A Dirty Multi-task Learning Method for Multi-modal Brain Imaging Genetics

Lei Du[1(✉)], Fang Liu[1], Kefei Liu[2], Xiaohui Yao[2], Shannon L. Risacher[3], Junwei Han[1], Lei Guo[1], Andrew J. Saykin[3], Li Shen[2], and for the Alzheimer's Disease Neuroimaging Initiative

[1] School of Automation, Northwestern Polytechnical University, Xi'an, China
dulei@nwpu.edu.cn
[2] Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
Li.Shen@pennmedicine.upenn.edu
[3] Indiana University School of Medicine, Indianapolis, IN, USA

Abstract. Brain imaging genetics is an important research topic in brain science, which combines genetic variations and brain structures or functions to uncover the genetic basis of brain disorders. Imaging data collected by different technologies, measuring the same brain distinctly, might carry complementary but different information. Unfortunately, we do not know the extent to which phenotypic variance is shared among multiple imaging modalities, which might trace back to the complex genetic mechanism. In this study, we propose a novel dirty multi-task SCCA to analyze imaging genetics problems with multiple modalities of brain imaging quantitative traits (QTs) involved. The proposed method can not only identify the shared SNPs and QTs across multiple modalities, but also identify the modality-specific SNPs and QTs, showing a flexible capability of discovering the complex multi-SNP-multi-QT associations. Compared with the multi-view SCCA and multi-task SCCA, our method shows better canonical correlation coefficients and canonical weights on both synthetic and real neuroimaging genetic data. This demonstrates that the proposed dirty multi-task SCCA could be a meaningful and powerful alternative method in multi-modal brain imaging genetics.

This work was supported by NSFC [61602384]; NSFC of Shaanxi [2017JQ6001]; CPSF [2017M613202]; PSF of Shaanxi [2017BSHEDZZ81]; Fundamental Research Funds for Central Universities [3102018zy029] at Northwestern Polytechnical University. This work was also supported by the National Institutes of Health [R01 EB022574, R01 LM011360, U01 AG024904, P30 AG10133, R01 AG19771, R01 AG 042437, R01 AG046171, R01 AG040770 and NSF IIS 1837964] at University of Pennsylvania and Indiana University.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

D. Shen et al. (Eds.): MICCAI 2019, LNCS 11767, pp. 447–455, 2019.
https://doi.org/10.1007/978-3-030-32251-9_49

Keywords: Brain imaging genetics · Multi-task sparse canonical correlation analysis · Multi-modal brain imaging

1 Introduction

Imaging genetics gains more and more attention recently. The primal aim of imaging genetics is to uncover the genetic basis of brain disorders [6]. Hence the genetic variations such as the single nucleotide polymorphisms (SNPs) and neuroimaging quantitative traits (QTs) are usually analyzed together. The imaging QTs obtained by different image technologies, measuring the brain from different perspectives, might carry complementary but different information. Therefore, combining multi-modal imaging QTs, and using them to study the modality-consistent biomarkers as well as the modality-dependent biomarkers could be beneficial to exploit meaningful genetic mechanism for brain disorders.

Both regression-based multi-task learning (MTL) and sparse canonical correlation analysis (SCCA) are widely used in brain imaging genetics [7]. The MTL methods only select features for the predicting variables [5,8], thereby are insufficient if we pursue a simultaneous feature selection for both SNPs and imaging QTs. SCCA improves the MTL methods, and yields a pair of canonical weights showing the relevance of SNPs and QTs simultaneously [2,10]. However, they could not load multi-modal imaging QTs in a unified model, resulting in the lack of the identification ability. The multi-view/multi-set SCCA studies the relationship among multiple sets of data, thereby could handle multi-modal imaging genetics. Unfortunately, similar to the multi-task SCCA (MTSCCA) [1,3], it could not identify diverse imaging genetic patterns such as the modality-consistent and modality-dependent associations.

