Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8797604, 14 pages
https://doi.org/10.1155/2022/8797604

Research Article

Deep Hash with Optimal Transport-Based Domain Adaptation for Multisite MRI Retrieval

Jingcheng Yang,[1,2] Qianqian Wang,[2,3] Tiwei Tao,[2] Sijie Niu,[1] and Mingxia Liu[2]

[1]*School of Information Science and Engineering, University of Jinan, Jinan 251000, China*
[2]*School of Information Science and Technology, Taishan University, Taian 271000, China*
[3]*School of Mathematics Science, Liaocheng University, Liaocheng, Shandong 252000, China*

Correspondence should be addressed to Sijie Niu; ise_niusj@ujn.edu.cn and Mingxia Liu; mxliu1226@gmail.com

Received 3 June 2022; Accepted 29 July 2022; Published 14 September 2022

Academic Editor: Geng Sun

The Internet of Things has a wide range of applications in the medical field. Due to the heterogeneity of medical data generated by different hospitals, it is very important to analyze and integrate data from different institutions. Functional magnetic resonance imaging (fMRI) is widely used in clinical medicine and cognitive neuroscience, while resting-state fMRI (rs-fMRI) can help reveal functional biomarkers of neurological disorders for computer-assisted clinical diagnosis and prognosis. Recently, how to retrieve similar images or case histories from large-scale medical image repositories acquired from multiple sites has attracted widespread attention in the field of intelligent diagnosis of diseases. Although using multisite data effectively helps increase the sample size, it also inevitably introduces the problem of data heterogeneity across sites. To address this problem, we propose a multisite fMRI retrieval (MSFR) method that uses a deep hashing approach and an optimal transport-based domain adaptation strategy to mitigate multisite data heterogeneity for accurate fMRI search. Specifically, for a given target domain site and multiple source domain sites, our approach uses a deep neural network to map the source and target domain data into the latent feature space and minimize their Wasserstein distance to reduce their distribution differences. We then use the source domain data to learn high-quality hash code through a global similarity metric, thereby improving the performance of cross-site fMRI retrieval. We evaluated our method on the publicly available Autism Brain Imaging Data Exchange (ABIDE) dataset. Experimental results show the effectiveness of our method in resting-state fMRI retrieval.

1. Introduction

With the rapid construction of digital hospitals, a large number of medical images are generated every day in hospitals. Fusion medical images from different hospitals and establishing an effective medical image retrieval system can reduce the workload of doctors to a certain extent, assist doctors in diagnosis, and strengthen cooperation and exchanges between different hospitals. Autism spectrum disorder (ASD) is a common neurodevelopmental disorder that usually occurs in early childhood. The main symptoms of this disease are social and verbal communication difficulties, narrow interests, stereotyped behavior, and impaired self-care ability [1]. According to the Centers for Disease Control and Prevention (CDC), one in every 59 American children was diagnosed with ASD in 2018, with the prevalence continuing to rise (https://www.cdc.gov/ncbddd/autism/data). The American Autism Association estimates that the lifetime treatment costs of ASD range from 3.5 to 5 million, bringing a heavy burden on patients and their families [2]. According to the World Health Organization, ASD has become one of the major diseases that seriously affect the quality of life and physical health. In order to auxiliary clinical diagnosis, it is important to develop an effective medical image retrieval system according to previous cases or medical images for better diagnosis and treatment of ASD. Several studies have already shown that neuroimaging can help find the imaging biomarkers and pathological changes in the brain, thus having great prospect in the diagnosis of ASD.

Magnetic resonance imaging (MRI) techniques have been widely used in the analysis of brain diseases such as ASD [3, 4], Alzheimer's disease [5–7], Parkinson's disease, and others [8–11]. Previous studies have revealed that behavioral and cognitive deficits in people with ASD are closely related to abnormal connectivity in brain networks, which include hyperconnectivity and underconnectivity [12–14]. Currently, the rs-fMRI technique is the best noninvasive tool to observe the changes of neural activities in brain networks. The most widely used method of rs-fMRI technology is the blood oxygen level-dependent (BOLD) method, which measures brain neural activities by detecting changes in blood flow of subjects without performing any specific task. Supervised learning usually has a strong assumption—independent and identically distributed, i.e., all training and test data are sampled independently from the same unknown data distribution. However, this is difficult to meet in most production practices. For multisite studies, there are differences in the distribution of data across these research institutions due to differences in the equipment used, parameters set, and demographic characteristics of the data collected by these institutions. Although many studies have demonstrated the effectiveness of rs-fMRI-based machine learning methods for automated diagnosis of ASD [15], most of them ignore the problem of fMRI heterogeneity across different imaging sites, which will significantly degrade generalization ability of models.

Conventional approaches using multisite data usually train data from each site separately and test them at different sites [16] or mix data from all sites [17]. These approaches clearly ignore the heterogeneity of data across sites and lead to poor model generalization. Other approaches of using multisite data take into account the data heterogeneity across different imaging sites and alleviate this problem by appropriate methods of domain adaptation [18, 19]. Typically, these approaches train the model on the source domain, perform domain adaptation on both source and target domain data, and finally test model using the target domain data. As shown in Figure 1, the model trained from source domain data may perform poorly in the target domain when the distribution difference between the source and target domain is significant. Thus, it is helpful to improve the model generalization by bringing the source and target domain distributions closer together through appropriate domain adaptation methods. Furthermore, these methods can be divided into semisupervised domain adaptation and unsupervised domain adaptation, which depends on whether the labels of the target domain can be used. Specifically, semisupervised domain adaptation approaches require labeling information of partial the target domain data besides the necessary labels of source domain data. However, labeled data is often labor-intensive, costly, and time-consuming. Therefore, unsupervised domain adaptation approaches are more widely used.

Existing studies on unsupervised domain adaptation methods can be broadly divided into two categories. Many approaches tend to use some predefined distribution divergence [20], such as maximum mean discrepancy (MMD) [21, 22], cosine similarity, Kullback–Leibler divergence,

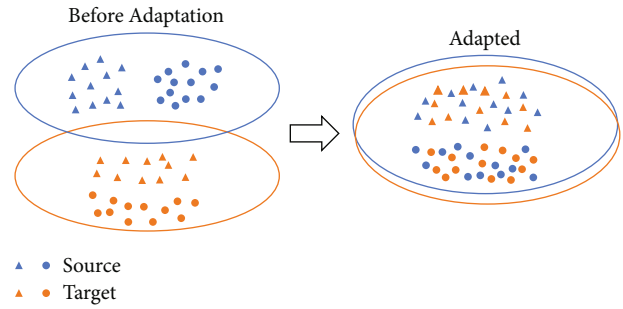

Figure 1: Motivation for performing domain adaptation. Blue indicates the source domain data distribution, orange indicates the target domain data distribution, and triangles and circles indicate samples from two categories. The left panel shows the large difference between the source and target domain distributions before domain adaptation, which can lead to poor model generalization performance. The right figure shows that after domain adaptation, the source and target domain distributions are significantly smaller than before, which is more conducive to obtaining a more generalized model.

mutual information, and higher-order moments. Many approaches are based on generative adversarial network (GAN), which learns domain-invariant features to confuse feature extractors and discriminators [23, 24]. Optimal transport (OT) is another popular approach [25, 26] that seeks a probabilistic coupling γ by minimizing the distribution between the source and target domains. The coupling γ is used to transform the source data through an estimated mapping (called barycentric mapping), which ultimately brings the source and target domain distributions closer together.

