A Transfer Model Based on Supervised Multi-Layer Dictionary Learning for Brain Tumor MRI Image Recognition

Yi Gu* and Kang Li

School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, China

Artificial intelligence (AI) is an effective technology for automatic brain tumor MRI image recognition. The training of an AI model requires a large number of labeled data, but medical data needs to be labeled by professional clinicians, which makes data collection complex and expensive. Moreover, a traditional AI model requires that the training data and test data must follow the independent and identically distributed. To solve this problem, we propose a transfer model based on supervised multi-layer dictionary learning (TSMDL) for brain tumor MRI image recognition in this paper. With the help of the knowledge learned from related domains, the goal of this model is to solve the task of transfer learning where the target domain has only a small number of labeled samples. Based on the framework of multi-layer dictionary learning, the proposed model learns the common shared dictionary of source and target domains in each layer to explore the intrinsic connections and shared information between different domains. At the same time, by making full use of the label information of samples, the Laplacian regularization term is introduced to make the dictionary coding of similar samples as close as possible and the dictionary coding of different class samples as different as possible. The recognition experiments on brain MRI image datasets REMBRANDT and Figshare show that the model performs better than competitive state of-the-art methods.

Keywords: brain tumor MRI image, supervised learning, transfer learning, Laplacian regularization, multi-layer dictionary learning

INTRODUCTION

Brain tumor is a common neurological disease. As a high incidence disease, its incidence rate has reached 1.34 per 100,000 in China, and over 200,000 patients diagnosed with primary or metastatic brain tumors in the United States every year. Among the incidence of systemic tumors, brain tumors are second only to those of the stomach, uterus, breast, and esophagus, accounting for approximately 2% of systemic tumors and the proportion of deaths has exceeded 2% (Sun et al., 2019; Sung et al., 2021). According to surveys, the incidence rate of brain tumors is the highest among children, and the highest incidence is 20–50-year-old young adults. Among childhood malignancies, brain tumors are the second most common, after leukemia. Brain tumors not only cause physical and mental suffering to patients, but also place a heavy financial burden on their families. As a standard technique for non-invasive brain tumor diagnosis, magnetic resonance

imaging (MRI) is an essential component of medical diagnosis and treatment. It uses magnetic resonance phenomena to obtain electromagnetic signals from the brain, so as to reconstruct brain information and provide a validated anatomical image of the brain. MRI can increase the diagnostic ability of medical diagnosticians. The wide application of MRI mainly benefits from the following characteristics (Amin et al., 2017; Bahadure et al., 2017): (1) no bony artifacts, good soft tissue resolution and clear visualization of soft tissue structures; (2) ability to image multiple aspects and multiple parameters, facilitating the acquisition of diagnostic information as a means of determining the various characteristics of the lesion; (3) no radiological damage and no ionizing radiation damage; (4) different profiles can be selected by adjusting the magnetic field, resulting in a three-dimensional image with different angles, which facilitates the identification of the lesion site; (5) has a flow-space effect and does not require an external contrast agent, allowing direct visualization of the vascular structure and facilitating the observation of the relationship between the vessel and the lesion. However, it is time consuming for radiologists to interpret the large number of MRI images and detect early brain tumors. These medical images need to be analyzed by doctors one by one, and the condition should be determined according to their experience.

Artificial intelligence (AI) technology, especially in particular medical image processing, is an effective way to address this challenge (Zeng et al., 2018; Sajjad et al., 2019; Mittal et al., 2019; Ge et al., 2020; Hua et al., 2021). In the process of brain disease diagnosis, firstly, the image features are extracted, and then the extracted image features are classified to complete the image classification and recognition. For example, Ismael and Abdel-Qader (2018) used Gabor filter and discrete wavelet transform to extract statistical features for brain tumor classification. Then this method used the tumor segmented as input and multi-layer perceptron (MLP) as the classifier. Liu et al. (2012) proposed a multi-level classification method for meningiomas. According to the type and growth rate of tumors, meningiomas are divided into three levels. In the classification step, the authors used a multiple logistic regression model. Mallick et al. (2019) proposed a brain MRI image classification method based on deep neural networks. Using encoding and decoding techniques, this method mainly used an automatic autoencoder to extract and classify brain images. To assist radiologists in MRI classification, Sachdeva et al. (2016) proposed a semi-automated classification method with multiple stages. To detect tumor regions, the first stage was the outline system detection of the content-based tumor regions, which can be manually indicated by the physician, called segmented regions of interest (SROI). Then, 71 texture and intensity features were extracted from the SROI regions, and the features were optimized by genetic algorithm. In the classification stage, support vector machine (SVM) and artificial neural network were used. Nikam and Shinde (2013) proposed a brain MRI image classification method based on distance learning. Firstly, the images were preprocessed, and many techniques such as gray transformation, median filtering, and high pass filtering were used to remove the noise of MRI brain image. The threshold segmentation method was used to segment the MRI brain image. Then the features are extracted

by correlation, entropy, contrast, homogeneity, and energy. Finally, a Euclidean distance classifier was used for classification. Ghassemi et al. (2020) proposed a CNN model for multi-class brain tumor classification. Firstly, the method was pre-trained as a discriminator in generative adversarial network to extract image features. Second, the softmax classifier was used to distinguish the three kinds of tumors. This model consists of six layers, which can be used together with various data augmentation techniques. Kiranmayee et al. (2016) proposed a brain MRI classification method using a SVM. In the data processing stage, a median adaptive filter was used to remove noise, and then the watershed method, fuzzy clustering method, and threshold method were used to segment MRI brain image. The kernel SVM was used as the classifier.

The dictionary learning method is widely used to solve various problems of computer vision and image analysis (Gu et al., 2020; Ni et al., 2020). Dictionary learning aims to find a suitable dictionary for the input data and transform it into a sparse representation, so as to mine the useful features of the data, simplify the learning task and reduce the complexity of the model. A kernel sparse representation was developed in Chen et al. (2017). It contained three key steps for multi-label brain tumor segmentation: component analysis-split for dictionary learning initialization, kernel dictionary learning and kernel sparse coding, and graph-cut method for image segmentation. A system combining an adaptive type-2 fuzzy system and dictionary learning was proposed in Ghasemi et al. (2020), in which the sparse coding step and dictionary learning step were executed alternately, and the fuzzy membership functions in the type-2 fuzzy system were used to represent model uncertainty and improve sparse representation. A learning method combining discriminate sub-dictionary and projective dictionary pair learning was developed for classifying proton magnetic resonance spectroscopy of brain gliomas tumor (Adebileje et al., 2017).

AI mainly uses intelligent methods to extract brain image features, which requires a large number of labeled data sets to understand the potential connections in the data. But in the field of medicine, because of the confidentiality and professionalism of patient information, medical data need to be marked by professional clinicians, and data collection is complex and expensive. Lack of labeled trainable data is one of the bottlenecks that restrict the development of medical image analysis. In addition, traditional AI methods require training data and test data to be independent and identically distributed. Transfer learning relaxes this restriction on training data and test data (Ni et al., 2018b; Jiang et al., 2020; Jiang et al., 2021). It can apply the knowledge or patterns learned from a related domain (source domain) to another target domain, and utilize the information shared by source domain samples and target domains, then finally build a model to adapt to the target domain.

