



In-house data adaptation to public data: Multisite MRI harmonization to predict Alzheimer's disease conversion

Sunghong Park^a, Sang Joon Son^a, Kanghee Park^b, Yonghyun Nam^c, Hyunjung Shin^{d,e,*}, Alzheimer's Disease Neuroimaging Initiative¹, Australian Imaging Biomarkers and Lifestyle Study of Ageing²

^a Department of Psychiatry, Ajou University School of Medicine, Suwon, 16499, Republic of Korea

^b Korea Institute of Science and Technology Information, Seoul, 02456, Republic of Korea

^c Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

^d Department of Artificial Intelligence, Ajou University, Suwon, 16499, Republic of Korea

^e Department of Industrial Engineering, Ajou University, Suwon, 16499, Republic of Korea

ARTICLE INFO

Keywords:

Alzheimer's disease
Mild cognitive impairment
Domain adaptation
Multisite MRI harmonization
Distribution matching
Connectivity mapping

ABSTRACT

For the machine learning-based prediction of the conversion from mild cognitive impairment to Alzheimer's disease, the collection of sufficient data to train a model is required, which involves a lot of time and expense. When data is not enough, combining public and in-house data may be appropriate by applying domain adaptation that alleviates inter-site heterogeneity. Existing methods simultaneously transform in-house and public data to represent them into a common feature space, and then train a classifier using labels in public data. However, this procedure causes the time- and cost-consuming re-training of classifier whenever in-house data changes, and also inheres the risk of information loss in public data. Motivated by this, we propose a method that only transforms in-house data while preserving public data, namely one-way domain adaptation. The proposed method represents in-house data similar with public data by matching the data distribution and the connectivity between brain regions with mean vectors and covariance matrices, respectively. Then, the pre-trained classifier in public data is applied to predict AD conversion for in-house data. The experiments, which use the Australian Imaging Biomarkers and Lifestyle Study of Aging and the Open Access Series of Imaging Studies as the in-house data and the Alzheimer's Disease Neuroimaging Initiative as the public data, show the effectiveness and efficiency of the proposed method, improving prediction performance about 34.8% on average without labels in the in-house datasets.

1. Introduction

Alzheimer's disease (AD) is the most common form of dementia affecting older people (Petrella, Coleman, & Doraiswamy, 2003). Due to the extension of life expectancy, the number of AD patients continues to increase, and the global Alzheimer's disease population is expected to

triple from approximately 50 million in 2015 to 131.5 million in 2050 (Cummings, Lee, Ritter, Sabbagh, & Zhong, 2019). As such, AD is emerging serious problem, but the cause is unclear and there is no therapy to revert it. Therefore, with the early detection of AD, it is important to predict the potential risk of disease progression and find an appropriate prevention strategy. According to clinical symptoms with

* Corresponding author.

E-mail addresses: pshong513@ajou.ac.kr (S. Park), sjsonpsy@ajou.ac.kr (S.J. Son), can17@kisti.re.kr (K. Park), yonghyun.nam@penntmedicine.upenn.edu (Y. Nam), shin@ajou.ac.kr (H. Shin).

¹ Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). 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.

² Data used in the preparation of this article was obtained from the Australian Imaging Biomarkers and Lifestyle flagship study of aging (AIBL) funded by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) which was made available at the ADNI database (www.loni.usc.edu/ADNI). The AIBL researchers contributed data but did not participate in analysis or writing of this report. AIBL researchers are listed at www.aibl.csiro.au.

mild memory loss or other cognitive loss, mild cognitive impairment (MCI) is considered a prodromal phase of AD with symptoms such as long-term memory loss, speech impairment, disorientation, and personality changes. Previous studies have suggested that approximately 12 % of subjects suffering from MCI progress to AD in the four years following the first symptoms (Petersen, et al., 1999). Consequently, early detection of the potential risk of AD boils down to the task of predicting whether MCI patients will convert to AD or not.

Most studies predicting the conversion from MCI to AD have founded on brain imaging such as magnetic resonance imaging (MRI). The difference between MCI and AD is associated with loss of brain volume and the development of localized lesions of white and gray matter. With the recent advance of brain imaging technologies, the computer-aided methods have implemented accurate prediction of the conversion (Arbabshirani, Plis, Sui, & Calhoun, 2017; Rathore, Habes, Iftikhar, Shacklett, & Davatzikos, 2017). Particularly, the application of machine learning (ML) has contributed significantly to these improvements (Pellegri, et al., 2018). A typical ML algorithm applied to brain MRI learns a set of preprocessed data, such as regional volumes and cortical thickness, to create a classifier which predicts the correct diagnostic outcome for new observations (Liu, Tosun, Weiner, Schuff, & Initiative, 2013).

The success of the prediction depends on the data provided to the ML algorithm. The more data we have, the more sophisticated our results will be. However, collecting medical data is usually time-consuming and expensive. As a result, an ML algorithm often suffers from a lack of data to learn. Besides the insufficiency of data, the amount of labeled data has a huge impact on the performance of classification. In general, the cases for normal subjects are dominant to those for AD subjects, so are the corresponding MR images. This incurs a lack of labels for learning. In order to obtain more labels for AD brain imaging, more cases of AD subjects need to be accumulated, which is also time-consuming and expensive.

The insufficiency in learning data can be supplemented by merging or integrating external data. Different data can be easily merged or integrated into one if they are homogeneous in features, measurement units, imaging protocols, etc. However, this rarely happens, and the different datasets are heterogeneous, which is referred to as inter-site heterogeneity. Meanwhile, domain adaptation can be used as one of the good alternatives to address the shortcomings (Ben-David, et al., 2010). It transforms the different datasets belonging to each domain to appear as one dataset. The features from different domains adapted to a common feature space by reducing the discrepancy between domains (J. Huang, Gretton, Borgwardt, Schölkopf, & Smola, 2006a). Here, datasets from different domains can be considered brain MR image sets from other sites, such as hospitals or research institutions. More simply, when we utilize the public data to analyze the in-house data of a certain site, they can be regarded as different domains. In general, the in-house data is relatively small in size, not rich in content, and highly biased compared to the public data. Therefore, adapting in-house data to public data can compensate for these shortcomings, and ML algorithms learn reliably from more sufficient data. Also, it is an additional gain that more labels can be provided to the algorithm, which leads to more accurate and robust prediction outcomes. Consequently, these benefits of domain adaptation using in-house data plus public data can be applied to AD conversion prediction, circumventing insufficiency in data and labels.

There have been interesting studies on AD conversion prediction with domain adaptation. Li et al. proposed the method that functional MRI features for source and target samples are extracted, and then samples are projected from both domains into a shared subspace to reduce the domain discrepancy (Li, Zhao, Chen, Xiao, & Qin, 2018). Moradi et al. utilized an information theoretic method for the unsupervised domain adaptation where the classifier is trained using MRIs represented by gray matter density features in different datasets (Moradi, Gaser, Huttunen, & Tohka, 2014). Cheng et al. suggested domain

transfer learning with a feature selection method to train a classifier reducing domain shift for MCI conversion prediction (Cheng, Liu, Zhang, Munsell, & Shen, 2015). As such, existing methods create a common feature space in which different domain samples are mixed and train a classifier for the task. Despite these successes, existing methods have the limitation in utilizing the public data. In previous studies, domain adaptation simultaneously transforms the in-house and public data, represents them into a common feature space, and train a classifier with labels in public data. For this reason, whenever the in-house data changes, the classifier must be newly trained. Such re-training is tedious and time- and cost-consuming.

In this study, we propose a novel method for domain adaptation to apply to the prediction problem of AD conversion from MCI patients using T1-weighted structural MRIs. To avoid the tedious re-training, this method employs a one-way transformation of the in-house data into larger, content-rich, and less-biased public data. It is also convenient to compare different in-house data from different sites.

In the next section, we introduce the related works and present the materials for this paper, and the following section includes the description for the detailed process of the proposed method, *distribution matching* and *connectivity mapping*, along with its mathematical implementation. The experiment section shows the various results on three different datasets: the Alzheimer's Disease Neuroimaging Initiative (ADNI), the Australian Imaging Biomarkers and Lifestyle Study of Aging (AIBL), and the Open Access Series of Imaging Studies (OASIS). Finally, we conclude this paper by mentioning discussions and contributions along with limitations of the proposed method.

2. Related works

Domain adaptation is a representative strategy for the data with a few information to perform training by utilizing other data with a lot of information (Ben-David, et al., 2010; J. Huang, Gretton, Borgwardt, Schölkopf, & Smola, 2006b). This allows the target domain to improve the performance of classification or regression from the information provided by the source domain (Chen, Weinberger, & Blitzer, 2011; Pan & Yang, 2009; Xue, Dai, Yang, & Yu, 2008). Accordingly, domain adaptation has been widely used in various fields. Among them, especially in the medical field, domain adaptation is a very useful method for collecting large amounts of patient data across multiple sites. In this case, since many hospitals, doctors, and medical devices have different standards and characteristics, it is essential for the consistent representation of heterogeneous datasets by domain adaptation. To implement this, various methods have been developed by existing studies, mostly being applied to the medical imaging data.