In this paper, we propose a novel multi-modal imaging data oriented imaging genetic learning method. Our method takes advantage of MTL and parameters decomposition. The MTL modeling strategy makes it easier and reasonable to incorporate multiple modalities of imaging QTs, and the parameters decomposition accommodates a flexible regularization. We call it the *dirty* MTSCCA following the terminology in [4]. The proposed dirty method decomposes the canonical weights into two parts, i.e. the task-consistent component which is shared among all tasks, and the task-dependent component which is close related to a specific task. By penalizing both components distinctly, the proposed method is able to identify both SNPs and imaging QTs that are revealed by all imaging technologies, as well as SNPs and QTs that could be only revealed using a specific imaging technology. We propose an efficient algorithm to solve the dirty MTSCCA which converges to a local optimum. The results on the real neuroimaging genetics data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [9] show that, the dirty MTSCCA learns improved bi-multivariate associations, the modality-shared SNPs and brain areas, as well as the modality-specific SNPs and brain areas, exhibiting a flexible and meaningful identification capability. Therefore, our dirty MTSCCA model is quite suitable to multi-modal imaging genetics, and further a significant addition to the existing method library.

2 The Dirty Multi-task SCCA

The Model. In this paper, we denote scalars as italic letters, column vectors as boldface lowercase letters, and matrices as boldface capitals. The i-th row and j-th column of $\mathbf{X} = (x_{ij})$ is denoted as $\mathbf{x}^i$ and $\mathbf{x}_j$ respectively. $\|\mathbf{x}\|_2$ denotes the Euclidean norm, $\|\mathbf{X}\|_{\mathrm{F}} = \sqrt{\sum_i \sum_j x_{ij}^2}$ denotes the Frobenius norm, and $\|\mathbf{X}\|_{2,1}$ denotes the $\ell_{2,1}$-norm. $\mathbf{X} \in \mathbb{R}^{n \times p}$ loads the genetic data with n subjects and p SNPs, and $\mathbf{Y}_c \in \mathbb{R}^{n \times q} (c = 1, \cdots, C)$ loads the c-th modality of phenotype data with q imaging QTs, where C is the number of imaging modalities (tasks).

To discover the flexible imaging genetic patterns in a multi-modal scene, we propose the dirty multi-task SCCA as follows

$$\begin{aligned} \min_{\mathbf{S},\mathbf{W},\mathbf{B},\mathbf{Z}} & \sum_{c=1}^{C} -(\mathbf{s}_c + \mathbf{w}_c)^\top \mathbf{X}^\top \mathbf{Y}_c (\mathbf{b}_c + \mathbf{z}_c) \\ & + \lambda_s \|\mathbf{S}\|_{\mathrm{G}_{2,1}} + \beta_s \|\mathbf{S}\|_{2,1} + \lambda_w \|\mathbf{W}\|_1 + \beta_b \|\mathbf{B}\|_{2,1} + \lambda_z \|\mathbf{Z}\|_1 \\ & s.t. \quad \|\mathbf{X}(\mathbf{s}_c + \mathbf{w}_c)\|_2^2 \leq 1, \|\mathbf{Y}_c(\mathbf{b}_c + \mathbf{z}_c)\|_2^2 \leq 1, \ \forall c. \end{aligned} \tag{1}$$

In this model, the conventional canonical weights $\mathbf{U}$ and $\mathbf{V}$ are decomposed as $\mathbf{U} = \mathbf{S} + \mathbf{W}$ and $\mathbf{V} = \mathbf{B} + \mathbf{Z}$. The $\mathbf{S}$ and $\mathbf{W}$ are associated with the SNP data, where $\mathbf{S}$ is the task-consistent component shared by multiple tasks, and $\mathbf{W}$ is the task-dependent component only associated with a single task. Similarly, $\mathbf{B}$ is the task-consistent component and $\mathbf{Z}$ is the task-dependent component for imaging QTs. The λ_s, β_s, λ_w, β_b, λ_z are nonnegative tuning parameters.

