MULTI-SOURCE DOMAIN ADAPTATION VIA OPTIMAL TRANSPORT FOR BRAIN DEMENTIA IDENTIFICATION

Hao Guan, Li Wang, Mingxia Liu**

Department of Radiology and BRIC
University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

ABSTRACT

Multi-site MRI data have been increasingly employed for automated identification of brain dementia, but are susceptible to large domain shift between different imaging sites/centers. Previous studies usually simply ignore the domain shift caused for instance by different scanners/protocols. Even though several studies proposed to reduce inter-domain discrepancy, they generally require a part of labeled target data and cannot well handle problems with multi-source domains. To this end, we propose a multi-source optimal transport (MSOT) framework for cross-domain Alzheimer's disease (AD) diagnosis with multi-site MRI data. Specifically, we first project data from multi-source domains to target domain through optimal transport in an unsupervised manner. Based on projected representation, we calculate the similarity between each source and target domains, and use this similarity as the source domain weight. We then train a support vector machine (SVM) classifier based on projected samples from each source domain. Finally, an ensemble learning strategy via weighted voting is used to predict labels of target samples. The proposed MSOT does not require labeled target data and can be efficiently optimized. Experiments were performed on three benchmark neuroimaging datasets for AD identification, with results suggesting the superiority of MSOT over several state-of-the-art methods.

***Index Terms*—** Data Adaptation, Optimal Transport, MRI, Brain Dementia, Classification

1. INTRODUCTION

Structural magnetic resonance imaging (MRI) provides an non-invasive solution to study the human brain, and has been widely used in computer-aided diagnosis of brain dementia, such as Alzheimer's disease (AD). Multi-site MRIs acquired from multiple imaging centers/domains have been increasingly employed to increase sample size and statistical power for AD-related disease diagnosis. Conventional methods typically assume that training data (*w.r.t.* source domain) and test data (*w.r.t.* target domain) share the same distribution. But this assumption does not always hold in practice, due to multi-site domain shift caused by different scanners/protocols and studied populations [1]. When directly applying a model (well trained on source data) to unseen target data, multi-site domain shift may lead to poor generalization performance.

To tackle this issue, some studies utilize domain adaptation techniques to reduce the discrepancy between source and target domains. Wachinger *et al.* [2] proposed an instance weighting method for AD classification. They first used a part of labeled target samples to estimate target data distribution, and then re-weighted source samples by calculating the probability of each source sample belonging to target domain. A classifier was finally trained using the re-weighted source samples and some labeled target data. In the instance weighting strategy, some samples may get very low weights. When the cross-domain discrepancy is large, most of source samples would be down-weighted, leading to a smaller set of effective instances for training. Li *et al.* [3] proposed a subspace alignment-based adaptation method for AD classification. They first extracted functional MRI features for source and target samples, and then projected samples from both domains into a shared subspace to reduce the domain discrepancy. Moradi *et al.* [4] proposed to utilize transductive support vector machine (TSVM) for domain adaptation using MRIs represented by gray matter density features. Cheng *et al.* [5] proposed a feature selection method to train a TSVM to reduce domain shift for mild cognitive impairment (MCI) prediction. Existing studies generally require a part of labeled target data, but it is very challenging to acquire labeled target samples. Besides, these methods usually assume that there is one single source domain and cannot well handle problems with multi-source domains.

In this paper, we propose a multi-source optimal transport (MSOT) framework for cross-domain AD diagnosis with multi-site MRIs. As shown in Fig. 1, we first project multi-source samples to target domain via optimal transport in an unsupervised manner. Based on the projected representation, we calculate the similarity/difference between each source domain and the target domain, and use this similarity as the weight of each source domain. We then train a support vector machine (SVM) classifier based on projected samples from each source domain. An ensemble learning strategy via weighted voting is used to predict labels of target samples.