To solve this problem, this paper focuses on solving the distribution differences between source and target domains. Through the feature mapping of source and the target domain samples, the source domain knowledge can be transferred to

target domain learning. Because dictionary learning can exploit the essential characteristics of the data, this paper uses Multi-layer dictionary learning (MDL) in transfer learning to exploit the shared knowledge between source and target domains. MDL first obtains the dictionary and sparse features of the first layer on the original samples, then obtains the dictionary and sparse features of the second layer based on the obtained sparse features of the first layer, and learns the dictionary and sparse features in turn to finally obtain the deep dictionary and sparse features. Finally, the new test data can be encoded by the multi-layer dictionary and the final classification results can be obtained. According to the difference of domain and task, transfer learning is divided into feature transfer, sample transfer and parameter transfer. In this paper, the target and source domain are images, and the task is to train the image, extract features, and realize the classification of different types of images, so this paper belongs to the parameter transfer mode. The advantages of this algorithm are as follows: (1) based on multi-layer learning, multi-layer dictionaries are obtained, and the discriminability of sparse representation coefficients can be enhanced in layer by layer dictionary learning; (2) through multi-layer shared dictionary learning, the sample reconstructions of source and target domains are constrained layer by layer, so as to minimize the error of sample reconstruction both in source and target domains; (3) by utilizing the label information, Laplacian regularization term is introduced, and the sparse coding of samples in the same class is as close as possible, while the sparse coding of samples in different classes is as different as possible. At the same time, in the last layer of the proposed model, the classification error term is introduced in the last layer of MDL to improve the discriminative performance of the model; (4) The recognition experiments on brain MRI image datasets REMBRANDT (Clark et al., 2013) and Figshare (Cheng et al., 2016) show that the proposed model performs satisfactory classification performance in terms of accuracy, precision, F1-score, and recall.

The rest of the paper is organized as follows: the related work is introduced in section "Backgrounds." The proposed method is given in section "Proposed Method", and experiments are performed in section "Experiment." Finally, conclusion and future work are summarized in section "Conclusion."

BACKGROUNDS

Dictionary Learning
Dictionary learning methods can basically be divided into unsupervised dictionary learning and supervised dictionary learning. The unsupervised dictionary learning does not make use of sample label information. The supervised dictionary learning makes use of sample label information and pays more attention to the discriminative ability of sparse representation coefficients.

KSVD (Jiang et al., 2013) is a famous supervised dictionary learning algorithm. KSVD introduces the classification error of a linear classifier into the objective function, while learning the representation and classification ability of the dictionary. The

objective function of K-SVD is

$$< \mathbf{D},\mathbf{W},\mathbf{Z}> = min||\mathbf{X}\text{-}\mathbf{D}\mathbf{Z}||_2^2 + \gamma||\mathbf{H}\text{-}\mathbf{W}\mathbf{Z}||_2^2 + \beta||\mathbf{W}||_2^2$$
$$s.t. \ ||\mathbf{Z}||_0 \leq T \tag{1}$$

where $\mathbf{Z}$ is the sparse representation coefficient, $\mathbf{W}$ is the parameter of the linear classifier, $\mathbf{H}$ is the label vector of the training data. To solve Eq. (1), the first two of these terms are combined and Eq. (1) is rewritten as

$$< \mathbf{D},\mathbf{W},\mathbf{Z} >= min \left|\left| \begin{pmatrix} \mathbf{X} \\ \sqrt{\gamma}\mathbf{H} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \sqrt{\gamma}\mathbf{W} \end{pmatrix} \mathbf{Z} \right|\right|_2^2 + \beta||\mathbf{W}||_2^2$$
$$s.t. \ ||\mathbf{Z}||_0 \leq T \tag{2}$$

Eq. (2) can be solved by using an iterative strategy. When $\mathbf{W}$ is fixed, the problem of <$\mathbf{D},\mathbf{Z}$> represents the same formulation as K-SVD, and it can therefore be solved using the K-SVD. When $\mathbf{D}$ and $\mathbf{Z}$ are fixed, Eq. (2) is a simple linear problem that can be solved by linear methods.

Multi-Layer Dictionary Learning
With the development of deep learning, researchers have found that the deeper the structure of a neural network, the better and more accurate the representation. MDL (also known as deep dictionary learning) refers to the idea of deep learning, and applies "deep structure" to layer-by-layer dictionary learning (Song et al., 2019; Gu et al., 2020). The dictionary and sparse representation obtained by the traditional single-layer dictionary learning method are shallow, which is not conducive to the task of recognition and classification when the data dimension is too high or the number of samples is too large. Singhal et al. (2017) proposed a deep dictionary learning model, which used the idea of deep learning to learn the multi-level dictionary and the deep features of the original samples. As an example, the two-layer dictionary learning is illustrated in **Figure 1**. $\mathbf{D}_1$ and $\mathbf{D}_2$ are dictionaries learned in the first and second layer. $\mathbf{Z}_2$ is the sparse coefficient learned in the second layer. The sample $\mathbf{X}$ can be represented as $X = \mathbf{D}_1\mathbf{Z}_1 = \mathbf{D}_1\mathbf{D}_2\mathbf{Z}_2$, where the sparse coding learning in the first layer $\mathbf{Z}_1 = \mathbf{D}_2\mathbf{Z}_2$. Specifically, the first layer is solved as a single layer of dictionary learning to
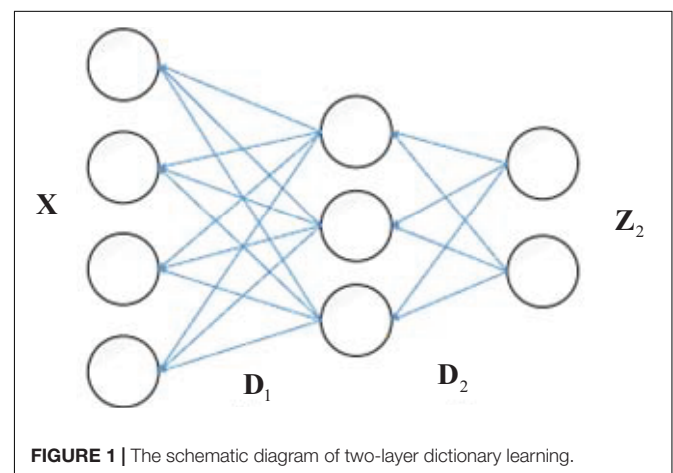

FIGURE 1 | The schematic diagram of two-layer dictionary learning.

obtain feature $\mathbf{Z}_1$ on dictionary $\mathbf{D}_1$, and $\mathbf{Z}_1$ is then used as input to the second layer, which is also solved as a single layer of dictionary learning to obtain feature $\mathbf{Z}_2$. The new test data can be encoded by the learned $\mathbf{D}_1$ and $\mathbf{D}_2$. In this way, after completing the L-layer dictionary learning, the final dictionary and sparse representations are obtained as $\mathbf{D}_L$ and $\mathbf{Z}_L$. In this case, the sample $\mathbf{X}$ can be represented as

$$\mathbf{X} = \mathbf{D}_1(\mathbf{D}_2(...(\mathbf{D}_L\mathbf{Z}_L))) \qquad (3)$$

Then the dictionaries in L-layers and the sparse coding can be solved by

$$\min_{\mathbf{D}_1,\mathbf{D}_2,...,\mathbf{D}_L,\mathbf{Z}_L} \left|\left|\mathbf{X}\text{-}\mathbf{D}_1(\mathbf{D}_2(...(\mathbf{D}_L\mathbf{Z}_L)))\right|\right|_F^2 + \left|\left|\mathbf{Z}_L\right|\right|_1 \qquad (4)$$