The existing domain adaptation studies for medical imaging data can be roughly divided into two types: singular value decomposition (SVD)-based methods and domain adversarial training-based methods. First, SVD decomposes a data matrix into three matrix products (two orthogonal matrices for left and right singular vectors and a diagonal matrix for singular values), which is an important method used in various fields such as data compression, dimensionality reduction, and data reconstruction. For domain adaptation, the main process of SVD-based methods is as follows: (1) SVD is performed on each data in the source and target domains. (2) Each data is represented on the same subspace by the orthogonal matrix for right singular vectors. (3) The orthogonal matrices are trained for the consistency of the different representation of the two data. (4) Each data is transformed by the optimized orthogonal matrices, and the target tasks are performed, such as classification or regression. This process has been successfully used to detect Alzheimer's disease on small datasets of neuroimaging (DADS) (Li, et al., 2018). Furthermore, the multi-site adaption framework via low-rank representation decomposition (MALRR) has also been developed to identify autism spectrum disorders. As a more advanced method, multi-site clustering and nested feature extraction (MCNFE) has recently been developed for identifying autism spectrum disorder with

Table 1
Demographics of subjects in the datasets.

Category	ADNI	AIBL	OASIS	
# of subjects (MCI-C / MCI-N)	147 (74/73)	30 (15/15)	42 (21/21)	
Age at baseline	74.6 ± 7.4	75.7 ± 5.7	73.5 ± 6.8	
Follow-up duration (months)	28.4 ± 10.1	27.0 ± 9.2	27.2 ± 10.2	
MMSE	baseline	27.4 ± 1.6	26.8 ± 1.5	27.6 ± 1.7
	follow-up	24.3 ± 4.8	24.3 ± 3.1	24.9 ± 3.7
CDR(0/0.5/1)	baseline	0/147/0	1/29/0	1/41/0
	follow-up	0/87/60	0/22/8	0/29/13

resting-state fMRI (Wang, Yao, Ma, & Liu, 2022).

Next, DAT-based methods are based on neural network typically consisting of feature extractor and classifier in which the former trains domain labels (source or target), and the latter trains class labels. The research that is the starting point of this approach is domain adversarial neural network (DANN), which developed a gradient reversal layer and trained the model so that the domain labels of each data could not be distinguished well (Ganin, et al., 2016). As a result, it was possible to extract well-matched features to make the data of different domains look like one, and at the same time to perform classification for the target domain by training the labeled data of source domain. The excellence of DANN has resulted in various follow-up studies. In the field of medical imaging, DANN was advanced by paired consistency and adversarial learning (PCAL) for multi-domain adaptation in brain MRI (Orbes-Arteaga, et al., 2019). Furthermore, conditional domain adversarial network (CDAN) has also been developed to classify ADHD with cross-site functional MRI datasets (Y.-L. Huang, Hsieh, Yang, & Lee, 2020). As a more advanced method, attention-guided deep domain adaptation (ADDA) has recently been developed for multi-site MRI harmonization to identify brain disorder (Guan, et al., 2021).

3. Materials

Datasets used in this paper was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (<https://www.loni.ucla.edu/ADNI>), the Australian Imaging Biomarkers and Lifestyle flagship study of aging (AIBL) (<https://www.aibl.csiro.au>), and the Open Access Series of Imaging Studies (OASIS) (<https://www.oasis-brains.org/>) (Marcus, Fotenos, Csernansky, Morris, & Buckner, 2010). Each dataset

contains two types of information of subjects, demographics and MR images, and each is described in the following subsections.

3.1. Demographics of subjects

Subjects were labeled according to the criteria in (Cheng, et al., 2015; Lee, Nho, Kang, Sohn, & Kim, 2019) as follows. (a) *MCI subjects*: MMSE scores between 24 and 30, a memory complaint, having objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, a Clinical Dementia Rating (CDR) of 0.5. (b) *AD subjects*: MMSE scores between 20 and 26, CDR of 0.5 or 1.0. As a result, 219 subjects including 110 MCI-C and 109 MCI-N were totally selected. Table 1 provides a summarized demographics.

3.2. MRI preprocessing

A total of 438 MR images were collected for the selected subjects. MR images were preprocessed as follows. First, we performed the anterior commissure (AC)-posterior commissure (PC) correction. Second, images were segmented into three different tissues, gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), by using the computational anatomy toolbox (CAT12: <https://www.neuro.uni-jena.de/cat/>) in SPM (Ashburner, et al., 2014; Penny, Friston, Ashburner, Kiebel, & Nichols, 2011). Next, we obtained the subject-labeled images with 95 ROIs based on HAMMER (Shen & Davatzikos, 2002). At last, the volume of GM tissue was normalized to the total intracranial volume by summing the GM, WM, and CSF volumes of all ROIs. The MRI preprocessing pipeline is schematically illustrated in Fig. 2, and this process is widely used in neuroimage-based application studies (Gaser, et al., 2022; Park, et al., 2023).

4. Methods

The proposed method consists of one-way domain adaptation and AD conversion prediction as shown in Fig. 1. First, the in-house data is transformed into the feature space of the public data via an adaptation matrix. To solve the heterogeneity problem of different domains, the matrix is designed to implement two properties, *distribution matching* and *connectivity mapping*. The former serves to reduce the discrepancy between regions by matching the center of the distribution. The latter,

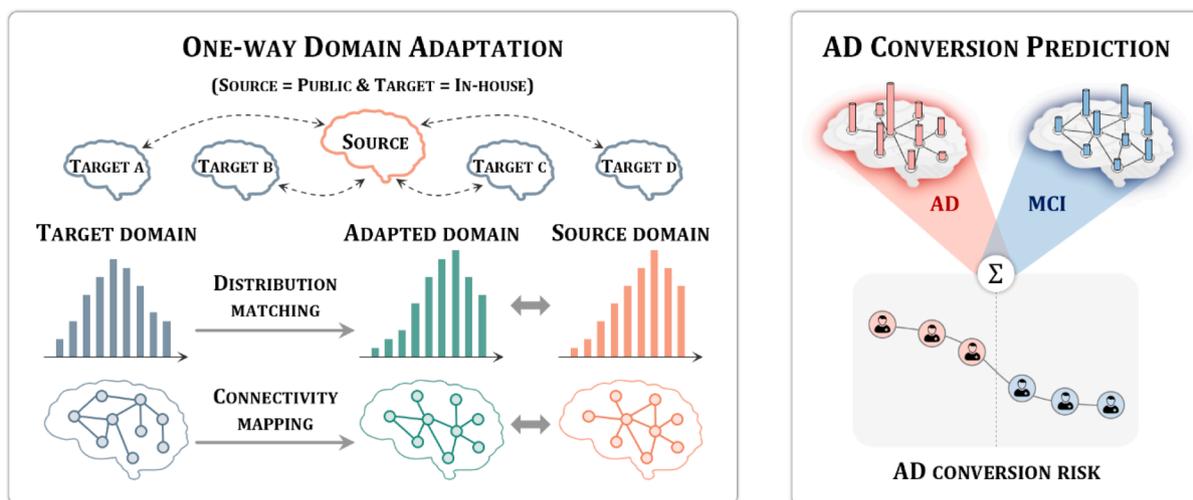


Fig. 1. Overview of the proposed method. The proposed method for domain adaptation aims to solve the prediction problem of AD conversion from MCI patients using T1-weighted structural MRIs. The proposed method involves one-way domain adaptation and prediction of AD conversion. It starts by transforming in-house data into the feature space of public data using an adaptation matrix. This matrix addresses domain heterogeneity through distribution matching and connectivity mapping, reducing differences between regions and establishing similar relationships between data features. Next, a machine learning algorithm performs classification on the combined public and transformed in-house data, leveraging available labels to define classification boundaries. This enables classification even in cases without labeled data, aligning the transformed in-house data in the feature space for accurate categorization using established class boundaries.

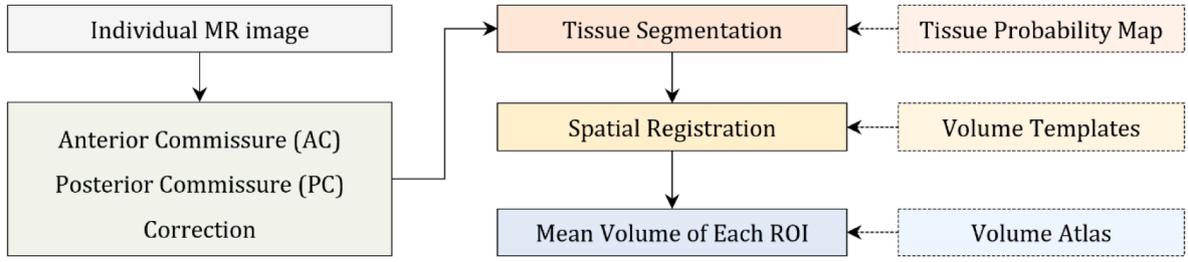


Fig. 2. MRI preprocessing pipeline. The preprocessing for MR images starts with AC-PC correction, followed by tissue segmentation (GM, WM, CSF) using CAT12 in SPM. Subject-labeled images with 95 ROIs were generated using HAMMER. Finally, GM volume was normalized by summing GM, WM, and CSF volumes across all ROIs and normalizing to total intracranial volume.

on the other hand, makes the relationship between the features of the in-house data similar with those of public data. To the AD conversion problem, the property of *distribution matching* has an effect of calibration for the brain volume data from different domains. Meanwhile, the property of *connectivity mapping* indicates that brain ROIs from in-house data maintain similar relationships even after adaptation, thus indicating that ROIs have strong connections and tend to have strong connections in the transformed feature space.