The model above encourages the modality-shared sparsity [4] and modality-dependent sparsity via distinct regularization terms. It penalizes the task-consist-ent component jointly by the block-sparse regularization, such as the $\mathrm{G}_{2,1}$-norm (definition is in [8]) and $\ell_{2,1}$-norm for SNPs and $\ell_{2,1}$-norm for QTs. This could help identify the shared SNPs and imaging QTs repeated by different imaging technologies. In addition, the task-dependent component is penalized differently by the ℓ_1-norm to induce element-wise sparsity for both SNPs and imaging QTs. This might uncover the SNPs and QTs that could only be identified using a specific imaging technology. Owing to this decomposition strategy, the proposed method is able to facilitate joint feature selection while allowing disparities as well. Since simultaneously demanding features being task-consistent and task-dependent is conflicting, the proposed dirty model is flexible and thus practical.

The Optimization Algorithm. The Eq. (1) is not a jointly convex function and thus can not be solved directly. Fortunately, it is convex in $\mathbf{S}$ if we fix $\mathbf{W}$, $\mathbf{B}$ and $\mathbf{Z}$. Likewise, Eq. (1) is also convex in $\mathbf{W}$, $\mathbf{B}$ and $\mathbf{Z}$ alternately with those remaining weight matrices being fixed. On this account, the dirty multi-task SCCA problem can be solved using the alternative iteration algorithm.

Updating* S *and* W*: When **B** and **Z** are fixed to constants, the objective with respect to **S** and **W** can be rewritten as

$$\min_{\mathbf{S},\mathbf{W}} \sum_{c=1}^{C} -(\mathbf{s}_c+\mathbf{w}_c)^\top \mathbf{X}^\top \mathbf{Y}_c(\mathbf{b}_c+\mathbf{z}_c)+\lambda_s \|\mathbf{S}\|_{\mathrm{G}_{2,1}}+\beta_s \|\mathbf{S}\|_{2,1}+\lambda_w \|\mathbf{W}\|_1 , \quad s.t. \|\mathbf{X}(\mathbf{s}_c+\mathbf{w}_c)\|_2^2 \le 1, \tag{2}$$

which can be solved by the following theorem.

Theorem 1. *The solution of Eq. (2) is given by* $\mathbf{s}_c^* = \frac{\hat{\mathbf{s}}_c}{\|\mathbf{X}(\hat{\mathbf{s}}_c+\hat{\mathbf{w}}_c)\|_2}$ *and* $\mathbf{w}_c^* = \frac{\hat{\mathbf{w}}_c}{\|\mathbf{X}(\hat{\mathbf{s}}_c+\hat{\mathbf{w}}_c)\|_2}$*, where* $\hat{\mathbf{s}}_c$ *is the solution of*

$$\min_{\mathbf{S}} \sum_{c=1}^{C} \frac{1}{2} \|\mathbf{X}\mathbf{s}_c - \mathbf{Y}_c(\mathbf{b}_c+\mathbf{z}_c)\|_2^2 + \lambda_s \|\mathbf{S}\|_{\mathrm{G}_{2,1}} + \beta_s \|\mathbf{S}\|_{2,1} , \tag{3}$$

and $\hat{\mathbf{w}}_c$ *is the solution of*

$$\min_{\mathbf{W}} \sum_{c=1}^{C} \frac{1}{2} \|\mathbf{X}\mathbf{w}_c - \mathbf{Y}_c(\mathbf{b}_c+\mathbf{z}_c)\|_2^2 + \lambda_w \|\mathbf{W}\|_1 . \tag{4}$$

Theorem 1 can be proved following the same procedure in [10] (Appendix A.2). Now Eq. (3) becomes a multi-task regression, and can be solved using the off-the-shelf methods. Given that $\mathbf{s}_c$'s can be solved jointly, we take the derivative of the objective of Eq. (3) with respect to **S**, and then let it be zero, i.e.

