Covariate Correcting Networks for Identifying Associations Between Socioeconomic Factors and Brain Outcomes in Children

Hyuna Cho[1], Gunwoong Park[2], Amal Isaiah[3], and Won Hwa Kim[1,4(✉)]

[1] POSTECH, Pohang, South Korea
{hyunacho,wonhwa}@postech.ac.kr
[2] University of Seoul, Seoul, South Korea
[3] University of Maryland School of Medicine, Baltimore, USA
[4] University of Texas at Arlington, Arlington, USA

Abstract. Brain development in adolescence is synthetically influenced by various factors such as age, education, and socioeconomic conditions. To identify an independent effect from a variable of interest (e.g., socioeconomic conditions), statistical models such as General Linear Model (GLM) are typically adopted to account for covariates (e.g., age and gender). However, statistical models may be vulnerable with insufficient sample size and outliers, and multiple tests for a whole brain analysis lead to inevitable false-positives without sufficient sensitivity. Hence, it is necessary to develop a unified framework for multiple tests that robustly fits the observation and increases sensitivity. We therefore propose a unified flexible neural network that optimizes on the contribution from the main variable of interest as introduced in original GLM, which leads to improved statistical outcomes. The results on group analysis with fractional anisotropy (FA) from Diffusion Tensor Images from Adolescent Brain Cognitive Development (ABCD) study demonstrate that the proposed method provides much more selective and meaningful detection of ROIs related to socioeconomic status over conventional methods.

1 Introduction

Experiences consistent with different levels of socioeconomic status (SES) are key to brain development [4,15]. Brain structure and functional connectivity were thought to be explained solely by their heritable aspects, consolidated by magnetic resonance imaging (MRI) based studies of twins with shared characteristics for brain volume and structure [3,16,31]. High correlation in genetically similar regions of interest (ROIs) support genetic aspects of intelligence [28]. However, recent evidence points to SES being a principal determinant of brain structural characteristics such as Pediatric Imaging, Neurocognition and Genetics (PING) study [17] that offer unique insights into the associations of SES and children's brain structure. However, most studies remain cross-sectional and are susceptible to confounds that impact delineation of associations between SES and brain

M. de Bruijne et al. (Eds.): MICCAI 2021, LNCS 12907, pp. 421–431, 2021.
https://doi.org/10.1007/978-3-030-87234-2_40

development [1,5]. Regardless of the method, most conventional approaches do not effectively differentiate meaningful associations between environmental exposures and brain morphometric or diffusion characteristics with MRIs.

In conventional approaches, statistical tests with neuroimaging measures under Gaussian assumption are independently performed at each ROI to obtain evaluative measures (e.g., p-value), and correcting for multiple comparisons delineates ROIs that are statistically associated with a variable of interest X (e.g., diseased vs. healthy) [8,9]. As several factors affect the brain synthetically, statistical models such as General Linear Model (GLM) are utilized to control for effects from nuisance covariates Z and identify the true effect from X [14,18,24]. GLM achieves this by comparing a pair of models, i.e., Full and Reduced Models, where the Full Model considers X and Z together while the Reduced Model takes only Z to fit given observations. Investigating the difference of errors explained by the two models, i.e., analysis of variance (ANOVA), yields the true effect from the variable of interest X corrected for covariates Z [26].

Notice that, while GLM offers a satisfactory formulation to control for covariates to describe marginal effects, it does not necessarily increase sensitivity of the method. It uses Ordinary Least Square (OLS) to fit a pair of linear models [24], and performs F-test based on residuals from OLS, which can often be misleading owing to multicollinearity and outliers [32]. Moreover, a whole brain analysis on several voxels or ROIs requires multiple tests which ends up with inevitable false-positives [2,27]. In machine learning, Domain Adaptation (DA) provides a way to control for an unwanted variable, i.e., domain. DA operates with a source and a target domain where separate neural networks for each domain are compared for a predictive task to transfer knowledge from source to target [7,25]. While the architectures of GLM and DA are similar, the objectives are quite different as the target variable in DA is a covariate in GLM. Also, DA may become very complex to control multiple covariates.