Second, an ML algorithm perform classification on the public data which now include the transformed in-house data. Notably, even if there is no labeled data such as no brain image for the case of AD conversion, we still perform classification for the in-house data, unless there is no label in the public data neither. Since public data have labels, it is available to form the region of the classification boundary. The transformed in-house data is then aligned into the feature space, and the data are classified using the class-boundaries.

4.1. One-way domain adaptation

4.1.1. Formulation

Here after, ‘in-house’ and ‘public’ as ‘target’ and ‘source’, respectively, by applying the terminology used for domain adaptation. Let the data matrices of source and target domain are denoted as $\mathbf{X}_S \in \mathbb{R}^{n_S \times k}$, $\mathbf{X}_T \in \mathbb{R}^{n_T \times k}$, respectively, where n_S and n_T are the number of subjects in each data, and k is the number of brain ROIs. The adaptation matrix to be derived is $\mathbf{A} \in \mathbb{R}^{k \times k}$. To implement the distribution matching, the proposed method represents the mean vector of the data matrix as the center of the distribution. The adaptation matrix \mathbf{A} is a bridge so that the mean vector $\bar{\mathbf{X}}_T$ from the target data \mathbf{X}_T matches the mean vector $\bar{\mathbf{X}}_S$ from the source data \mathbf{X}_S . Therefrom, for the two mean vectors, \mathbf{A} minimizes the discrepancy defined by L2 norm as follows.

$$\|\bar{\mathbf{X}}_T \mathbf{A} - \bar{\mathbf{X}}_S\|_2^2 \quad (1)$$

The adaptation matrix \mathbf{A} simultaneously implements the connectivity mapping. In general, the brain connectivity is represented as the correlation between brain ROIs, but in this study, it is represented as the covariance. With the derived covariance matrices $\mathbf{C}_S \in \mathbb{R}^{k \times k}$ and $\mathbf{C}_T \in \mathbb{R}^{k \times k}$ for the source and target domain, respectively, \mathbf{A} reduces the difference between \mathbf{C}_S and \mathbf{C}_T by the term below.

$$\|\mathbf{A}^T \mathbf{C}_S \mathbf{A} - \mathbf{C}_T\|_2^2 \quad (2)$$

The objective function for \mathbf{A} is defined by linearly combining Eq. (1) and (2),

$$\operatorname{argmin}_{\mathbf{A}} \frac{\gamma_D}{2} \|\bar{\mathbf{X}}_T \mathbf{A} - \bar{\mathbf{X}}_S\|_2^2 + \frac{\gamma_C}{2} \|\mathbf{A}^T \mathbf{C}_S \mathbf{A} - \mathbf{C}_T\|_2^2 + \frac{1}{2} \|\mathbf{A}\|_2^2 \quad (3)$$

where the γ_D and γ_C are combining coefficients ($\gamma_* \geq 0$), and the last term $\|\mathbf{A}\|_2^2$ stands for the regularization. The purpose of \mathbf{A} is to transform \mathbf{X}_T to make it as similar to \mathbf{X}_S as possible by regularizing all elements evenly, so it is defined as L2-norm rather than L1-norm which can cause

the bias or the sparsity towards specific elements. The optimized solution for \mathbf{A} is derived by applying the alternating direction method of multipliers (ADMM) algorithm (Boyd, Parikh, Chu, Peleato, & Eckstein, 2011; Parikh & Boyd, 2014).

4.1.2. Optimization

At first, the original objective function in Eq. (3) is redefined into the following equivalent formulation by introducing an additional matrix \mathbf{U} and an equality constraint:

$$\operatorname{argmin}_{\mathbf{A}, \mathbf{U}} \frac{\gamma_D}{2} \|\bar{\mathbf{X}}_T \mathbf{U} - \bar{\mathbf{X}}_S\|_2^2 + \frac{\gamma_C}{2} \|\mathbf{A}^T \mathbf{C}_S \mathbf{U} - \mathbf{C}_T\|_2^2 + \frac{1}{2} \|\mathbf{U}\|_2^2 \quad (4)$$

s.t. $\mathbf{A} = \mathbf{U}$

Eq. (4) is a quadratic minimization problem in terms of \mathbf{A} and \mathbf{U} separately, subjecting to the equality constraint. Therefrom, the augmented Lagrangian function \mathbb{L} is

$$\mathbb{L}(\mathbf{A}, \mathbf{U}, \boldsymbol{\theta}) = \frac{\gamma_D}{2} \|\bar{\mathbf{X}}_T \mathbf{U} - \bar{\mathbf{X}}_S\|_2^2 + \frac{\gamma_C}{2} \|\mathbf{A}^T \mathbf{C}_S \mathbf{U} - \mathbf{C}_T\|_2^2 + \frac{1}{2} \|\mathbf{U}\|_2^2 + \operatorname{Tr}(\boldsymbol{\theta}^T (\mathbf{A} - \mathbf{U})) + \frac{\rho}{2} \|\mathbf{A} - \mathbf{U}\|_2^2$$

where $\boldsymbol{\theta}$ is the dual variable matrix and ρ is the penalty parameter, both for the equality constraint. The optimization process iteratively minimizes the augmented Lagrangian over the primal variable matrices \mathbf{A} and \mathbf{U} separately, while updating the dual variable matrix. Specifically, in the $(i+1)$ -th iteration, the following three steps are performed with the $(\mathbf{A}^{(i)}, \mathbf{U}^{(i)}, \boldsymbol{\theta}^{(i)})$.

(Step 1: Minimization over \mathbf{U}) Given the current fixed $\mathbf{A}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$, \mathbf{U} can be updated by

$$\mathbf{U}^{(i+1)} := \operatorname{argmin}_{\mathbf{U}} \mathbb{L}(\mathbf{A}^{(i)}, \mathbf{U}, \boldsymbol{\theta}^{(i)})$$

and the solution is

$$\mathbf{U}^{(i+1)} = \left(\gamma_D \bar{\mathbf{X}}_T^T \bar{\mathbf{X}}_T + \gamma_C \mathbf{C}_S^T \mathbf{A}^{(i)} \mathbf{A}^{(i)T} \mathbf{C}_S + (1 + \rho) \mathbf{I} \right)^{-1} \left(\gamma_D \bar{\mathbf{X}}_T^T \bar{\mathbf{X}}_S + \gamma_C \mathbf{C}_S^T \mathbf{A}^{(i)} \mathbf{C}_T + \boldsymbol{\theta}^{(i)} + \rho \mathbf{A}^{(i)} \right) \quad (5)$$

(Step 2: Minimization over \mathbf{A}) With the current values of $\mathbf{U}^{(i+1)}$ and $\boldsymbol{\theta}^{(i)}$ being fixed, \mathbf{A} is updated by

$$\mathbf{A}^{(i+1)} := \operatorname{argmin}_{\mathbf{A}} \mathbb{L}(\mathbf{A}, \mathbf{U}^{(i+1)}, \boldsymbol{\theta}^{(i)})$$

and the solution is

$$\mathbf{A}^{(i+1)} = \left(\gamma_C \mathbf{C}_S \mathbf{U}^{(i+1)} \mathbf{U}^{(i+1)T} \mathbf{C}_S^T + \rho \mathbf{I} \right)^{-1} \left(\mathbf{C}_S \mathbf{U}^{(i+1)} \mathbf{C}_T^T - \boldsymbol{\theta}^{(i)} + \rho \mathbf{U}^{(i+1)} \right) \quad (6)$$

(Step 3: Update of the dual variable matrix $\boldsymbol{\theta}$) The dual matrix $\boldsymbol{\theta}$ is lastly updated as below:

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} + \rho (\mathbf{A}^{(i+1)} - \mathbf{U}^{(i+1)}), \quad \rho = \min(\delta \rho, \rho_{\max}) \quad (7)$$