$$(\mathbf{X}^\top\mathbf{X} + \lambda_s\tilde{\mathbf{D}} + \beta_s\mathbf{D})\mathbf{S} = \mathbf{X}^\top[\mathbf{Y}_1(\mathbf{b}_1+\mathbf{z}_1) \ \cdots \ \mathbf{Y}_C(\mathbf{b}_C+\mathbf{z}_C)], \tag{5}$$

where $\tilde{\mathbf{D}}$ is a block diagonal matrix with the k-th block being $\frac{1}{\|\mathbf{S}^k\|_{\mathrm{F}}}\mathbf{I}_k$, and $\mathbf{I}_k$ is an identity matrix of size equaling to the size of the k-th group. The grouping information could be previously defined according to the linkage disequilibrium (LD) structure of SNPs. **D** is a diagonal matrix with the i-th diagonal element being $\frac{1}{\|\mathbf{s}^i\|_2}(i = 1, \cdots, p)$. Then we can easily obtain $\hat{\mathbf{S}}$ from Eq. (5), which can be efficiently solved via the iterative algorithm [8].

Due to the ℓ_1-norm regularization, $\mathbf{w}_c$'s are not coupled and thus can be updated separately. By first taking the derivative of Eq. (4) regarding each $\mathbf{w}_c$, and letting it be zero, we arrive at

$$(\mathbf{X}^\top\mathbf{X} + \lambda_w\breve{\mathbf{D}}_c)\mathbf{w}_c = \mathbf{X}^\top\mathbf{Y}_c(\mathbf{b}_c+\mathbf{z}_c), \tag{6}$$

with $\breve{\mathbf{D}}_c$ being a diagonal matrix whose i-th element is $\frac{1}{|w_{ic}|}(i = 1, \cdots, p)$.

***Updating* B *and* Z:** Once we obtain **S** and **W**, we can fix them to solve **B** and **Z**. Since each $\mathbf{b}_c$ and $\mathbf{z}_c$ are associated with each modality of imaging QTs, i.e. $\mathbf{Y}_c$, they are not closely coupled too. Thus $\mathbf{b}_c$ and $\mathbf{z}_c$ for different task can be solved separately.

With this observation, we follow the same procedure to that of updating $\mathbf{w}_c$ according to Theorem 1. Taking the derivative of the objective with respect to every $\mathbf{b}_c$, we easily obtain

$$(\mathbf{Y}_c^\top \mathbf{Y}_c + \beta_b \mathbf{Q})\mathbf{b}_c = \mathbf{Y}_c^\top \mathbf{X}(\mathbf{s}_c + \mathbf{w}_c), \tag{7}$$

with $\mathbf{Q}$ being a diagonal matrix and the j-th element is $\frac{1}{\|\mathbf{b}^j\|_2}(j = 1, \cdots, q)$.

Similarly, we have

$$(\mathbf{Y}_c^\top \mathbf{Y}_c + \lambda_z \breve{\mathbf{Q}}_c)\mathbf{z}_c = \mathbf{Y}_c^\top \mathbf{X}(\mathbf{s}_c + \mathbf{w}_c), \tag{8}$$

where $\breve{\mathbf{Q}}_c$ is a diagonal matrix whose j-th diagonal element is $\frac{1}{|z_j|}(j = 1, \cdots, q)$.

Equations (5)–(8) pave the way to solve the problem Eq. (1), and then we present the pseudo-code in Algorithm 1. The algorithm iteratively updates $\mathbf{S}$, $\mathbf{W}$, $\mathbf{B}$ and $\mathbf{Z}$ till the predefined stopping conditions are satisfied. Further, this algorithm is guaranteed to converge to a local optimum since Eqs. (5–8) converge according to the Theorem 1 in [8].

Algorithm 1. The Dirty Multi-task SCCA Algorithm

Require:

$\mathbf{X} \in \mathcal{R}^{n \times p}$, $\mathbf{Y}_c \in \mathcal{R}^{n \times q}$, $c = 1, \cdots, C$, λ_s, β_s, λ_w, β_b, λ_z

Ensure:

Output $\mathbf{S}$, $\mathbf{W}$, $\mathbf{B}$, $\mathbf{Z}$.

