Attentive Deep Canonical Correlation Analysis for Diagnosing Alzheimer's Disease Using Multimodal Imaging Genetics

Rong Zhou[1], Houliang Zhou[1], Brian Y. Chen[1], Li Shen[2], Yu Zhang[3], and Lifang He[1(✉)]

[1] Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA
lih319@lehigh.edu
[2] Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA, USA
[3] Department of Bioengineering, Lehigh University, Bethlehem, PA, USA

Abstract. Integration of imaging genetics data provides unprecedented opportunities for revealing biological mechanisms underpinning diseases and certain phenotypes. In this paper, a new model called attentive deep canonical correlation analysis (ADCCA) is proposed for the diagnosis of Alzheimer's disease using multimodal brain imaging genetics data. ADCCA combines the strengths of deep neural networks, attention mechanisms, and canonical correlation analysis to integrate and exploit the complementary information from multiple data modalities. This leads to improved interpretability and strong multimodal feature learning ability. The ADCCA model is evaluated using the ADNI database with three imaging modalities (VBM-MRI, FDG-PET, and AV45-PET) and genetic SNP data. The results indicate that this approach can achieve outstanding performance and identify meaningful biomarkers for Alzheimer's disease diagnosis. To promote reproducibility, the code has been made publicly available at https://github.com/rongzhou7/ADCCA.

Keywords: Brain imaging genetics · Canonical correlation analysis · Self-attention · Alzheimer's disease

1 Introduction

Alzheimer's disease (AD) is an irreversible neurodegenerative disorder that affects millions of people worldwide [5]. In recent years, brain imaging genetics has emerged as a promising field for the diagnosis and prediction of AD and its prodromal stage – mild cognitive impairment (MCI). This approach

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43895-0_64.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14221, pp. 681–691, 2023.
https://doi.org/10.1007/978-3-031-43895-0_64

largely focuses on using neuroimaging techniques, such as MRI and PET, to identify brain regions that are associated with specific genetic variants such as single nucleotide polymorphisms (SNPs). Such analyses have produced a wealth of research findings [23,26,28] that have demonstrated significant associations between imaging characteristics and genetics in AD, and have the great potential to identify new multimodal biomarkers affecting specific brain systems and provide an enormous impetus for drug discovery.

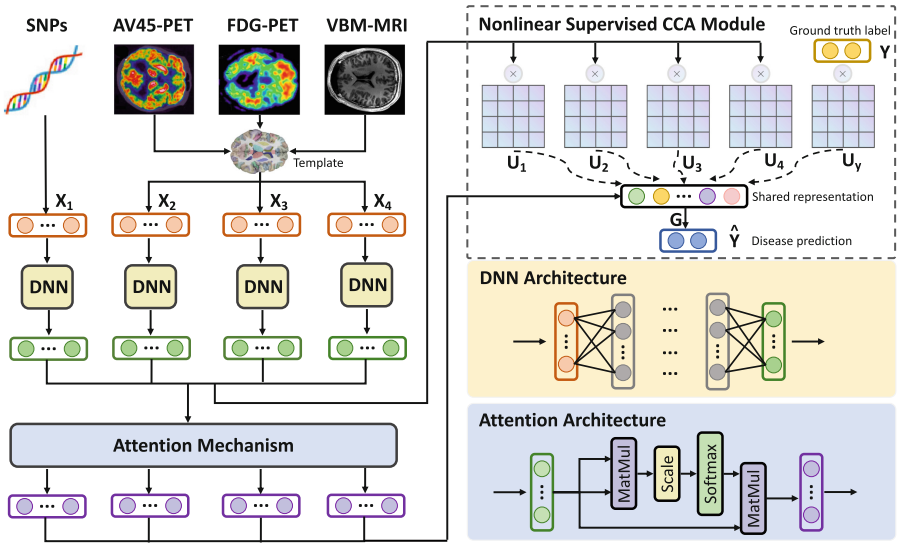

Fig. 1. An overview of the proposed framework. $\mathbf{X}_1, \cdots, \mathbf{X}_4$ are input modality data, and $\mathbf{Y}$ is the label information. DNNs first operate on each modality, generating hidden representations for each modality. These hidden representations go through a self-attention mechanism generating improved self-attention representations. At the same time, the hidden representations and label $\mathbf{Y}$ are multiplied by individual projection matrices $\mathbf{U}_1, \ldots, \mathbf{U}_4, \mathbf{U}_y$ based on CCA, thus mapping them to a shared representation $\mathbf{G}$. Finally, the disease prediction is calculated by self-attention representations with projection matrices and shared representation $\mathbf{G}$.