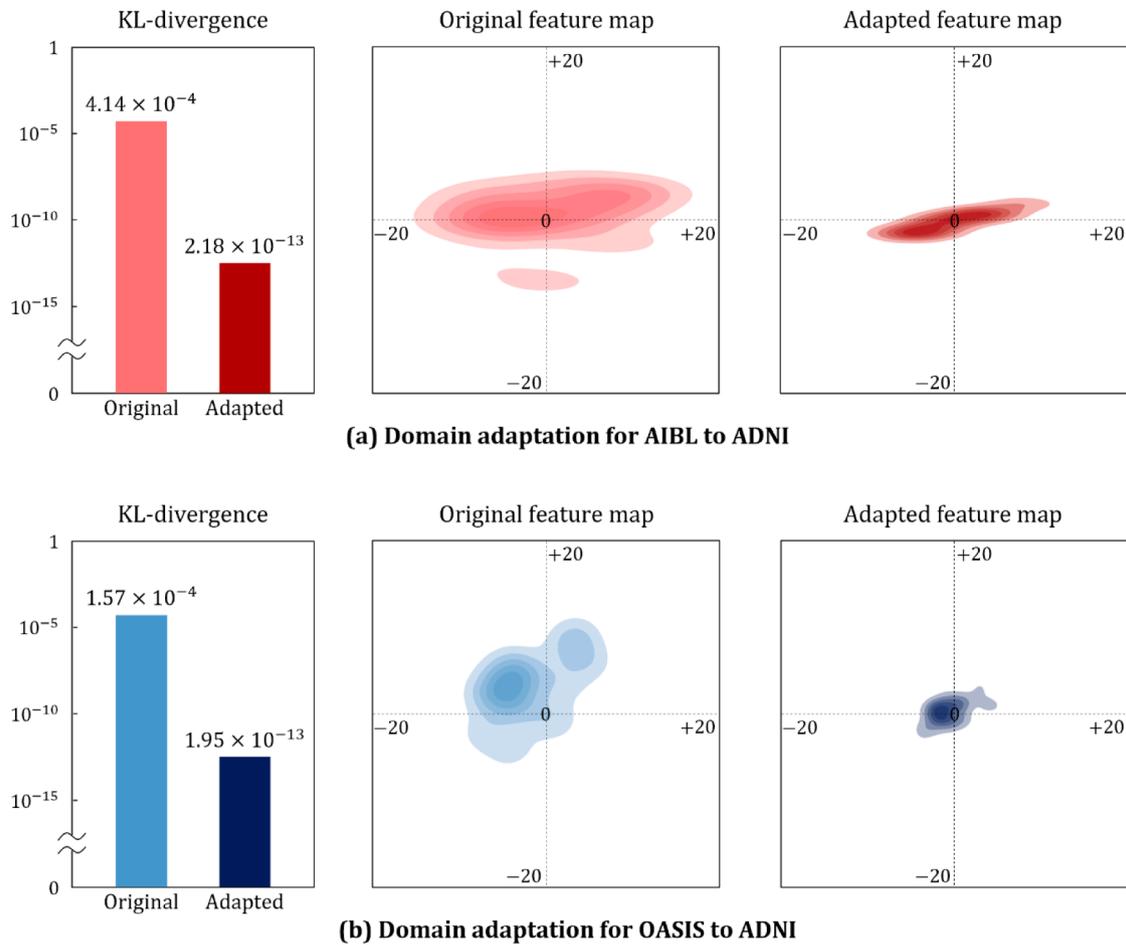


Fig. 3. Results for distribution matching. The left plots in both (a) and (b) represent KL-divergence of raw and adapted feature in AIBL and OASIS, for the ADNI. The middle and right plots show the original and the adapted feature maps visualized through t-SNE. With scaling, the mean vector of the ADNI data represents the origin.

In addition, \mathbf{A} and \mathbf{U} are initially set as

$$\begin{aligned} \mathbf{A}^{(1)} &= \mathbf{U}^{(1)} = \underset{\mathbf{U}}{\operatorname{argmin}} \|\bar{\mathbf{X}}_{\mathcal{F}}\mathbf{U} - \bar{\mathbf{X}}_{\mathcal{T}}\|_2^2 + \lambda\|\mathbf{U}\|_2^2 \\ &= (\bar{\mathbf{X}}_{\mathcal{F}}^T\bar{\mathbf{X}}_{\mathcal{F}} + \lambda\mathbf{I})^{-1}(\bar{\mathbf{X}}_{\mathcal{F}}^T\bar{\mathbf{X}}_{\mathcal{T}}) \end{aligned} \quad (8)$$

where λ is the regularization trade-off parameter, and all elements in the initial dual matrix are zeros. Algorithm 1 summarizes the procedure of one-way domain adaptation. For the tabular formed data, $\mathbf{X}_{\mathcal{F}}$ and $\mathbf{X}_{\mathcal{T}}$, of the preprocessed MR images, the mean vector and covariance matrix of each are used as inputs to the proposed method, which are utilized to implement the distribution matching and connectivity mapping, respectively. The proposed method performs iterative optimization to train the adaptation matrix \mathbf{A} for transforming $\mathbf{X}_{\mathcal{F}}$ as similar as possible to $\mathbf{X}_{\mathcal{T}}$, by applying the ADMM.

Algorithm 1. One-way domain adaptation

Input:

Mean vectors $\bar{\mathbf{X}}_{\mathcal{F}}$ and $\bar{\mathbf{X}}_{\mathcal{T}}$
Covariance matrices $\mathbf{C}_{\mathcal{F}}$ and $\mathbf{C}_{\mathcal{T}}$
Hyperparameters $\rho_{\max} = 10^6$, $\delta = 1.01$, $\epsilon = 10^{-7}$

Output: Adaptation matrix \mathbf{A}^*

Initialize $\mathbf{A}^{(1)}$, $\mathbf{U}^{(1)}$ by Eq. (8) and $\theta^{(1)} = 0$

repeat

Update $\mathbf{U}^{(i+1)}$, $\mathbf{A}^{(i+1)}$, $\theta^{(i+1)}$ by Eq. (5) - (7)

until $\|\mathbf{A}^{(i+1)} - \mathbf{U}^{(i+1)}\|_{\infty} < \epsilon$

return Adaptation matrix \mathbf{A}^*

4.2. AD conversion prediction

For subjects in the target domain, AD conversion prediction is performed by applying the classifier f pre-trained in the source domain. To train f , various machine learning algorithms can be used by selecting an appropriate error function \mathcal{E} . In the proposed method, the cross-entropy loss is used for training f as follows:

$$\mathcal{E}(\mathbf{Y}_{\mathcal{T}}, \hat{\mathbf{Y}}_{\mathcal{T}}) = \mathbf{Y}_{\mathcal{T}}^T \log \hat{\mathbf{Y}}_{\mathcal{T}} + (\mathbf{1} - \mathbf{Y}_{\mathcal{T}})^T \log (\mathbf{1} - \hat{\mathbf{Y}}_{\mathcal{T}})$$

where $\mathbf{Y}_{\mathcal{T}}$ is the labels with $Y_{\mathcal{T}(i)} = 1$ for MCI-C subjects and $Y_{\mathcal{T}(i)} = 0$ for MCI-N subjects, and $\hat{\mathbf{Y}}_{\mathcal{T}}$ is the predicted values. Finally, the prediction result for subjects in the target domain is derived as below:

$$\hat{\mathbf{Y}}_{\mathcal{T}} = f(\mathbf{X}_{\mathcal{F}}\mathbf{A}^*)$$

5. Results

5.1. Results for one-way domain adaptation

In this subsection, we describe results for one-way domain adaptation. The proposed method has two objectives to reduce the discrepancy between distributions of brain volume data (distribution matching) and to make the relationship between brain ROIs in the two domains similar (connectivity mapping). We represent the result for each objective in order.

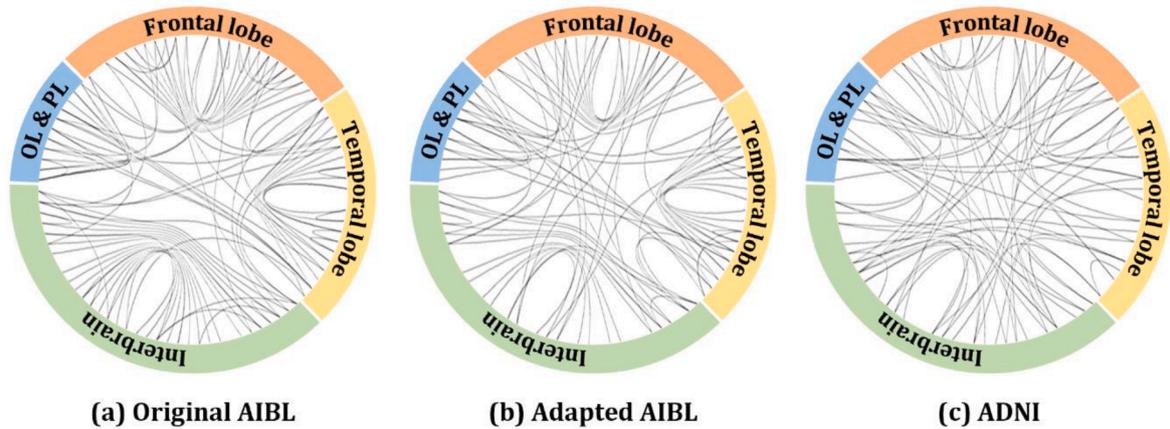


Fig. 4. Results for connectivity mapping. Each node represents an ROI, and nodes are sorted by dividing the frontal lobe, temporal lobe, interbrain, and occipital lobe (OL) and parietal lobe (PL). The original AIBL in (a) reveals more connections within regions and fewer between them. In contrast, (c) displays increased inter-region edges in the ADNI. The outcome of the proposed method in (b) minimizes the difference of edge pattern between (a) and (c). Particularly, the adapted AIBL illustrates reduced edges within the frontal lobe and the interbrain, while edges between them are increased. These findings highlight the successful performance of the proposed connectivity mapping.

