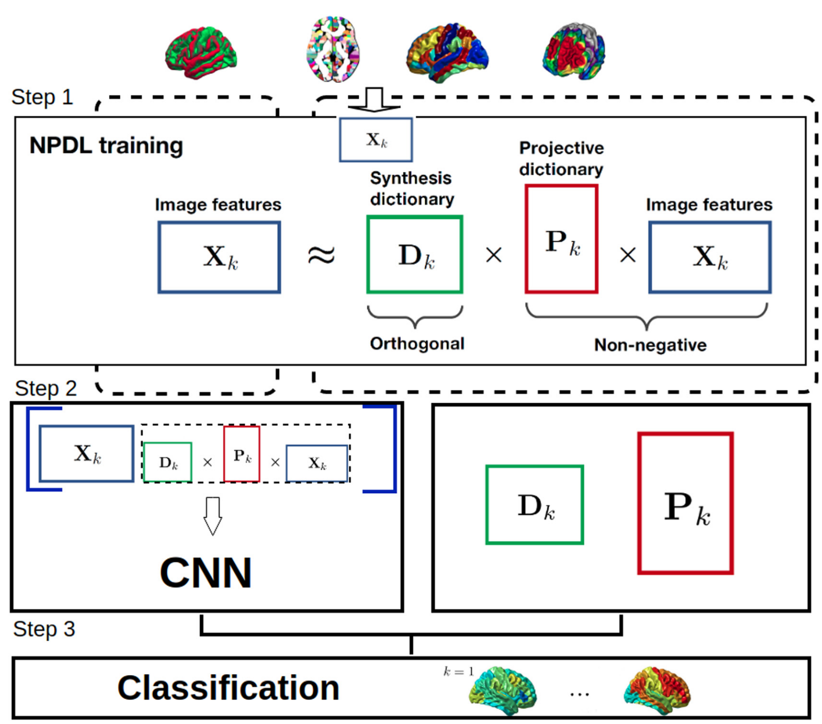


Fig. 1. The Scheme of the proposed MMDL method. The input is tensor format data with T1 MRI volume, surface area, thickness and fMRI connectome.

[1] https://paris-saclay-cds.github.io/autism_challenge/.

2.1 Initialize with Discriminative Learning

The modeling of ASD identification is first treated as a 3D tensor based class-wise discriminative dictionary learning problem, we set $\mathcal{X} = fold(\mathbf{X})$ and the operation of *fold* is to fold up each column of the matrix in to the corresponding subject of the tensor, then $\mathbf{X} = [\mathbf{X}_1, \cdots, \mathbf{X}_k, \cdots, \mathbf{X}_K]$ is the data samples and k is the class number but in this case, it is the binary classification for identification of Autism Spectrum Disorder.

Following [2,9], we introduced a linear feature selection dictionary $\mathbf{P}_k \in \mathbb{R}^{M\times S}$ and a reconstruction dictionary $\mathbf{D}_k \in \mathbb{R}^{S\times M}$ for the class k, where M is the number of subject in class k and S is the dimension of the feature for each subject, with $\mathbf{P} = [\mathbf{P}_1, \cdots, \mathbf{P}_k, \cdots, \mathbf{P}_K]$ and $\mathbf{D} = [\mathbf{D}_1, \cdots, \mathbf{D}_k, \cdots, \mathbf{D}_K]$, performing data modeling in two layer fully connected neural network format with low-rank constrain on the selected features of each group:

$$\underset{\mathbf{P},\mathbf{D}}{\arg\min} \quad \sum_{k=1}^{K} \|\mathbf{X}_k - \mathbf{D}_k\mathbf{P}_k\mathbf{X}_k\|_F^2 + \lambda_1\|\mathbf{P}_k\mathbf{X}_k\|_* + \lambda_2\|\mathbf{P}_k\mathbf{X}_{\overline{k}}\|_F^2, \tag{1}$$

