WSSADN: A Weakly Supervised Spherical Age-Disentanglement Network for Detecting Developmental Disorders with Structural MRI

Pengcheng Xue[1], Dong Nie[2], Meijiao Zhu[3], Ming Yang[3], Han Zhang[4], Daoqiang Zhang[1], and Xuyun Wen[1(✉)]

[1] College of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, Key Laboratory of Brain-Machine Intelligence Technology, Ministry of Education, Nanjing, China
[2] Meta Inc., California, USA
[3] Department of Radiology, Children's Hospital of Nanjing Medical University, Nanjing, China
[4] School of Biomedical Engineering, ShanghaiTech University, Shanghai, China
wenxuyun@nuaa.edu.cn

Abstract. Structural magnetic resonance imaging characterizes the morphology and anatomical features of the brain and has been widely utilized in the diagnosis of developmental disorders. Given the dynamic nature of developmental disorder progression with age, existing methods for disease detection have incorporated age as either prior knowledge to be integrated or as a confounding factor to be disentangled through supervised learning. However, the excessive focus on age information in these methods restricts their capability to unearth disease-related features, thereby affecting the subsequent disease detection performance. To address this issue, this work introduces a novel weakly supervised learning-based method, namely, the Weakly Supervised Spherical Age Disentanglement Network (WSSADN). WSSADN innovatively combines an attention-based disentangler with the Conditional Generative Adversarial Network (CGAN) to remove normal developmental information from the brain representation of the patient with developmental disorder in a weakly supervised manner. By reducing the focus on age information during the disentanglement process, the effectiveness of the extracted disease-related features is enhanced, thereby increasing the accuracy of downstream disease identification. Moreover, to ensure effective convergence of the disentanglement and age information learning modules, we design a consistency regularization loss to align the age-related features generated by the disentangler and CGAN. We evaluated our method on three different tasks, including the detection of preterm neonates, infants with congenital heart disease, and autism spectrum disorders. The experimental results demonstrate that our method significantly outperforms existing state-of-the-art methods across all tasks. The codes will be publicly available in https://github.com/xuepengcheng1231/WSSADN.

Keywords: Developmental disorder · Weakly supervised learning · Disentanglement · Conditional Generative Adversarial Network.

1 Introduction

Structural magnetic resonance imaging (sMRI) provides detailed patterns of brain morphology and anatomical features, leading to its widespread application in the study of developmental disorders caused by diverse factors, such as preterm birth and cardiac diseases, etc [15,24,14,22]. Recently, sMRI-based deep learning methods, due to their powerful representation learning capabilities, have been proposed to extract discriminative features from imaging data for brain disorder diagnosis [18,23].

Compared to the common brain disorder, developmental disorder detection is more challenging, due to the complex and dynamic evolution of the brain's structure throughout the developmental process, especially in early stages [20,7,2,8]. As indicated by relevant studies, patients with developmental disorders such as autism spectrum disorder (ASD) [20,7] and congenital heart disease (CHD) [21] exhibit brain structural changes influenced not only by normal developmental processes but also significantly by the pathological processes. However, these two kinds of processes are closely intertwined with age, presenting a dynamic complexity that evolves over age. This complexity undoubtedly increases the difficulty of accurately extracting disease-specific brain changes (i.e., discriminative features) from the multitude of changes. Moreover, given the dynamic nature of diseases, diagnostic models developed for one age group may not be applicable to other age groups. To tackle these problems, existing studies mainly utilized two strategies: 1) Fusion-based methods, incorporating age as prior knowledge into the model to enhance the diagnosis; and 2) Disentangling-based methods, removing age-related normal development from brain representation to extract brain developmental deviations only induced by the disorders.

Although these methods have achieved certain success, they still have obvious limitations. Fusion-based methods, attempting to enhance diagnostic accuracy by incorporating age information as an extra input, fail to effectively extract disease-specific characteristics, thereby impacting the potential interpretability of the analysis results [1,19]. To tackle this challenge, researchers have employed a disentangling-based approach, usually utilizing a supervised learning-based age predictor to learn age information and employing disentanglement techniques to eliminate age-related normal development from brain representation [12,4,28]. However, the supervised learning approach might overly prioritize extracting age-related information, particularly for a continuous variable like age. Consequently, this excessive focus could interfere with the extraction of disease-related features.