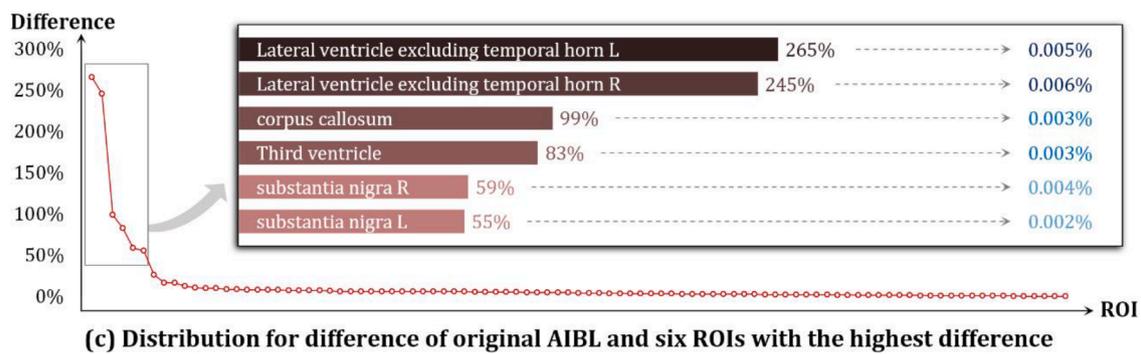
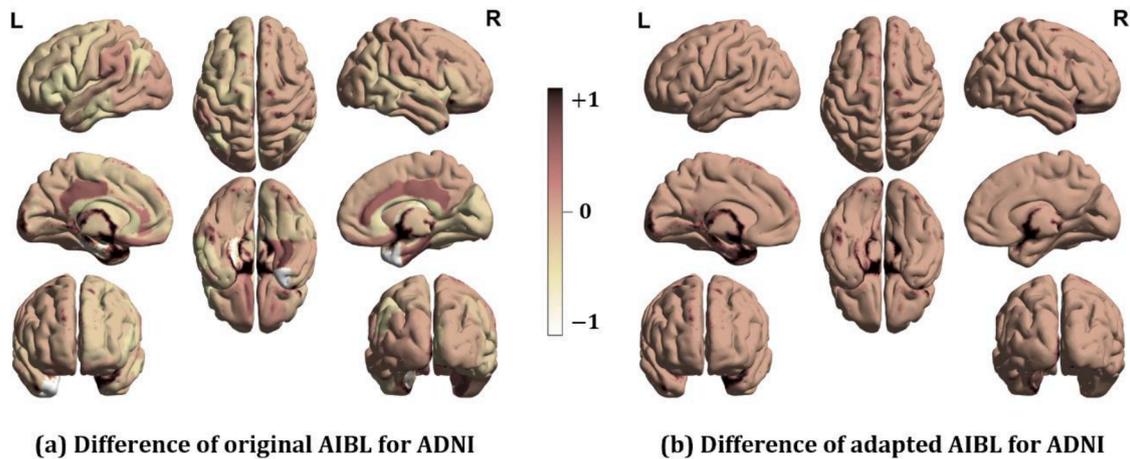


Fig. 5. Comparison results of brain volume features of AIBL for ADNI. (a) indicates the difference of brain volume feature between raw AIBL for ADNI, and (b) depicts the difference between adapted AIBL for ADNI. (c) shows the distribution for difference of raw AIBL by entire ROI and presents the result of reducing the difference via domain adaptation for six ROIs with the highest.

5.1.1. Distribution matching

At first, AIBL and OASIS were set as target domains and new feature representations of those data were derived according to ADNI, respectively. We measured KL-divergence (Kullback, 1959; Kullback & Leibler, 1951) to evaluate whether the data distribution of the target and the source coincided with each other. The results for distribution matching are shown in Fig. 3. From the left plots of Fig. 3(a) and (b), it seems that the KL-divergence with the ADNI data distribution was significantly

reduced in both AIBL and OASIS. The value of KL-divergence indicates that the two distributions become more similar to each other as they get smaller and closer to zero. Therefore, the distribution matching aimed by the proposed method was performed well. In addition, the middle and right plots show the original and the adapted feature maps for AIBL and OASIS visualized through t-SNE (Van Der Maaten, 2014; Van der Maaten & Hinton, 2008). With scaling, the mean vector of the ADNI data represents the origin. Those plots show that for both AIBL and OASIS,

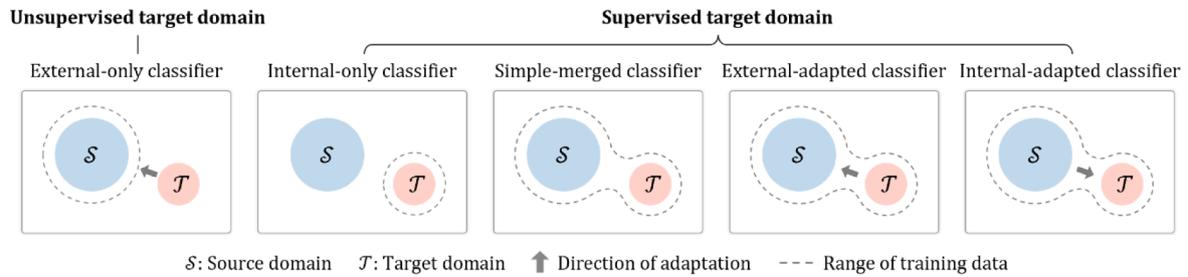


Fig. 6. Classifier training methods for comparison experiments. The external-only classifier only trains the source label and performs prediction of target domain by domain adaptation. The internal-only classifier is a case in which the target domain constructs a model using only its own label. The simple-merged classifier utilizes the labels of both domains but does not perform domain adaptation, but in the external-adapted classifier, the target domain is adapted to the source, and in the internal-adapted classifier, the source domain is adapted to the target.

the adapted features are closer to the origin. This consolidates that domain adaptation is well matched to the ADNI data distribution.

5.1.2. Connectivity mapping

Next, as the second objective of domain adaptation, the proposed method performs connectivity mapping so that the relationship between brain ROIs of the target domain is similar to that of the source domain. Fig. 4 shows the result in the form of a brain connectivity network. Each node represents an ROI, and nodes were sorted by dividing the frontal lobe, temporal lobe, interbrain, and occipital lobe (OL) and parietal lobe (PL). In Fig. 4(a), the original AIBL shows many intra-region edges and few inter-region edges. On the contrary, in Fig. 4(c), there are many inter-regions edges in ADNI data. The proposed method shows the result of Fig. 4(b) by reducing the difference in these edge patterns. Particularly, the adapted AIBL indicates that edges within the frontal lobe and the interbrain are decreased while between these regions increased. These results present that the connectivity mapping of the proposed method was well performed.

5.1.3. Overall comparison

At last, Fig. 4 shows the results of comparing the adapted brain volume features of AIBL derived by applying the proposed method with ADNI. Fig. 5(a) and (b) are diagrams of the differences between ADNI and original AIBL and adapted AIBL. Before domain adaptation, AIBL

showed a wide difference in volume for each ROI compared to ADNI. The colors of the ROI shown in Fig. 5(a) are also varied. After domain adaptation, however, the difference in brain volume between the two data approaches zero in almost all ROIs, with Fig. 5(b) in a single color. Fig. 5(c) shows the difference from the original AIBL with respect to ADNI as a percentage for each ROI and sorted in descending order. There are six notable ROIs with a large difference over 50% as large as 265%, and most of them are related to the ventricle and substantia. This seems to be because these regions are located in the midline between the two hemispheres, so it is the deep part of the brain where a large deviation can occur depending on the imaging equipment. Nevertheless, in the proposed method, features were derived so that even these ROIs were the same as the source domain, showing differences close to zero. This result suggests that the proposed method performs robust domain adaptation.

5.2. Results for AD conversion prediction

In this subsection, we describe results for AD conversion prediction. For the diversified analysis of the effect of domain adaptation, we conducted various comparison experiments assuming two scenarios. One is the case of unsupervised target domain where there is no label information for prediction in the target data, so it is completely dependent on the source domain. The other is the case of supervised target domain

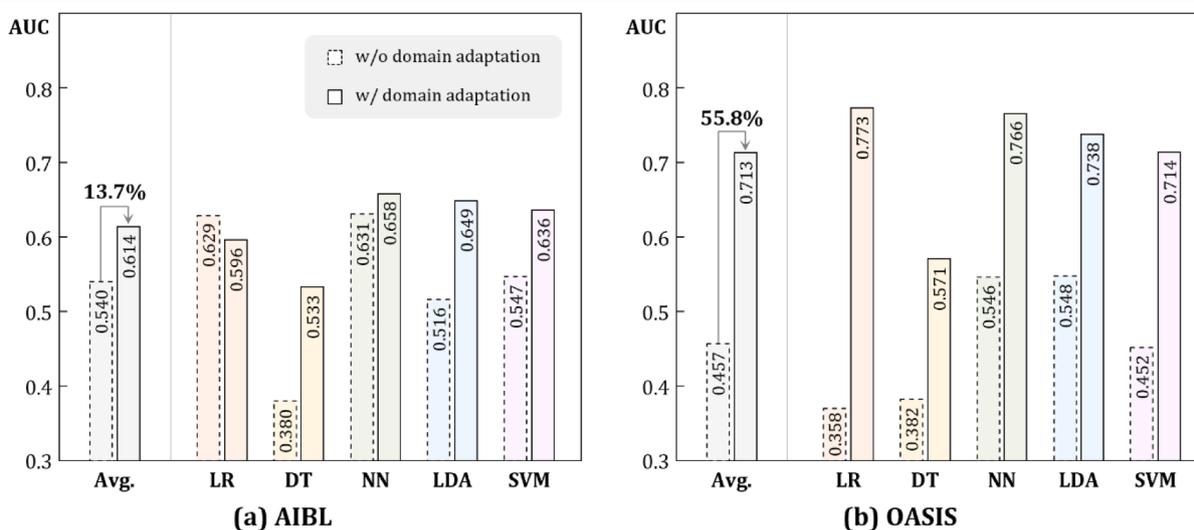


Fig. 7. Results for comparison experiments on the unsupervised target domain. With the unlabeled target domain, the proposed method employs the external-only classifier. This includes domain adaptation of the target domain, utilizing classifiers trained on the source domain. The experiment compares the performance with and without domain adaptation. By applying the proposed method to the AIBL and OASIS as target domains using classifiers trained on the ADNI as the source domain, it is revealed that domain adaptation improves classification performance by 13.7% and 55.8% on average in the AIBL and OASIS, respectively. This highlights that the proposed method enhances classification performance on unsupervised target domains.