where $\overline{k}$ ($\overline{k} \in \{\overline{k} : |k-\overline{k}| \neq 0\}$), $\|\cdot\|_F$ is the Frobenius norm, $\lambda_1, \lambda_2 > 0$ control the trade-off between the reconstruction accuracy and regularization terms, and $\mathbf{X}_{\overline{k}}$ is the data matrix not belonging to $\mathbf{X}_k$. The regularization term $\|\mathbf{P}_k\mathbf{X}_{\overline{k}}\|_F^2$ is used for forcing $\mathbf{P}_k\mathbf{X}_{\overline{k}}$ towards zero, projecting the samples of non-class to a nearly null space. In this model, $\mathbf{P}_k$ projects the samples $\mathbf{X}_k$ into an encoding coefficient matrix $\mathbf{A}_k = \mathbf{P}_k\mathbf{X}_k$, it can reconstruct $\mathbf{X}_k$ with the reconstruct dictionary $\mathbf{D}_k$, such as Fig. 2.

Ideally, the dictionary $\mathbf{D}$ follows orthogonality constraint with $\mathbf{D}_k^\top\mathbf{D}_k = \mathbf{I}$ to avoid overfitting. Hence, $\mathbf{X}_k$ can be taken as a combination of these similar components by enforcing the encoding coefficients $\mathbf{A}_k = \mathbf{P}_k\mathbf{X}_k$ to be non-negative and low rank. To boost the discrimination of $\mathbf{D}$ and $\mathbf{A}$, we explore weighted nuclear norm [3,8] on $\mathbf{A}$, since the features of subjects within the same class have low rank performance. This leads to the following discriminative learning (DL) problem.

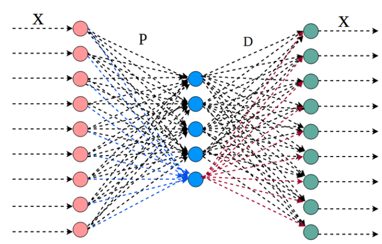


Fig. 2. The flowchart of two layer fully connected neural network based discriminative learning.

$$\begin{aligned} \underset{\mathbf{P},\mathbf{D}}{\arg\min} \quad & \sum_{k=1}^{K} \|\mathbf{X}_k - \mathbf{D}_k\mathbf{P}_k\mathbf{X}_k\|_F^2 + \lambda_1\|\mathbf{A}_k\|_{w,*} + \lambda_2 \sum_{\overline{k}\in\{\overline{k}:|k-\overline{k}|>T\}} \|\mathbf{P}_k\mathbf{X}_{\overline{k}}\|_F^2 \\ \text{s.t.} \quad & \mathbf{D}_k^\top\mathbf{D}_k = \mathbf{I},\ \mathbf{A}_k = \mathbf{P}_k\mathbf{X}_k,\ \mathbf{A}_k \geq 0,\ k = 1, ..., K. \end{aligned} \tag{2}$$

where, the first term is reconstruction error, the second regularization makes the representation low rank, since the components of $\mathbf{X}_k$ are similar and have low-rank performance, $\mathbf{P}_k\mathbf{X}_k \geq 0$ makes representation non-negative and thus, creating sparsity in this way.

To learn dictionary sets $\mathbf{A}$, $\mathbf{D}$ and $\mathbf{P}$, we applied an alternating direction method of multipliers (ADMM) based algorithm as [9].

2.2 Learn the Classification Priors with CNN

Once the $\mathbf{D}$ and $\mathbf{A}$ is acquired, we can get more similar input data with labels as the input of the deep CNN learning. In this sub-section, we will describe the input-output and the architecture of the CNN.

Input-Output: Once the estimated features are achieved, we instead train the CNN with all the multi-model features directly. We use both the estimated multi-model features and the original features $\{\mathbf{DA}\}$ as an input of CNN. This has the advantage of 1) having the estimated features similar to the original multi-model features. 2) with more training data, it can help improve the training accuracy by boosting the learning performance. The intermediate classifiers as the output of CNN. 3) It can work rapidly since the CNN works on features instead of images.