In the literature, various methods have been proposed to brain imaging genetics analysis [3,9–11,13,19,27,29]. In particular, canonical correlation analysis (CCA) [12] is a powerful multivariate statistical technique for quantifying the associations between different sets of data. CCA and its variations have been widely applied in imaging genetics studies because of its advantages in biological interpretation. For example, Du et al. [8] proposed a joint multitask sparse canonical correlation analysis and classification (MTSCCALR) for identifying imaging genetics biomarkers of AD. Kim et al. [16] introduced a multi-task learning-based structured sparse canonical correlation analysis (MTS2CCA) for identifying brain imaging genetics related to sleep. Moon et al. [20] proposed

a supervised deep generalized canonical correlation analysis (SDGCCA) for improving the classification of phenotypes and revealing biomarkers associated with phenotypes in the context of AD. Despite much progress made in this area, CCA-based traditional shallow models assume that the relationships between genetic and imaging data are linear. However, this may not always be the case, and nonlinear relationships may exist in brain imaging genetics data, leading to biased results. On the other hand, the existing CCA-based deep models do not provide a direct interpretation of the underlying biological mechanisms driving the observed associations between genetic and imaging data. Most of them explored post-hoc explanations as justifications for model predictions. This can limit the ability to translate findings into clinically relevant insights.

In this paper, we propose a novel attentive deep canonical correlation analysis (ADCCA) model for diagnosing AD disease and discovering biomarkers using multimodal brain imaging genetics data. As illustrated in Fig. 1, the proposed framework comprises three key components: (i) deep neural network (DNN) modeling for generating latent representations of each modality to capture intra-modality correlations; (ii) attention update mechanism for focusing on the most salient regions of input data; and (iii) nonlinear supervised CCA modeling for integrating multiple modalities to discriminate phenotypic groups. By combining the power of these techniques, the ADCCA approach effectively models nonlinear relationships among multimodal imaging genetics data and provides simultaneous predictions and interpretations. The model is trained end-to-end using a combination of classification and correlation losses.

Through extensive experiments on the real-world ADNI dataset with three imaging modalities (VBM-MRI, FDG-PET, and AV45-PET) and genetic SNP data, we show that our model achieves outstanding performance for classifying AD vs. HC, AD vs. MCI, and MCI vs. HC groups. Also, it is demonstrated that the model explanation can reveal disorder-specific biomarkers coinciding with neuroscience findings. Last, we show that the combination of classification and correlation models can boost disease prediction performance.

2 Method

Suppose that the problem includes N subjects with M modalities. Let $\mathbf{X}_m \in \mathbb{R}^{N \times d_m}$ denote the m-th modality data, where d_m represents the dimension of features in the m-th modality, $m = 1, 2, \cdots, M$. Let $\mathbf{Y} \in \mathbb{R}^N$ denote the label information of all subjects. In this work, we seek to learn a disease prediction model that estimates $\hat{\mathbf{Y}}$ from $\{\mathbf{X}_m\}_{m=1}^M$ by making full use of all M modalities, as well as identify disease-specific biomarkers for clinical interpretation.

The proposed ADCCA aims to combine the strengths of DNN, attention mechanism, and CCA to integrate and exploit the complementary information from multiple data modalities (Fig. 1). First, we use a separate DNN containing several fully-connected hidden layers to learn hidden representations for each modality, denoted as $f_m(\mathbf{X}_m) \in \mathbb{R}^{N \times l_m}$, where l_m represents the dimension of last layer of DNN corresponding to the m-th modality. Second, we employ the

attention mechanism [25] on the basis of the DNN model. With the help of the attention mechanism, our method can explicitly capture the important features hidden in the input data. Specifically, we use self-attention, sometimes called intra-attention, which is regarded as an improvement in attention that focuses on internal links of features and reduces external data dependency to compute its representation. Suppose there are three linear transformation matrices for the m-th modality: $\mathbf{W}_Q^m, \mathbf{W}_K^m, \mathbf{W}_V^m$. Mathematically, the self-attention representation of $f_m(\mathbf{X}_m)$ can be calculated as:

$$\mathrm{Att}(f_m(\mathbf{X}_m)) = \mathrm{Softmax}\left(\frac{f_m(\mathbf{X}_m)\mathbf{W}_Q^m(f_m(\mathbf{X}_m)\mathbf{W}_K^m)^\top}{\sqrt{l_m}}\right) f_m(\mathbf{X}_m)\mathbf{W}_V^m. \quad (1)$$