Table 2
Results for comparison experiments on the supervised target domain.

Dataset	Algorithm	Classifier			
		Internal-only	Simple-merged	External-adapted	Internal-adapted
AIBL	LR	0.532	0.542	0.555	0.494
	DT	0.406	0.485	0.483	0.637
	NN	0.658	0.565	0.694	0.488
	LDA	0.497	0.499	0.597	0.461
	SVM	0.478	0.580	0.554	0.470
OASIS	LR	0.778	0.794	0.663	0.642
	DT	0.724	0.650	0.766	0.749
	NN	0.769	0.682	0.827	0.821
	LDA	0.782	0.637	0.837	0.787
	SVM	0.735	0.750	0.808	0.499

that the data has labels but additionally utilizes the source domain with a larger amount of information to improve prediction efficiency.

5.2.1. Comparison methods for evaluation

For the performance comparison, we designed five classifier training methods for these scenarios, as shown in Fig. 6. In this figure, the external-only classifier corresponds to the former scenario, and the other four to the latter. First, the external-only classifier indicates a model trained with the label of the source domain, which is external information of the target domain. After the data in target domain is transformed by domain adaptation, it is applied to the classifier to derive the prediction result. Next, the internal-only classifier is a model in which the target domain learns with its own labels without utilizing the source domain. On the contrary, the simple-merged classifier uses labels in both the target domain and the source, but domain adaptation for the target is not performed. The external-adapted classifier is a model in which the labels of both domains are utilized as the target domain is adapted to the source domain, and the internal-adapted classifier is a model in which the source domain is adapted to the target by reversing the direction of domain adaptation. Each classifier training method was applied to five machine learning algorithms mainly used for AD conversion prediction: logistic regression (LR) (Johnson, et al., 2014; Xiao, Cui, Qiao, Zheng, & Zhang, 2021), decision tree (DT) (Jin & Deng, 2018; Ritter, et al., 2015),

neural network (NN) (Basaia, et al., 2019; Gao, et al., 2020), linear discriminant analysis (LDA) (Cho, Seong, Jeong, Shin, & Initiative, 2012; Coupé, et al., 2015; Wolz, et al., 2011), and support vector machine (SVM) (Schmitter, et al., 2015; Wei, Li, Fogelson, & Li, 2016; Zhang, et al., 2015). The performance was measured by the area under receiving operating characteristic curve (AUC), and the entire experiment was 100 times repeated.

5.2.2. Results for prediction on the unsupervised target domain

When there is no label in the target domain, prediction can be performed using the external-only classifier. At this time, the target domain is transformed through domain adaptation, and the classifier learned from the source domain can be used for AD conversion prediction. In this experiment, the performance of the proposed method is compared to the case where domain adaptation is not applied, and the results are depicted in Fig. 7. AIBL and OASIS were applied as target domains to the classifier learned from ADNI, the source domain, respectively. As a result, it was found that the prediction results with domain adaptation for ADNI were 13.7 % and 55.8 % better on average in AIBL and OASIS, respectively, than without domain adaptation. Therefore, when prediction is performed on the unsupervised target domain, high performance can be derived by applying the proposed method.

5.2.3. Results for prediction on the supervised target domain

On the other hand, when there are labels in the target domain, prediction can be performed alone, but it can also receive the advantage of source information through domain adaptation. In this experiment, we compared internal-only classifier with three other classifiers which are trained in both source and target domains. Table 2 shows the results for comparison experiments on the supervised target domain. As a result, the internal-only classifier showed lower performance than other classifiers. Especially, the performance of the external-adapted classifier was the best in most experimental settings. Fig. 8 depicts the result for AUC comparison of internal-only and external-adapted classifiers. At first, Fig. 8(a) shows AUC comparison by algorithms used for prediction on the AIBL and OASIS. The diagonal line is a criterion for performance comparison, and when a point is located above the line, the vertical axis performs better. The external-adapted classifier performed better in all

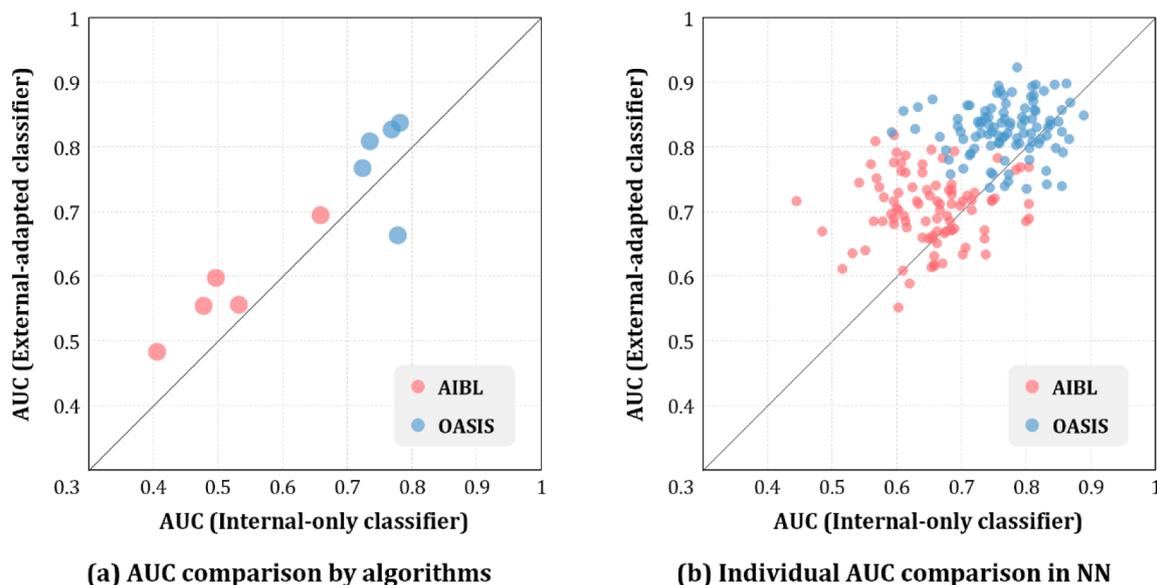


Fig. 8. AUC comparison of internal-only and external-adapted classifiers. The diagonal line is a criterion for performance comparison, and when a point is located above the line, the vertical axis performs better. (a) shows the AUC comparisons for the prediction on the AIBL and OASIS, indicating that the external-adapted classifier outperforms the internal-only classifier in most algorithms. (b) focuses on the neural network, demonstrating consistent superiority of the external-adapted classifier in both datasets, particularly in OASIS. These results highlight that the proposed method enhance the classification performance through domain adaptation, even if the label information exists in the target domain.

Table 3
Comparison results for domain adaptation.

Dataset	SVD-based method			DAT-based method			Ours	
	DADS	MALRR	MCNFE	DANN	PCAL	CDAN		ADDA
AIBL	6.82×10^{-5}	5.42×10^{-5}	1.10×10^{-5}	8.06×10^{-6}	3.14×10^{-6}	8.05×10^{-7}	2.90×10^{-8}	2.18×10^{-13}
OASIS	3.02×10^{-5}	5.19×10^{-5}	1.09×10^{-5}	2.71×10^{-6}	9.84×10^{-7}	2.71×10^{-7}	1.04×10^{-9}	1.95×10^{-13}

Table 4
Comparison results for AD conversion prediction.

(a) Unsupervised target domain						
Method	AIBL			OASIS		
	AUROC	Accuracy	AUPRC	AUROC	Accuracy	AUPRC
DADS	0.558	0.601	0.573	0.643	0.662	0.655
MALRR	0.603	0.625	0.617	0.648	0.674	0.659
MCNFE	0.627	0.659	0.631	0.691	0.736	0.713
DANN	0.600	0.655	0.614	0.679	0.721	0.707
PCAL	0.614	0.669	0.626	0.692	0.743	0.721
CDAN	0.631	0.686	0.643	0.754	0.791	0.776
ADDA	0.644	0.698	0.655	0.769	0.802	0.783
Ours	0.658	0.707	0.683	0.773	0.810	0.796
(b) Supervised target domain						
Method	AIBL			OASIS		
	AUROC	Accuracy	AUPRC	AUROC	Accuracy	AUPRC
DADS	0.627	0.674	0.663	0.739	0.789	0.760
MALRR	0.661	0.691	0.684	0.764	0.820	0.799
MCNFE	0.671	0.714	0.707	0.768	0.824	0.802
DANN	0.653	0.683	0.671	0.813	0.855	0.827
PCAL	0.677	0.724	0.706	0.824	0.872	0.854
CDAN	0.689	0.732	0.713	0.833	0.885	0.861
ADDA	0.717	0.755	0.734	0.897	0.932	0.921
Ours	0.694	0.740	0.729	0.837	0.890	0.863

but one case. Next, Fig. 8(b) presents the individual AUC comparison of internal-only classifier and external-adapted classifier for neural networks which showed overall superior results compared to other algorithms. In both data, the external-adapted classifier shows a general performance advantage, and its superiority is more stable in OASIS than AIBL. Therefore, even if there is label information in the target domain, it is possible to perform more successful prediction by performing domain adaptation to the source domain through the proposed method.