Most current retrieval works based on medical images such as MRI and CT data rely on natural image retrieval techniques. Due to the powerful nonlinear representational capabilities of deep learning and the advantages of hash codes in data storage and fast searching, deep hash learning for image retrieval has achieved promising performance in terms of retrieval accuracy and speed. Deep hashing methods can generally be classified into (1) unsupervised hashing [27] and (2) supervised hashing [28, 29] according to whether supervised information is used. Unsupervised hashing usually uses topological information and data distribution to learn hash functions but often requires longer hash code to obtain better retrieval accuracy. Supervised hashing can learn higher-quality hash functions through supervised information, so it is better than unsupervised hashing methods. Existing supervised hashing methods typically learn hashing functions from pairwise or triplet relations, which only captures data similarity locally. A recent paper proposed a global similarity metric method to compute a central similarity metric by introducing the concept of hash center, which encourages hash codes of similar sample pairs to approach the same hash center and those of dissimilar sample pairs to converge to different hash centers, thus improving hash learning efficiency and retrieval accuracy [28]. However, achieving true end-to-end deep hash training remains a difficult task since the use of sign function usually

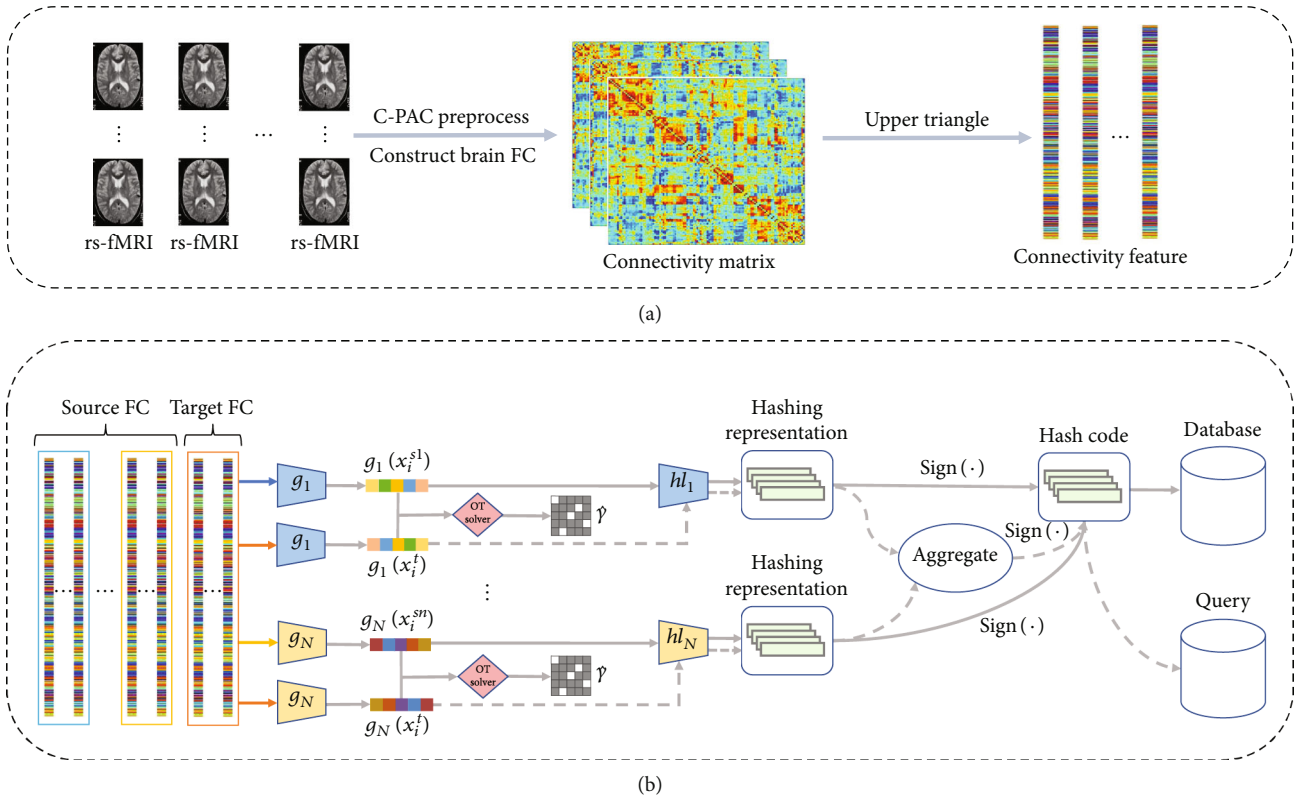

Figure 2: Illustration of the proposed multisite fMRI retrieval (MSFR) method for resting-state functional MRI (rs-fMRI) retrieval. It includes two major components: (a) generating functional connectivity features and (b) retrieval model learning. The raw rs-fMRI was first preprocessed using C-PAC, then the brain was divided into 116 brain regions using the AAL brain template, then the mean time series of each brain region were extracted, the Pearson correlation coefficient between the time series of pairwise brain regions was calculated to obtain the functional connectivity matrix, and finally, the upper triangular elements were retained to obtain the functional connectivity features. The proposed retrieval framework consists of submodels corresponding to the source domains, each of which includes domain adaptation based on optimal transport and hash learning based on a central similarity metric. Blue indicates the first source domain, yellow indicates the N-th source domain, and orange indicates the target domain. The inputs to each submodel are source and target domain data, where the neural network g assigned to each source domain is represented twice as a way of distinguishing between the source and target domain inputs. The labeled source domain data is used to learn the hash function, and the unlabeled target domain data is used only in the domain adaptation phase. The dotted line in the figure indicates that the target domain data only uses the hash coding layer hl in the testing stage. The retrieval database consists of a hash code obtained from the hash representation of the source domain in its corresponding hash coding layer, followed by the sign function. The query samples are obtained from the hash representation of the target domain data in each submodel and then processed by the aggregate module before being obtained by the sign function. The aggregate block is responsible for summing up the hash representations generated by each target domain sample in each hash coding layer to obtain the final hash representation of the sample.

lead to gradient vanishing problems. Thus, most approaches employ a relaxation scheme in place of strict discrete constraints [28–30], but these relaxation strategies usually lead to suboptimal hash codes due to quantization errors. Su et al. [31] designed a hash coding layer to learn hash codes directly, which uses a strictly sign function in forward propagation to maintain discrete constraints and intactly transmits the gradients to the front layer in backward propagation to avoid gradient vanishing.

In this article, we propose a multisite fMRI retrieval (MSFR) method for multisite fMRI retrieval. To achieve better retrieval results, the MSFR will firstly reduce the heterogeneity of rs-fMRI between different sites and then employ a deep hashing method for fMRI retrieval where a novel hash coding layer is introduced to use classification loss to assist in better hash code learning. As shown in Figure 2(b), we design a multiway parallel framework for domain adaptation retrieval. Specifically, we design a structurally identical deep neural network to each source domain. On this basis, the optimal transfer module is used to reduce the discrepancy between each source and target domain, while hash function learning is performed. Experimental results on rs-fMRI scans from the publicly Autism Brain Imaging Data Exchange (ABIDE) dataset demonstrate the effectiveness of the proposed method. To our knowledge, this is among the first studies to apply domain adaptation methods to multisite functional MRI retrieval.

Major contributions of this work can be summarized as follows: (1) We propose to reduce the marginal distribution between source domain sites and target domain sites through optimal transport theory to alleviate the problem of data heterogeneity between different sites. (2) We develop

a method to learn high-quality hash functions using a hashing method with a global similarity metric combined with a classification loss introduced by a novel hash coding layer. (3) We introduce a parallel framework using an inconsistency model to further improve multisite fMRI retrieval.

The rest of this paper is organized as follows. In Section 2, we review the relevant studies. In Section 3, we introduce the materials and present our proposed method. In Section 4, we present the experimental settings and experimental results. In Section 5, we perform ablation study, analyze the influence of two hyperparameters, and discuss the limitations of the current work. This paper is finally concluded in Section 6.

2. Related Works

Unlike the general natural image dataset, medical images are more difficult to collect due to their privacy and the high level of expertise required to support the various annotations. Data from multiple institutions are typically pooled together to form larger data sets and to facilitate research. However, multisite data introduces the problem of data heterogeneity, so domain adaptation has attracted extensive research in recent years in the diagnosis and analysis of disease in order to reduce data heterogeneity and improve model performance. For example, Wang et al. [32] proposed a multisite ASD detection algorithm. To eliminate intersite data heterogeneity, they first divided the multisite training data into two groups according to whether the sample was ASD or healthy control (HC). Then to learn the latent representations and cluster them within each group, similarity-based multiview linear reconstruction model was used. A nested singular value decomposition (SVD) method was designed to mitigate the heterogeneity of the data across sites. Zhang et al. [33] aligned feature and label distributions based on optimal transmission theory to reduce multisite heterogeneity in ASD diagnosis and achieve better classification performance. Many studies have demonstrated that domain adaptation can promote the efficiency of medical image analysis, so it is natural that for multisite learning tasks, domain adaptation becomes an essential component.

Medical image retrieval can help medical practitioners to quickly and accurately retrieve relevant data from the vast amount of medical images and to detect abnormalities in brain information in a timely manner, enabling automatic prediction and diagnosis of brain diseases in order to make effective preventive measures. There has been a lot of work done using medical images for retrieval. For example, Peng et al. [34] designed a retrieval model that combines triple loss with existing deep Cauchy hashing methods to accelerate nearest neighbor search in Hamming space and tested it on colorectal cancer (CRC) histology dataset.

3. Materials and Method

3.1. Materials and Image Preprocessing. In this work, we used rs-fMRI data from the International Autism Brain Imaging Data Exchange (ABIDE) repository (http://preprocessed-connectomes-project.org/abide/). It consists of 17 international sites containing 1112 subjects, including 539 subjects with ASD and 573 healthy controls (HCs). The dataset is a publicly available online project that provides imaging data from ASD and healthy control participants and information on their phenotypes [35]. Considering the limited amount of data at some sites, we selected 4 sites with a sample size of more than 50 among the 17 sites for the experiment, namely Leuven, NYU, UCLA, and UM. So we use the 4 imaging sites with 408 subjects, including 228 ASD patients and 180 HCs. Table 1 shows the demographic information and imaging parameters used in the four sites.