Third, following [20], we learn cross-modality features and incorporate the label information of samples for supervised learning based on CCA. The correlation loss function is defined as follows:

$$L_{cor} = \left\|\mathbf{G} - \mathbf{U}_y^\top \mathbf{Y}\right\|_F^2 + \sum_{m=1}^{M} \left\|\mathbf{G} - \mathbf{U}_m^\top f_m(\mathbf{X}_m)\right\|_F^2, \quad s.t. \ \ \mathbf{G}\mathbf{G}^\top = \mathbf{I}. \quad (2)$$

where $\mathbf{U}_1, \cdots, \mathbf{U}_4, \mathbf{U}_y$ are projection matrices for each modality and label information, respectively. $\mathbf{I}$ denotes the identity matrix.

According to Eq. (2), we have $\mathbf{G} \approx \mathbf{U}_m^\top f_m(\mathbf{X}_m) \approx \mathbf{U}_y^\top \mathbf{Y}$. Thus, the label $\mathbf{Y}$ can be approximated as follows: $\mathbf{Y} \approx \left(\mathbf{U}_y^\top\right)^\dagger \mathbf{U}_m^\top f_m(\mathbf{X}_m)$, where $\mathbf{U}_y^\dagger$ denotes the pseudo-inverse of $\mathbf{U}_y$. Then, we substitute self-attention representations that are more representative of each modality into the above equation and let $\hat{\mathbf{Y}}_m = \left(\mathbf{U}_y^\top\right)^\dagger \mathbf{U}_m^\top \mathrm{Att}(f_m(\mathbf{X}_m))$. Further, the conventional supervised cross-entropy loss [7] is used to enable the propagation of label information directly to the DNN of each modality.

$$L_{cls} = \sum_{m=1}^{M} \mathrm{CrossEntropy}\left(\mathbf{Y}, \mathrm{Softmax}(\hat{\mathbf{Y}}_m)\right). \quad (3)$$

The final label prediction of ADCCA can be obtained using the following soft voting of the label presentation of each modality: $\hat{\mathbf{Y}} = \mathrm{Softmax}((\sum_{m=1}^{M} \hat{\mathbf{Y}}_m)/M)$. Overall, our final training objective can be defined as:

$$L = L_{cls} + \lambda L_{cor}, \quad (4)$$

where L_{cls} is the supervised cross-entropy disease prediction loss, L_{cor} is the correlation loss, and λ is a tunable hyperparameter that scales the numerical value of each loss item to the same order of magnitude to balance their influence. The solution on loss function L is similar to the SGDCCA method except for substituting the outputs of DNN models to their self-attention representations.

3 Experiments and Results

3.1 Data Acquisition and Preprocessing

Brain imaging genetic data used in this study were obtained from the public ADNI database [22]. There is a total of 597 participants with both geno-

type and brain imaging data, including 104 AD, 305 MCI, and 188 healthy control (HC) subjects. The image data consisted of three modalities including structural Magnetic Resonance Imaging (VBM-MRI), 18 F-fluorodeoxyglucose Positron Emission Tomography (FDG-PET), and 18 F-florbetapir PET (AV45-PET). These three imaging modalities allowed us to examine brain structure, glucose metabolism, and amyloid plaque deposition, respectively.

Following the previous studies [2,30], we preprocessed neuroimaging data to extract ROI-based features. Specifically, the multi-modality imaging scans were aligned to each participant's same visit. All imaging scans were aligned to a T1-weighted template image, and segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) maps. They were normalized to the standard Montreal Neurological Institute (MNI) space as $2 \times 2 \times 2$ mm^3 voxels, being smoothed with an 8 mm FWHM kernel. We preprocessed the structural MRI scans with voxel-based morphometry (VBM) by using the SPM software [1], and registered the FDG-PET and AV45-PET scans to the MNI space by SPM. We subsampled the whole brain imaging and contained 90 ROIs (excluding the cerebellum and vermis) based on the AAL-90 atlas [24]. ROI-level measures were calculated by averaging all the voxel-level measures within each ROI.