Table 5
Parameter sensitivity analysis on one-way domain adaptation.

(a) AIBL								
Parameter	γ_D							
	Value	10^{-3}	10^{-2}	10^{-1}	10^0	10^1	10^2	10^3
γ_C	10^{-3}	4.46×10^{-7}	4.46×10^{-7}	4.46×10^{-7}	3.62×10^{-7}	1.79×10^{-7}	2.54×10^{-8}	5.09×10^{-10}
	10^{-2}	4.41×10^{-12}	4.44×10^{-12}	4.46×10^{-12}	4.49×10^{-12}	4.51×10^{-12}	4.46×10^{-12}	3.90×10^{-12}
	10^{-1}	1.35×10^{-12}	1.36×10^{-12}	1.38×10^{-12}	1.39×10^{-12}	1.40×10^{-12}	1.39×10^{-12}	1.24×10^{-12}
	10^0	4.78×10^{-13}	4.86×10^{-13}	4.94×10^{-13}	5.02×10^{-13}	5.09×10^{-13}	5.11×10^{-13}	4.62×10^{-13}
	10^1	2.18×10^{-13}	2.30×10^{-13}	2.42×10^{-13}	2.54×10^{-13}	2.66×10^{-13}	2.75×10^{-13}	2.59×10^{-13}
	10^2	3.44×10^{-7}	3.44×10^{-7}	3.44×10^{-7}	3.44×10^{-7}	2.57×10^{-8}	3.18×10^{-10}	3.22×10^{-12}
	10^3	2.60×10^{-6}	2.60×10^{-6}	2.60×10^{-6}	1.53×10^{-6}	1.29×10^{-8}	3.48×10^{-10}	3.63×10^{-12}
	(b) OASIS							
Parameter	γ_D							
	Value	10^{-3}	10^{-2}	10^{-1}	10^0	10^1	10^2	10^3
γ_C	10^{-3}	3.92×10^{-7}	3.87×10^{-7}	3.69×10^{-7}	3.19×10^{-7}	2.03×10^{-7}	5.16×10^{-8}	2.10×10^{-9}
	10^{-2}	2.69×10^{-11}	2.69×10^{-11}	2.69×10^{-11}	2.69×10^{-11}	2.69×10^{-11}	2.67×10^{-11}	2.51×10^{-11}
	10^{-1}	1.44×10^{-11}	1.44×10^{-11}	1.44×10^{-11}	1.44×10^{-11}	1.44×10^{-11}	1.43×10^{-11}	1.35×10^{-11}
	10^0	7.95×10^{-12}	7.95×10^{-12}	7.96×10^{-12}	7.96×10^{-12}	7.96×10^{-12}	7.91×10^{-12}	7.46×10^{-12}
	10^1	9.04×10^{-13}	9.07×10^{-13}	9.09×10^{-13}	9.12×10^{-13}	9.14×10^{-13}	9.12×10^{-13}	8.61×10^{-13}
	10^2	1.61×10^{-8}	1.61×10^{-8}	1.61×10^{-8}	1.61×10^{-8}	1.72×10^{-9}	1.97×10^{-11}	1.95×10^{-13}
	10^3	4.89×10^{-7}	4.89×10^{-7}	5.05×10^{-7}	2.85×10^{-7}	5.08×10^{-9}	1.06×10^{-10}	1.11×10^{-12}

5.3. Comparison experiments

In this subsection, we compare the results for domain adaptation and AD conversion prediction with the existing studies which are described in Section 2.

5.3.1. Experimental settings

Comparison experiments were conducted in the same settings as the proposed method for a total of seven comparison methods including three SVD-based methods (DADS, MALRR, and MCNFE) and four DAT-based methods (DANN, PCAL, CDAN, and ADDA). Performance on domain adaptation was measured by KL-Divergence (KLD), and performance on AD conversion prediction was measured by AUROC, accuracy, and area under the precision-recall curve (AUPRC).

5.3.2. Comparison results for domain adaptation

Comparison results for domain adaptation are shown in Table 3. All methods, including the proposed method, yielded better KLDs when domain adapting OASIS to ADNI than AIBL. Among the comparison methods, it was found that the DAT-based methods transform the target domain more similarly to the source domain than the SVD-based methods. Nevertheless, the proposed method showed the best results for domain adaptation in both AIBL and OASIS datasets than the comparison methods.

5.3.3. Comparison results for AD conversion prediction

Comparison experiments for AD conversion prediction were performed for both unsupervised and supervised target domains. The performance of the comparison methods was compared with the best results obtained by the proposed method (NN for AIBL and LR for OASIS for the unsupervised target domains and NN for AIBL and LDA for OASIS for the supervised target domains). Table 4 shows the comparison results for AD conversion prediction. First, in Table 4(a), the proposed method shows better prediction performance than all comparison methods for the

unsupervised target domains. Comparing the average values of all evaluation metrics measured in both AIBL and OASIS, the proposed method shows about 10 % better performance than the comparison methods. Among the comparison methods, the average performance for the DAT-based methods was measured to be about 8 % higher than the average performance for the SVD-based methods. On the other hand, the results for the supervised target domains shown in Table 4(b) indicate a slightly different aspect. The method that shows the best performance is the DAT-based ADDA, and the proposed method derives the next best performance. When comparing the average values of the evaluation metrics, ADDA shows about 4 % better results than the proposed method. Similar to the results for unsupervised target domains, the DAT-based method yielded about 7 % better prediction results on average than the SVD-based methods.

6. Discussion

In this section, the proposed method is discussed in terms of the performance, supplementation, and application. At first, in this study, the performance of each of the two components, *one-way domain adaptation* and *AD conversion prediction*, of the proposed method was evaluated through various experiments. In one-way domain adaptation, we aimed at *distribution matching* and *connectivity mapping* so that the data distribution of the target domain and the connectivity between brain regions are similar to those of the source domain. These aims were implemented by training the adaptation matrix that reduces the difference between mean vectors and covariance matrices. As a result, the KL-divergence between source and target domains was dramatically decreased, and therefrom, the performance of AD conversion prediction could be improved as well. Even if there is no label information in the AIBL and OASIS datasets used as target domains, it was possible to derive 13.7 % and 55.8 % improved prediction results, respectively, through the proposed one-way domain adaptation. Unsurprisingly, higher performance was shown when the target domain contained labels. Through comparison experiments, it was confirmed that the performance of the proposed method for domain adaptation was superior to other methods. This result leads to the best prediction performance for the unsupervised target domains. On the other hand, for the supervised target domains, the proposed method showed the second-best performance compared to the DAT-based method by a slight difference.

Next, there were three supplementary points of the proposed method that we came to recognize. *First*, several hyperparameters in the objective function of the one-way domain adaptation need to be optimized. γ_D and γ_C in Equation (3), which are respectively combined with terms for implementing distribution matching and connectivity mapping, were experimentally selected within $\{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2, 10^3\}$. Table 5 shows the results for the parameter sensitivity analysis on one-way domain adaptation. From the result, it can be seen that the trend of performance change according to the change of each hyperparameter is not consistent. Therefore, the value of γ_D was extremely different in deriving the best domain adaptation results from the two datasets. Considering that the result of one-way domain adaptation has a significant effect on AD conversion prediction performance, the robustness of hyperparameters is the first issue to be addressed in the future work. *Second*, the proposed method was designed to apply a single public data. To supplement the insufficient in-house data more effectively, this method should be extended to enable utilizing multiple public data. *Last*, the one-way domain adaptation is limited to the linear transformation. For the more sophisticated matching with the public data, the non-linear transformation would be better than the linear.

Although this study was conducted for a specific task on medical imaging data, the proposed method can be applied to various fields. The core of our method is to transform the information-poor target domain to be similar to the information-rich source domain while preserving the original properties of the source domain. This feature means that the

transformed target domain is combined with the original source domain, so that any task previously performed in the source domain can be performed, without restrictions. This suggests that there is no need for restrictions or additional works caused by simultaneous transformation of two domains according to the pre-defined task. Therefore, the proposed method is a well-generalized algorithm that can be flexibly applied to various fields or expert systems.

7. Conclusion

In this paper, we propose a novel method for domain adaptation to predict the AD conversion from MCI patients using MR images. The most pronouncing feature of our method is to only transform the in-house data (target domain) so that preserve the public data (source domain). This process prevents the re-training and the loss of rich information of the source domain, and therefrom, can guarantee the robustness of classification performance. Consequently, the prediction for the transformed data of the AIBL and OASIS was performed on the classifier which trained by the labels of the ADNI, and the results showed the good performance including the significant improvement comparing with the prediction without the proposed method. Therefore, the proposed method can be applied to datasets in multiple sites, which have not sufficient information, and it is considered a useful method to perform successful AD conversion prediction with only a small amount of data.

In sum, our main contributions are summarized as follows. (a) We tackle the inter-site heterogeneity problem comes from merging or integrating multisite MRI data to supplement the insufficiency of in-house data. (b) To tackle this problem, we propose the one-way domain adaptation which only transforms the in-house data to match the data distribution and the brain region connectivity with the public data. (c) The proposed method avoids the tedious re-training and the information loss of public data, for the change of in-house data. (d) We could validate that our method enables the small data to conduct the predictive task more efficiently and effectively by using the large dataset, in the prediction of AD conversion from MCI patients on the famous brain MRI public datasets.