To improve the signal-noise ratio of rs-fMRI and to better extract blood oxygen level-dependent (BOLD) time series, in this work, we used rs-fMRI provided by the Preprocessing Connectome Project Program (http://preprocessed-connectomes-project.org/abide/) and preprocessed it using the Configurable Pipeline for the Analysis of Connectomes (C-PAC) [36]. The specific processes of this preprocessing pipeline include slice-timing, motion correction, nuisance signal regression, temporal filtering, and spatial normalization of Montreal Neurological Institute (MNI) templates. The brain was then divided into 116 regions of interest (ROIs) using anatomical automatic labeling (AAL) atlas [37]. Subsequently, for each sample, the mean BOLD time series of a set of brain regions were extracted, and the Pearson correlation coefficients between the individual ROIs were calculated, resulting in a symmetric 116×116 matrix, which is the resting-state functional connectivity matrix. To facilitate the use of this data as input to the model, we obtained the functional connectivity features by retaining only the elements on the upper triangle of the functional connectivity matrix and converting the retained triangles into a 6,670-dimensional feature vector for representing each subject. The construction process of the functional connectivity features is shown in Figure 2(a).

3.2. Proposed Method. The overall framework of our proposed MSFR is illustrated in Figure 2. It consists of two major parts: (a) constructing the functional connectivity matrix and extracting functional connectivity features for each subject and (b) learning the entire hash retrieval model and acquiring the retrieval database and query sample hash codes. The training of the hash retrieval model consists of two important processes: (1) reducing the marginal distribution discrepancy between source and target domain based on optimal transport theory to alleviate the problem of data heterogeneity between different sites and (2) learning high-quality hash functions using a hashing method with a global similarity metric combined with a classification loss introduced by a novel hash coding layer. More detailed descriptions are given below.

3.2.1. Optimal Transport-Based Multisite fMRI Adaptation. Given a dataset of N-labeled source domains $(X_s^k, Y_s^k)_{k=1}^N$, where $X_s^k = \{x_i^{sk}\}_{i=1}^{n_{sk}}$ denotes the functional connectivity features of all subjects and $Y_s^k = \{y_i^{sk}\}_{i=1}^{n_{sk}}$ is the category label associated with the k-th source domain data, the unlabeled target domain is represented as $X_t = \{X_j^t\}_{j=1}^{n_t}$, where n_t is the number of target subjects. In this work, we use a

Table 1: Demographic information and scanning parameters of four image sites in the ABIDE database. Ages are reported in terms of mean ± standard deviation; M/F: male/female; TR: time of repetition; TE: echo time.

Site	Healthy controls		Autism Spectrum disorder		Scanner	Voxel size (mm)	Flip angle (deg)	TR (ms)	TE (ms)	Bandwidth (HZ/Px)
	M/F	Age	M/F	Age						
Leuven	24/8	18.80 ± 9.00	21/4	13.10 ± 4.79	Philips (Intera)	3.59 × 3.59 × 4	90	—	33	—
NYU	79/14	16.49 ± 7.68	66/5	17.59 ± 7.84	Siemens Magnetom (Allegra)	3 × 3 × 4	90	2000	15	3906
UCLA	31/7	14.65 ± 4.97	28/8	16.27 ± 6.48	Siemens Magnetom (Trio Trim)	3 × 3 × 4	90	3000	28	2442
UM	56/9	17.35 ± 7.12	43/5	17.05 ± 8.36	GE (Signa)	3.438 × 3.438 × 3	90	2000	30	—

traditional assumption in unsupervised multisource domain adaptation, namely that the conditional probability distribution $Q_s(Y_s^k|X_s^k) = Q_t(Y_t|X_t)$, but the marginal probability distribution $P_s(X_s^k) \neq P_t(X_t)$ [38].

Recently, Damodaran et al. [39] combined deep learning with optimal transform- (OT-) based domain adaptation for a classification task, by learning a new representation for the source and target domains in the same distribution and preserve the discriminative ability of the classifier. DeepJDOT uses a deep convolutional neural network g to map the input to a latent feature space and then uses a classifier f to map the data in that space to a label space over the target domain. By jointly optimizing the feature space and label space to solve the coupling γ and the model g, f, DeepJDOT can achieve good performance on the target domain.

Inspired by this, our work draws on DeepJDOT's approach but differs in that we consider only the marginal distribution. This is because getting a robust hash function is more important in our retrieval task than getting a robust classification model. However, the alignment conditional distribution in DeepJDOT supposes that the target domain samples are predicted accurately enough, which requires a more robust classifier, which is different from our task. So on this basis, in each submodel, we embed the feature extractor g_k into the optimization of the optimal transport coupling γ^k between the marginal distributions p_s^k and p^t by aligning the marginal distributions.

So the loss function for our domain adaptation component between source domain k and target domain is defined as follows:

$$\mathscr{L}_{dot} = \sum_{i=1}^{n_{sk}} \sum_{j=1}^{n_t} \gamma_{ij}^k \left\| g_k\left(x_i^{sk}\right) - g_k\left(x_j^t\right) \right\|^2. \tag{1}$$

3.2.2. Central Similarity Metric-Based Hashing Learning. While existing supervised hashing methods typically capture data similarity only locally, we build on the central similarity metric proposed by Yuan et al. [28] to capture global similarity between data while using a novel hash coding layer to introduce classification loss to accelerate the convergence of the retrieval model.

Specifically, we define the hash centers as a set of points $C = \{c_1,\cdots,c_m\}$ in the L-dimensional Hamming space, where $c_i \in \{0,1\}^L$. Note that the hash centers have sufficient distance in Hamming space. Similar samples should be clustered into the same center, and different samples should be clustered into different centers so that we can aggregate similar samples together and separate dissimilar samples better. In the L-dimensional Hamming space, the average pairwise distance between hash centers satisfies:

$$\frac{1}{T}\sum_{i\neq j}^{m} D_H\left(c_i, c_j\right) \geq \frac{l}{2}, \tag{2}$$

where T denotes the number of different combinations of hash centers c_i and c_j, m is the number of hash centers, and D_H is the Hamming distance. For single-label data, a corresponding number of hash centers $\{c_1,\cdots,c_q\}$ are generated based on the number of categories q. Specifically, each bit of the hash center c_i is sampled from the Bernoulli distribution Bern (0.5). After that, each sample is assigned a hash center corresponding to its category to obtain the semantic hash center $C' = \{c_1', \cdots, c_N'\}$, where c_i' is the hash center of sample x_i.

Maximizing the logarithm posterior of the hash codes, the hashing learning objective can be obtained. Specifically, given the semantic hash center C', the logarithm maximum a posteriori probability estimation of the hash code $H = \{h_1,\cdots,h_N\}$ for training data can be obtained by maximizing $\log P(C'|H)P(H)$.

where $h_i \in \{0,1\}^L$, obtained by a hash function mapping the data from the input space to the L-dimensional Hamming space; $P(H)$ is the a priori distribution of hash codes; and $P(C'|H)$ is the likelihood function modelled as the Gibbs distribution. Then we obtain the central similarity loss $\mathscr{L}_{CS}$ for source domain k:

$$\mathscr{L}_{CS} = \frac{1}{L}\sum_{i}^{n_{sk}} \sum_{l\in L} \left(c_{ki,l}' \log h_{ki,l} + \left(1 - c_{ki,l}'\right) \log\left(1 - h_{ki,l}\right)\right), \tag{3}$$

where $c_{ki,l}'$ is the l-th bit of the hash center of the i-th sample in the k-th source domain. It is worth noting that h_{ki} in the training phase is not really a binary code. This is because the binary hash code is implemented with a sign function appended after the hash encoding layer, but since the use

of the sign function causes the gradient to vanish, this can make the optimization NP-hard. A common practice is to use the tanh function instead of the discrete constraint of the sign function. However, the use of a relaxation scheme generates quantization errors, so the bimodal Laplace prior proposed in deep hashing network for efficient similarity retrieval (DHN) [40] is introduced here for quantization, and since $\mathscr{L}_Q$ is a nonsmooth function and it is difficult to calculate the derivative, we use the smooth function cosh to replace it. The quantization loss $\mathscr{L}_Q$ is defined as

$$\mathscr{L}_Q = \sum_{i}^{n_{sk}} \sum_{l=1}^{L} \left(\log\cosh\left|2h_{ki,l} - 1\right| - 1\right). \tag{4}$$

The hashing method based on the central similarity metric will lead to quantization error, so we introduce a novel hash coding layer in Greedy Hash [31] and utilize the classification loss to assist in learning a better hash code. Specifically, the sign function (i.e., $b_{ki} = \text{sign}\ (h_{ki})$) is used strictly for the output of the hash coding layer in forward propagation, while the gradient is transmitted directly to the previous layer during backward propagation, effectively preventing the gradient from vanishing. Thus, the classification loss using cross-entropy in source domain k can be defined as

$$\mathscr{L}_{CLS} = \sum_{i=1}^{n_{sk}} \mathscr{L}_{CE}\left(y_i^{sk}, f_k(b_{ki})\right), \tag{5}$$

where $\mathscr{L}_{CE}$ is the cross-entropy and f_k is the classifier. We define the overall loss function as

$$\mathscr{L}_{total} = \alpha \mathscr{L}_{dot} + \mathscr{L}_{CS} + \mathscr{L}_Q + \beta \mathscr{L}_{CLS}. \tag{6}$$

α and β are parameters that control the contribution of OT domain adaptation loss and classification loss, respectively.