To overcome these challenges, we developed the Weakly Supervised Spherical Age Disentanglement Network (WSSADN). This method employs a weakly supervised learning strategy, rather than a supervised learning strategy, to effectively remove representations related to normal development from brain representation, so that this approach can extract more accurate disease-induced brain changes, thereby enhancing the performance of developmental disorder diagnosis. Specifically, we incorporate a Conditional Generative Adversarial Network [16][3](CGAN)-based age information generator, namely the Weakly Supervised Age Information Learning module, into an attention-based disentangler. This

setup utilizes age-related embeddings generated by the CGAN to steer the optimization of the disentangler. By diminishing the focus on age learning in a weakly supervised manner, we enhance the identification capabilities of disease-related features. Furthermore, we design a consistency regularization loss to align the age-related representations generated by the disentangler and CGAN, ensuring the effective convergence of the disentanglement process. The main contributions of our work can be summarized as follows:

- We propose a weakly supervised disentanglement-based brain representation learning method, which can effectively remove normal developmental information from the brain features of patients with developmental disorders. By retaining only disease-related features, the diagnosis accuracy can be significantly improved.
- We introduce a consistency regularization loss, a crucial component that guarantees the effective operation of the disentanglement module. This mechanism is specifically engineered to prevent trivial solutions, ensuring the integrity and reliability of the extracted features.
- Our extensive experimental evaluation confirms the superior performance of our method in accurately diagnosing developmental disorders at various age stages, underscoring its effectiveness and potential for widespread applications in the clinic.

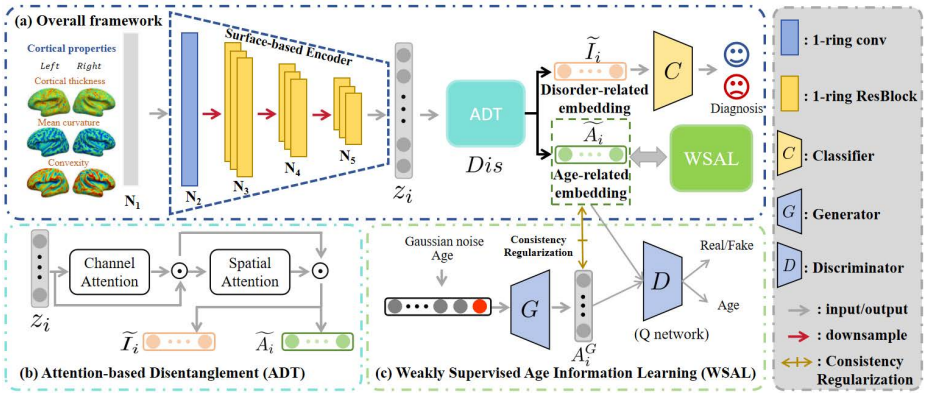

Fig. 1. Overview of our proposed method, consisting of three components: (a) Surface-based encoder E_n to extract cortical features; (b) Attention-based disentanglement (ADT) for disentangling the cortical features into age-related and disorder-related categories; (c) Weakly supervised age information learning (WSAL) based on CGAN for guiding the age-related embedding learning in ADT module. N_k denotes the number of vertices in the spherical feature maps.

2 Method

2.1 Network Architecture

The architecture of our proposed method is shown in Fig. 1, which comprises of three modules. The first module is the surface-based encoder for extracting cortical features from three different cortical mesh data. The second module is Attention-based DisenTanglement (ADT) to decompose the extracted cortical features into disease-related and age-related embeddings. The third module is Weakly Supervised Age information Learning (WSAL), designed to guide the optimization of the ADT module.