The main contributions are listed as follows. *First*, we

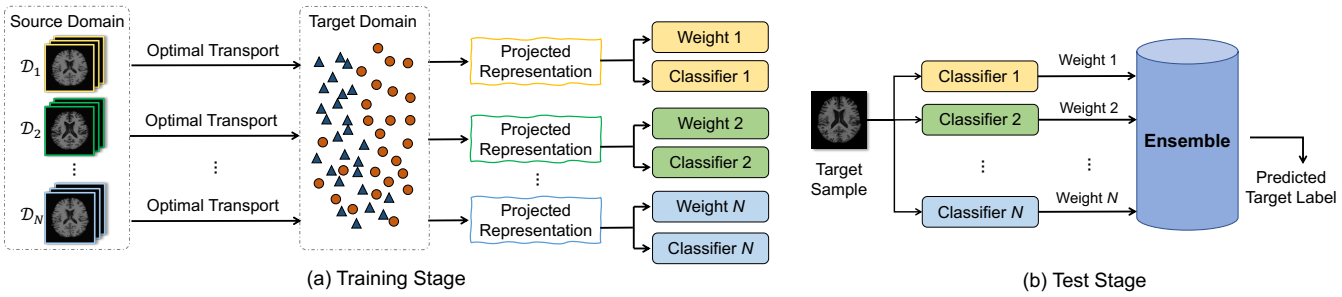


Fig. 1: Illustration of the proposed multi-source optimal transport (MSOT) framework for unsupervised domain adaptation and structural MRI-based brain dementia classification.

design an MSOT framework to adapt multi-site MRI data for AD diagnosis, through which the difference/similarity in data distribution between domains can be quantitatively measured. *Second*, our MSOT does not require labeled target data and can be optimized efficiently. When a new source domain/center is involved, we do not need to re-train models for existing multi-source domains. *Besides*, we conduct experiments on 3 benchmark MRI datasets and compare MSOT with several representative domain adaptation methods.

2. METHODOLOGY

2.1. Problem Setting

We study the problem of unsupervised multi-source domain adaptation for structural MRI-based brain dementia classification. Suppose there is a collection of N labeled source domains (*i.e.*, $\mathcal{D}_S = \{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_N\}$), and one unlabeled target domain (*i.e.*, $\mathcal{D}_T$). All the domains are defined on the same input space $\mathcal{X}$ but with unknown different marginal distributions. The goal is to utilize all the samples from multiple source domains to train a model that can precisely predict the labels of target samples.

2.2. Optimal Transport

Suppose Ω_1 and Ω_2 are two domains, and $\mu_s = \mathcal{P}(\Omega_1)$ and $\mu_t = \mathcal{P}(\Omega_2)$ are two probability measures for these two domains. Let $\mathbf{T}$ denote a mapping: $\Omega_1 \rightarrow \Omega_2$. $\mathbf{T}$ is said to be measure preserving if the measure of any subset S of Ω_1 remains the same after it has been mapped into Ω_2, *i.e.*, $\mu_s(\mathbf{T}^{-1}(S)) = \mu_t(S), \forall S \in \Omega_2$. Given a transportation/mapping cost function c: $\Omega_2 \times \Omega_2 \rightarrow \mathbb{R}_0^+$, the optimal transportation (OT) is defined as the measure preserving mapping $\mathbf{T}^* : \Omega_1 \rightarrow \Omega_2$ with the minimal total cost:

$$\mathbf{T}^* = \underset{\mathbf{T}}{\operatorname{argmin}} \int_{\Omega_1} c(\mathbf{x}, T(\mathbf{x})) d\mu(\mathbf{x}) \tag{1}$$