To address the issues above, we propose to develop a novel Artificial Neural Network (ANN) architecture that constructs an ensemble of multiple pairs of linear models that correct for covariates and optimize on statistical sensitivity. The framework shares the same hypothesis space bias with GLM with a loss inspired by F-test; we choose to maximize the overall F-statistics with constraints from domain knowledge. Optimizing such an ANN model in a constrained space improves sensitivity corrected for covariates together with selective identification of task-specific ROIs. Our contributions here are summarized as: **1)** Devising a novel ANN that marginalizes effects from covariates, **2)** The framework can simultaneously optimize multiple models for whole brain analyses, **3)** Various experiments identify associations between socioeconomic status and brain outcomes, which were not available with conventional approaches.

2 Preliminary: General Linear Model

GLM is a generalization of multiple linear regression to the case of more than one target variable. Hence, it can be written as

$$Y_j = \beta_{1j} X_1 + \cdots + \beta_{pj} X_p + \epsilon_j$$

Y_j $(j = 1, \cdots, m)$ is target variable (e.g., ROI measures), $\{X_k\}_{k=1}^p$ is a set of independent variables (e.g., group), and ϵ_j is a Gaussian distributed error.

GLM provides a general statistical framework for testing various associations and hypotheses such as ANOVA, [26], analysis of covariance (ANCOVA) [26] and the multivariate analysis of covariance (MANCOVA) [30]. Hence, GLM is an appropriate method of determining the effect of variables of interest $X = (X_1, ..., X_p)$ on target variables Y_j even with nuisance covariates $Z = (Z_1, \cdots, Z_q)$ (e.g., age, gender, and so on). Specifically, GLM compares a Full Model (with X and Z) and a Reduced Model (with Z), which perform linear operations on Y_j:

$$\text{Full Model: } Y_j = Z\lambda_j^F + X\beta_j^F + \epsilon_j, \quad \text{Reduced Model: } Y_j = Z\lambda_j^R + \epsilon_j$$

where coefficients $\lambda_j^F, \lambda_j^R \in \mathbb{R}^q$ and $\beta_j^F \in \mathbb{R}^p$. In the finite sample setting, the influence of X for Y_j can be measured by comparing the sum of squared of errors (SSE) for the Full and Reduced Models:

$$SSE_j^R := \sum_{i=1}^n \left(y_{ij} - z_i\lambda_j^R\right)^2 \quad \text{and} \quad SSE_j^F := \sum_{i=1}^n \left(y_{ij} - z_i\lambda_j^F - x_i\beta_j^F\right)^2 \tag{1}$$

where $(y_{ij})_{i=1}^n$, $(x_i)_{i=1}^n$, and $(z_i)_{i=1}^n$ are set of observed values of Y_j, X, and Z for the i-th subject, respectively.

F-test is one of the popular statistical approaches to determine whether the influence of X is significant. More precisely, it considers the following hypotheses:

$$H_{0j} : \beta_j^F = \mathbf{0} \quad \text{vs.} \quad H_{1j} : \beta_j^F \neq \mathbf{0} \tag{2}$$

where $\mathbf{0} = (0, \cdots, 0)' \in \mathbb{R}^p$ and resultant statistics F_j^* is defined as

$$F_j^* := \left(\frac{SSE_j^R - SSE_j^F}{df_R - df_F}\right)\left(\frac{SSE_j^F}{df_F}\right)^{-1} \tag{3}$$

where $df_R = n-q$ and $df_F = n-p-q$ are the degrees of freedom from the Reduced and Full Model, respectively. The F-test rejects H_{0j} if F_j^* is sufficiently large. This is feasible because a larger F_j^* comes from a larger $SSE_j^R - SSE_j^F$ which means that X is providing auxiliary information on Y_j that covariates Z cannot explain. In general, coefficients β_j^F, λ_j^F and λ_j^R are estimated by OLS which can be vulnerable to outliers and lack of sample size. Also, m-multiple hypothesis tests often suffer from a multicollinearity and lack of sensitivity problems. These issues motivate to develop a new robust framework, to be described shortly.

3 CoCoNet: Covariate Correcting Network

Consider the same setting as in GLM (under Gaussain assumption) where a set of measurements Y_j at location j, a main variable of interest X (e.g., group),

and covariates Z (e.g., age and gender) across n samples are given. The aim here is to quantify the true effect from X on Y_j with *high sensitivity* across all j, while correcting for effects from confounding Z. For this, we design a novel framework, Covariate Correcting Network (CoCoNet), whose architecture is motivated from performing independent hypothesis tests across all j using GLM. Unlike conventional approaches, CoCoNet optimizes the multiple GLM at the same time using specialized ANN architecture with regularizers rather than performing multiple independent ANOVA.