1: Initialize $\mathbf{S} \in \mathcal{R}^{p \times C}$, $\mathbf{W} \in \mathcal{R}^{p \times C}$, $\mathbf{B} \in \mathcal{R}^{q \times C}$ and $\mathbf{Z} \in \mathcal{R}^{q \times C}$;
2: **while** not convergence **do**
3: Update $\hat{\mathbf{S}}$ according to Eq. (5), and update $\hat{\mathbf{w}}_c$ according to Eq. (6);
4: Compute $\mathbf{S}^*$ and $\mathbf{W}^*$ according to Theorem 1;
5: Update $\hat{\mathbf{b}}_c$ according to Eq. (7), and update $\hat{\mathbf{z}}_c$ according to Eq. (8);
6: Compute $\mathbf{B}^*$ and $\mathbf{Z}^*$ according to $\mathbf{b}_c^* = \frac{\hat{\mathbf{b}}_c}{\|\mathbf{Y}_c(\hat{\mathbf{b}}_c + \hat{\mathbf{z}}_c)\|_2}$, and $\mathbf{z}_c^* = \frac{\hat{\mathbf{z}}_c}{\|\mathbf{Y}_c(\hat{\mathbf{b}}_c + \hat{\mathbf{z}}_c)\|_2}$;
7: **end while**

3 Experiments

We choose the most related MTSCCA (multi-task SCCA) [1] and mSCCA (multi-view/multi-set SCCA) [10] as benchmarks. The experiments are conducted via a nested 5-fold cross-validation method where the inner loop is for parameter tuning. In this study, all methods run on the same platform and data partition, and employ the same stopping condition, i.e. both $\max_c |(\mathbf{s}_c + \mathbf{w}_c)^{t+1} - (\mathbf{s}_c + \mathbf{w}_c)^t| \leq 10^{-5}$ and $\max_c |(\mathbf{b}_c + \mathbf{z}_c)^{t+1} - (\mathbf{b}_c + \mathbf{z}_c)^t| \leq 10^{-5}$.

Real Neuroimaging Genetics Data. The genotying and brain imaging data used in this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). One primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission

tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org. The SNP, MRI and PET data were downloaded from the LONI website (adni.loni.usc.edu). Table 1 shows the details of 755 non-Hispanic Caucasian participants, including 281 AD, 292 MCI and 182 healthy control (HC). There were three modalities of imaging data, i.e. the 18-F florbetapir PET (AV45) scans, fluorodeoxyglucose positron emission tomography (FDG) scans, and structural MRI (sMRI) scans. These multi-modality imaging data had been aligned to every participant's same visit. The sMRI data were processed with voxel-based morphometry (VBM) via SPM. These scans were aligned to a T1-weighted template image, segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) maps, normalized to the standard MNI space, and smoothed with an 8 mm FWHM kernel. The FDG and AV45 scans were also registered into the same MNI space. Then a subsample step was conducted and 116 regions of interest (ROI) level measurements were generated based on the MarsBaR automated anatomical labeling (AAL) atlas. These imaging QTs were pre-adjusted to remove the effects of the baseline age, gender, education, and handedness by the regression weights generated from HCs. We investigated 1011 SNPs from chromosome 19 including the well-known AD risk genes such as *APOE*. The LD block information is also used as prior knowledge. Our goal was to examine correlations between the multiple modalities of QTs (GM densities for sMRI scans, amyloid values for AV45 scans and glucose utilization for FDG scans) and SNPs.

Table 1. Participant characteristics.