In the context of brain MR image analysis which is under discrete settings, given two MR image datasets/domains $\mathcal{D}_1$ and $\mathcal{D}_2$, the solution of an optimal transport is to find the measure preserving mapping $\mathcal{D}_1 \rightarrow \mathcal{D}_2$ that minimizes the cost on all the samples. In this work, we employ the Wasserstein distance (also known as Earth Mover Distance) to measure the similarity/difference between two MR image domains. Accordingly, the minimal cost for OT is defined as the following **Wasserstein distance** between $\mathcal{D}_1$ and $\mathcal{D}_2$:

$$\mathbf{W}(\mathcal{D}_1, \mathcal{D}_2) = inf \sum_{\mathbf{x} \in \mathcal{D}_1} c(\mathbf{x}, T(\mathbf{x})) \mu(\mathbf{x}) \tag{2}$$

2.3. Multi-Source Unsupervised Domain Adaptation

As shown in Fig. 1 (a), the MSOT contains 3 steps: 1) multi-source data projection via optimal transport, 2) multi-source domain weight learning, and 3) ensemble classification.

Stage 1): Multi-Source Data Projection. Given the collection of N source domains $\mathcal{D}_S = \{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_N\}$ and the unlabeled target domain $\mathcal{D}_T$, we first project each source domain into the target domain while keeping their conditional distributions. Let $\mathbf{OT}$ be the optimal transport operator. All data in the i-th ($i = 1, \cdots, N$) source domain $\mathcal{D}_i$ will be mapped into the target domain via $\mathcal{D}_i^* = \mathbf{OT}(\mathcal{D}_i)$, with the optimal solution defined in Eq. 1. In this way, the projected source data and target data will share the same distribution space, so they can be compared directly. Note that the projection is conducted in the feature space and the features are extracted from each MRI scan before projection.

Stage 2): Multi-Source Domain Weight Learning. With projected source data (*i.e.*, $\mathcal{D}_S^* = \{\mathcal{D}_1^*, \mathcal{D}_2^*, \cdots, \mathcal{D}_N^*\}$) and their ground-truth labels, we train N classifiers, with each one corresponding to a source domain.

In multi-source learning problems, some source domains may have more similar data distribution with the target domain, thus contributing more to the learning in the target domain. Therefore, we further propose to quantitatively measure the similarity/difference in data distribution between each source and the target domains. Given the i-th projected source domain $\mathcal{D}_i^*$ and target domain $\mathcal{D}_T$, their similarity can be defined as $I(\mathcal{D}_i^*, \mathcal{D}_T) = \frac{1}{W(\mathcal{D}_i^*, \mathcal{D}_T)}$, where $W(\cdot)$ is the Wasserstein distance. The domain weight for the i-th source

domain is calculated as:

$$\alpha_i = \frac{I(\mathcal{D}_i^*, \mathcal{D}_T)}{\sum_{j=1}^{N}(I(\mathcal{D}_j^*, \mathcal{D}_T))} \quad (3)$$

When a new source domain (*e.g.*, a new imaging center) is involved, we do not need to re-project existing multi-source data. This makes our method convenient to use in practice.

Stage 3): Ensemble Classification. As shown in Fig. 1 (b), given a new target sample, we first input its MRI feature into N SVM classifiers (trained on N source domains), and can obtain N probability scores. These scores are further combined via weighted voting based on the source domain weights (*i.e.*, $\{\alpha_i, \cdots, \alpha_N\}$) to output the predicted label for the target sample. Such an ensemble classification strategy takes advantage of the distribution similarity between each source domain and target domain, which helps reduce the negative impact of irrelevant or even noisy source domains.

2.4. Implementation

We use Python to implement the proposed MSOT framework. For the optimal transport mapping, we employ the Sinkhorn-Knopp algorithm [6] in the POT library[1]. The coefficient of the regularization term in the Sinkhorn algorithm is empirically set as 30 in the experiments. We use the linear SVM (with C = 10) as the classifier for classification.