From Sect. 2, as larger F_j^* is desired for those j's rejecting the H_{0j} showing *significant group differences*, the optimization should maximize the sum of F_j^* across j (i.e., minimizing its inverse). One can easily see from (3) that increasing SSE_j^R will increase F_j^*, hence naively optimizing on (3) will lead to loose fitting of the Reduced Model. To avoid this, Mean Squared Error (MSE) of predicting Y with X, Z and their coefficients from both models, i.e., MSE^F and MSE^R, are considered such that the regressions for both models become tight. Lastly, from a neuroscience perspective, it is intuitive to include a sparsity constraint as changes in the brain due to certain conditions may manifest on selective regions. We therefore include ℓ_1-norm penalty term on the $F^* = (F_1^*, \cdots, F_m^*)$ indicating sparse detection of ROIs that reject the null. Formulating the ideas discussed above, the initial loss to minimize is defined as

$$\frac{1}{\sum_j^m F_j^*} + \gamma_{\ell_1} \|F^*\|_{\ell_1} + \sum_j^m (\gamma_R MSE_j^R + \gamma_F MSE_j^F) \tag{4}$$

where γ_{ℓ_1}, γ_R and γ_F are hyperparameters to balance individual terms.

Carefully observing (3), maximizing F_j^* is equivalent to maximizing the ratio of SSE^R and SSE^F (as in the natural ANOVA) as

$$F_j^* = \left(\frac{SSE_j^R - SSE_j^F}{SSE_j^F}\right)\left(\frac{df_R - df_F}{df_F}\right)^{-1} \propto \left(\frac{SSE_j^R - SSE_j^F}{SSE_j^F}\right) \propto \left(\frac{SSE_j^R}{SSE_j^F}\right) \tag{5}$$

as the degrees of freedoms are constants. Moreover, applying ℓ_1 penalty on (3) is equivalent to imposing the ℓ_1 constraint on $SSE_j^R - SSE_j^F = 0$, denoting $\beta_j^F = \mathbf{0}$. In the end, the loss function (4) is revised as

$$L(\lambda, \beta) = \frac{1}{\sum_j^m \frac{SSE_j^R}{SSE_j^F}} + \gamma_{\ell_1} \|\beta^F\|_{\ell_1} + \sum_j^m \left(\frac{\gamma_R}{n - df_R} SSE_j^R + \frac{\gamma_F}{n - df_F} SSE_j^F\right) \tag{6}$$

where γ_{ℓ_1}, γ_R, and γ_F are hyperparameters that balance contributions from each term. It is important to set γ_R and γ_F properly to avoid underfitting of individual models, as the Full and Reduced Models converge at different speeds.

The overall architecture of CoCoNet is illustrated in Fig. 1, that consists of m pairs of Full and Reduced Models. Reduced Model takes q covariates, i.e., Z,

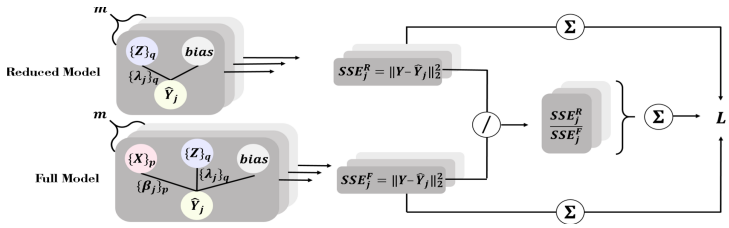

Fig. 1. Overview of CoCoNet structure, consisting of m Full Models and m Reduced Models for all target variables. The residuals from each model pairs are diversely used to define the loss function, which lead to maximizing statistical outcomes.

with a bias term. The input of the Full Model includes p variables of interest (i.e., X) along with Z and bias. The outputs from a single Reduced and Full Model pair are estimations $\hat{Y}_j^R$ and $\hat{Y}_j^F$ of the same target variable Y_j computed from pertinently trained weights and inputs so that the model ultimately minimizes (6). The loss L given in (6) to be minimized is computed from $\hat{Y}$'s based on SSE_j^R's and SSE_j^F's together with a ℓ_1 penalty on β.