PROPOSED METHOD

Objective Function

We assume that there is a corresponding association between source and target domains in transfer learning. From this point, based on the framework of MDL, we try to learn the common shared dictionary between source and target domains to exploit the shared knowledge among different related domains. At the same time, by making full use of the label information of the samples, the classification error term is introduced in the last layer of the multi-layer dictionary, which makes the sparse representation of the target domain more discriminative. According to this idea, we propose a transfer model based on supervised multi-layer dictionary learning (TSMDL), and its objective function is

$$\min_{\mathbf{Z},\mathbf{w},b} \left|\left|\mathbf{X}_s - \mathbf{D}_1\mathbf{Z}_1^s\right|\right|_F^2 + \left|\left|\mathbf{X}_t - \mathbf{D}_1\mathbf{Z}_1^t\right|\right|_F^2 + \alpha_1 \sum_{i,j} Q_{1,i,j}^{C,s}\left|\left|\mathbf{z}_{1,i}^s - \mathbf{z}_{1,j}^s\right|\right|_F^2$$

$$-\beta_1 \sum_{i,j} Q_{1,i,j}^{M,s}\left|\left|\mathbf{z}_{1,i}^s - \mathbf{z}_{1,j}^s\right|\right|_F^2 + \alpha_1 \sum_{i,j} \mathbf{Q}_{1,i,j}^{C,t}\left|\left|\mathbf{z}_{1,i}^t - \mathbf{z}_{1,j}^t\right|\right|_F^2$$

$$-\beta_1 \sum_{i,j} Q_{1,i,j}^{M,t}\left|\left|\mathbf{z}_{1,i}^t - \mathbf{z}_{1,j}^t\right|\right|_F^2 + ... + \left|\left|\mathbf{Z}_{L-1}^s - \mathbf{D}_L\mathbf{A}_L^s\right|\right|_F^2$$

$$+\left|\left|\mathbf{z}_{L-1}^t - \mathbf{D}_L\mathbf{A}_L^t\right|\right|_F^2 + \alpha_L \sum_{i,j} Q_{L,i,j}^{C,s}\left|\left|\mathbf{z}_{L,i}^s - \mathbf{z}_{L,j}^s\right|\right|_F^2$$

$$-\beta_L \sum_{i,j} Q_{L,i,j}^{M,s}\left|\left|\mathbf{z}_{L,i}^s - \mathbf{z}_{L,j}^s\right|\right|_F^2 + \alpha_L \sum_{i,j} \mathbf{Q}_{L,i,j}^{C,t}\left|\left|\mathbf{z}_{L,i}^t - \mathbf{z}_{L,j}^t\right|\right|_F^2$$

$$-\beta_L \sum_{i,j} \mathbf{Q}_{L,i,j}^{M,t}\left|\left|\mathbf{z}_{L,i}^t - \mathbf{z}_{L,j}^t\right|\right|_F^2 + \lambda \sum_{c=1}^{C_s} f(\mathbf{z}_L^s, \mathbf{y}_c^s, \mathbf{w}_c^s, \mathbf{b}_c^s)$$

$$+\lambda \sum_{c=1}^{C_t} f(\mathbf{z}_L^t, \mathbf{y}_c^t, \mathbf{w}_c^t, \mathbf{b}_c^t) + \lambda_l \sum_{l=1}^{L} \left|\left|\mathbf{z}_l\right|\right|_F^2,$$

s.t. $\left|\left|d_l^j\right|\right|_2^2 = 1, l = 1, 2, ..., L, \quad j = 1, 2, ..., K_l$

$$(5)$$

where K_l is the size of dictionary in the lth layer. $\mathbf{Z}_l^s$ and $\mathbf{Z}_l^s$ are the spares coding matrixes in the l-th layer. $\mathbf{Q}_l^{C,(\cdot)}$ and $\mathbf{Q}_l^{M,(\cdot)}$ are the matrixes of weights for data samples of the same class and data samples of different classes in the l-th layer. $Q_{l,ij}^{C,(\cdot)}$ and $Q_{l,ij}^{M,(\cdot)}$ can

be defined as

$$Q_{l,ij}^{C,(\cdot)} = \begin{cases} \frac{1}{|C|}, & (\mathbf{z}_{l,i}^{(\cdot)}, \mathbf{z}_{l,j}^{(\cdot)}) \text{ belonging to the same class} \\ 0, & \text{otherwise} \end{cases} \qquad (6)$$

$$Q_{l,ij}^{M,(\cdot)} = \begin{cases} \frac{1}{|M|}, & (\mathbf{z}_{l,i}^{(\cdot)}, \mathbf{z}_{l,j}^{(\cdot)}) \text{ belonging to different classes} \\ 0, & \text{otherwise} \end{cases} \qquad (7)$$

where $(\cdot)$ means s or t.

We explain the above Eq. (5) as follows:

1. The first two terms $\left|\left|\mathbf{X}_s - \mathbf{D}_1\mathbf{Z}_1^s\right|\right|_F^2$ and $\left|\left|\mathbf{X}_t - \mathbf{D}_1\mathbf{Z}_1^t\right|\right|_F^2$ are the reconstruction error terms of source domain and target domain data in the first layer of the learning framework.

2. The third and fourth terms $\sum_{i,j} Q_{1,i,j}^{C,s}\left|\left|\mathbf{z}_{1,i}^s - \mathbf{z}_{1,j}^s\right|\right|_F^2$ and $\sum_{i,j} \mathbf{Q}_{1,i,j}^{M,s}\left|\left|\mathbf{z}_{1,i}^s - \mathbf{z}_{1,j}^s\right|\right|_F^2$ are the Laplacian regularization terms of the source domain in the first layer, which, respectively, constrain the dictionary codes of the same class in the source domain to be as close as possible, and the dictionary codes of different classes to be as different as possible. The fifth and sixth terms $\sum_{i,j} \mathbf{Q}_{1,i,j}^{C,t}\left|\left|\mathbf{z}_{1,i}^t - \mathbf{z}_{1,j}^t\right|\right|_F^2$ and $\sum_{i,j} \mathbf{Q}_{1,i,j}^{M,t}\left|\left|\mathbf{z}_{1,i}^t - \mathbf{z}_{1,j}^t\right|\right|_F^2$ are the Laplacian regularization terms of the target domain in the first layer. Similarly to the third and fourth terms, their goal is to, respectively, constrain the dictionary codes of the same class in the target domain to be as close as possible, and the dictionary codes of different classes to be as different as possible.

3. Following the generation rules for the first six terms, the corresponding reconstruction error terms and Laplacian regularization terms for the source and target domains are constructed for layers 2 to L.

4. $\sum_{c=1}^{C_s} f(\mathbf{Z}_L^s, \mathbf{y}_c^s, \mathbf{w}_c^s, \mathbf{b}_c^s)$ and $\sum_{c=1}^{C_t} f(\mathbf{Z}_L^t, \mathbf{y}_c^t, \mathbf{w}_c^t, \mathbf{b}_c^t)$ are classification error terms for the last layer of the source domain and target domain, respectively. Its goal is to improve the discriminative ability of the model. In this study, we use SVM multi-class classifier. The parameters $w_c^{(\cdot)}$ and $b_c^{(\cdot)}$ are hyperplane parameters in the SVM.

Define Laplacian matrix in the same class $\mathbf{P}_l^{C,(\cdot)}$ as $\mathbf{P}_l^{C,(\cdot)} = \tilde{\mathbf{Q}}_l^{C,(\cdot)} - \mathbf{Q}_l^{C,(\cdot)}$, where $\tilde{\mathbf{Q}}_{l,ii}^{C,(\cdot)} = \sum_j \mathbf{Q}_{l,ij}^{C,(\cdot)}$, Laplacian matrix

TABLE 1 | The basic information of transfer learning tasks in the experiment.

Tasks	Source domain		Target domain	
T1	Normal images from the REMBRANDT dataset	Tumors images from the Figshare dataset	Normal images from the REMBRANDT dataset	Tumors images from the REMBRANDT dataset
T2	Normal images from the REMBRANDT dataset	Tumors images from the REMBRANDT dataset	Normal images from the REMBRANDT dataset	Tumors images from the Figshare dataset

in the different classes $\mathbf{P}_l^{M,(\cdot)}$ as $\mathbf{P}_l^{M,(\cdot)} = \tilde{\mathbf{Q}}_l^{M,(\cdot)} - \mathbf{Q}_l^{M,(\cdot)}$, where $\tilde{\mathbf{Q}}_{l,ii}^{M,(\cdot)} = \sum_j Q_{l,ij}^{M,(\cdot)}$. Let $\mathbf{X} = [\mathbf{X}_s, \mathbf{X}_t]$, $\mathbf{Z} = [\mathbf{Z}_s, \ \mathbf{Z}_t]$, $\mathbf{P}_l^C = \begin{bmatrix} \mathbf{P}_l^{C,s} & 0 \\ 0 & \mathbf{P}_l^{C,t} \end{bmatrix}$, $\mathbf{P}_l^M = \begin{bmatrix} \mathbf{P}_l^{M,s} & 0 \\ 0 & \mathbf{P}_l^{M,t} \end{bmatrix}$, Eq. (5) can be written as