Surface-based Spherical Encoder Given the uniform spherical structure of the cortical surface derived from an icosahedron, we developed a surface-based encoder using spherical convolution [29] to capture spatially fine-grained information by exploiting its intrinsic spherical topology. The encoder, denoted as E_n and implemented as Spherical Res-Net [11,29], incorporates a sequence of spherical convolution and spherical pooling layers to extract spatial patterns from a set of surface maps with cortical properties. Ultimately, the latent variables $z_i \in \mathbb{R}^{162*M}$ capture mixed features (such as mean cortical thickness, mean curvature, and convexity) of the input cortical surface $x_i \in \mathbb{R}^{10242*3}$, where M denotes the number of channels and 162 and 10242 are the number of vertices.

Attention-based Disentanglement. Considering the developmental biases inherent in developmental disorders, it becomes imperative to eliminate age-related information from brain structural data to further generate purer disorder-related representations. To achieve this goal, we introduced the ADT module as the disentangler (Dis) to decompose the latent variables z_i into age-related representation A_i and disorder-related representation I_i. To be specific, the ADT module consists of channel attention and spatial attention [13,26]. The age-related and disorder-related representations are defined as $A_i = z_i \odot (1 - \psi(z_i))$ and $I_i = z_i \odot \psi(z_i)$, respectively. Herein, the operation $\odot$ denotes element-wise multiplication, and ψ indicates the attention module aforementioned. Finally, we employed a 1-D convolution f_c to fuse information from all channels and respectively map A_i and I_i to the latent variables $\widetilde{z}_i^{\,l}, \widetilde{z}_i^{\,r} \in \mathbb{R}^{162*1}$ (l represents the left hemisphere, r represents the right hemisphere).

Weakly Supervised Age Information Learning To guide the learning of disentangled age-related representation without compromising the diagnostic capacity, we designed a new age information learning method based on CGAN [16][3], namely the WSAL module. In contrast to conventional age information learning methods that rely on age predictors, WSAL employs imprecise age representations generated by CGAN as labels, which is proposed weakly supervised method. This technique guides the disentanglement process (i.e., optimization of the ADT module) in a weakly supervised manner, aiming to mitigate the model's

excessive emphasis on age. Specifically, as shown in Fig. 1, WSAL consists of a generator (G) and a discriminator (D). The G is used for generating age-related representation A_i^G with the subject's age a and Gaussian noise z as input. The generated A_i^G is used for guiding the disentanglement of age-related representation $\widetilde{A_i}$ in the ADT module. The discriminator D is designed for distinguishing the A_i^G from G in WSAL module and $\widetilde{A_i}$ from ADT module. Based on the discriminative capability of D, the distribution $P(G(z,a))$ gradually approximates the real distribution $P(\widetilde{A_i})$ to maximize the mutual information boundary between A_i^G and the disentangled representation $\widetilde{A_i}$ by optimizing the following loss:

$$\min_G \max_D L_{adv} = \mathbb{E}_{x\sim P_{\widetilde{A_i}}}[logD(x)] + \mathbb{E}_{z\sim P_{noise}}[1 - log(D(G(z,a)))] \tag{1}$$

To retain age information to the greatest extent when generating age representations, we further employed a Q network to predict age from A_i^G, where the Q network shares parameters with D except for the last layer. The objective is achieved by minimizing L_{WSAL}, implemented as L2 norm:

$$\min_{G,Q} L_{WSAL} = 1/B \sum_{i=1}^{B} (Q(A_i^G) - a)^2 \tag{2}$$

where B is batch size.

Consistency Regularization Although we attempt to achieve the alignment between disentangled and generated age-related representation through discriminator D, the prerequisite is that the discriminator D should possess superior discriminative capacity, which mandates multiple iterations in the model training process and is prone to yielding trivial solutions. Therefore, to ensure effective convergence and prevent ineffective solutions during the disentanglement, we designed a consistency regularization (CR) constraint. We adopted the L2 norm to reinforce the consistency between the disentangled and generated age-related latent embedding by minimizing the following loss:

$$\min_{Dis} L_{CR} = 1/B \sum_{i=1}^{B} (A_i^G - \widetilde{A_i})^2 \tag{3}$$

Loss Function In our work, we adopt a simple Multi-layer Perceptron (MLP) as the final classifier. We adopt binary cross-entropy as the classification loss, denoted as L_C. To ensure that the generator G produces representations approximating the real distribution, it is imperative to train G based on the discriminative capability of the discriminator D. Consequently, we decompose equation (1) into two parts, L_D and L_G, constituting an adversarial game.