For genetic SNP data, according to the AlzGene database[1], only the SNPs belonging to top AD gene candidates were selected. The genetic data were genotyped by the Human 610-Quad or OmniExpress Array platform (Illumina, Inc., San Diego, CA, USA), and preprocessed following standard quality control and imputation procedures. There were 54 SNPs included which were collected from the neighbor of AD risk gene APOE according to the ANNOVAR annotation.

3.2 Evaluation of Disease Classification Performance

In our experiments, the whole data were separated into three groups, including AD vs. HC, AD vs. MCI, and MCI vs. HC. To quantitatively evaluate the performance of different methods, we considered four commonly-used evaluation metrics: accuracy (ACC), F1-score (F1), area under receiver operating characteristic curve (AUC), and Matthews correlation coefficient (MCC) [6]. Since the number of subjects was limited, we calculated the mean and standard deviation of all metrics using 5-fold cross-validation (CV). Many researchers have successfully adopted multimodal brain imaging data into CCA. We carefully choose five related methods for comparison: 1) vanilla DNN [18], 2) generalized CCA (GCCA) [15], 3) deep generalized CCA (DGCCA) [4], 4) MTSCCALR [8], and 5) SDGCCA [20]. Note that GCCA and DGCCA are unsupervised learning methods, and the others are supervised learning methods. The proposed model includes four DNNs, one for each modality, with three fully-connected layers and a Tanh activation function, which is trained with Adam optimizer with the learning rate set to 0.0001 and weight decay set to 0.001.

Table 1 presents the classification results, where $\pm$ represents the standard deviation of evaluation scores across the 5 folds. From the results, it can be

[1] www.alzgene.org.

observed that the proposed ADCCA method significantly outperforms all other methods in terms of all four metrics. The higher AUC and MCC scores indicate that our method is able to accurately identify both positive and negative cases of AD. The smaller standard deviations of ADCCA illustrated the overall stability and reproducibility of the experiment.

Table 1. Classification performance comparison. The best results are in bold.

Task	Measures	DNN	GCCA	DGCCA	MTSCCALR	SDGCCA	ADCCA
AD vs. HC	ACC	.866 ± .037	.812 ± .037	.837 ± .028	.828 ± .047	.914 ± .029	**.932 ± .010**
	F1	.873 ± .049	.811 ± .054	.833 ± .041	.862 ± .046	.883 ± .034	**.901 ± .025**
	AUC	.943 ± .030	.930 ± .015	.939 ± .013	.893 ± .051	.978 ± .013	**.979 ± .015**
	MCC	.720 ± .080	.652 ± .079	.688 ± .060	.629 ± .087	.822 ± .057	**.895 ± .043**
AD vs. MCI	ACC	.689 ± .035	.618 ± .059	.638 ± .017	.746 ± .049	.812 ± .063	**.825 ± .011**
	F1	.579 ± .032	.583 ± .038	.535 ± .037	.679 ± .041	.683 ± .079	**.823 ± .032**
	AUC	.811 ± .025	.726 ± .050	.756 ± .022	.836 ± .039	.880 ± .043	**.925 ± .050**
	MCC	.413 ± .046	.256 ± .054	.281 ± .048	.482 ± .104	.569 ± .110	**.625 ± .024**
MCI vs. HC	ACC	.523 ± .026	.499 ± .024	.519 ± .044	.594 ± .029	.647 ± .058	**.758 ± .033**
	F1	.529 ± .031	.543 ± .084	.513 ± .044	.513 ± .025	.702 ± .058	**.799 ± .030**
	AUC	.570 ± .030	.540 ± .032	.574 ± .054	.637 ± .022	.796 ± .074	**.816 ± .051**
	MCC	.103 ± .058	.105 ± .075	.109 ± .100	.172 ± .045	.273 ± .110	**.407 ± .073**

3.3 The Most Discriminative Brain Regions and SNPs

Identifying the most discriminative brain regions (*i.e.*, ROIs) and SNPs is crucial for AD diagnosis. Here, we employed the integrated gradients interface provided by Captum [17] to assign importance scores to each feature of different modalities by analyzing the pre-trained model, which can provide a comprehensive explanation of how the input features of a deep learning model contribute to the model's output. The reason why not using the self-attention weights is that we use the self-attention to assign attention scores to hidden representations instead of the original features, thus it may not fully capture the importance of the original features in the input data. Figure 2(a-c) shows the top 20 discriminative ROIs identified by the proposed method from each individual brain imaging modality. Figure 2(d) shows the top 20 discriminative ROIs selected by the average importance scores of ROIs from the three modalities. We found that the hippocampal, amygdala, uncus, and gyrus regions are only identified by using the three modalities together. These selected regions are known to be highly related to AD and MCI in previous studies [21]. Besides, the result shows that the selected ROIs exhibited differences across different classification groups, indicating that our model can effectively differentiate the important ROIs for specific diseases. Figure 3 shows the most frequently selected SNPs with importance scores. The result indicates that rs6448453, rs3865444, and rs2718058 are the most discriminative SNPs which is consistent with previous evidence [14].