3.2.3. Alternating Optimization Algorithm. In the training phase, we randomly sample a batch n_b of samples in the target domain and one of the N source domains and assign a submodel to that source domain. Here, each submodel consists of a feature extractor g_k, a hash coding layer hl_k, and a classification layer f_k. The following optimization is then performed:

Step A: by fixing the parameters of g_k in the submodel, this domain adaptation part becomes a standard OT problem, and the coupling γ^k can be obtained by optimizing Equation (7) with the network simplex flow algorithm:

$$\min_{\gamma^k \in \prod(p_s, p_t)} \sum_{i,j=1}^{n_b} \gamma_{ij}^k \left\| g_k\left(x_i^{sk}\right) - g_k\left(x_j^t\right) \right\|^2 \tag{7}$$

Step B: we set θ to denote all parameters of model to be optimized. With a fixed γ^k obtained in the previous step, the submodel (g_k, hl_k, f_k) can be updated using stochastic gradient descent (SGD) by objective function:

$$\begin{aligned} \min_{\theta} \alpha & \sum_{i,j=1}^{n_b} \gamma_{ij}^k \left\| g_k\left(x_i^{sk}\right) - g_k\left(x_j^t\right) \right\|^2 \\ & + \frac{1}{L} \sum_{i=1}^{n_b} \left(c'_{ki,l} \log h_{ki,l} + \left(1 - c'_{ki,l}\right) \log\left(1 - h_{ki,l}\right) \right) \\ & + \sum_{i=1}^{n_b} \left(\log\cosh\left|2h_{ki,l} - 1\right| - 1\right) \\ & + \beta \sum_{i=1}^{n_b} \mathscr{L}_{CE}\left(y_i^{sk}, f_k(b_{ki})\right) \end{aligned} \tag{8}$$

3.2.4. Implementation. Our proposed MSFR model is implemented using the Python based on PyTorch. A corresponding number of submodules are assigned according to the number of source domains. Each submodel has exactly the same structure, including a feature extractor g_k, a hash coding layer hl_k, and a classification layer f_k. Here, g_k consists of two fully connected layers containing 1,024 and 128 neurons, respectively, each followed by a ReLU activation function. hl_k consists of a fully connected layer, and the number of neurons is determined by the length of the hash code (we set four sets of hash code lengths of 24 bits, 32 bits, 48 bits, and 64 bits), followed by a tanh function or a sign function according to the needs of the task. Specifically, the hash coding layer hl_k is followed by the tanh activation function when learning the global similarity of samples using the central similarity metric, and the hash coding layer hl_k is followed by the sign function when training the classifier directly using binary hash codes. Note that the sign function is only used in forward propagation, and the gradient is transmitted directly to the previous layer during backward propagation to prevent the gradient from vanishing. f_k consists of a fully connected layer of two neurons and a softmax layer. We set the batch size to 30, used the SGD optimizer and set the learning rate to 0.002, and fixed the hyperparameters α to 0.001 and β to 4.

4. Experiments

4.1. Experimental Setup. We used data from four sites in the ABIDE database, Leuven, NYU, UCLA, and UM, to evaluate the effectiveness of our method. In the training stage, we take turns selecting one site as the target domain, and the remaining three sites are used as the source domain. The test stage retrieval database consists of hash codes of all source domain data, and the query samples consist of hash codes of target domain data.

Following previous work [30, 41, 42], we use the precision recall curve (PR curve), precision curves with respect to different numbers of top returned samples, and mean average precision (mAP) as evaluation metrics to measure the retrieval effectiveness. The mAP is the widely used metric to measure the accuracy of the Hamming ranking protocol. The mAP is computed as the mean value of the average precision (AP) for all queries, and AP is computed as

$$\mathrm{AP} = \frac{\sum_{t=1}^{T} P(t)\delta(t)}{\sum_{t=1}^{T} \delta(t)}, \tag{9}$$

where T is the top-T search results returned and $P(t)$ denotes the precision of the t-th result returned. $\delta(t) = 1$ means that the t-th result queried is correct, and vice versa $\delta(t) = 0$. Here, the mAP value is calculated based on the first half returned neighbors for source domain sites. Namely, assuming that the number of samples in the database is N, then top $- T = N/2$.

4.2. Competing Methods. In our experiments, we compare our MSFR with the following 12 methods, including six deep hash methods and four traditional hash methods, as well as deep and traditional domain adaptation hash methods. The following is a brief description of each comparison method.

(1) DHN [40] used pairwise cross-entropy losses to ensure similarity learning between samples and quantization losses to control the quality of hash learning. To eliminate errors due to inner product instead of the Hamming distance, the authors used bimodal Laplacian prior. With the value $\{-1, 1\}$, the prior probability density is the largest, which means the hash function can learn the hash code of $\{-1, 1\}$ with maximum probability. ε is a diversity parameter of the bimodal Laplacian prior, and we use the default value of 10 for ε

(2) HashNet [43] improved on DHN in two main ways, firstly by balancing positive and negative sample pairs and secondly by using different β in the relaxation quantization phase so that the output keeps approaching 1, i.e., $\lim_{\beta \longrightarrow \infty} \tanh(\beta x)$

(3) LCDSH [44] pointed out that directly reducing the quantization error would change the feature distribution of the neural network which in turn changes the similarity between the query and retrieved images, so the authors proposed a locally constrained deep supervised hashing algorithm to solve this problem. Since λ is a trade-off parameter that balances the discriminability and the locality constrain, we use the default value of 3 for λ

(4) DCH [30] also used loss of control similarity and loss of control quantization error, differing by proposing the use of a sharply varying Cauchy distribution to convert distance into similarity, thus making the learned hash codes more discriminative. γ is an important Cauchy distribution scale parameter, used to control the balance between precision and recall, and we use the default value of 20 for γ

(5) QSMIH [45] uses the quadratic mutual information (QMI) in information-theoretic measures to learn hash codes. In order to meet the large-scale hash retrieval and further improve the retrieval accuracy, quadratic spherical mutual information (QSMI) is proposed on this basis. Since α is a parameter to force the learned hash code to be close to 1 or -1, we use the default value of 0.01 for α

(6) DTSH [46] used a triple-label likelihood function to learn hash codes, and maximizing the triple-label likelihood can make the query samples more similar to the positive samples and more distinguishable from the negative samples. Here, λ is a hyperparameter to balance the negative log triplet likelihood and the quantization error. And we use the default value of 1 for λ

(7) TAH [42] is a deep domain adaptation hashing algorithm that combines pairwise t-distribution cross-entropy loss to learn concentrated hash code and an adversarial network to align the data distribution between the source and target domains. Here, μ is the trade-off parameter between maximum A posterior loss and adversarial learning loss. The penalty of adversarial networks μ is increased from 0 to 1 gradually

(8) SH [47]. In this method, the process of encoding image feature vectors by SH can be viewed as a graph segmentation problem, where a relaxation solution to the graph segmentation problem can be provided with the help of the analysis of the eigenvalues and feature vectors of the Laplacian matrix of similar graphs

(9) ITQ [41]. In this method, the data in the original space is first downscaled using PCA, and then the data points in this dataset are mapped onto the vertices of a binary hypercube such that the corresponding quantization error is minimized, resulting in an excellent binary encoding for this dataset

(10) LFH [48] proposed a latent factor-based model that uses the Hamming distance to model the similarity between pairs of samples. To solve the time-consuming problem of finding hyperparameters in different datasets, specialized hyperparameters are automatically assigned based on the number of samples in the dataset and the number of similarity labels

(11) SDH [49] proposed a method to directly learn binary hash codes without relaxation, and the learning objective is to generate the best hash code for linear classification. Specifically, the training data is first mapped into the Hamming space, and then the transformed data is classified in this space. One of the key steps of the algorithm is to utilize discrete cyclic coordinate descent (DCC) to generate the hash code bit by bit, which solves the NP-hard binary optimization problem. Here, λ is the regularization parameter, and we use the default value of 1 for λ

(12) GTH [50] is a traditional unsupervised domain adaptation hashing algorithm, which starts from the perspective of the hash projection errors in the source and target domains and seeks the maximum likelihood estimate of the errors to reduce the domain discrepancy. At the same time, the hash projections of the source and target domains are optimized iteratively, with the two influencing each other to facilitate the final optimal state

The first seven competing methods (DHN, HashNet, LCDSH, DCH, QSMIH, DTSH, and TAH) are deep hashing approaches, and three methods (SH, ITQ, and GTH) are unsupervised traditional hashing methods, while two methods (LFH, and SDH) are supervised traditional hashing methods. In all deep hashing methods, to match the input to our data and to ensure fairness, we replace the convolutional neural network with one that is consistent with the structure of g_k, and the hash encoding layer is consistent with the setting of hl_k. All deep hash retrieval methods are supervised learning with labels, and the models all consist of two parts: the image feature extraction part and the hash coding layer that generates the hash codes.