	HC	MCI	AD
Num	182	292	281
Gender(M/F, %)	47.16/52.84	54.52/45.48	47.37/52.63
Handedness(R/L, %)	90.91/9.09	87.35/12.65	91.50/8.50
Age (mean±std)	72.97±6.00	71.81±7.62	72.38±7.31
Education (mean±std)	16.52±2.58	15.97±2.78	16.14±2.78

Bi-multivariate Associations. Table 2 contains both training and testing canonical correlation coefficients (CCCs), showing the identified bi-multivariate associations. There are three CCCs for each method since we have three imaging modalities, thereby three SCCA tasks. The proposed method obtains better CCCs than both mSCCA and MTSCCA, which is further confirmed by the p-values (p-values for SNP-AV45 on testing set look strange, but are normal due to the directionality of the paired t-test) between our method and benchmarks. This demonstrates that, by decomposing the canonical weights and penalizing them distinctly, our method has the superior modeling capability in multi-modal scenes, thereby exhibits improved bi-multivariate associations.

Table 2. CCCs (mean±std.) between SNP and three modalities of imaging QTs. p-values between our method and benchmarks are also shown in parentheses.

	Training CCCs			Testing CCCs		
	SNP-AV45	SNP-FDG	SNP-VBM	SNP-AV45	SNP-FDG	SNP-VBM
mSCCA	0.44 ± 0.01 (8.69E−09)	0.33 ± 0.01 (1.80E−07)	0.25 ± 0.02 (2.60E−05)	0.41 ± 0.07 (1.4E−01)	0.29 ± 0.07 (3.2E−01)	0.21 ± 0.07 (4.1E−01)
MTSCCA	0.47 ± 0.01 (3.50E−09)	0.35 ± 0.01 (1.4E−01)	0.29 ± 0.01 (1.09E−02)	0.43 ± 0.07 (1.02E−03)	0.30 ± 0.06 (2.00E−01)	0.19 ± 0.08 (2.90E−02)
Our method	0.48 ± 0.01	0.36 ± 0.01	0.29 ± 0.01	0.44 ± 0.07	0.30 ± 0.06	0.21 ± 0.07

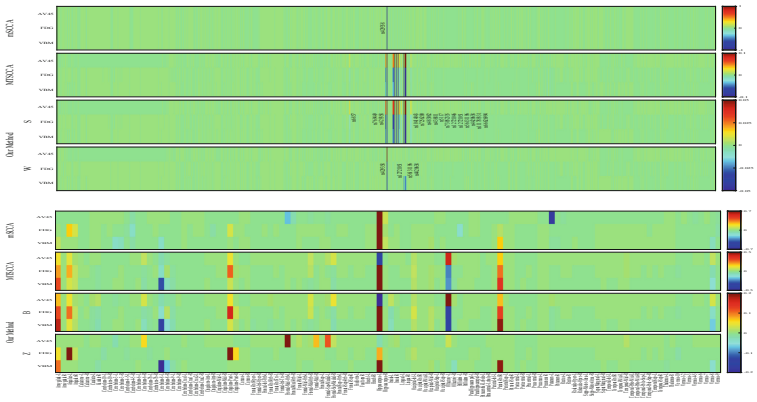

Fig. 1. Comparison of canonical weights. The weights for SNPs are shown on top, and those of imaging QTs are on bottom. Row 1–4: (1) mSCCA; (2) MTSCCA; (3) our method. Our method has two weights for SNPs and QTs owing to the parameter decomposition. Within each panel, there are three rows corresponding to three type of imaging QTs, i.e. AV45, FDG and VBM.

Task-Consistent and Task-Dependent Feature Selection. Now we investigate the identified SNPs and imaging QTs based on the absolute values of the estimated canonical weights. The upper half part of Fig. 1 shows the feature selection for SNPs while that of imaging QTs is presented on the lower half part. Since our model has two separate components for SNPs, i.e. the task-consistent component **S** and the task-dependent **W**, we show both of them here. mSCCA yields one canonical weight vector other than a weight matrix for SNPs, and thus we repeatedly stack its canonical weight vector three times to make its heat map available. We observe that all SNPs with nonzero values of our method have been shown to be relevant to the progression of AD. For example, rs429358 is identified by both **S** and **W**, demonstrating its strong association with AD. In addition, the proposed dirty MTSCCA shows a clear task-consistent pattern, indicating that these SNPs, e.g. rs12721051, rs56131196, rs438811, rs483082, rs56131196, rs5117 etc., could be correlated with AD no matter which imaging