3. EXPERIMENTS

3.1. Studied Subjects

We evaluate the proposed MSOT framework on three public benchmark neuroimaging datasets, including 1) Alzheimer's Disease Neuroimaging Initiative (ADNI-1) [7], 2) ADNI-2, and 3) ADNI-3. ADNI-1 contains 428 structural MR images (1.5T T1-weighted), with 199 AD and 229 cognitively normal (CN) subjects. ADNI-2 consists of 360 structural MR images (3T T1-weighted), with 159 AD and 201 CN subjects. As for ADNI-3, we use 120 structural MR images (3T T1-weighted), with 60 AD and 60 CN subjects. Since several subjects participated in both ADNI-1 and ADNI-2, we simply remove them from ADNI-2 for independent evaluation. Domain heterogeneity of these datasets comes from different scanning parameter (e.g., 1.5T vs. 3T) or updated scanners. All the brain MR images are preprocessed through a standard pipeline, including skull stripping, intensity correction, and spatial normalization. We utilize a 3D CNN [8] to extract MRI features for subjects in the three datasets, and obtain a 256-dimension feature vector for representing each subject.

3.2. Experimental Setup

With three datasets as independent domains, we conduct three groups of experiments. In each group, one dataset is treated as the target domain, while the other two are used as multi-source domains. Four metrics are used for performance evaluation, *i.e.*, accuracy (ACC), sensitivity (SEN), specificity (SPE), and the area under the ROC curve (AUC).

We compare our MSOT with a baseline method (denoted as **Baseline**) without domain adaptation, and two single-source adaptation methods, *i.e.*, **TCA** [9] and **CORAL** [10]. We also compare MSOT with its single-source variant (denoted as **MSOT-S**) to investigate the influence of our multi-source projection strategy. The competing methods share the same MRI features as the proposed MSOT method.

1. **Baseline.** We use the support vector machine (SVM), a widely used classifier in neuroimaging analysis, as the baseline model. A linear SVM (with $C = 10$) is trained on a combination of samples from all source domains, and then directly applied to the target domain. This method does not consider the domain shift problem.

2. **Single-Source Adaptation Methods.** In this group of methods, all multi-source samples are first combined to construct a new single-source domain. Then, a specific method (*e.g.*, TCA, CORAL or MSOT-S) are used to adapt the new single-source and the target domains. For the fair comparison, a linear SVM (with $C = 10$) is used in these three methods for classification.

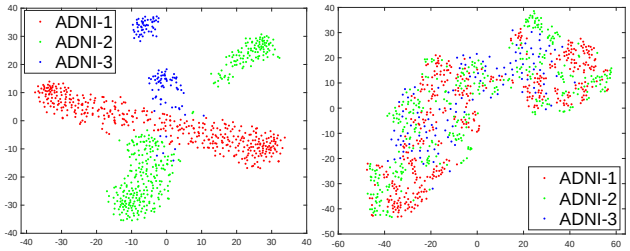
Table 1: Results (%) of cross-domain brain dementia diagnosis (*i.e.*, AD vs. CN classification) achieved by five different methods in three multi-source learning tasks.

Task (Source → Target Domains)	Method	ACC	SEN	SPE	AUC
ADNI-1+ADNI-2 → ADNI-3	Baseline	56.67	80.00	33.33	67.94
	TCA	70.83	50.00	**91.66**	76.72
	CORAL	63.33	80.00	46.67	70.83
	MSOT-S	**85.00**	**81.66**	88.33	**87.33**
	MSOT	**85.83**	**81.67**	**90.00**	**87.86**
ADNI-1+ADNI-3 → ADNI-2	Baseline	68.94	58.02	77.56	65.70
	TCA	70.93	46.91	**86.34**	74.04
	CORAL	70.02	70.99	69.27	74.72
	MSOT-S	**80.92**	**80.86**	80.97	**87.80**
	MSOT	**82.02**	**81.48**	**82.44**	**89.00**
ADNI-2+ADNI-3 → ADNI-1	Baseline	61.92	72.29	50.24	67.71
	TCA	71.55	**78.79**	63.41	78.16
	CORAL	68.35	72.19	64.93	74.42
	MSOT-S	**72.47**	74.63	**70.56**	**80.94**
	MSOT	**73.39**	**78.04**	**70.12**	**81.92**