$$
\begin{aligned}
\min_{\mathbf{D}_l, \mathbf{P}_l^C, \mathbf{P}_l^M, \mathbf{Z}_l, w, b} & \left| \mathbf{X}\text{-}\mathbf{D}_1\mathbf{Z}_1 \right|_F^2 + Tr(\mathbf{Z}_1(\alpha_1\mathbf{P}_1^C + \lambda_1\mathbf{I})\mathbf{Z}_1^T) \\
& -\beta_1 Tr(\mathbf{Z}_1\mathbf{P}_1^M\mathbf{Z}_1^T) + \ldots + \\
& \left\| \mathbf{Z}_{L-1} - \mathbf{D}_L\mathbf{Z}_L \right\|_F^2 + Tr(\mathbf{Z}_L(\alpha_L\mathbf{P}_L^C + \lambda_L\mathbf{I})\mathbf{Z}_L^T) \\
& -\beta_L Tr(\mathbf{Z}_L\mathbf{P}_L^M\mathbf{Z}_L^T) + \lambda\sum_{c=1}^{\mathbf{C}}f(\mathbf{Z}_L, \mathbf{y}_c, \mathbf{w}_c, \mathbf{b}_c),
\end{aligned} \quad (8)
$$
$$
\text{s.t. } \left\| d_l^j \right\|_2^2 = 1, l = 1, 2, \ldots, L, \ \ j = 1, 2, \ldots, K_l,
$$

Again, we simplify the function above and obtain that

$$
\begin{aligned}
\min_{\mathbf{D}_l, \mathbf{P}_l^C, \mathbf{P}_l^M, \mathbf{Z}_l, w, b} & \left\| \mathbf{X}\text{-}\mathbf{D}_1\mathbf{Z}_1 \right\|_F^2 + Tr(\mathbf{Z}_1(\alpha_1\mathbf{P}_1^C - \beta_1\mathbf{P}_1^M + \lambda_1\mathbf{I}) \\
& \mathbf{Z}_1^T + \ldots + \\
& \left\| \mathbf{Z}_{L-1} - \mathbf{D}_L\mathbf{Z}_L \right\|_F^2 + Tr(\mathbf{Z}_L(\alpha_L\mathbf{P}_L^C - \beta_L\mathbf{P}_L^M \\
& +\lambda_L\mathbf{I})\mathbf{Z}_L^T) + \lambda\sum_{c=1}^{\mathbf{C}}f(\mathbf{Z}_L, \mathbf{y}_c, \mathbf{w}_c, \mathbf{b}_c),
\end{aligned} \quad (9)
$$
$$
\text{s.t. } \left\| d_l^j \right\|_2^2 = 1, l = 1, 2, \ldots, L, \ \ j = 1, 2, \ldots, K_l,
$$

Optimization

We use the alternating optimization approach to solve Eq. (9). The parameters to be solved include $\mathbf{D}_1$, $\mathbf{P}_1^C$, $\mathbf{P}_1^M$, $\mathbf{Z}_1, \ldots, \mathbf{D}_L$, $\mathbf{P}_L^C$, $\mathbf{P}_L^M$, $\mathbf{Z}_L$, w and b. In the following, we divide the solution of these variables into three parts.

a. Update parameters $\mathbf{D}_1$, $\mathbf{P}_1^C$, $\mathbf{P}_1^M$, $\mathbf{Z}_1, \ldots, \mathbf{D}_L$, $\mathbf{P}_L^C$ and $\mathbf{P}_L^M$

First, we update parameters $\mathbf{D}_1$, $\mathbf{P}_1^C$, $\mathbf{P}_1^M$ and $\mathbf{Z}_1$ in the first layer. When fixed the other parameters, the objective function of TSMDL is

$$
\begin{aligned}
\min_{\mathbf{D}_1, \mathbf{P}_1^C, \mathbf{P}_1^M, \mathbf{Z}_1} & \left\| \mathbf{X}\text{-}\mathbf{D}_1\mathbf{Z}_1 \right\|_F^2 + Tr(\mathbf{Z}_1(\alpha_1\mathbf{P}_1^C - \beta_1\mathbf{P}_1^M + \lambda_1\mathbf{I})\mathbf{Z}_1^T \\
& +\left\| \mathbf{Z}_1 - \mathbf{D}_2\mathbf{Z}_2 \right\|_F^2, \\
\text{s.t. } & \left\| d_1^j \right\|_2^2 = 1
\end{aligned} \quad (10)
$$

Further, the parameters except for $\mathbf{D}_1$ are fixed, the optimization problem can be written as

$$
\begin{aligned}
\min_{\mathbf{D}_1, \mathbf{P}_1^C, \mathbf{P}_1^M, \mathbf{Z}_1} & \left\| \mathbf{X}\text{-}\mathbf{D}_1\mathbf{Z}_1 \right\|_F^2, \\
\text{s.t. } & \left\| d_1^j \right\|_2^2 = 1
\end{aligned} \quad (11)
$$

Following (Boyd et al., 2011), the optimal value of $\mathbf{D}_1$ can be computed by an alternating direction method of multipliers. Then the Laplacian matrixes $\mathbf{P}_1^C$ and $\mathbf{P}_1^M$ can be computed according to Eqs.(6, 7).

The optimal value of $\mathbf{Z}_1$ can be obtained by taking the derivation of Eq.(8) as the following formulation, i.e.,

$$
\mathbf{Z}_1 = (\mathbf{D}_1^T\mathbf{D}_1 + \alpha_1\mathbf{P}_1 - \beta_1\mathbf{P}_1^M + (\lambda_1 + 1)\mathbf{I})^{-1}(\mathbf{D}_1^T\mathbf{X} + \mathbf{D}_2\mathbf{Z}_2) \quad (12)
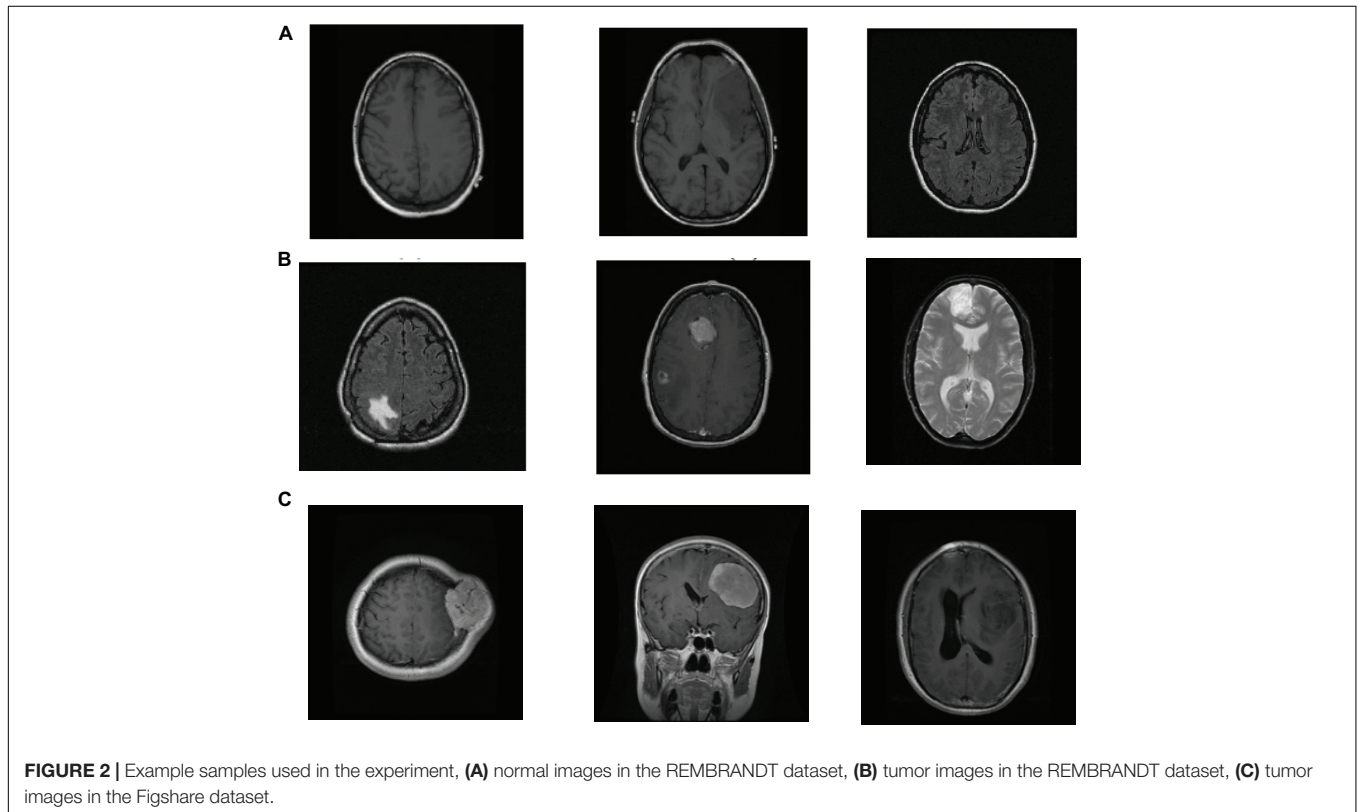$$