Our proposed model adopts an end-to-end iterative training approach. Firstly, we optimize all module parameters through the L_{Global} loss function, formulated as: $L_{Global} = L_C + \lambda_1 * L_D + \lambda_2 * L_{WSAL}$. Subsequently, we optimize the

parameters of G and Dis by optimizing the $L_{G,Dis}$, formulated as: $L_{G,Dis} = L_{CR} + \lambda_3 * L_G$. The $\lambda_1, \lambda_2, \lambda_3$ are employed to adjust the relative significance of each loss term.

3 Experiments and Results

3.1 Datasets

To validate the effectiveness of the proposed method, experiments were conducted on three different datasets, including the third release of the Developing Human Connectome Project (DHCP), an in-house collected dataset of CHD infants which has received ethical approval and Autism Brain Imaging Data Exchange (ABIDE I). The DHCP dataset encompasses 281 preterm and 514 full-term neonates with ages ranging from 26 to 45 weeks. The CHD dataset includes 96 CHD infants and 101 age-matched healthy controls from Children's hospital of Nanjing Medical University, with their ages varying from 10 to 72 months. The ABIDE I dataset consists of 489 patients with ASD and 509 age-matched healthy controls, covering an age range from 6 to 64 years. All three datasets were downsampled to 10,242 vertices, and each vertex has three cortical morphological attributes, including cortical thickness, mean curvature, and convexity. The 5-fold cross-validation strategy was employed.

3.2 Experiments Setup

In this work, all methods were implemented using PyTorch, utilizing the Adam optimizer with a weight decay of 1e-5. A cosine decay strategy was applied for learning rate scheduling during training, with an initial learning rate set at 0.001. The batch size was set to 40. The maximum training epochs were set to 200. Based on the parameter analysis experiments (see Fig. 1 in appendix), the hyperparameters λ_1, λ_2, and λ_3 were set to 0.1, 0.001, and 10, respectively. We employed accuracy and AUC as evaluation metrics in disease diagnosis tasks and used mean square error for evaluating age prediction accuracy.

3.3 Experimental Results

Comparison with SOTA Methods. We compared the proposed method with four SOTA methods, including one method without considering age information (i.e., NeuroExplainer, NE [27]), two age-fusion-based methods (i.e., Spherical Multi-task CNN, SMTCNN [10] and Spherical CNN++, SCNN++ [10]), and one age-disentangling-based method (i.e., Spherical Disentanglement CNN, SDCNN [28]). These methods follow the same optimizer settings and training epochs as our method.

As shown in Table 1, our proposed method outperforms the SOTA methods in all experimental tasks, validating the effectiveness of weakly supervised age learning in enhancing the learning of disease-specific features. Firstly, incorporating age factors resulted in an average improvement of 1% in diagnostic

Table 1. Comparison results of disease diagnosis accuracy on three different datasets.

Methods	DHCP		CHD		ABIDE I	
	ACC	AUC	ACC	AUC	ACC	AUC
NE*	0.894±0.02	0.946±0.02	0.776±0.04	0.75±0.03	0.593±0.02	0.57±0.02
SCNN++†	-	-	0.781±0.03	0.719±0.06	0.605±0.01	0.564±0.02
SMTCNN†	0.916±0.02	0.935±0.02	0.782±0.03	0.76±0.04	0.598±0.01	0.584±0.03
SDCNN‡	0.925±0.01	0.932±0.03	0.79±0.07	0.76±0.10	0.613±0.02	0.58±0.01
Ours	**0.940±0.02**	**0.965±0.02**	**0.826±0.06**	**0.82±0.06**	**0.624±0.02**	**0.628±0.01**

* represents methods ignoring the age information. † represents fusion-based methods. ‡ represents disentangling-based methods with supervised learning.

performance compared to age-agnostic strategies, underscoring the superiority of considering age information in the diagnosis of developmental disorders. Secondly, through a comprehensive comparison of three age-handling approaches, we found that weakly supervised disentanglement learning performed the best. It exhibited an average improvement of 1.7% over supervised disentanglement learning on the three datasets and an average improvement of 3.3% over fusion-based methods. This finding once again confirms that an excessive focus on age information may not be conducive to the effective capture of disease-related features by the model.