4 Experimental Result

We performed several group analyses based on household income criteria; from high to subtle effect-size groups. The results show quantitative improvements with CoCoNet over baselines, which yield clinically meaningful results (Figs. 2 and 3).

4.1 Experiment on ABCD Dataset

Dataset. The Adolescent Brain Cognitive Development (ABCD) study [11] is an ongoing observational assessment of brain development in children recruited from 21 sites in the U.S. at 9–10 years of age. The rationale for study are provided by [6,29]. Enrolled children underwent cognitive, behavioral, demographic, health, and sleep assessments annually and brain imaging every two years. DTI from individuals were parcellated by co-registration with the Destrieux Atlas [10] comprising 148 regions, and average fractional anisotropy (FA) was computed at each ROI. The current study used the baseline dataset (v2.0.1) approved by institutional review boards. This fully processed data for our experiment is hosted by the National Institutes of Health Data Archive (NDA), and was downloaded and analyzed following a data use agreement.

Setup. Baselines are as follows: 1) conventional GLM, 2) GLM with Neural Network for regression (GLM_{NN}), 3) GLM with Lasso for the Full Model (GLM_{Lasso}). All of these methods analyze variance at each ROI independently and yield corresponding F-statistics and p-values. For GLM_{Lasso}, sparsity coefficient was set to 0.01. For CoCoNet, hyperparameters were set as: $m = 148$

Table 1. Demographics of the ABCD Dataset.

Category	BP	NP	High	Mid	Low
# of Subjects	952	8876	4205	2792	2831
Gender (M/F)	487/465	4608/4268	2207/1998	1448/1344	1440/1391
Age (Mean ± std)	118.5 ± 7.3	119.1 ± 7.5	119.4 ± 7.5	118.8 ± 7.5	118.7 ± 7.5
Scanner (Siemens/GE/Philips)	623/215/114	5719/2024/1133	2717/891/597	1866/591/335	1759/757/315

Table 2. ROIs identified using GLM from BP vs. NP groups and comparisons of statistics (p-values with Bonferroni correction at $\alpha = 0.01$ and F-statistics).

idx	ROI	p-value				F-statistic			
		GLM	GLM_{NN}	GLM_{Lasso}	CoCoNet	GLM	GLM_{NN}	GLM_{Lasso}	CoCoNet
1	left g.temp. sup.plan.polar	5.45E-05	5.45E-05	5.45E-05	1.34E-11	16.30	16.30	16.30	45.86
2	right g.and.s. cingul.ant	3.92E-05	3.92E-05	3.92E-05	1.11E-16	16.92	16.92	16.92	94.82
3	right g.front. sup	3.06E-05	3.06E-05	5.05E-05	6.66E-16	17.39	17.39	16.44	65.42
4	right g.oc.temp. med.parahip	2.15E-05	2.15E-05	2.21E-05	1.11E-16	18.06	18.07	18.02	111.76
5	right g.temp.sup. plan.polar	1.86E-06	1.86E-06	1.86E-06	1.11E-16	22.76	22.76	22.76	92.41
6	right s.front. sup	3.46E-06	3.46E-06	5.98E-06	1.18E-08	21.57	21.57	20.52	32.58

(i.e., # of ROIs), $p = 1$ (i.e., group), $q = 4$ (i.e., age, gender, $\text{scanner}_{1,2}$), $\gamma_{stat} = 1$, $\gamma_{\ell_1} = 10$, $\gamma_R = 0.1$, $\gamma_F = 0.1$, and learning rate was 0.01. CoCoNet was implemented with Pytorch and optimized using Adam optimizer [19]. All the p-values were corrected for multiple comparisons using Bonferroni correction at $\alpha = 0.01$.

Two sets of experiments based on several group criteria were performed. In the first experiment, the subjects were divided into two groups, i.e., Below Poverty and Non-Poverty, using the poverty criteria from U.S. Census Bureau ($16,910) [21]. In the latter, the subjects were categorized into three groups: 1) Low (<$50,000), 2) Mid (between $50,000 and $100,000), and High (≥$100,000) income groups based on [22] to look at more detailed development differences between Low vs. Middle and Middle vs. High income groups.