FIGURE 2 | Example samples used in the experiment, **(A)** normal images in the REMBRANDT dataset, **(B)** tumor images in the REMBRANDT dataset, **(C)** tumor images in the Figshare dataset.

For $2 \leq l \leq L$, when fixing the other parameters, the objective function of $\mathbf{D}_l$ is

$$\min_{\mathbf{D}_l} \left\| \mathbf{Z}_{l-1} - \mathbf{D}_l \mathbf{Z}_l \right\|_F^2,$$
$$\text{s.t. } \left\| d_2^j \right\|_2^2 = 1, j = 1, 2, ..., K_L \quad (13)$$

After obtaining the $\mathbf{D}_l$, the optimal value of $\mathbf{Z}_l (2 \leq l \leq L - 1)$ can be obtained by,

$$\mathbf{Z}_l = (\mathbf{D}_l^T \mathbf{D}_l + \alpha_l \mathbf{P}_l^C - \beta_l \mathbf{P}_l^M + (\lambda_l + 1)\mathbf{I})^{-1}(\mathbf{D}_l^T \mathbf{X} + \mathbf{D}_{l+1} \mathbf{Z}_{l+1}) \quad (14)$$

b. Update parameter $\mathbf{Z}_L$:

When the other parameters are fixed, the objective function of TSMDL related to $\mathbf{Z}_L$ is

$$\min_{\mathbf{Z}_L} \ \left\| \mathbf{Z}_{L-1} - \mathbf{D}_L \mathbf{Z}_L \right\| + Tr(\mathbf{Z}_L(\alpha_L \mathbf{P}_L^C - \beta_L \mathbf{P}_L^M + \lambda_L \mathbf{I})\mathbf{Z}_L^T)$$

$$+ \lambda \sum_{c=1}^C f(\mathbf{Z}_L, \mathbf{y}_c, \mathbf{w}_c, \mathbf{b}_c) \quad (15)$$

Let $\mathbf{z}_L^i (i = 1, 2, ..., N)$ be the ith column of $\mathbf{Z}_L$. We rewrite Eq. (15) related to $\mathbf{z}_L^i$ as

$$\min_{\mathbf{Z}_L} \ \left\| \mathbf{z}_{L-1}^i - \mathbf{D}_L \mathbf{Z}_L \right\| + Tr(\mathbf{z}_L^i(\alpha_L \mathbf{P}_L^C - \beta_L \mathbf{P}_L^M + \lambda_L \mathbf{I})\mathbf{z}_L^{iT})$$

$$+ \lambda \sum_{c=1}^C f(\mathbf{z}_L^i, \mathbf{y}_c^i, \mathbf{w}_c, \mathbf{b}_c) \quad (16)$$

In this study, we use standard L1-SVM for term $f(\mathbf{z}_L^i, \mathbf{y}_c^i, \mathbf{w}_c, \mathbf{b}_c)$, thus we can set $\mathbf{y}_c^i = 1$ if class label $\mathbf{y}_c^i = c$ and otherwise $y_c^i = -1$. In this case, the optimal value of $\mathbf{z}_L^i$ can be computed by a least square problem.

c. Update parameters $\mathbf{w}$ and b

When the other parameters are fixed, the objective function of TSMDL related to $\mathbf{w}$ and b is

$$\min_{\mathbf{w}_c, \mathbf{b}_c} \sum_{c=1}^C f(\mathbf{Z}_L, \mathbf{y}_c, \mathbf{w}_c, b_c), \quad (17)$$
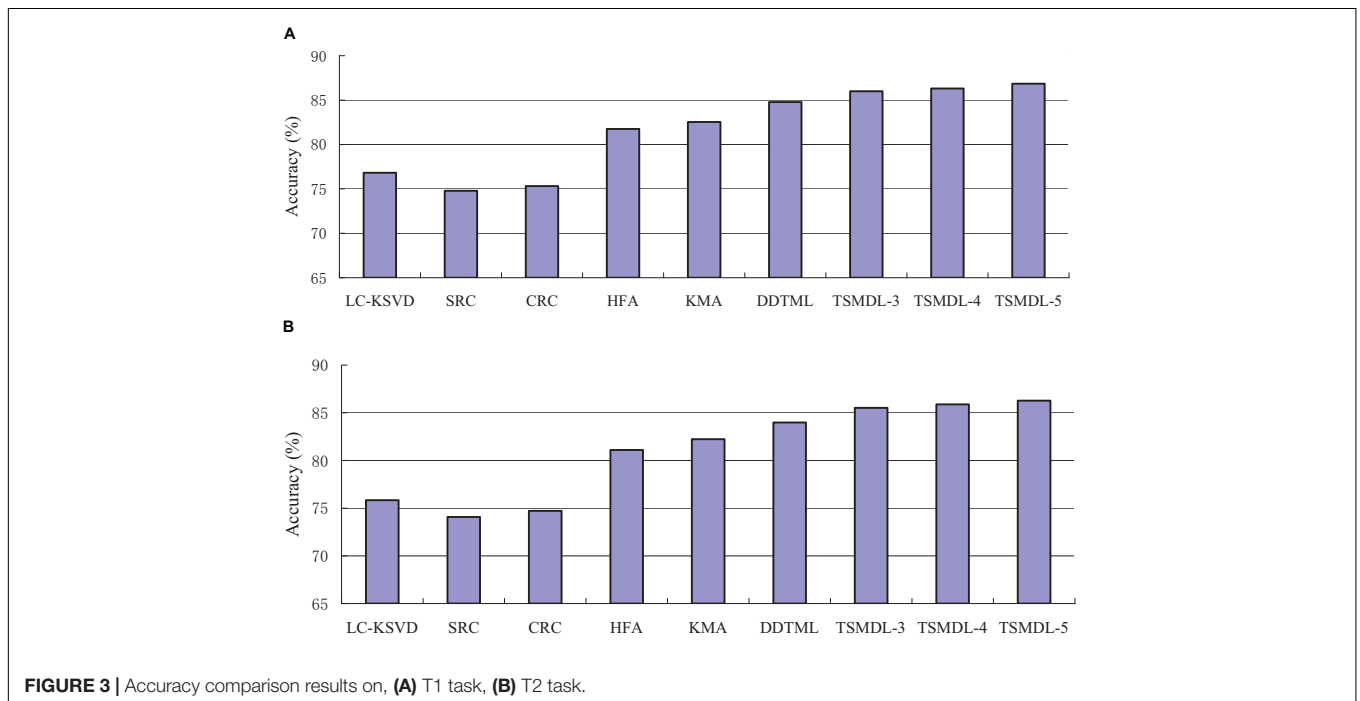
Obviously, Eq. (17) can be solved by various SVM solvers.