3.3. Result and Analysis

We perform three cross-domain classification tasks, with the quantitative result reported in Table 1. From this table, we can have the following observations. *First*, compared by four domain adaptation methods (*i.e.*, TCA, CORAL, MSOT-S and MSOT), the Baseline method cannot achieve good performance when trained and tested in different domains. This implies that the domain shift poses a huge challenge to conventional machine learning methods, because they assume

[1] https://pythonot.github.io



(a) Before adaptation (b) After adaptation via MSOT

Fig. 2: Visualization of feature distributions of three domains.

that the source and target domains have the same data distribution. *Second*, single-source adaptation methods (*i.e.*, TCA, CORAl, and MSOT-S) achieve inferior overall performance, compared with our MSOT. This suggests that performing domain adaptation on the combination of all source domains (*i.e.*, ignoring their domain-specific characteristics) cannot well handle the domain shift problem. *Finally*, compared with four competing approaches, the proposed MSOT yields relative balanced prediction performance for both categories (i.e., AD and CN) in three cross-domain classification tasks.

We further use t-SNE [11] to visualize feature distributions of ADNI-1, ADNI-2 and ADNI-3 before and after domain adaptation via our MSOT method, with results shown in Fig. 2. From Fig. 2 (a), we can see that the original distribution difference between three datasets is large. After adaptation via MSOT, the distribution gap among three datasets is significantly reduced as shown in Fig. 2 (b). This clearly suggests the effectiveness of the proposed MSOT method for multi-source data adaptation.

4. CONCLUSION

In this paper, we propose a multi-source optimal transport (MSOT) framework for cross-domain diagnosis of brain dementia with multi-site MRI data. Our framework can project multi-source data into target domain via optimal transport, without requiring labeled target data for training. Based on projected source data, we construct multiple classifiers, and also compute the source domain weight by calculating the similarity between each source domain and target domain. An ensemble classification strategy is designed to estimate labels of target samples. Experimental results on three datasets suggest the superiority of MSOT in cross-domain AD diagnosis.

5. COMPLIANCE WITH ETHICAL STANDARDS

This is a numerical simulation study for which no ethical approval was required.

6. ACKNOWLEDGMENTS

This work was partly supported by NIH grants (Nos. AG041721 and MH117943). Data used in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI).

7. REFERENCES

[1] J. Quionero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, *Dataset Shift in Machine Learning*, The MIT Press, 2009.

[2] C. Wachinger and M. Reuter, "Domain adaptation for Alzheimer's disease diagnostics," *NeuroImage*, vol. 139, pp. 470–479, 2016.

[3] W. Li, Y. Zhao, X. Chen, Y. Xiao, and Y Qin, "Detecting Alzheimer's disease on small dataset: A knowledge transfer perspective," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1234–1242, 2019.

[4] E. Moradi, C. Gaser, H. Huttunen, and J. Tohka, "MRI based dementia classification using semi-supervised learning and domain adaptation," in *MICCAI Workshop Proceedings, Challange on Computer-Aided Diagnosis of Dementia, based on Structural MRI Data*, 2014.

[5] B. Cheng, M. Liu, D. Zhang, B. C. Munsell, and D. Shen, "Domain transfer learning for MCI conversion prediction," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1805–1817, 2015.

[6] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," in *NeurIPS*, 2013, pp. 2292–2300.

[7] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, 2008.

[8] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, "Residual and plain convolutional neural networks for 3D brain MRI classification," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2017, pp. 835–838.

[9] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2010.

[10] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," in *AAAI*, 2016, pp. 2058–2065.

[11] L. V. D. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. 11, pp. 2579–2605, 2008.