Below Poverty vs. Non-Poverty. We first performed a group analysis on Below Poverty (BP) vs. Non-Poverty (NP) groups. All models yielded several ROIs that showed statistically significant group differences between the BP and NP (6, 7, 6, and 52 ROIs for GLM, GLM_{NN}, GLM_{Lasso}, and CoCoNet respectively) and CoCoNet showed decrease in the p-values.

Quantitative results with F-statistics and p-values on the 6 ROIs that were detected with conventional GLM are compared in Table 2. All these ROIs were commonly detected across all models tested on the ABCD data. We first used a Linear NN, i.e., GLM_{NN}, to fit the same data with the same objective function of OLS, and used its regression result for F^* in (3) and p-values. This baseline

was designed to confirm that NN can replace OLS in GLM. As expected, we observed that both GLM with OLS and NN performed almost the same with very marginal difference as seen in the F^*'s and p-values. Lasso with penalty on β (i.e., GLM_{Lasso}) yielded increase in SSE^F as the parameter space was restricted by the regularizer. However, it still resulted in very similar results with GLM and GLM_{NN} in their F^* and p-values on the ROIs detected with GLM. On the other hand, notice that both F^* and resultant p-values from CoCoNet got significantly better over the three baseline approaches. This result was obtained by leveraging a slightly different model fitting; we observed $\sim$0.01 decrease on average in R^2 for the Reduced Model but overall statistical outcomes have improved. Some of the F^* went negative for unimportant ROIs with GLM_{Lasso} and CoCoNet due to regularizers, but it was not a problem for the ROIs with β's close to 0.

Low vs. Middle vs. High Income Group Analyses. The Low vs. Middle/ Middle vs. High group comparisons are more challenging than the BP vs. NP analysis as their effect sizes are more subtle in the income spectrum. The results are summarized in Table 3 with the identified ROI labels and their p-values.

Table 3. Identified ROIs and p-values from Low vs. Mid (Left) and Mid vs. High income (Right) analyses. (detected ROIs surviving Bonferroni at α=0.01 in bold)

(a) Low vs. Middle income

idx	ROI	GLM	GLM_{NN}	GLM_{Lasso}	CoCoNet
1	left g.and.s.transv.frontopol	(4.45E-01)	(4.45E-01)	(1.00)	**8.03E-13**
2	left s.front.sup	(4.22E-01)	(4.22E-01)	(4.41E-01)	**1.11E-16**
3	left s.interm.prim.jensen	(2.59E-03)	**4.26E-05**	(2.59E-03)	**1.11E-16**
4	left s.postcentral	(2.30E-01)	(2.30E-01)	(1.00)	**5.45E-14**
5	left s.precentral.inf.part	(2.00E-01)	(2.00E-01)	(2.00E-01)	**1.11E-16**
6	right g.oc.temp.med.parahip	**3.63E-05**	**3.63E-05**	**3.63E-05**	**1.11E-16**
7	right g.temp.sup.plan.polar	(7.81E-04)	**3.30E-05**	(7.81E-04)	**8.68E-10**
8	right s.calcarine	(1.18E-01)	(1.18E-01)	(1.00)	**6.99E-15**
9	right s.front.middle	(1.68E-01)	(1.68E-01)	(1.00)	**1.10E-05**
10	right s.orbital.h.shaped	(1.55E-01)	(1.55E-01)	(1.00)	**1.11E-16**
11	right s.parieto.occipital	(1.78E-03)	(1.78E-03)	(1.78E-03)	**1.11E-16**

(b) Middle vs. High income

idx	ROI	GLM_{NN}	CoCoNet
1	left g.front.inf.triangul	(1.00)	**1.10E-08**
2	left g.oc.temp.med.lingual	(1.02E-01)	**1.11E-16**
3	left g.pariet.inf.supramar	(1.31E-01)	**1.85E-06**
4	left g.parietal.sup	(1.35E-01)	**2.59E-08**
5	left g.postcentral	**3.65E-05**	**1.11E-16**
6	left g.temp.sup.lateral	(1.95E-01)	**1.02E-11**
7	left s.orbital.med.olfact	(7.96E-01)	**5.62E-08**
8	right g.front.middle	(9.58E-02)	**4.82E-13**
9	right g.front.sup	**6.38E-05**	**1.11E-16**
10	right g.temp.sup.plan.tempo	(9.70E-01)	**1.11E-16**
11	right s.occipital.ant	(4.98E-01)	**2.13E-09**
12	right s.oc.temp.lat	(7.38E-01)	**1.11E-16**
13	right s.temporal.transverse	(6.53E-01)	**1.11E-16**