Ablation Study. To demonstrate the necessity of each component within our proposed method, we separately removed the WSAL and CR modules from WSSADN, constructing two baseline methods: w/o WSAL and w/o CR, with the results displayed in Table 2. We observe that: 1) Methods without the WSAL or CR modules exhibit significantly lower classification accuracy compared to the complete WSSADN model; 2) our proposed method, which includes the CR module, shows higher accuracy than the model without it (i.e., w/o CR). These findings effectively validate the importance of incorporating both the WSAL and CR modules into our method to enhance its diagnostic performance.

Comparison between Supervised and Weakly Supervised Learning. To assess the effect of minimizing age information focus on disease feature recognition, we compared SDCNN, a supervised learning method, with our weakly

Table 2. Comparison results of ablation study on three different datasets.

Methods	DHCP		CHD		ABIDE I	
	ACC	AUC	ACC	AUC	ACC	AUC
w/o WSAL	0.907±0.01	0.946±0.01	0.795±0.03	**0.868±0.04**	0.6±0.02	0.617±0.03
w/o CR	0.92±0.06	0.946±0.05	0.802±0.08	0.815±0.09	0.623±0.04	0.608±0.04
ours	**0.94±0.02**	**0.965±0.02**	**0.826±0.06**	0.82±0.06	**0.624±0.02**	**0.628±0.01**

Table 3. Comparative results between supervised and weakly supervised learning methods in age prediction accuracy and disease diagnostic accuracy.

Methods	DHCP (week)		CHD (month)		ABIDE I (year)	
	MSE	Accuracy	MSE	Accuracy	MSE	Accuracy
SDCNN	**34.2±23.4**	0.925±0.01	**61.0±7.9**	0.79±0.07	**67.3±52.5**	0.613±0.02
w/o CR	74.2±19.1	0.92±0.06	176.6±103.6	0.802±0.08	314.4±200.3	0.623±0.04
ours	53.3±18.2	**0.94±0.02**	69.2±12.4	**0.826±0.06**	88.4±14.85	**0.624±0.02**

supervised approach in terms of age prediction and disease diagnosis accuracy. Additionally, a baseline model without the CR module (w/o CR) was introduced to evaluate the impact of the proposed CR loss on age prediction. The results summarized in Table 3 indicate that while SDCNN achieves the highest age prediction accuracy, it exhibits the lowest diagnostic accuracy, supporting our hypothesis that reducing age information focus improves diagnostic accuracy. Furthermore, our final method, WSSADN, demonstrates a decrease in age prediction accuracy compared to w/o CR, indicating the effectiveness of CR loss in promoting convergence during disentanglement.

Contribution of Brain Regions. Given the convolution block f_c fuses the multi-level properties of cortical surface, we explored the gradient weights of disorder-related representation produced by f_c to generate the active maps through Grad-CAM. The results are shown in Fig. 2. We observe that the distribution of brain regions with significant contributions varies greatly among the three developmental disorders. Specifically, preterm neonates primarily involve the rostral middle frontal, superior temporal, fusiform, and lateral occipital areas, consistent with previous studies [17]. CHD is concentrated in the inferior parietal, cingulate, and temporal areas, consistent with [5][6]. ASD focuses on the medial frontal, precentral, postcentral, superior temporal, and paracentral areas. The results are in line with the reports in related to studies to some extent [9][25]. Furthermore, we also note areas of overlap across all three diseases, such