We show the optimization procedure of TSMDL in algorithm 1.

Input: Training data matrix **X**, parameters α_l, β_l and λ_l, $\forall l$

1: Initialize **D** using K-SVD algorithm on each class, initialize **P** using principal component analysis (PCA) algorithm;
2: **while** not converged **do**
3:   Compute $\mathbf{D}_l (1 \leq l \leq L)$ by Eq. (11);
4:   Compute the Laplacian matrixes $\mathbf{P}_l^C$ and $\mathbf{P}_l^M$ ($1 \leq l \leq L$) by Eqs. (6, 7);
5:   Compute $\mathbf{Z}_l (1 \leq l \leq L - 1)$ by Eqs. (12–14);
6:   Compute $\mathbf{z}_L^i$ by Eq. (16);
7:   Compute **w** and b by Eq. (17);
8: **end while**
Output: $\mathbf{D}_1, \mathbf{P}_1^C, \mathbf{P}_1^M, \mathbf{Z}_1, ..., \mathbf{D}_L, \mathbf{P}_L^C, \mathbf{P}_L^M, \mathbf{Z}_L$, **w** and b.

Learning a Classifier

We compute $\Theta_1 = (\mathbf{D}_l^T \mathbf{D}_l + \alpha \mathbf{P}_l^C - \beta \mathbf{P}_l^M + \lambda_l \mathbf{I})^{-1} \mathbf{D}_l^T (l = 1, 2, ..., L)$. The test sample $\mathbf{X}_{\text{new}}$, we compute its sparse coding as $z_{\text{new}} = \Theta_1 ... \Theta_L \mathbf{x}_{\text{new}}$. Finally, we can use the following formulation to predict the class label of $\mathbf{x}_{\text{new}}$

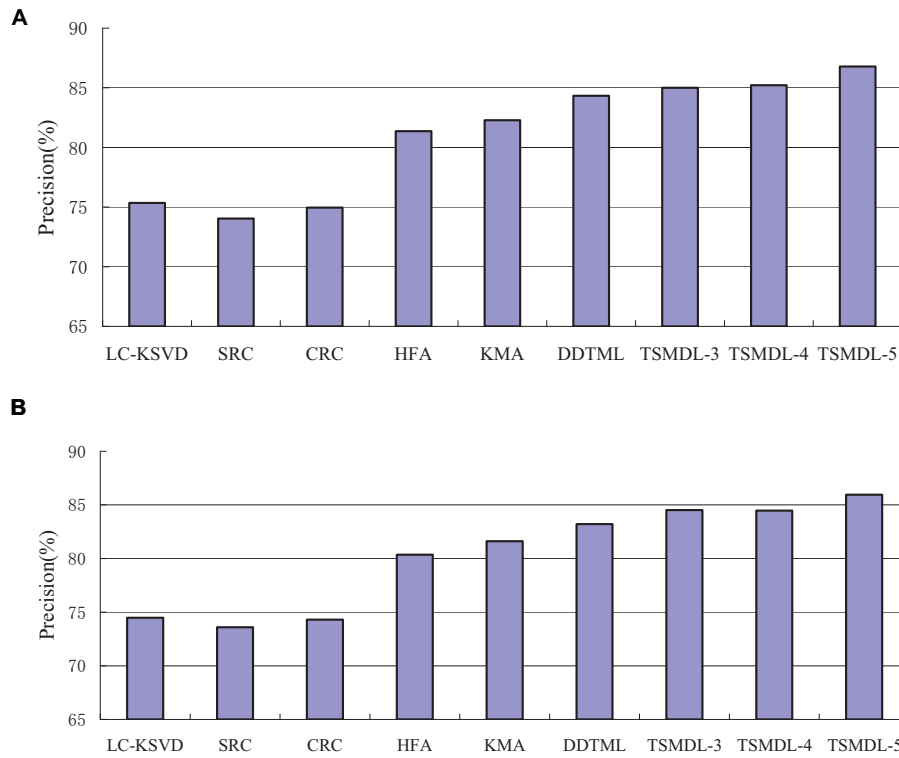$$\mathbf{y} = \arg \max_c \ \mathbf{w}_c^T \mathbf{z}_{\text{new}} + b_c, \ \forall c \quad (18)$$

FIGURE 3 | Accuracy comparison results on, **(A)** T1 task, **(B)** T2 task.

FIGURE 4 | Precision comparison results on tasks, **(A)** T1, **(B)** T2.
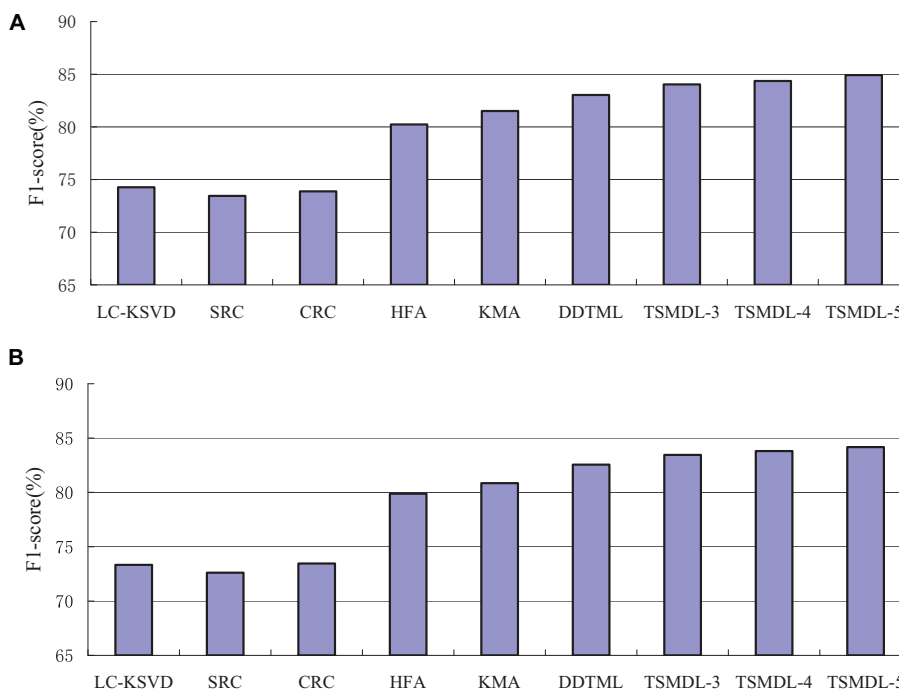

FIGURE 5 | F-Score comparison results on tasks, **(A)** T1, **(B)** T2.

EXPERIMENT

Experiment Settings

The datasets used in the study are taken from (Clark et al., 2013) and Figshare (Cheng et al., 2016) datasets. Figshare dataset is collected from two hospitals in China between 2005 and 2010. It contains a vast number of MRI images from 233 patients with brain tumors, including meningiomas, pituitary and gliomas. All images are digitized at a resolution of 512×512 pixels. REMBRANDT dataset contains a vast number of MRI images collected from 130 brain tumor patients. The patients' ages ranged from 15 to 89 years, with a mean of 47.5 years. The MRI images included astrocytomas from 47 patients; oligodendrogliomas from 21 patients; GBMs from 40 patients and the truth of the tumor in the remaining patients is unknown. All images are digitized at a resolution of 256×256 pixels. Some slices in the REMBRANDT dataset, where the tumor lesions are found, are considered normal samples. In the experiment, we design two transfer learning tasks, and show their information in **Table 1**. **Figure 2** shows sample images in the REMBRANDT and Figshare datasets. The main objective of two tasks is to classify the brain MRI images into normal and tumor classes. In task T1, we randomly select 200 normal images and 200 tumor images from the REMBRANDT dataset as source domain, and randomly select 200 normal images from the REMBRANDT dataset and 200 tumor images from Figshare dataset as target domain. There are no duplicate images in source and target domains. We use all images in the source domain and 10% images in the target domain as training data, and use the rest of the images in the target domain as testing data. We use wavelet transform wavelets and gray level co-occurrence matrix (GLCM) method for feature extraction (Mohankumar, 2016). Each image is extracted onto a 540 dimensional vector.