Architecture: We adopt a general CNN architecture. We can apply any CNN models, however here we just adopted the architecture of the CNN has three blocks, which are listed as followed:

- *Conv + BN + ReLU + max pooling*: For the first block, we use 8 filters with size $3 \times 3 \times 8$, the max pooling is done by applying a 2×2 max filter.
- *Conv + BN + ReLU + max pooling*: For the second block, 16 filters with size $3 \times 3 \times 16$.
- *Conv + BN + ReLU + max pooling*: For the third block, 32 filters with size $3 \times 3 \times 32$.
- *AverPooling + FC + Softmax + Classification*

In the CNN, batch normalization (BN) is to accelerate the training, rectified linear units (ReLU) is the activation function, the max pooling layer performs down-sampling and to compute the maximum of each region. The average pooling layer(AverPooling) is for down-sampling and averaging the values of each region. FC is Fully Connected Layer.

2.3 Classification

In the classification process, we input the multi-model testing data into the well trained CNN and the learned $\mathbf{P}$ and $\mathbf{D}$, which can be used to classify samples by measuring the reconstruction error for each class as the approach. Instead of using the classification results via CNN as the final result. It can be as intermediate to further improve the classification performance. Thus the final classification result can be obtained by fusing the results of two classifiers.

To get the intermediate classification results of the initial discriminative learning. Here, we set $\mathbf{x}^i \in \mathbb{R}^{S_i}$ be the features of type i for the subject to classify. We define as $e_k^i = \|\mathbf{x}^i - \mathbf{D}_k^i \mathbf{P}_k^i \mathbf{x}\|_2$ the error of reconstructing $\mathbf{x}^i$ with the dictionaries of class k for feature type i. We then assign the sample to the class whose dictionary gives the lowest error $\widehat{k}_i = \arg\min_k e_k^i$.

To combine the information of the two classifiers to further improve the classification result, the final classification result can be obtained by solving the optimization problem as followed.

$$\arg\min_{\boldsymbol{\alpha}} \left(k_{\text{real}} - \sum_j \alpha_j \widehat{k}_j\right)^2, \quad \text{s.t. } \sum_j \alpha_j = 1,\ \alpha_j \geq 0, \forall j. \tag{3}$$

$\widehat{k}_j$ is the classification result of each classifier j ($j = 2$) and the final output is class label k_{real}. Constraints on regression coefficients α_i enforce the final prediction to be a convex combination of classification results from the classifiers of CNN and discriminative dictionary learning.

3 Experiments

In this section, experiments are conducted on the public set of IMPAC[2] to evaluate the effectiveness of the proposed MMDL approach. We use 1127 subjects with 590 subjects as control and 537 subjects with Autism Spectrum Disorder. The model is evaluated with structural MRI using measures of cortical thickness, surface area and volume and resting state fMRI with 17.01($\pm$10) years old. The structural MRI is preprocessed with FreeSurfer and FSL, then the features are averaged following an adapted Desikan protocol, giving a total of 70 features per type of measure for both brain hemispheres. Connectomes were derived from fMRI using the correlation matrix of each subject, we use the singular values vector of the connectomes of fMRI as the input features. Then, the input is a tensor format data with a subject (subject with the label), volume, surface area, thickness and singular values of fMRI connectome matrix of each subject. For functional MRI in this study, we use the MSDL functional atlas [6], we reconstructed connectivity matrices using 70 brain discriminative regions by applying singular value decomposition (SVD) on these connectivity matrices, the singular values are then rearranged as vectors of 70 features.

For these tasks, We split available 1127 examples into a training set and validation set, the latter containing 10% of examples. The validation set was used to tune the regularization parameters and the size M of synthesis dictionary $\mathbf{D}$. Afterward, the 8-fold cross-validation is applied on these experiments to measure performance in terms of prediction accuracy (ACC), Specificity, Sensitivity, area under the curve (AUC) and root mean square error (RMSE).