The use of data (including preprocessing and functional connectivity feature extraction) in all comparison methods is consistent with our MSFR. In the training phase, we treat one site as the test set, and the remaining three sites are combined to be used as the training set. In the test phase, the retrieval database consists of hash codes of all source domain data, and the query samples consist of hash codes of the target domain data. We use the default values of hyperparameters in their corresponding articles for all comparison methods.

4.3. Results on ABIDE with Multisite rs-fMRI Data. We evaluate our MSFR method using data from four real sites in the ABIDE database, Leuven, NYU, UCLA, and UM, and compare it with other state-of-the-art hash retrieval methods. In all experiments, data from each site is used as target domain (test set) in turn, and data from the remaining sites are used as source domains (training set). The hash codes of the target domain data are used as the query set, and the hash codes of all source domain data are used as the retrieval database for test, with the aim of retrieving samples similar to the query samples in the database. The comparison methods include seven deep hashing algorithms (DHN, HashNet, LCDSH, DCH, QSMIH, DTSH, and TAH) and five traditional hashing algorithms (SH, ITQ, LFH, SDH, and GTH). The experimental results are shown in Table 2 and Figure 3.

As shown in Table 2, our method has the best mAP values for different bits on almost all target domain sites. Taking the hash code length of 48 bits as an example, when Leuven is used as the target domain, our MSFR (74.56%) is 7.78% higher than the best performance achieved by QSMIH (66.78%). When NYU is used as the target domain, our MSFR (69.25%) is 3.85% higher than the best performance of DCH (65.40%). Our MSFR (75.42%) outperforms the best comparison method (i.e., DCH with an mAP of 74.23%) by 1.19% when UCLA is used as the target domain, and our MSFR (75.45%) is 6.84% higher than the best performance yielded by TAH (68.61%) when UM is used as the target domain. This shows that our MSFR can be effective for cross-site retrieval. Furthermore, we can see that TAH has an advantage over other deep comparison methods on almost all target domain sites, especially when the target domain site is UM. This suggests that a proper domain adaptation method can be helpful for multisite fMRI retrieval.

Figure 3 shows that our method achieves the best average results with different hash code lengths when each of the four sites is used as a target domain for retrieval of other source domains. Specifically, compared to the deep hashing methods, our MSFR has a higher average mAP for different hash code lengths when retrieved as a query set against the source domain database at each of the four sites. Compared to the traditional hash retrieval methods, the average mAP of our MSFR is higher for different hash code lengths for each of the four sites as query sets retrieved from the source domain database. In addition, we can notice that the traditional hashing methods are generally less effective than deep hashing methods. The possible reason is that shallow hashing models may be not able to learn discriminative features and compact hash codes, while deep hashing methods can learn fMRI features and hash codes in an end-to-end manner.

In Figures 4 and 5, we further show the PR curves, recall curves, and precision curves (from left to right) with a hash code length of 48 bits with NYU and UM as the target domain, respectively, for retrieving the source domain samples and returning top n samples based on the Hamming distance ranking. Figures 4 and 5 show that our MSFR achieves almost best results in terms of PR curve (a), recall curve (b), and precision curve (c) compared to all other comparative methods. Specifically, from the PR curves in Figures 4(a) and 5(a), we can observe that MSFR basically achieves the highest precision at all recall levels. Besides, MSFR achieves higher precision at lower recall than other methods, which is important for accuracy-oriented medical image retrieval systems. Precision curves with respect to different numbers of top returned samples are shown in Figures 4(c) and 5(c). As can be seen, our MSFR achieves almost the best precision, especially when the number of samples retrieved is within 100. This indicates that our MSFR achieves more precise retrieval. And in medical image retrieval tasks, users tend to pay more attention to the top ranked samples retrieved, and MSFR has significantly better accuracy than other methods when the number of samples returned is small.

5. Discussion

5.1. Ablation Study. We introduced three variants (denoted as MSFR-1, MSFR-2, and MSFR-3) of MSFR for ablation experiments. Specifically, MSFR-1 denotes a hash approach that retains only the central similarity metric, MSFR-2 denotes a hash approach that retains only the central similarity metric and classification loss, and MSFR-3 denotes a hash approach that retains only the central similarity metric

Table 2: The mAP results achieved by different methods with different bits of hash code in retrieving data from the four sites of the ABIDE database. Data from each site are used in turn as the target domain (test set and query set) and data from the remaining sites are used as the source domain (training set and retrieval database).

2*method	Leuven (target domain)				NYU (target domain)				UCLA (target domain)				UM (target domain)			
	24 bits	32 bits	48 bits	64 bits	24 bits	32 bits	48 bits	64 bits	24 bits	32 bits	48 bits	64 bits	24 bits	32 bits	48 bits	64 bits
DHN	0.6362	0.6360	0.6349	0.6524	0.6423	0.6773	0.6240	0.7058	0.7129	0.7409	0.7129	0.7139	0.6528	0.6758	0.6274	0.6706
HashNet	0.6485	0.6061	0.6369	0.6346	0.6453	0.6310	0.6118	0.6567	0.7039	0.7441	0.7264	0.7400	0.6455	0.6205	0.6012	0.6316
LCDSH	0.6311	0.5984	0.6609	0.6630	0.5964	0.5562	0.6352	0.5999	0.7151	0.7256	0.7127	0.7112	0.5849	0.5762	0.6481	0.6375
DCH	0.6337	0.6419	0.6502	0.6548	0.6395	0.6614	0.6540	0.5996	0.6529	0.6727	0.7423	0.6820	0.5976	0.6222	0.6365	0.6596
QSMIH	0.6576	0.6350	0.6678	0.6894	0.6413	0.6644	0.6234	0.6129	0.6627	0.7017	0.7412	0.7396	0.5897	0.6254	0.6473	0.5841
DTSH	0.6349	0.6360	0.6349	0.6670	0.6474	0.6649	0.5989	0.6472	0.7774	0.6883	0.7189	0.7379	0.6833	0.6003	0.5828	0.5891
TAH	0.6486	0.6918	0.6118	0.6478	0.6738	0.6614	0.6260	0.6511	0.6669	0.6625	0.6753	0.7034	0.6779	0.6877	0.6861	0.7021
SH	0.5343	0.5291	0.5352	0.5231	0.5253	0.5378	0.5404	0.5346	0.5240	0.5216	0.5315	0.5266	0.5377	0.5329	0.5404	0.5303
ITQ	0.5421	0.5504	0.5458	0.5553	0.5476	0.5366	0.5465	0.5512	0.5369	0.5417	0.5383	0.5421	0.5396	0.5542	0.5439	0.5297
LFH	0.4867	0.4867	0.4867	0.4867	0.5787	0.5787	0.5787	0.5787	0.6085	0.6085	0.6085	0.6085	0.6025	0.6025	0.6025	0.6025
SDH	0.5789	0.5707	0.5800	0.6039	0.5854	0.5765	0.5967	0.5830	0.5539	0.5254	0.5154	0.5214	0.6520	0.6249	0.5402	0.5941
GTH	0.5255	0.5312	0.5324	0.5339	0.5415	0.5406	0.5367	0.5391	0.5321	0.5300	0.5288	0.5353	0.5304	0.5321	0.5285	0.5374
MSFR (ours)	0.7027	0.7188	0.7456	0.7054	0.6930	0.6837	0.6925	0.6880	0.7684	0.7269	0.7542	0.7236	0.6924	0.7220	0.7545	0.7209