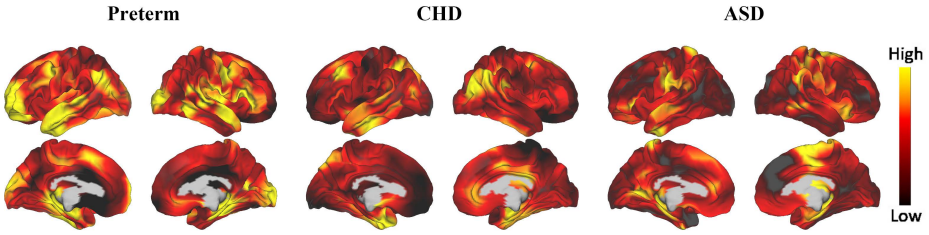

Fig. 2. Heatmaps of the brain regions' contributions in classification tasks on three different developmental disorders.

as the superior temporal lobe, which plays a significant role in the classification of each condition. This suggests commonalities between the diseases, meriting further investigation in the future.

4 Conclusion

In this paper, we proposed a Spherical Weakly Supervised Age Disentanglement Network to improve the diagnostic performance of developmental disorders by eliminating the developmental biases in disease-related feature learning in a weakly supervised manner. Evaluated on three benchmark datasets, the proposed method consistently achieves superior performance in developmental disorder diagnosis at various age stages. Our approach introduces an innovative design strategy for the detection of developmental disorders, paving the way for advancements in diagnostic methodologies.

Acknowledgments. This research was supported by the STI 2030-Major Projects (2022ZD0209000); Fundamental Research Funds for the Central Universities (NZ2024040); National Natural Science Foundation of China(62136004); and China Postdoctoral Science Foundation funded project(2021TQ0150, 2021M701699).

Disclosure of Interests. No potential disclosure of interests was reported by the authors.

References

1. Al-Saei, A.N.J., Nour-Eldine, W., Rajpoot, K., Arshad, N., Al-Shammari, A.R., Kamal, M., Akil, A.A.S., Fakhro, K.A., Thornalley, P.J., Rabbani, N.: Validation of plasma protein glycation and oxidation biomarkers for the diagnosis of autism. Molecular Psychiatry pp. 1–7 (2023)
2. Arpi, E., Ferrari, F.: Preterm birth and behaviour problems in infants and preschool-age children: A review of the recent literature. Developmental Medicine & Child Neurology **55**(9), 788–796 (2013)
3. Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., Abbeel, P.: Infogan: Interpretable representation learning by information maximizing generative adversarial nets. Advances in neural information processing systems **29** (2016)
4. Cheng, J., Zhang, X., Zhao, F., Wu, Z., Yuan, X., Wang, L., Lin, W., Li, G.: Prediction of infant cognitive development with cortical surface-based multimodal learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 618–627. Springer (2023)
5. Clouchoux, C., Du Plessis, A., Bouyssi-Kobar, M., Tworetzky, W., McElhinney, D., Brown, D., Gholipour, A., Kudelski, D., Warfield, S., McCarter, R., et al.: Delayed cortical development in fetuses with complex congenital heart disease. Cerebral cortex **23**(12), 2932–2943 (2013)
6. Cordina, R., Grieve, S., Barnett, M., Lagopoulos, J., Malitz, N., Celermajer, D.S.: Brain volumetrics, regional cortical thickness and radiographic findings in adults with cyanotic congenital heart disease. NeuroImage: Clinical **4**, 319–325 (2014)