3.1 Prediction of Autism Spectrum Disorder

We first demonstrate the proposed framework's performance by predicting the autism spectrum disorder, based on cortical thickness, cortical surface area and

[2] https://paris-saclay-cds.github.io/autism_challenge/.

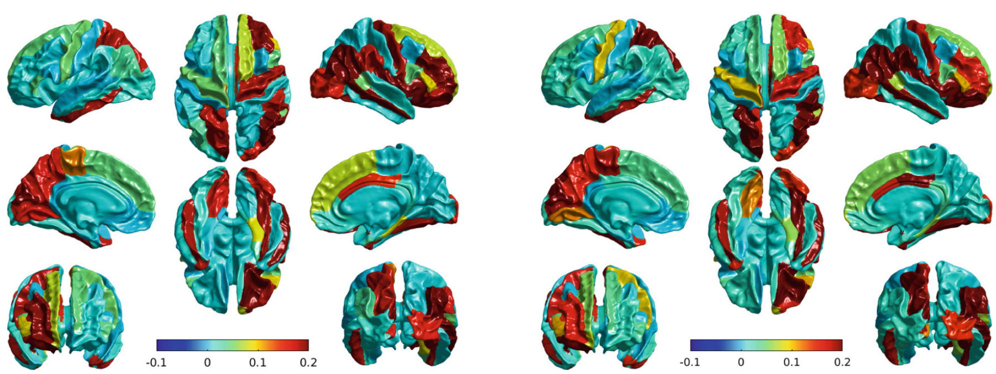
Table 1. Classification results on the database of IMPAC on 8-fold cross-validation.

Method	ACC	Sens.	Spec.	AUC
SVM	0.621 ± 0.027	0.616 ± 0.071	0.629 ± 0.060	0.622 ± 0.028
SVM+CNN	0.664 ± 0.042	0.892 ± 0.030	0.476 ± 0.056	0.684 ± 0.039
RF	0.525 ± 0.034	0.391 ± 0.142	0.628 ± 0.147	0.509 ± 0.022
RF+CNN	0.661 ± 0.019	0.860 ± 0.062	0.481 ± 0.046	0.670 ± 0.030
DL_{In}	0.648 ± 0.040	0.742 ± 0.077	0.543 ± 0.045	0.643 ± 0.041
MMDL	$\mathbf{0.690} \pm 0.055$	0.790 ± 0.049	$\mathbf{0.689} \pm 0.048$	$\mathbf{0.733} \pm 0.051$

cortical volumes of T1 structure MRI and functional connectivity. For functional MRI in this study, we use the MSDL functional atlas [6].

The average ACC, Sensitivity (Sens.), Specificity(Spec.) and AUC of the proposed methods with the comparisons are on 8-fold cross-validation (CV) reported in Table 1, the proposed MMDL method outperforms the SVM and random forest (RF)[5] based methods, as shown in the Table 1, the proposed method has the highest ACC, Specificity and AUC. By fusioning the result of SVM and RF with CNN (i.e., 'SVM+CNN' and 'RF+CNN' in Table 1) separately, the results have improved. Compared to the competed methods in Table 1, our approach yields improvements of about 0.026 in ACC, 0.061 in Specificity (Spec.) and 0.049 in AUC.

In the proposed model, we show the features that are predicted with the discriminative learning model of Eq. (2) in Fig. 3 is an example of with the cortical volume and predicted one, they are quite similar and the RMSE between them is 0.09.



a) Tested volume.   b) predicted volume.

Fig. 3. The tested volume and the predicted one.

4 Conclusion

In this paper, a MMDL method is proposed by fusion discriminative learning and priors of deep CNN to regularize the classification problem. Specifically, in the non-negative discriminative dictionary learning model, this approach learns discriminative features by imposing both orthogonality on the synthesis dictionary, non-negativity low-rank constraints on projective coefficients. We initialize more multi-model data from dictionary learning model as the input of CNN, which can improve the training accuracy. Then, both training priors are returned to the fusion framework to improve the performance. Experiments on the tasks of identifying the ASD showed the benefit of our approach compared to state-of-the-art methods. The proposed method can be used for synthesizing the neural patterns of cortical.