technology is employed. Our method and MTSCCA identify more AD-related loci than mSCCA, demonstrating the multi-task modeling possesses comprehensive feature selection capacity. The heat maps of imaging QTs exhibit interesting task-consistent and task-dependent profiles. Our method shows that the left hippocampus, the left olfactory sulcus, the inferior parietal lobule and the left amygdala show clearly task-consistent patterns, indicating that these brain areas could be revealed by all imaging technologies, i.e. the sMRI, FDG and AV45 scans. Besides, task-dependent $\mathbf{Z}$ shows that the right medial orbitofrontal cortex and the left medial frontal gyrus are highlighted using the AV45-PET scans. The left and right angular gyrus, and the cingulum are identified by using the FDG-PET scans. Both left and right of the eighth cerebelum are highlighted when using the VBM measures of sMRI scans. MTSCCA and mSCCA can also identify several meaningful brain areas, however, they could not uncover the diverse complex association between SNPs and imaging QTs of multiple modalities. This real study demonstrates that our proposed dirty multi-task SCCA could be very promising and meaningful in multi-modal brain imaging genetics.

4 Conclusions

Imaging data collected by different technologies, measuring the same brain distinctly, might carry complementary but different information. In this paper, we propose a dirty multi-task SCCA method which incorporates multiple modalities of imaging data into a unified model. By decomposing the SCCA's canonical weights into the task-consistent component and the task-dependent component, and penalizing them distinctly, our method has the ability of identifying diverse meaningful bi-multivariate associations between SNPs and imaging QTs. We derive an efficient optimization algorithm to solve the dirty model. The neuroimaging genetics study demonstrates that the proposed method obtains better canonical correlation coefficients and canonical weights than multi-view SCCA and multi-task SCCA. A future direction is to extend this flexible model to be guided by the diagnosis status since currently it is unsupervised.

References

1. Du, L., et al.: Fast multi-task SCCA learning with feature selection for multi-modal brain imaging genetics. In: BIBM, pp. 356–361 (2018)
2. Du, L., et al.: A novel SCCA approach via truncated ℓ_1-norm and truncated group lasso for brain imaging genetics. Bioinformatics **34**(2), 278–285 (2018)
3. Du, L., et al.: Identifying progressive imaging genetic patterns via multi-task sparse canonical correlation analysis: a longitudinal study of the adni cohort. Bioinformatics **35**(14), i474–483 (2019)
4. Jalali, A., Ravikumar, P., Sanghavi, S.: A dirty model for multiple sparse regression. IEEE Trans. Inf. Theory **59**(12), 7947–7968 (2013)
5. Lee, S., Zhu, J., Xing, E.P.: Adaptive multi-task lasso: with application to eQTL detection. In: NIPS, pp. 1306–1314 (2010)

6. Potkin, S.G., et al.: Genome-wide strategies for discovering genetic influences on cognition and cognitive disorders: methodological considerations. Cogn. Neuropsychiatry **14**(4–5), 391–418 (2009)
7. Shen, L., Thompson, P.M., Potkin, S.G., Bertram, L., Farrer, L.A., et al.: Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. Brain Imaging Behav. **8**(2), 183–207 (2014)
8. Wang, H., et al.: Identifying quantitative trait loci via group-sparse multitask regression and feature selection: an imaging genetics study of the ADNI cohort. Bioinformatics **28**(2), 229–237 (2012)
9. Weiner, M.W., et al.: The Alzheimer's disease neuroimaging initiative: progress report and future plans. Alzheimer's Dement. **6**(3), 202–211 (2010)
10. Witten, D.M., Tibshirani, R., Hastie, T.: A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. Biostatistics **10**(3), 515–34 (2009)