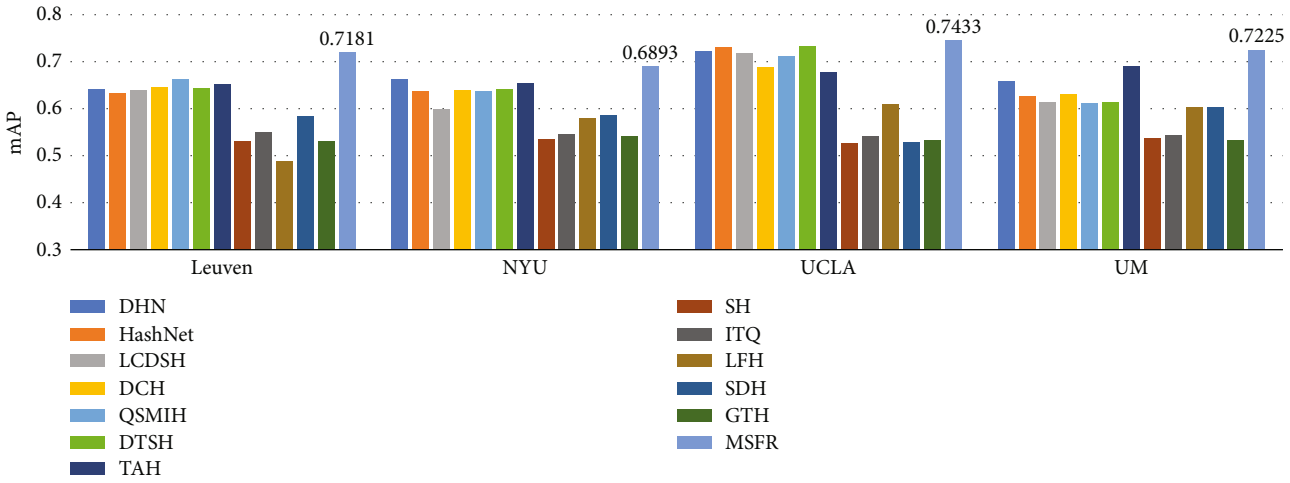

Figure 3: Average mAP results of different methods with different hash code lengths, where each of the four sites, Leuven, NYU, UCLA, and UM, is used as a query set for retrieval from the source domain data.

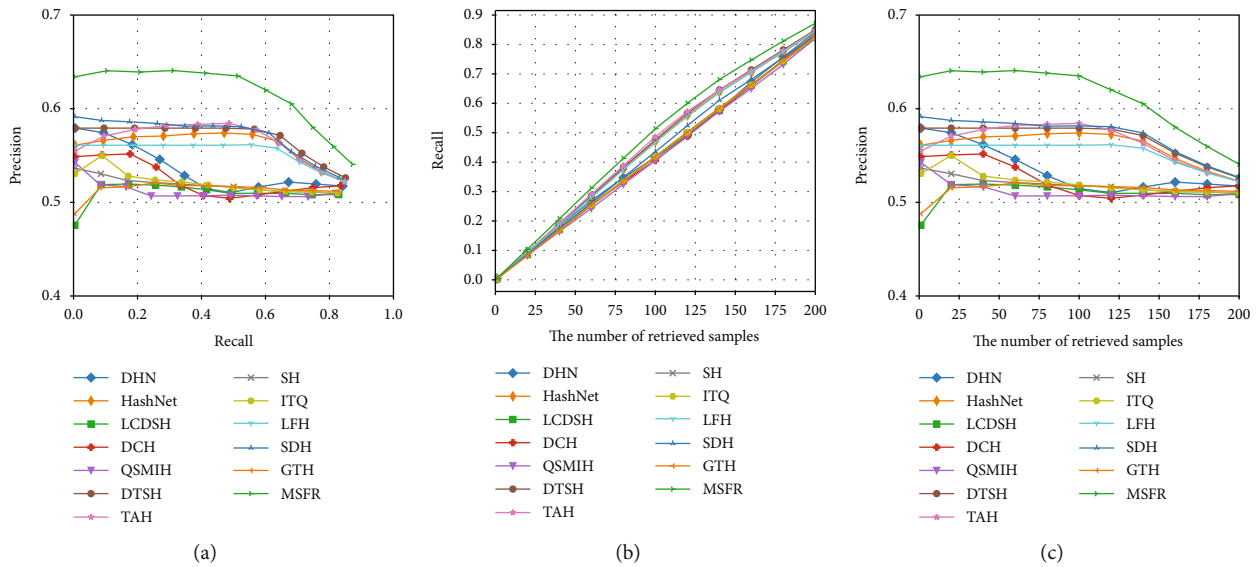

Figure 4: PR curves, recall curves, and precision curves (from left to right) for different methods that retrieve the source domain samples and return top n samples based on Hamming distance ranking when the target domain is NYU and the hash code length is 48 bits.

and OT domain adaptation approach. In Table 3, we report the mAP results achieved by the MSFR and its three variants with different hash code lengths when the data from the four sites of the ABIDE database were retrieved. As shown in Table 3, we can observe that MSFR generally outperformed the other three variants, which also implies that our optimal transport domain adaptation loss and classification loss both contribute to the model performance. Specifically, the performance of MSFR-2 is generally higher than MSFR-1 when the target domain is Leuven, NYU, and UCLA, and there is a sharp drop in performance only when UM is the target domain and the hash code length is 48 bits and 64 bits. And the results of MSFR-1 are 2.23%, 10.42%, and 6.75% lower than MSFR-2 at 24 bits, 48 bits, and 64 bits, respectively, probably due to the large data heterogeneity. When we deal with the data heterogeneity problem for MSFR-1 and MSFR-2, that is, MSFR-3 and MSFR, we find that the results for MSFR-1 and MSFR-2 are significantly improved in all target domains. Besides, we can observe that MSFR-3 and MSFR usually outperform MSFR-1 and MSFR-2 in most cases, implying that OT-based domain adaptation helps improve the retrieval performance. With UM as the target domain, the MSFR result is 5.18% higher than MSFR-3 for a hash code length of 48 bits, and the gap between MSFR and MSFR-3 is reduced to 1.56% and 4.78% for hash code lengths of 24 bits and 64 bits, respectively.

We also conducted experiments on single-source domain, with the UM site as the target domain and the hash code length set to 48 bits. The experimental results are shown in Figure 6, from which we can observe that our approach achieves the best overall results. By comparing MSFR-2 and MSFR, it can be observed that domain adaptation for both source and target domains can significantly improve the retrieval results, which further illustrates the

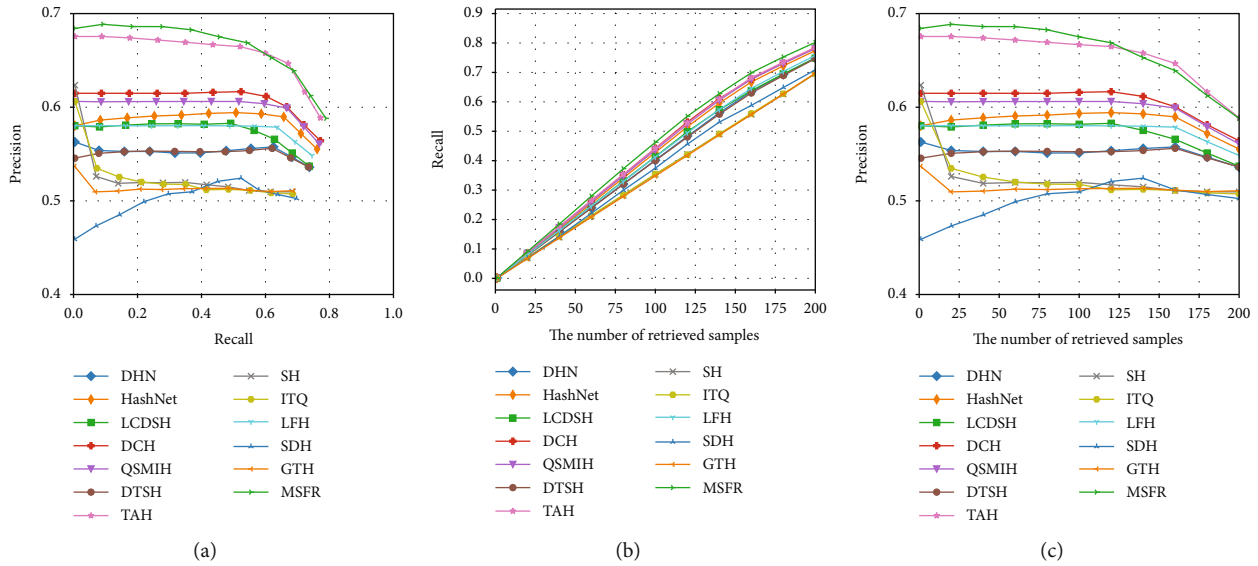

Figure 5: PR curves, recall curves, and precision curves (from left to right) for different methods that retrieve the source domain samples and return top n samples based on Hamming distance ranking when the target domain is UM and the hash code length is 48 bits.