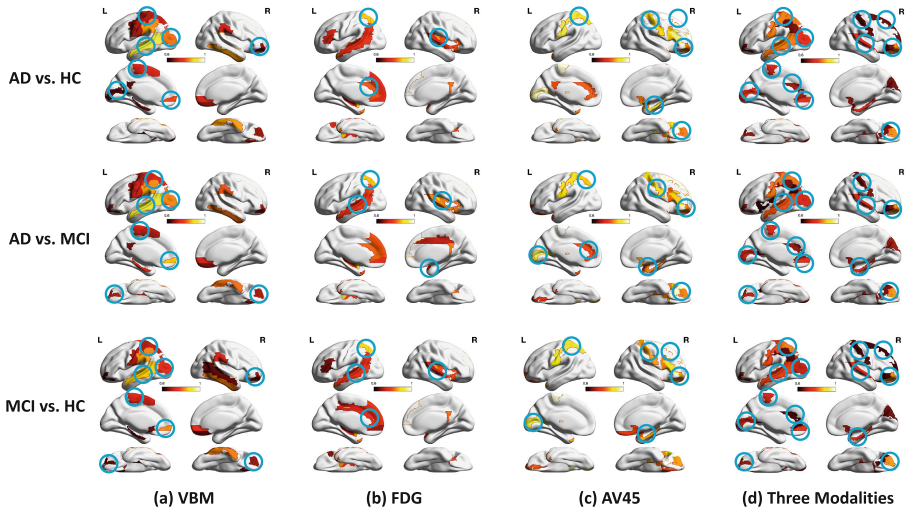

Fig. 2. Top 20 discriminative ROIs identified by ADCCA from three brain imaging modalities for three different classification groups in lateral, medial, and ventral view. The color bar indicates the importance score. The commonly selected ROIs across different modalities are circled in blue. (Color figure online)

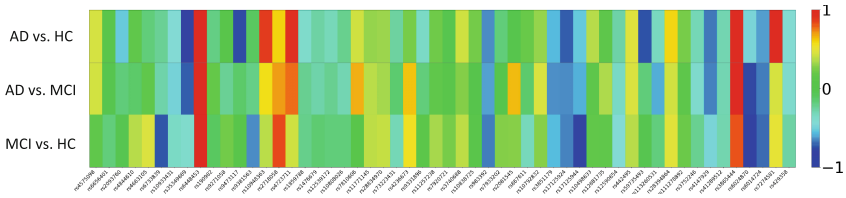

Fig. 3. The importance scores of SNPs. The red color indicates a high score. (Color figure online)

3.4 Ablation Study

The proposed ADCCA is trained using both correlation and classification losses. To understand the impact of each loss on classification, we conducted ablation studies by evaluating the performance of two additional models: the ADCCA model trained without the correlation loss (w/o L_{cor}) and without the classification loss (w/o L_{cls}). The results presented in Table 2 indicate that ADCCA outperforms the other two models for all evaluation metrics on all three classification tasks, suggesting that both correlation and classification losses contribute to ADCCA's improved performance. Removing either loss leads to decreased performance, and the impact will be particularly significant if the classification loss is eliminated.