In the comparison of Low vs. Middle income groups, CoCoNet identified 11 ROIs of which p-values survive Bonferroni correction at $\alpha = 0.01$, while GLM and GLM_{Lasso} yielding only 1 and GLM_{NN} detecting 3 ROIs subsumed by the 11 ROIs from CoCoNet. These results are visually compared in Fig. 2. Notably, the overall p-values from CoCoNet are distributed at smaller values than those from baselines, meaning that CoCoNet was able to intensively detect valid ROIs.

In the Middle vs. High group analysis, GLM and GLM_{Lasso} did not yield any ROIs surviving Bonferroni correction even after hyperparameter tuning. However, GLM_{NN} and CoCoNet detected 2 and 13 regions respectively. Comparing the p-values of the detected ROIs, the associations between socioeconomic characteristics and brain outcomes were much more sensitively identified with CoCoNet. These descriptions can be visually seen in Fig. 3 whose first row is the p-value map from GLM_{NN} and the second row is that from CoCoNet.

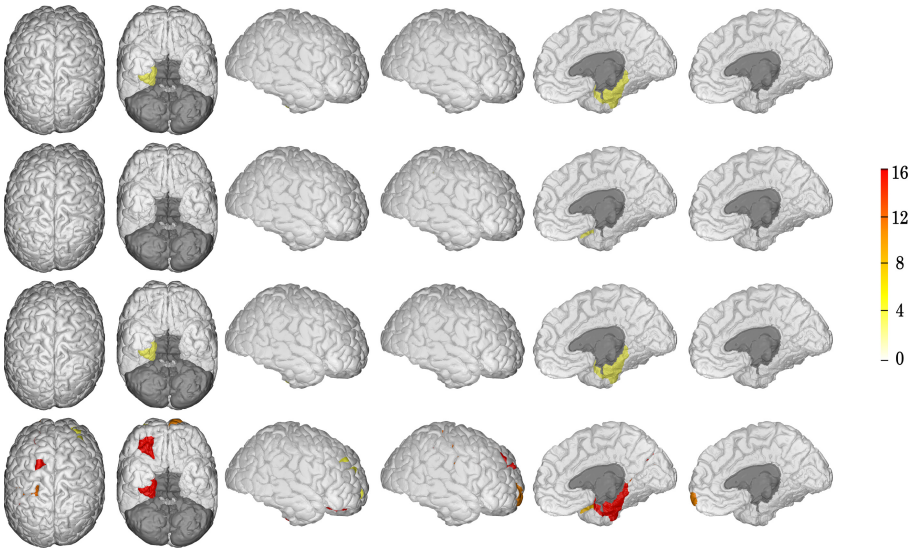

Fig. 2. p-value maps in $-\log_{10}$ scale from Low vs. Mid income group analysis. First row: GLM, Second row: GLM_{NN}, Third row: GLM_{Lasso}, Fourth row: CoCoNet. CoCoNet detects more ROIs with lower p-values than the baselines.

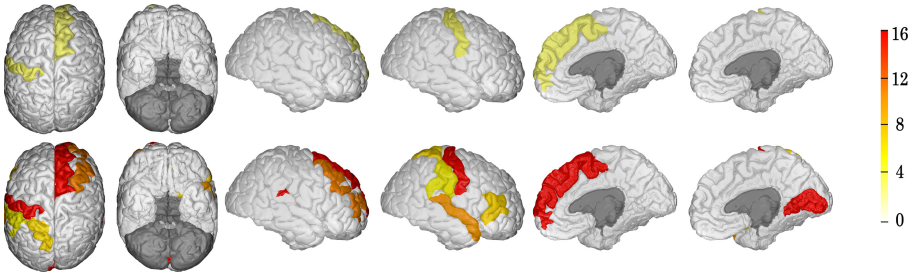

Fig. 3. Middle vs. High income group analysis and resultant p-value maps in $-\log_{10}$ scale. First row: GLM_{NN}, Second row: CoCoNet. GLM and GLM_{Lasso} yielded no ROIs.