Table 3: The mAP results of three variants of MSFR with different hash code lengths when each of the four sites of the ABIDE database is treated as target domain (test and query data) for retrieval.

2*method	Leuven (target domain)				NYU (target domain)				UCLA (target domain)				UM (target domain)			
	24 bits	32 bits	48 bits	64 bits	24 bits	32 bits	48 bits	64 bits	24 bits	32 bits	48 bits	64 bits	24 bits	32 bits	48 bits	64 bits
MSFR-1	0.6644	0.6422	0.6338	0.6678	0.6156	0.6356	0.6144	0.6282	0.6898	0.6588	0.7035	0.6802	0.6191	0.5915	0.6639	0.6231
MSFR-2	0.6524	0.6850	0.6807	0.6678	0.6448	0.6372	0.6318	0.6180	0.7139	0.6990	0.6879	0.6598	0.5968	0.5940	0.5597	0.5556
MSFR-3	0.7323	0.6746	0.7011	0.6395	0.6793	0.6974	0.7042	0.6505	0.7028	0.6868	0.7385	0.7524	0.7080	0.7544	0.7027	0.7687
MSFR (ours)	0.7027	0.7188	0.7456	0.7054	0.6930	0.6837	0.6925	0.6880	0.7684	0.7269	0.7542	0.7236	0.6924	0.7220	0.7545	0.7209

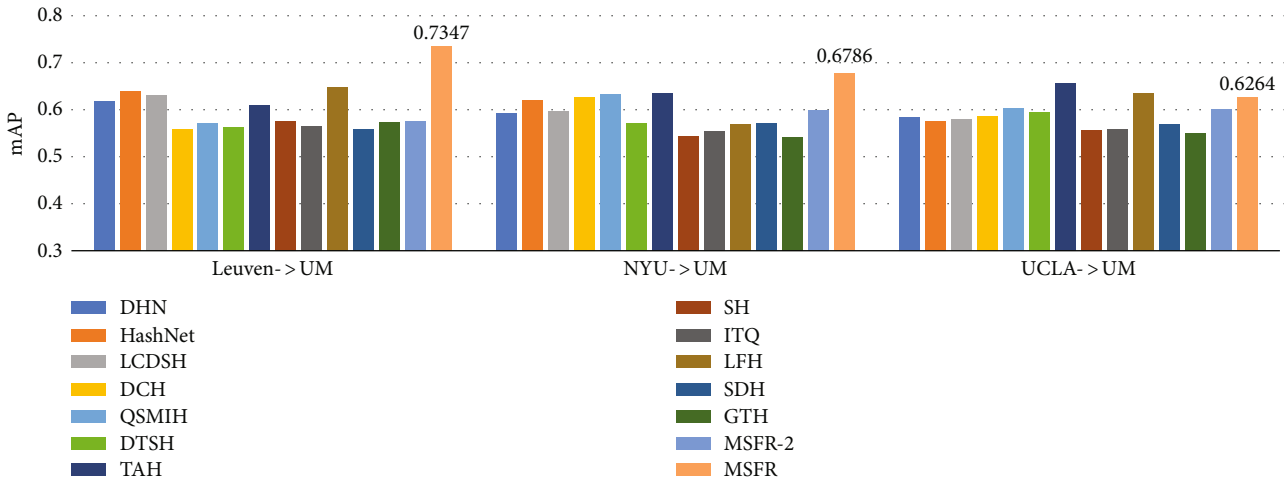

Figure 6: Single-source domain retrieval mAP for different methods with hash code length of 48, with UM as the target domain (query set) and Leuven, NYU, and UCLA as the source domains (retrieval database), respectively. Experimental results for MSFR and MSFR-2 are obtained from the submodels corresponding to the source domains.

effectiveness of our domain adaptation method in alleviating the data heterogeneity between different sites. Thus, the proposed fMRI retrieval framework is expected to be applied to multisource domains to reduce data heterogeneity across sites and improve retrieval performance.

5.2. Parameter Analysis. We further analyzed the influence of two hyperparameters in the proposed method. Figure 7 reports the mAP results of our method using different values of the hyperparameter α (for the domain adaptation component) and the hyperparameter β (for the classification loss).

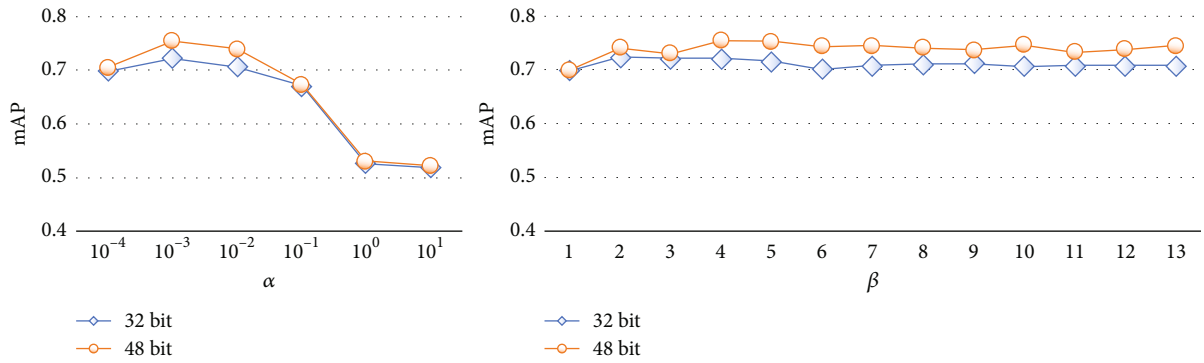

Figure 7: Influence of two hyperparameters on the mAP results of the proposed method.

Figure 7(a) shows that our MSFR can yield stable results when $\alpha \in [10^{-4}, 10^{-1}]$ but cannot generate good results when $\alpha > 1$. It can be seen from Figure 7(b) that the results of MSFR fluctuate within a small range of [0.70,0.75] with different values of β, suggesting that our method is not very sensitive to the hyperparameter β.

5.3. Limitations and Future Work. This work still has some limitations to be addressed in the future. First, to address the data heterogeneity of multiple source and target domains, we designed a multiway parallel framework based on the number of source domains. This means that the larger the number of source domains, the larger the number of submodels, which may ultimately lead to higher overall complexity and training costs. Therefore, a unified model needs to be investigated in the future to solve the problem of high computational cost under a large number of source domains. Second, our model only considers data heterogeneity in the source and target domains, ignoring the fact that there is also data heterogeneity between different source domains. It is interesting to employ multisource domain adaptation strategies to further promote the performance of our method. Third, there is insufficient exploitation of the constructed brain functional connectivity data. In general, each brain network contains not only node features but also topological information between different nodes, and the high-level topological information of brain networks cannot be captured using traditional neural networks alone. As graph convolutional networks (GCN) can automatically learn node features and topological information between nodes, we will design GCN-based models to fully exploit the information in fMRI data, which will also be our future work.

6. Conclusion

In this paper, we proposed a multisite fMRI retrieval (MSFR) method that uses a deep hashing approach and a domain adaptation strategy to mitigate multisite data heterogeneity for accurate fMRI search. For a given target domain site and multiple source domain sites, our MSFR used a deep neural network to map the source and target domain data into the latent feature space and minimize their Wasserstein distance to reduce their distribution differences. We then used the source domain data to learn high-quality hash code through a global similarity metric, thereby improving the performance of cross-site fMRI retrieval. We validated the MSFR method on a real ASD multisite dataset, with results demonstrating its effectiveness in rs-fMRI retrieval.

Data Availability

The dataset used in this work can be found on the public Autism Brain Imaging Data Exchange (ABIDE) website (http://preprocessed-connectomes-project.org/abide/). The investigators within the ABIDE contributed to the design and implementation of ABIDE and provided data but did not participate in the analysis or writing of this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

J. Yang, Q. Wang, and T. Tao were supported in part by the Taishan Scholar Program of Shandong Province and National Natural Science Foundation of China (No. 62176112). J. Yang and S. Niu were supported in part by the Development Program Project of Youth Innovation Team of Institutions of Higher Learning in Shandong Province.

References

[1] E. Anagnostou and M. J. Taylor, "Review of neuroimaging in autism spectrum disorders: what have we learned and where we go from here," *Molecular Autism*, vol. 2, no. 1, pp. 1–9, 2011.

[2] F. R. Volkmar and L. A. Wiesner, *Essential clinical guide to understanding and treating autism*, John Wiley Sons, 2017.

[3] Z. Yao, B. Hu, Y. Xie et al., "Resting-state time-varying analysis reveals aberrant variations of functional connectivity in autism," *Frontiers in Human Neuroscience*, vol. 10, p. 463, 2016.