Table 2. Classification performance comparison with and without L_{cor} and L_{cls}

Task	Method	ACC	F1	AUC	MCC
AD vs. HC	ADCCA	$\mathbf{.932 \pm .010}$	$\mathbf{.901 \pm .025}$	$\mathbf{.979 \pm .015}$	$\mathbf{.895 \pm .043}$
	(w/o) L_{cor}	$.924 \pm .025$	$.892 \pm .034$	$.963 \pm .016$	$.837 \pm .067$
	(w/o) L_{cls}	$.876 \pm .037$	$.853 \pm .029$	$.928 \pm .020$	$.791 \pm .058$
AD vs. MCI	ADCCA	$\mathbf{.825 \pm .011}$	$\mathbf{.823 \pm .032}$	$\mathbf{.925 \pm .050}$	$\mathbf{.625 \pm .024}$
	(w/o) L_{cor}	$.806 \pm .029$	$.795 \pm .032$	$.897 \pm .048$	$.589 \pm .033$
	(w/o) L_{cls}	$.758 \pm .059$	$.723 \pm .038$	$.856 \pm .050$	$.466 \pm .054$
MCI vs. HC	ADCCA	$\mathbf{.758 \pm .033}$	$\mathbf{.799 \pm .030}$	$\mathbf{.816 \pm .051}$	$\mathbf{.407 \pm .073}$
	(w/o) L_{cor}	$.692 \pm .033$	$.713 \pm .056$	$.761 \pm .072$	$.317 \pm .092$
	(w/o) L_{cls}	$.619 \pm .024$	$.683 \pm .084$	$.599 \pm .032$	$.176 \pm .075$

3.5 Hyperparameter Analysis

We investigated the impact of two important hyperparameters in the ADCCA model: λ, which appears in the loss function to balance the classification and correlation losses, and the dimension of the shared representation $\mathbf{G}$. In order to explore the effects of these hyperparameters on the performance of the model, we conducted experiments using different values of λ and the shared representation dimensionality. Due to the space limit, we only report the classification results in AD vs. HC group, as shown in Fig. 4. The results in other groups can be found in the supplementary material. We observed that decreasing the value of λ generally leads to improved model performance across various tasks, but a lambda value of zero causes the model's performance to deteriorate. This may indicate that for the ADCCA model, L_{cls} is more important than L_{cor}. Furthermore, combining these two loss functions to jointly guide the model can lead to improved model performance. We also found that for the AD vs. HC group, the model achieves good performance even with a low-dimensional shared representation. However, for other groups, the impact of the shared representation dimension on the model's performance seems not significant. One explanation for this could be that the AD vs. HC group exhibits distinct feature differences, allowing the original features to be well represented even when mapped into a low-dimensional shared representation.

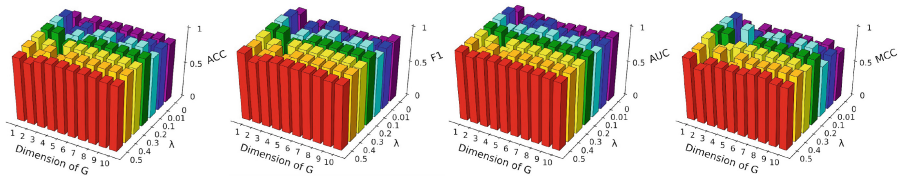

Fig. 4. Sensitivity analysis of hyperparameters on AD vs. HC

4 Conclusion

In this work, we propose a novel deep canonical correlation analysis method for multimodal Alzheimer's disease diagnosis that leverages attention mechanisms to enhance interpretability and multimodal feature learning. Experimental results on the real-world imaging-genetics dataset demonstrate that our approach achieves better classification performance than the existing state-of-the-art methods in terms of both classification accuracy and correlation between the modalities. In an exploratory analysis, we further show that the biomarkers identified by our model are closely associated with Alzheimer's disease. Our proposed approach is applicable to other diseases with multimodal data available. However, the limited size of medical datasets may restrict the effectiveness and generalization ability of such deep learning models. To address this issue, a potential future direction is to employ pre-training and transfer learning techniques that facilitate learning across datasets.

Acknowledgements. This work is partially supported by the National Science Foundation (MRI-2215789 and IIS-1909879), National Institutes of Health (U01AG068057, U01AG-066833, R01LM013463, R01MH129694, and R21MH130956), Alzheimer's Association grant (AARG-22-972541), and Lehigh's grants under Accelerator (S00010293), CORE (001250), and FIG (FIGAWD35).