7. Daniels, A.M., Mandell, D.S.: Explaining differences in age at autism spectrum disorder diagnosis: A critical review. Autism **18**(5), 583–597 (2014)
8. Dibble, M., Ang, J.Z., Mariga, L., Molloy, E.J., Bokde, A.L.: Diffusion tensor imaging in very preterm, moderate-late preterm and term-born neonates: a systematic review. The Journal of pediatrics **232**, 48–58 (2021)
9. Ecker, C., Bookheimer, S.Y., Murphy, D.G.: Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. The Lancet Neurology **14**(11), 1121–1134 (2015)
10. Han, K., Li, G., Fang, Z., Yang, F.: Multi-template meta-information regularized network for alzheimer's disease diagnosis using structural mri. IEEE Transactions on Medical Imaging (2023)
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
12. Hu, D., Wang, F., Zhang, H., Wu, Z., Wang, L., Lin, W., Li, G., Shen, D., Consortium, U.B.C.P.: Disentangled intensive triplet autoencoder for infant functional connectome fingerprinting. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23. pp. 72–82. Springer (2020)
13. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)
14. Ji, W., Li, G., Jiang, F., Zhang, Y., Wu, F., Zhang, W., Hu, Y., Wang, J., Wei, X., Li, Y., et al.: Preterm birth associated alterations in brain structure, cognitive functioning and behavior in children from the abcd dataset. Psychological Medicine pp. 1–10 (2023)
15. Katuwal, G.J., Cahill, N.D., Baum, S.A., Michael, A.M.: The predictive power of structural mri in autism diagnosis. In: 2015 37th annual international conference of the ieee engineering in medicine and biology society (EMBC). pp. 4270–4273. IEEE (2015)
16. Mirza, M., Osindero, S.: Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 (2014)
17. Nam, K.W., Castellanos, N., Simmons, A., Froudist-Walsh, S., Allin, M.P., Walshe, M., Murray, R.M., Evans, A., Muehlboeck, J.S., Nosarti, C.: Alterations in cortical thickness development in preterm-born individuals: Implications for high-order cognitive functions. NeuroImage **115**, 64–75 (2015)
18. Nogay, H.S., Adeli, H.: Diagnostic of autism spectrum disorder based on structural brain mri images using, grid search optimization, and convolutional neural networks. Biomedical Signal Processing and Control **79**, 104234 (2023)
19. Nogay, H.S., Adeli, H.: Multiple classification of brain mri autism spectrum disorder by age and gender using deep learning. Journal of Medical Systems **48**(1), 15 (2024)
20. O'Hearn, K., Lynn, A.: Age differences and brain maturation provide insight into heterogeneous results in autism spectrum disorder. Frontiers in Human Neuroscience **16**, 957375 (2023)
21. Rollins, C.K., Ortinau, C.M., Stopp, C., Friedman, K.G., Tworetzky, W., Gagoski, B., Velasco-Annis, C., Afacan, O., Vasung, L., Beaute, J.I., et al.: Regional brain growth trajectories in fetuses with congenital heart disease. Annals of neurology **89**(1), 143–157 (2021)
22. Salama, A.A., Alarabawy, R.A., El-Shehaby, W., El-Amrousy, D., Baghdadi, M.S., Rizkallah, M.F.: Brain volumetrics, regional cortical thickness and radiographic

findings in children with cyanotic congenital heart disease using quantitative magnetic resonance imaging. The Egyptian Journal of Radiology and Nuclear Medicine **47**(4), 1617–1627 (2016)
23. Saurabh, S., Gupta, P.: Deep learning-based modified bidirectional lstm network for classification of adhd disorder. Arabian Journal for Science and Engineering pp. 1–18 (2023)
24. Shan, X., Uddin, L.Q., Xiao, J., He, C., Ling, Z., Li, L., Huang, X., Chen, H., Duan, X.: Mapping the heterogeneous brain structural phenotype of autism spectrum disorder using the normative model. Biological Psychiatry **91**(11), 967–976 (2022)
25. Wallace, G.L., Robustelli, B., Dankner, N., Kenworthy, L., Giedd, J.N., Martin, A.: Increased gyrification, but comparable surface area in adolescents with autism spectrum disorders. Brain **136**(6), 1956–1967 (2013)
26. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV). pp. 3–19 (2018)
27. Xue, C., Wang, F., Zhu, Y., Li, H., Meng, D., Shen, D., Lian, C.: Neuroexplainer: Fine-grained attention decoding to uncover cortical development patterns of preterm infants. arXiv preprint arXiv:2301.00815 (2023)
28. Yuan, X., Cheng, J., Zhao, F., Wu, Z., Wang, L., Lin, W., Zhang, Y., Li, G.: Multitask joint prediction of infant cortical morphological and cognitive development. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 545–554. Springer (2023)
29. Zhao, F., Xia, S., Wu, Z., Duan, D., Wang, L., Lin, W., Gilmore, J.H., Shen, D., Li, G.: Spherical u-net on cortical surfaces: methods and applications. In: Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26. pp. 855–866. Springer (2019)