[4] Y. Chu, G. Wang, L. Cao, L. Qiao, and M. Liu, "Multi-scale graph representation learning for autism identification with

functional MRI," *Frontiers in Neuroinformatics*, vol. 15, p. 802305, 2022.

[5] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67–77, 2010.

[6] W. Lieb, A. S. Beiser, R. S. Vasan et al., "Association of plasma leptin levels with incident Alzheimer disease and MRI measures of brain aging," *Jama*, vol. 302, no. 23, pp. 2565–2572, 2009.

[7] M. Liu, D. Cheng, K. Wang, and Y. Wang, "Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis," *Neuroinformatics*, vol. 16, no. 3, pp. 295–308, 2018.

[8] C. McDaniel and S. Quinn, "Developing a graph convolution-based analysis pipeline for multi-modal neuroimage data: an application to Parkinson's disease," in *Proceedings of the Python in Science Conference*, pp. 42–49, Austin, Texas, USA, 2019.

[9] X. Zhang, L. He, K. Chen, Y. Luo, J. Zhou, and F. Wang, "Multi-view graph convolutional network and its applications on neuroimage analysis for Parkinson's disease," *AMIA Annual Symposium Proceedings*, vol. 2018, p. 1147, 2018.

[10] L. Zhang, M. Wang, M. Liu, and D. Zhang, "A survey on deep learning for neuroimaging-based brain disorder analysis," *Frontiers in Neuroscience*, vol. 14, p. 779, 2020.

[11] M. Liu and D. Zhang, "Sparsity score: a novel graph-preserving feature selection method," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 28, no. 4, p. 1450009, 2014.

[12] R.-A. M. Â. Uller, P. Shih, B. Keehn, J. R. Deyoe, K. M. Leyden, and D. K. Shukla, "Underconnected, but how? A survey of functional connectivity MRI studies in autism spectrum disorders," *Cerebral Cortex*, vol. 21, no. 10, pp. 2233–2243, 2011.

[13] M. E. Vissers, M. X. Cohen, and H. M. Geurts, "Brain connectivity and high functioning autism: a promising path of research that needs refined models, methodological convergence, and stronger behavioral links," *Neuroscience Biobehavioral Reviews*, vol. 36, no. 1, pp. 604–625, 2012.

[14] L. Q. Uddin, K. Supekar, and V. Menon, "Reconceptualizing functional brain connectivity in autism from a developmental perspective," *Frontiers in Human Neuroscience*, vol. 7, p. 458, 2013.

[15] H. Chen, X. Duan, F. Liu et al., "Multivariate classification of autism spectrum disorder using frequency-specific resting-state functional connectivity-a multicenter study," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 64, pp. 1–9, 2016.

[16] S. I. Ktena, S. Parisot, E. Ferrante et al., "Metric learning with spectral graph convolutions on brain connectivity networks," *Neuro Image*, vol. 169, pp. 431–442, 2018.

[17] A. Abraham, M. P. Milham, A. Di Martino et al., "Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example," *Neuro Image*, vol. 147, pp. 736–745, 2017.

[18] W. M. Kouw and M. Loog, "A review of domain adaptation without target labels," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 3, pp. 766–785, 2019.

[19] H. Guan and M. Liu, "Domain adaptation for medical image analysis: a survey," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 3, pp. 1173–1185, 2022.

[20] C. Yu, J. Wang, C. Liu et al., "Learning to match distributions for domain adaptation," 2020, http://arxiv.org/abs/2007.10791.

[21] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. S. Â. Olkopf, and A. Smola, "A kernel two-sample test," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 723–773, 2012.

[22] M. Long, Y. Cao, J. Wang, and M. Jordan, "Learning transferable features with deep adaptation networks," in *International Conference on Machine Learning. PMLR*, pp. 97–105, Lille, France, 2015.

[23] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7167–7176, Honolulu, 2017.

[24] Y. Zhang, T. Liu, M. Long, and M. Jordan, "Bridging theory and algorithm for domain adaptation," in *International Conference on Machine Learning. PMLR*, pp. 7404–7413, Baltimore, Maryland, USA, 2019.

[25] N. Courty, R. Flamary, A. Habrard, and A. Rakotomamonjy, "Joint distribution optimal transportation for domain adaptation," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[26] R. Flamary, N. Courty, D. Tuia, and A. Rakotomamonjy, "Optimal transport for domain adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, 2016.

[27] E. Yang, C. Deng, T. Liu, W. Liu, and D. Tao, "Semantic structure-based unsupervised deep hashing," in *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pp. 1064–1070, Stockholm, Sweden, 2018.

[28] L. Yuan, T. Wang, X. Zhang et al., "Central similarity quantization for efficient image and video retrieval," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3083–3092, Seattle, USA, 2020.

[29] H. Liu, R. Wang, S. Shan, and X. Chen, "Deep supervised hashing for fast image retrieval," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2064–2072, Las Vegas, Nevada, USA, 2016.

[30] Y. Cao, M. Long, B. Liu, and J. Wang, "Deep cauchy hashing for hamming space retrieval," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1229–1237, Salt Lake City, Utah, USA, 2018.

[31] S. Su, C. Zhang, K. Han, and Y. Tian, "Greedy hash: towards fast optimization for accurate hash coding in CNN," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[32] N. Wang, D. Yao, L. Ma, and M. Liu, "Multi-site clustering and nested feature extraction for identifying autism spectrum disorder with restingstate fMRI," *Medical Image Analysis*, vol. 75, article 102279, 2022.

[33] J. Zhang, P. Wan, and D. Zhang, "Transport-based joint distribution alignment for multi-site autism spectrum disorder diagnosis using resting-state fMRI," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Lecture Notes in Computer Science, pp. 444–453, Springer, Cham, 2020.

[34] T. Peng, M. Boxberg, W. Weichert, N. Navab, and C. Marr, "Multi-task learning of a deep k-nearest neighbour network for histopathological image classification and retrieval," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.Medical Image Computing*

and Computer Assisted Intervention – MICCAI 2019, Lecture Notes in Computer Science, pp. 676–684, Springer, Cham, 2019.

[35] A. Di Martino, C.-G. Yan, Q. Li et al., "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism," *Molecular Psychiatry*, vol. 19, no. 6, pp. 659–667, 2014.

[36] C. Craddock, S. Sikka, B. Cheung et al., "Towards automated analysis of connectomes: the configurable pipeline for the analysis of connectomes (C-PAC)," *Frontiers Neuroinformatics*, vol. 42, pp. 10–3389, 2013.

[37] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou et al., "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuro Image*, vol. 15, no. 1, pp. 273–289, 2002.

[38] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2066–2073, Providence, RI, USA, 2012.

[39] B. B. Damodaran, B. Kellenberger, R. Flamary, D. Tuia, and N. Courty, "Deepjdot: deep joint distribution optimal transport for unsupervised domain adaptation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 447–463, Munich, Germany, 2018.

[40] H. Zhu, M. Long, J. Wang, and Y. Cao, "Deep hashing network for efficient similarity retrieval," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016.

[41] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, "Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, pp. 2916–2929, 2012.

[42] Z. Cao, M. Long, C. Huang, and J. Wang, "Transfer adversarial hashing for hamming space retrieval," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[43] Z. Cao, M. Long, J. Wang, and P. S. Yu, "Hashnet: deep learning to hash by continuation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5608–5617, Venice, Italy, 2017.

[44] H. Zhu and S. Gao, "Locality constrained deep supervised hashing for image retrieval," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*, pp. 3567–3573, Melbourne, Australia, 2017.

[45] N. Passalis and A. Tefas, "Deep supervised hashing using quadratic spherical mutual information for efficient image retrieval," *Signal Processing: Image Communication*, vol. 93, article 116146, 2021.

[46] X. Wang, Y. Shi, and K. M. Kitani, "Deep supervised hashing with triplet labels," in *Computer Vision ACCV 2016*, pp. 70–84, Springer, 2016.

[47] Y. Weiss, A. Torralba, and R. Fergus, "Spectral hashing," *Advances in Neural Information Processing Systems*, vol. 21, 2008.

[48] P. Zhang, W. Zhang, W.-J. Li, and M. Guo, "Supervised hashing with latent factor models," in *Proceedings of the 37th International ACM SIGIR Conference on Research Development in Information Retrieval*, pp. 173–182, New York, 2014.

[49] F. Shen, C. Shen, W. Liu, and H. Tao Shen, "Supervised discrete hashing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 37–45, Boston, Massachusetts, USA, 2015.

[50] J. Liu and L. Zhang, "Optimal projection guided transfer hashing for image retrieval," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 8754–8761, 2019.