In the experiment, we compare our model with LC-KSVD (Jiang et al., 2013), SRC (Wright et al., 2009), CRC (Zhang et al., 2011), HFA (Long et al., 2013), KMA (Tuia and Camps-Valls, 2016), and DDTML (Ni et al., 2018a). Following the authors, all parameters in comparative methods are set in their default settings. The parameters β, λ_1, and λ in TSMDL are set in the grid {0.01, 0.05, 0.1,...,2}. The number of layers is set in {3, 4, 5}, and the TSMDL model is accordingly named as TSMDL-3, TSMDL-4, and TSMDL-5, respectively. The sizes of dictionaries are 500, 450, 400, 350, and 300 corresponding to layer 1 to layer 5, respectively. In order to ensure the stability and effectiveness of the experimental results, for the proposed model and other comparative experimental methods, we run each task 10 times. All the methods are implemented in MATLAB, and the environment that we used in the experiments is a computer with Intel Core i5-3317U 1.70 GHz CPU, 16 GB RAM.

Experiment Results

In this subsection, we present the effect of TSMDL on T1 and T2 tasks. We summarize the performance of all comparative methods in terms of accuracy, precision, F1-score, and recall. The experiment results are shown in **Figures 3–6**, respectively. According to **Figures 3–6**, we can draw the following results:

In terms of accuracy, precision, F1-score, and recall, the proposed TSMDL achieves the best results. In addition,

FIGURE 6 | Recall comparison results on tasks, **(A)** T1, **(B)** T2
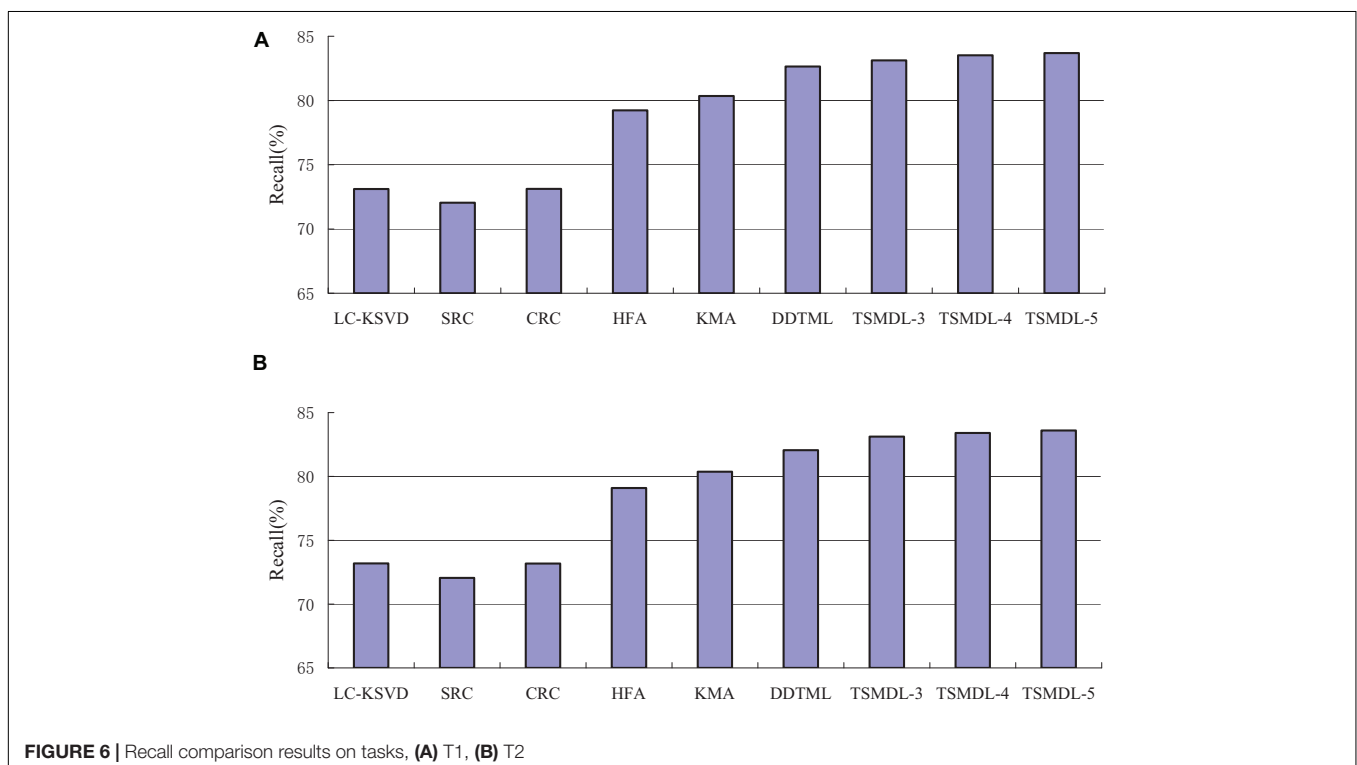
the performance of TSMDL-5 is better than TSMDL-3 and TSMDL-4. It is indicated that the multi-layer framework of dictionary learning can exploit the instinct structure of data samples and can build a relationship between source and target domains. Thus, TSMDL is suitable for the application of brain tumor MRI image recognition.

In the experiments, except for the LC-KSVD, SRC, and CRC algorithms, all other algorithms are transfer learning-based classification methods, which show that transfer learning strategy is helpful for brain tumor MRI image classification in the target domain. The classification knowledge in the source domain can be effectively transferred to the target domain to help the target domain achieve better classification results.

The proposed TSMDL in this paper is obviously superior to other transfer learning methods, which shows that multiple layer transfer learning dictionary learning can truly restore the brain MRI images of source and target domains, and reduce the distribution difference between domains. Thus, it can strengthen the domain adoption between source and target domains in the sparse representation space. The reason is that TSMDL is based on MDL; it can learn a more complex and accurate dictionary to represent the original data, and obtain more discriminative representation coefficients. In addition, TSMDL is a supervised learning model, in which the label information can be exploited, so TSMDL can obtain higher discrimination performance.

CONCLUSION

With the popularity of MRI equipment, a large number of new MRI brain images emerge, but obtaining labeled data is very time-consuming and expensive. Therefore, the goal of this paper is to use a large number of labeled data from the source domain to learn a classifier with strong generalization ability, and to classify the target domain with only a small number of labeling samples. Therefore, based on the MDL framework, we learn the common dictionary on each layer

of the network to minimize the sample reconstruction error of the constrained source domain and target domain. At the same time, the Laplacian regularization term is introduced in each layer of the network to make the sparse coding of similar samples as close as possible, while the sparse coding of different classes of samples is as different as possible. The experimental results on brain MRI image datasets REMBRANDT and Figshare show that our model achieves the state-of-the-art methods. Future works will include studying the effect of using unlabeled samples in the target domain while training, and other relevant problems like large-scale and online adaptation of dictionaries.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The download URLs of the REMBRANDT and Figshare datasets are, respectively, [https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT] and [https://figshare.com/articles/dataset/brain_tumor_dataset/1512427].

AUTHOR CONTRIBUTIONS

YG developed the theoretical framework and model in this work and drafted the manuscript. YG and KL implemented the algorithm and performed experiments and result analysis. Both contributed to the article and approved the submitted version.