Acknoledgements. This work was supported, in part, by the Fonds de recherche du Quebec (CCC 246110, 271636), National Nature Science Foundation of China (NSFC: 61902220), Shandong Province grant (ZR2018BF009) and the Science and Technology Innovation Fund Project of Dalian, China (No.2019J13SN100).

J.-B.P. was partially funded by National Institutes of Health (NIH) NIH-NIBIB P41 EB019936 (ReproNim) NIH-NIMH R01 MH083320 (CANDIShare) and NIH RF1 MH120021 (NIDM), the National Institute Of Mental Health of the NIH under Award Number R01MH096906 (Neurosynth), as well as the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives initiative and the Brain Canada Foundation with support from Health Canada.

References

1. Chaddad, A., Rathore, S., Zhang, M., Desrosiers, C., Niazi, T.: Deep radiomic features from mri scans predict survival outcome of recurrent glioblastoma (2019). arXiv preprint arXiv:1911.06687
2. Gu, S., Zhang, L., Zuo, W., Feng, X.: Projective dictionary pair learning for pattern classification. In: Advances in Neural Information Processing Systems, pp. 793–801 (2014)
3. Gu, S., Zhang, L., Zuo, W., Feng, X.: Weighted nuclear norm minimization with application to image denoising. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2862–2869 (2014)
4. Li, J., Wu, Y., Zhao, J., Lu, K.: Low-rank discriminant embedding for multiview learning. IEEE Trans. Cybern. **47**(11), 3516–3529 (2017)
5. Tong, T., Gao, Q., Guerrero, R., Ledig, C., Chen, L., Rueckert, D., Alzheimer's Disease Neuroimaging Initiative et al.: A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease. IEEE Trans. Biomed. Eng. **64**(1), 155–165 (2017)
6. Varoquaux, G., Gramfort, A., Pedregosa, F., Michel, V., Thirion, B.: Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 562–573. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22092-0_46

7. Wang, M., Zhang, D., Huang, J., Shen, D., Liu, M.: Low-rank representation for multi-center autism spectrum disorder identification. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 647–654. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_73
8. Zhang, M., Desrosiers, C.: High-quality image restoration using low-rank patch regularization and global structure sparsity. IEEE Trans. Image Process. **28**(2), 868–879 (2019)
9. Zhang, M., Desrosiers, C., Guo, Y., Zhang, C., Khundrakpam, B., Evans, A.: Brain status prediction with non-negative projective dictionary learning. In: Shi, Y., Suk, H.-I., Liu, M. (eds.) MLMI 2018. LNCS, vol. 11046, pp. 152–160. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00919-9_18
10. Zhang, M., Desrosiers, C., Zhang, C.: Atlas-based reconstruction of high performance brain MR data. Pattern Recogn. **76**, 549–559 (2018)
11. Zhang, M., Guo, Y., Zhang, C., Poline, J.-B., Evans, A.: Modeling and analysis brain development via discriminative dictionary learning. In: Knoll, F., Maier, A., Rueckert, D., Ye, J.C. (eds.) MLMIR 2019. LNCS, vol. 11905, pp. 80–88. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33843-5_8
12. Zhao, Yu., Ge, F., Zhang, S., Liu, T.: 3D deep convolutional neural network revealed the value of brain network overlap in differentiating autism spectrum disorder from healthy controls. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11072, pp. 172–180. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00931-1_20
13. Zheng, W., Eilamstock, T., Wu, T., Spagna, A., Chen, C., Hu, B., Fan, J.: Multi-feature based network revealing the structural abnormalities in autism spectrum disorder. IEEE Trans. Affect. Comput. (2019)