4.2 Discussions on the Results

Statistical Parametric Mapping (SPM) is a conventional approach. While SPM performs multiple ROI/voxel-wise independent regressions, we proposed to solve them simultaneously as a unified framework. Almost identical results between GLM and GLM_{NN} show this is doable, which share the same inductive bias with CoCoNet without ℓ_1-penalty. With ℓ_1-penalty in CoCoNet, we observed marginal decrease in R^2 and the residuals should be very close to F-distribution.

The results demonstrate substantially improved performance over conventional regression-based tools in isolating distinct ROIs with structural alterna-

tions. Notably, these results support the refinement of current methods used to identify meaningful effects associated with environmental exposures that impact brain development. As effect sizes are tied to sample size in imaging-based datasets such as the ABCD study [11], there is a need to extend conventional approaches prone to errors tied to differences detected by chance alone.

Our results show that the right parahippocampal gyrus has the highest F-statistic with CoCoNet when compared to GLM-based methods. This is an ROI implicated in memory encoding and retrieval, which are key to overall cognition in children. Other meaningful ROIs include the anterior cingulate and the planum temporale, which are regions critical to motivational learning and executive function, as well as language development. Our method consistently identified key ROIs such as the post-central sulcus [33] and the calcarine sulcus [20] in comparisons of children from low and middle income groups. These ROIs, responsible for the somatosensory and visual attributes of cortical processing, appear to be consistent with regional substrates for specific aspects of sensory processing thought to be impacted by environmental exposures [12]. Fusiform gyrus [23] and parts of the parietal cortex [13] (e.g. supramarginal gyrus) are detected from middle and high income comparison using CoCoNet. Biological specificity of structural alterations in the brain is emphasized from the results.

5 Conclusion

In this paper, we developed a novel framework, i.e., CoCoNet, which is an ANN that finds appropriate regressions to enhance multiple ANOVA. The results demonstrate substantially improved performance over conventional tools in isolating brain regions with structural alternations potentially impacted by environmental factors. We see very high potential of CoCoNet being adopted in various studies that require higher sensitivity accounting for confounding effects.

Acknowledgement. This research was supported by NSF IIS CRII 1948510, NSF IIS Core 2008602, NIH R03 AG070701, NIH RF1 AG059312, and IITP-2019- 0-01906 funded by MSIT (AI Graduate School Program at POSTECH), and partially supported by NRF-2021R1C1C1004562 funded by MSIT. The ABCD Study is supported by NIH and federal partners under various awards: U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters at https://abcdstudy.org/federal-partners, A listing of participating sites and study investigators at https://abcdstudy.org/consortium_members. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