FUNDING

REFERENCES

Adebileje, S. A., Ghasemi, K., Aiyelabegan, H. T., and Rad, H. S. (2017). Accurate classification of brain gliomas by discriminate dictionary learning based on projective dictionary pair learning of proton magnetic resonance spectra. *Magn. Reson. Chem.* 55, 318–322. doi: 10.1002/mrc.4532

Amin, J., Sharif, M., Yasmin, M., and Fernandes, S. L. (2017). A distinctive approach in brain tumor detection and classification using MRI. *Pattern Recognit. Lett.* 139, 118–127. doi: 10.1016/j.patrec.2017.10.036

Bahadure, N. B., Ray, A. K., and Thethi, H. P. (2017). Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *Int. J. Biomed. Imaging* 2017:9749108. doi: 10.1155/2017/9749108

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends.* 3, 1–128. doi: 10.1561/2200000016

Chen, X., Nguyen, B. P., Chui, C. K., and Ong, S. H. (2017). An automatic framework for multi-label brain tumor segmentation based on kernel sparse

representation. *Acta Polytech. Hung.* 14, 25–43. doi: 10.12700/APH.14.1.2017.1.3

Cheng, J., Yang, W., Huang, M., Huang, W., Jiang, J., Zhou, Y., et al. (2016). Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation. *PLoS One* 11:e0157112. doi: 10.1371/journal.pone.0157112

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., et al. (2013). The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J. Digit. Imaging* 26, 1045–1057. doi: 10.1007/s10278-013-9622-7

Ge, C., Gu, I. Y. H., Jakola, A. S., and Yang, J. (2020). Deep semi-supervised learning for brain tumor classification. *BMC Med. Imaging* 20:87. doi: 10.1186/s12880-020-00485-0

Ghasemi, M., Kelarestaghi, M., Eshghi, F., and Sharifi, A. (2020). T2-FDL: a robust sparse representation method using adaptive type-2 fuzzy dictionary learning for medical image classification. *Expert Syst. Appl.* 158:113500. doi: 10.1016/j.eswa.2020.113500

Ghassemi, N., Shoeibi, A., and Rouhani, M. (2020). Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. *Biomed. Signal Process. Control* 57:101678. doi: 10.1016/j.bspc.2019.101678

Gu, X., Zhang, C., and Ni, T. (2020). A hierarchical discriminative sparse representation classifier for EEG signal detection. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 18, 53–61. doi: 10.1109/TCBB.2020.2973978

Hua, L., Gu, Y., Gu, X. Q., Xue, J., and Ni, T. G. (2021). A novel brain MRI image segmentation method using an improved multi-view fuzzy c-means clustering algorithm. *Front. Neurosci.* 15:662674. doi: 10.3389/fnins.2021.662674

Ismael, M. R., and Abdel-Qader, I. (2018). "Brain tumor classification via statistical features and back-propagation neural network," in *Proceedings of the 2018 IEEE International Conference on Electro/Information Technology (EIT)*, Rochester, MI, 252–257. doi: 10.1109/EIT.2018.8500308

Jiang, Y., Gu, X., Ji, D., Qian, P., Xue, J., Zhang, Y., et al. (2020). Smart diagnosis: a multiple-source transfer TSK fuzzy system for EEG seizure identification. *ACM Trans. Multimed. Comput. Commun. Appl.* 16:59. doi: 10.1145/3340240

Jiang, Y., Gu, X., Wu, D., Hang, W., Xue, J., Qiu, S., et al. (2021). A novel negative-transfer-resistant fuzzy clustering model with a shared cross-domain transfer latent space and its application to brain CT image segmentation. *IEEE/ACM Trans. Comput.Biol. Bioinformatics* 18, 40–52. doi: 10.1109/TCBB.2019.2963873

Jiang, Z., Lin, Z., and Davis, L. S. (2013). Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 2651–2664. doi: 10.1109/TPAMI.2013.88

Kiranmayee, B. V., Rajinikanth, T. V., and Nagini, S. (2016). Enhancement of SVM based MRI brain image classification using pre-processing techniques. *Indian J. Sci. Technol.* 9, 1–7. doi: 10.17485/ijst/2016/v9i29/91042

Liu, Y. H., Muftah, M., Das, T., Bai, L., Robson, K., and Auer, D. (2012). Classification of MR tumor images based on Gabor wavelet analysis. *J. Med. Biol. Eng.* 32, 22–28. doi: 10.5405/jmbe.813

Long, M., Wang, J., Ding, G., Sun, J., and Yu, P. S. (2013). "Transfer feature learning with joint distribution adaptation," in *Proceedings of the International Conference on Computer Vision*, (New York, NY: IEEE), 2200–2207. doi: 10.1109/ICCV.2013.274

Mallick, P. K., Ryu, S. H., Satapathy, S. K., Mishra, S., Nguyen, G. N., and Tiwari, P. (2019). Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network. *IEEE Access* 15, 46278–46287. doi: 10.1109/ACCESS.2019.2902252

Mittal, M., Goyal, L. M., Kaur, S., Kaur, I., Verma, A., and Hemanth, D. J. (2019). Deep learning based enhanced tumor segmentation approach for MR brain images. *Appl. Soft Comput. J.* 78, 346–354. doi: 10.1016/j.asoc.2019.02.036

Mohankumar, S. (2016). Analysis of different wavelets for brain image classification using support vector machine. *Int. J. Adv. Signal Image Sci.* 30, 1–4. doi: 10.29284/ijasis.2.1.2016.1-4

Ni, T., Gu, X., and Jiang, Y. (2020). Transfer discriminative dictionary learning with label consistency for classification of EEG signals of epilepsy. *J. Ambient Intell. Humaniz. Comput.* doi: 10.1007/s12652-020-02620-9 [Epub ahead of Print].

Ni, T., Gu, X., Wang, H., Zhang, Z., Chen, S., and Jin, C. (2018a). Discriminative deep transfer metric learning for cross-scenario person reidentification. *J. Electron. Imaging* 27:043026. doi: 10.1117/1.JEI.27.4.043026

Ni, T., Gu, X., Wang, J., Zheng, Y., and Wang, H. (2018b). Scalable transfer support vector machine with group probabilities. *Neurocomputing* 273, 570–582. doi: 10.1016/j.neucom.2017.08.049

Nikam, P. B., and Shinde, V. D. (2013). MRI brain image classification and detection using distance classifier method in image processing. *Int. J. Eng. Res. Technol.* 2, 1980–1985.

Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N., and Ahuja, C. K. (2016). A package-SFERCB-"segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors". *Appl. Soft Comput.* 47, 151–167. doi: 10.1016/j.asoc.2016.05.020

Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., and Baik, S. W. (2019). Multi-grade brain tumor classification using deep CNN with extensive data. *J. Comput. Sci.* 30, 174–182. doi: 10.1016/j.jocs.2018.12.003

Singhal, V., Aggarwal, H. K., Tariyal, S., and Majumdar, A. (2017). Discriminative robust deep dictionary learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 55, 5274–5283. doi: 10.1109/TGRS.2017.2704590

Song, J., Xie, X., Shi, G., and Dong, W. (2019). Multi-layer discriminative dictionary learning with locality constraint for image classification. *Pattern Recognit.* 91, 135–146. doi: 10.1016/j.patcog.2019.02.018

Sun, R., Wang, K., Guo, L., Yang, C., Chen, J., Ti, Y., et al. (2019). A potential field segmentation based method for tumor segmentation on multi-parametric MRI of glioma cancer patients. *BMC Med. Imaging* 19:48. doi: 10.1186/s12880-019-0348-y

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J. Clin.* 71, 1–41. doi: 10.3322/caac.21660

Tuia, D., and Camps-Valls, G. (2016). Kernel manifold alignment for domain adaptation. *PLoS One* 11:e0148655. doi: 10.1371/journal.pone.0148655

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., and Ma, Y. (2009). Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 210–227. doi: 10.1109/TPAMI.2008.79

Zeng, K., Zheng, H., Cai, C., Yang, Y., Zhang, K., and Chen, Z. (2018). Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network. *Comput. Biol. Med.* 99, 133–141. doi: 10.1016/j.compbiomed.2018.06.010

Zhang, L., Yang, M., and Feng, X. (2011). "Sparse representation or collaborative representa- tion: which helps face recognition?," in *Proceedings of the IEEE Conference on Computer Vision*, Vol. 2011, Barcelona, 471–478. doi: 10.1109/ICCV.2011.6126277