References

1. Ashburner, J., Friston, K.J.: Voxel-based morphometry-the methods. Neuroimage **11**(6), 805–821 (2000)
2. Barshan, E., Fieguth, P.: Stage-wise training: An improved feature learning strategy for deep models. In: Feature extraction: modern questions and challenges, pp. 49–59. PMLR (2015)
3. Batmanghelich, N.K., Dalca, A., Quon, G., Sabuncu, M., Golland, P.: Probabilistic modeling of imaging, genetics and diagnosis. IEEE Trans. Med. Imaging **35**(7), 1765–1779 (2016)
4. Benton, A., Khayrallah, H., Gujral, B., Reisinger, D.A., Zhang, S., Arora, R.: Deep generalized canonical correlation analysis. In: Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019), pp. 1–6 (2019)
5. Catania, M., et al.: A novel bio-inspired strategy to prevent amyloidogenesis and synaptic damage in Alzheimer's disease. Mol. Psych. 1–8 (2022)
6. Chicco, D., Jurman, G.: The advantages of the Matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. BMC Genomics **21**, 1–13 (2020)
7. De Boer, P.T., Kroese, D.P., Mannor, S., Rubinstein, R.Y.: A tutorial on the cross-entropy method. Ann. Oper. Res. **134**, 19–67 (2005)

8. Du, L., et al.: Identifying diagnosis-specific genotype-phenotype associations via joint multitask sparse canonical correlation analysis and classification. Bioinformatics **36**, i371–i379 (2020)
9. Du, L., et al.: Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach. Med. Image Anal. **61**, 101656 (2020)
10. Ghosal, S., et al.: Bridging imaging, genetics, and diagnosis in a coupled low-dimensional framework. In: Shen, D., et al. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV, pp. 647–655. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32251-9_71
11. Ghosal, S., et al.: A biologically interpretable graph convolutional network to link genetic risk pathways and imaging phenotypes of disease. In: ICLR (2022)
12. Hotelling, H.: Relations between two sets of variates. Biometrika **28**(3/4), 321–377 (1936)
13. Hu, W., et al.: Adaptive sparse multiple canonical correlation analysis with application to imaging (epi) genomics study of schizophrenia. IEEE Trans. Biomed. Eng. **65**(2), 390–399 (2017)
14. Jansen, I.E., et al.: Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat. Genet. **51**(3), 404–413 (2019)
15. Kettenring, J.R.: Canonical analysis of several sets of variables. Biometrika **58**(3), 433–451 (1971)
16. Kim, M., et al.: Multi-task learning based structured sparse canonical correlation analysis for brain imaging genetics. Med. Image Anal. **76**, 102297 (2022)
17. Kokhlikyan, N., et al.: Captum: A unified and generic model interpretability library for pytorch. arXiv preprint arXiv:2009.07896 (2020)
18. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
19. Liu, J., Calhoun, V.D.: A review of multivariate analyses in imaging genetics. Front. Neuroinform. **8**, 29 (2014)
20. Moon, S., Hwang, J., Lee, H.: SDGCCA: supervised deep generalized canonical correlation analysis for multi-omics integration. J. Comput. Biol. **29**(8), 892–907 (2022)
21. Mu, Y., Gage, F.H.: Adult hippocampal neurogenesis and its role in Alzheimer's disease. Mol. Neurodegener. **6**(1), 1–9 (2011)
22. Muller, S.G., et al.: The Alzheimer's disease neuroimaging initiative. Neuroimaging Clin. **15**(4), 869–877 (2005)
23. Shen, L., Thompson, P.M.: Brain imaging genetics: integrated analysis and machine learning. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1–1. IEEE Computer Society (2021)
24. Tzourio-Mazoyer, N., et al.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage **15**(1), 273–289 (2002)
25. Vaswani, A., et al.: Attention is all you need. Adv. Neural Inform. Process. Syst. **30** (2017)
26. Viding, E., Williamson, D.E., Forbes, E.E., Hariri, A.R.: The integration of neuroimaging and molecular genetics in the study of developmental cognitive neuroscience. MIT press (2008)
27. Wang, M.L., Shao, W., Hao, X.K., Zhang, D.Q.: Machine learning for brain imaging genomics methods: a review. Mach. Intell. Res. **20**(1), 57–78 (2023)

28. Xin, Y., Sheng, J., Miao, M., Wang, L., Yang, Z., Huang, H.: A review of imaging genetics in Alzheimer's disease. J. Clin. Neurosci. **100**, 155–163 (2022)
29. Zhou, H., Zhang, Yu., Chen, B.Y., Shen, L., He, L.: Sparse interpretation of graph convolutional networks for multi-modal diagnosis of Alzheimer's disease. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII, pp. 469–478. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16452-1_45
30. Zhu, Y., et al.: Graphene and graphene oxide: synthesis, properties, and applications. Adv. Mater. **22**(35), 3906–3924 (2010)