References

1. Ahirwar, A.: Study of techniques used for medical image segmentation and computation of statistical test for region classification of brain MRI. IJ Inf. Technol. Comput. Sci. **5**(5), 44–53 (2013)
2. Barnes, G.R., Litvak, V., Brookes, M.J., et al.: Controlling false positive rates in mass-multivariate tests for electromagnetic responses. Neuroimage **56**(3), 1072–1081 (2011)
3. Bowyer, R.C., Jackson, M.A., Le Roy, C.I., et al.: Socioeconomic status and the gut microbiome: a Twinsuk cohort study. Microorganisms **7**(1), 17 (2019)
4. Brito, N.H., Noble, K.G.: Socioeconomic status and structural brain development. Front. Neurosci. **8**, 276 (2014)
5. Bullmore, E.T., Suckling, J., Overmeyer, S., et al.: Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. IEEE Trans. Med. Imaging **18**(1), 32–42 (1999)
6. Casey, B., Cannonier, T., Conley, M.I., et al.: The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. Dev. Cogn. Neurosci. **32**, 43–54 (2018)
7. Daumé III, H.: Frustratingly easy domain adaptation. arXiv preprint arXiv:0907.1815 (2009)
8. Della Nave, R., Ginestroni, A., Tessa, C., et al.: Brain white matter damage in sca1 and sca2. an in vivo study using voxel-based morphometry, histogram analysis of mean diffusivity and tract-based spatial statistics. Neuroimage **43**(1), 10–19 (2008)
9. Della Nave, R., Ginestroni, A., Tessa, C., et al.: Brain white matter tracts degeneration in Friedreich ataxia. An in vivo MRI study using tract-based spatial statistics and voxel-based morphometry. Neuroimage **40**(1), 19–25 (2008)
10. Destrieux, C., Fischl, B., Dale, A., et al.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage **53**(1), 1–15 (2010)
11. Dick, A.S., Lopez, D.A., Watts, A.L., et al.: Meaningful associations in the adolescent brain cognitive development study. BioRxiv (2021). https://doi.org/10.1101/2020.09.01.276451
12. Farah, M.J., Shera, D.M., Savage, J.H., et al.: Childhood poverty: specific associations with neurocognitive development. Brain Res. **1110**(1), 166–174 (2006)
13. Fogassi, L., Ferrari, P.F., Gesierich, B., et al.: Parietal lobe: from action organization to intention understanding. Science **308**(5722), 662–667 (2005)
14. Glueck, D.H., Muller, K.E.: Adjusting power for a baseline covariate in linear models. Stat. Med. **22**(16), 2535–2551 (2003)
15. Hackman, D.A., Farah, M.J.: Socioeconomic status and the developing brain. Trends Cogn. Sci. **13**(2), 65–73 (2009)
16. Ivanovic, D.M., Leiva, B.P., Pérez, H.T., et al.: Nutritional status, brain development and scholastic achievement of Chilean high-school graduates from high and low intellectual quotient and socio-economic status. Br. J. Nutr. **87**(1), 81–92 (2002)
17. Jernigan, T.L., Brown, T.T., Hagler Jr., D.J., et al.: The pediatric imaging, neurocognition, and genetics (ping) data repository. Neuroimage **124**, 1149–1154 (2016)
18. Kim, W.H., Adluru, N., Chung, M.K., et al.: Multi-resolution statistical analysis of brain connectivity graphs in preclinical Alzheimer's disease. Neuroimage **118**, 103–117 (2015)

19. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
20. Lambert, S., Sampaio, E., Scheiber, C., et al.: Neural substrates of animal mental imagery: calcarine sulcus and dorsal pathway involvement-an fMRI study. Brain Res. **924**(2), 176–183 (2002)
21. Lee, A.: Us poverty thresholds and poverty guidelines: What's the difference. Population Reference Bureau (2018)
22. Marshall, A.T., Betts, S., Kan, E.C., et al.: Association of lead-exposure risk and family income with childhood brain outcomes. Nat. Med. **26**(1), 91–97 (2020)
23. McCarthy, G., Puce, A., Gore, J.C., et al.: Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. **9**(5), 605–610 (1997)
24. Oakes, T.R., Fox, A.S., Johnstone, T., et al.: Integrating VBM into the general linear model with voxelwise anatomical covariates. Neuroimage **34**(2), 500–508 (2007)
25. Pan, S.J., Tsang, I.W., Kwok, J.T., et al.: Domain adaptation via transfer component analysis. IEEE Trans. Neural Networks **22**(2), 199–210 (2010)
26. Rutherford, A.: ANOVA and ANCOVA: A GLM Approach. Wiley (2011)
27. Scarpazza, C., Tognin, S., Frisciata, S., et al.: False positive rates in voxel-based morphometry studies of the human brain: should we be worried? Neurosci. Biobehav. Rev. **52**, 49–55 (2015)
28. Thompson, P.M., Cannon, T.D., Narr, K.L., et al.: Genetic influences on brain structure. Nat. Neurosci. **4**(12), 1253–1258 (2001)
29. Volkow, N.D., Koob, G.F., Croyle, R.T., et al.: The conception of the ABCD study: From substance use to a broad NIH collaboration. Dev. Cogn. Neurosci. **32**, 4–7 (2018)
30. Weerahandi, S.: Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models, vol. 500. Wiley (2004)
31. Yang, F., Isaiah, A., Kim, W.H.: COVLET: covariance-based wavelet-like transform for statistical analysis of brain characteristics in children. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12267, pp. 83–93. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59728-3_9
32. Zimmerman, D.W.: Increasing the power of the ANOVA F test for outlier-prone distributions by modified ranking methods. J. Gen. Psychol. **122**(1), 83–94 (1995)
33. Zlatkina, V., Amiez, C., Petrides, M.: The postcentral SULCAL complex and the transverse postcentral sulcus and their relation to sensorimotor functional organization. Eur. J. Neurosci. **43**(10), 1268–1283 (2016)