CRedit authorship contribution statement

Sunghong Park: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Sang Joon Son:** Conceptualization, Validation, Writing – review & editing, Funding acquisition. **Kanghee Park:** Conceptualization, Methodology, Software, Formal analysis, Validation. **Yonghyun Nam:** Conceptualization, Methodology, Software, Formal analysis, Validation. **Hyunjung Shin:** Conceptualization, Methodology, Investigation, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (NRF-2022R1A6A3A01086784), the BK21 FOUR program of the NRF funded by the MOE (NRF5199991014091), and the Ajou University research fund. This research was also supported by the NRF grant funded by the Ministry of Science & ICT (MSIT) (NRF-

2019R1A5A2026045 and NRF-2021R1A2C2003474), the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the MSIT (No. 2022-0-00653, Voice Phishing Information Collection and Processing and Development of a Big Data Based Investigation Support System), the grant funded by the MSIT (KISTI Project No. K-23-L03-C02-S01 and J-23-RD-CR02-S01), and the grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, Republic of Korea (HR21C1003 and HR22C1734).

References

- Arbabshirani, M. R., Plis, S., Sui, J., & Calhoun, V. D. (2017). Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *NeuroImage*, *145*, 137–165.
- Ashburner, J., Barnes, G., Chen, C.-C., Daunizeau, J., Flandin, G., Friston, K., ... Moran, R. (2014). SPM12 manual. *Wellcome Trust Centre for Neuroimaging, London, UK*, 2464, 4.
- Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., ... s. D. N. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical*, *21*, Article 101645.
- Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. *Machine learning*, *79*, 151–175.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, *3*, 1–122.
- Chen, M., Weinberger, K. Q., & Blitzer, J. (2011). Co-Training for Domain Adaptation. In *NIPS* (Vol. 24, pp. 2456–2464). Citeseer.
- Cheng, B., Liu, M., Zhang, D., Munsell, B. C., & Shen, D. (2015). Domain transfer learning for MCI conversion prediction. *IEEE Transactions on Biomedical Engineering*, *62*, 1805–1817.
- Cho, Y., Seong, J.-K., Jeong, Y., Shin, S. Y., & Initiative, A. S. D. N. (2012). Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *NeuroImage*, *59*, 2217–2230.
- Coupé, P., Fonov, V. S., Bernard, C., Zandifar, A., Eskildsen, S. F., Helmer, C., ... Allard, M. (2015). Detection of Alzheimer's disease signature in MR images seven years before conversion to dementia: Toward an early individual prognosis. *Human Brain Mapping*, *36*, 4758–4770.
- Cummings, J., Lee, G., Ritter, A., Sabbagh, M., & Zhong, K. (2019). Alzheimer's disease drug development pipeline: 2019. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *5*, 272–293.
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... Lempitsky, V. (2016). Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, *17*, 2096.
- Gao, F., Yoon, H., Xu, Y., Goradia, D., Luo, J., Wu, T., Su, Y., & Initiative, A. S. D. N. (2020). AD-NET: Age-adjust neural network for improved MCI to AD conversion prediction. *NeuroImage: Clinical*, *27*, 102290.
- Gaser, C., Dahnke, R., Thompson, P. M., Kurth, F., Luders, E., & Initiative, A. S. D. N. (2022). CAT—A computational anatomy toolbox for the analysis of structural MRI data. *bioRxiv*, 2022.2006. 2011.495736.
- Guan, H., Liu, Y., Yang, E., Yap, P.-T., Shen, D., & Liu, M. (2021). Multi-site MRI harmonization via attention-guided deep domain adaptation for brain disorder identification. *Medical Image Analysis*, *71*, Article 102076.
- Huang, J., Gretton, A., Borgwardt, K., Schölkopf, B., & Smola, A. (2006a). Correcting sample selection bias by unlabeled data. *Advances in Neural Information Processing Systems*, *19*.
- Huang, J., Gretton, A., Borgwardt, K., Schölkopf, B., & Smola, A. (2006b). Correcting sample selection bias by unlabeled data. *Advances in Neural Information Processing Systems*, *19*, 601–608.
- Huang, Y.-L., Hsieh, W.-T., Yang, H.-C., & Lee, C.-C. (2020). Conditional domain adversarial transfer for robust cross-site ADHD classification using functional MRI. In *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1190–1194). IEEE.
- Jin, M., & Deng, W. (2018). Predication of different stages of Alzheimer's disease using neighborhood component analysis and ensemble decision tree. *Journal of Neuroscience Methods*, *302*, 35–41.
- Johnson, P., Vandewater, L., Wilson, W., Maruff, P., Savage, G., Graham, P., ... Martins, R. N. (2014). Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease. *BMC Bioinformatics*, *15*, 1–14.
- Kullback, S. (1959). *Information theory and statistics*. New York: John Riley and Sons Inc.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22*, 79–86.
- Lee, G., Nho, K., Kang, B., Sohn, K.-A., & Kim, D. (2019). Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Scientific reports*, *9*, 1–12.
- Li, W., Zhao, Y., Chen, X., Xiao, Y., & Qin, Y. (2018). Detecting Alzheimer's disease on small dataset: A knowledge transfer perspective. *IEEE journal of biomedical and health informatics*, *23*, 1234–1242.
- Liu, X., Tosun, D., Weiner, M. W., Schuff, N., & Initiative, A. S. D. N. (2013). Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. *NeuroImage*, *83*, 148–157.
- Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2010). Open access series of imaging studies: Longitudinal MRI data in nondemented and demented older adults. *Journal of cognitive neuroscience*, *22*, 2677–2684.
- Moradi, E., Gaser, C., Huttunen, H., & Tohka, J. (2014). MRI based dementia classification using semi-supervised learning and domain adaptation. In *MICCAI 2014 Workshop Proceedings, Challenge on Computer-Aided Diagnosis of Dementia, based on Structural MRI Data* (Vol. 20).
- Orbes-Arteaga, M., Varsavsky, T., Sudre, C. H., Eaton-Rosen, Z., Haddow, L. J., Sørensen, L., Nielsen, M., Pai, A., Ourselin, S., & Modat, M. (2019). Multi-domain adaptation in brain MRI through paired consistency and adversarial learning. In *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data: First MICCAI Workshop, DART 2019, and First International Workshop, MIL3ID 2019, Shenzhen, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings 1* (pp. 54–62): Springer.
- Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*, 1345–1359.
- Parikh, N., & Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in optimization*, *1*, 127–239.
- Park, S., Hong, C. H., Lee, D.-G., Park, K., Shin, H., Initiative, A., & s. D. N. (2023). Prospective classification of Alzheimer's disease conversion from mild cognitive impairment. *Neural Networks*, *164*, 335–344.
- Pellegrini, E., Ballerini, L., Hernandez, M. d. C. V., Chappell, F. M., González-Castro, V., Anblagan, D., Danso, S., Muñoz-Maniega, S., Job, D., & Pernet, C. (2018). Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *10*, 519–535.
- Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of neurology*, *56*, 303–308.
- Petrella, J. R., Coleman, R. E., & Doraiswamy, P. M. (2003). Neuroimaging and early diagnosis of Alzheimer disease: A look to the future. *Radiology*, *226*, 315–336.
- Rathore, S., Habes, M., Ifitkhar, M. A., Shacklett, A., & Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage*, *155*, 530–548.
- Ritter, K., Schumacher, J., Weygandt, M., Buchert, R., Allefeld, C., Haynes, J.-D., ... s. D. N. (2015). Multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *1*, 206–215.
- Schmitter, D., Roche, A., Maréchal, B., Ribes, D., Abdulkadir, A., Bach-Cuadra, M., ... Maeder, P. (2015). An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical*, *7*, 7–17.
- Shen, D., & Davatzikos, C. (2002). HAMMER: Hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, *21*, 1421–1439.
- Van Der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. *The Journal of Machine Learning Research*, *15*, 3221–3245.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*.
- Wang, N., Yao, D., Ma, L., & Liu, M. (2022). Multi-site clustering and nested feature extraction for identifying autism spectrum disorder with resting-state fMRI. *Medical Image Analysis*, *75*, Article 102279.
- Wei, R., Li, C., Fogelson, N., & Li, L. (2016). Prediction of conversion from mild cognitive impairment to Alzheimer's Disease using MRI and structural network features. *Frontiers in aging neuroscience*, *8*, 76.
- Wolz, R., Julkunen, V., Koikkalainen, J., Niskanen, E., Zhang, D. P., ... Rueckert, D. s. D. N. (2011). Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. *PLoS One*, *6*, e25446.
- Xiao, R., Cui, X., Qiao, H., Zheng, X., & Zhang, Y. (2021). Early diagnosis model of Alzheimer's disease based on sparse logistic regression. *Multimedia Tools and Applications*, *80*, 3969–3980.
- Xue, G.-R., Dai, W., Yang, Q., & Yu, Y. (2008). Topic-bridged PLSA for cross-domain text classification. In *In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 627–634).
- Zhang, Y., Wang, S., Phillips, P., Dong, Z., Ji, G., & Yang, J. (2015). Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomedical Signal Processing and Control*, *21*, 58–73.