Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Kernel sparse representation for MRI image analysis in automatic brain tumor segmentation*

Ji-jun TONG[†1], Peng ZHANG[1], Yu-xiang WENG[2], Dan-hua ZHU[3]

[1]*School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China*

[2]*Department of Neurosurgery, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China*

[3]*State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China*

[†]E-mail: jijuntong@zstu.edu.cn

Received July 27, 2016; Revision accepted Dec. 23, 2016; Crosschecked Apr. 14, 2018

Abstract: The segmentation of brain tumor plays an important role in diagnosis, treatment planning, and surgical simulation. The precise segmentation of brain tumor can help clinicians obtain its location, size, and shape information. We propose a fully automatic brain tumor segmentation method based on kernel sparse coding. It is validated with 3D multiple-modality magnetic resonance imaging (MRI). In this method, MRI images are pre-processed first to reduce the noise, and then kernel dictionary learning is used to extract the nonlinear features to construct five adaptive dictionaries for healthy tissues, necrosis, edema, non-enhancing tumor, and enhancing tumor tissues. Sparse coding is performed on the feature vectors extracted from the original MRI images, which are a patch of $m\times m\times m$ around the voxel. A kernel-clustering algorithm based on dictionary learning is developed to code the voxels. In the end, morphological filtering is used to fill in the area among multiple connected components to improve the segmentation quality. To assess the segmentation performance, the segmentation results are uploaded to the online evaluation system where the evaluation metrics dice score, positive predictive value (PPV), sensitivity, and kappa are used. The results demonstrate that the proposed method has good performance on the complete tumor region (dice: 0.83; PPV: 0.84; sensitivity: 0.82), while slightly worse performance on the tumor core (dice: 0.69; PPV: 0.76; sensitivity: 0.80) and enhancing tumor (dice: 0.58; PPV: 0.60; sensitivity: 0.65). It is competitive to the other groups in the brain tumor segmentation challenge. Therefore, it is a potential method in differentiation of healthy and pathological tissues.

Key words: Brain tumor segmentation; Kernel method; Sparse coding; Dictionary learning
https://doi.org/10.1631/FITEE.1620342 **CLC number:** TP181; R739.41

1 Introduction

The segmentation of brain tumor plays an important role in diagnosis, treatment planning, and surgical simulation. It is important to develop a valuable and robust tool for brain tumor segmentation. Because of high spatial resolution and superior soft tissue contrast in magnetic resonance imaging (MRI) images, compared with computed tomography (CT) images, MRI images have been used widely in brain tumor segmentation for its non-invasion and harmlessness. Brain tumor segmentation is challenging due to the complex structure, fuzzy boundaries, and similarity in the intensities of abnormal and normal brain tissue regions. The corresponding segmentation methods can be classified into three categories based on the required degrees of human interaction (Olabarriaga and Smeulders, 2001), i.e., manual and semi-automatic/fully automatic segmentations. In practice, the segmentation of brain tumor is delineated by experienced clinicians. It not only is time-

* Project supported by the National Natural Science Foundation of China (No. 31200746), the Zhejiang Provincial Key Research and Development Plan, China (No. 2015C03023), and the '521' Talent Project of ZSTU, China

ORCID: Ji-jun TONG, http://orcid.org/0000-0002-6209-6605

© Zhejiang University and Springer-Verlag GmbH Germany, part of Springer Nature 2018

consuming, but also relies on subjective experience. Therefore, many researchers have devoted themselves to improve the accuracy and efficiency of semi-automatic and fully automatic segmentations on MRI images.

The challenges of automatic brain tumor segmentation have gained much attention. In the early stage, the segmentation methods can be threshold-based (Juergens et al., 2008; Taheri et al., 2010), region-growing-based (Rousson et al., 2004; Wang and Vemuri, 2004), or outlier-detection-based (Prastawa et al., 2004). A threshold-based method with one or more thresholds is simple and effective. The threshold values are usually estimated through local or global statistic properties such as mean intensity values (Shanthi and Kumar, 2007) and the valley between the two modal peaks in a histogram. Due to the high complexity of brain structure and excessive intensity similarity in the edge of normal and abnormal brain tissues, the threshold-based method is usually used in the first step to determine the location of a brain tumor (Gibbs et al., 1996; Atkins and Mackiewich, 1998). Region-based segmentation approaches examine pixels in an image, and form disjoint regions by merging neighborhood pixels with homogeneity properties based on a predefined similarity criterion (Wong, 2005). Region growing has been proved to be an effective approach and used commonly (Mittelhäußer and Kruggel, 1995; Chong et al., 2004; Salman et al., 2005). The outlier detection framework proposed by Prastawa et al. (2004), which combines the model of the normal tissues and the geometric and spatial model of tumor and edema in T1 and T2 images, seems to be robust and promising.

In these studies, the hand-designed features including intensity and texture have been used, which has greatly influenced the segmentation performance. An alternative approach is to learn potential features from the original dataset directly. Sparse coding has been shown to be promising in brain tumor classification by learning over-complete dictionaries that facilitate a sparse representation of the MRI images as a linear combination of a few atoms (Duarte-Carvajalino and Sapiro, 2009).

In contrast to previous studies, we present a voxel-based sparse representation approach for automatic brain tumor segmentation from three-dimensional (3D) MRI images. It performs clustering analysis based on dictionary learning, where the features are extracted directly from MRI images. Due to the high similarity between healthy and non-healthy tissues, kernel trick is adopted to map the low-level features into a high-dimensional space. Finally, the performance of the proposed method is evaluated with a dataset from the brain tumor segmentation (BRATS) challenge of BRATS 2015.

2 Materials

The MRI dataset used in this study is obtained from BRATS 2015 challenges in conjunction with the MICCAI 2015 conference. The BRATS 2015 training dataset is available from the virtual skeleton database (VSD) (Kistler et al., 2013; Menze et al., 2015), which consists of multi-contrast magnetic resonance (MR) scans of 274 patients (220 high-grade tumors and 54 low-grade tumors), while the testing dataset contains 110 subjects with low-grade glioma (LGG) and high-grade glioma (HGG). For each patient, T1-weighted, contrast enhanced T1 (T1C) weighted, T2-weighted, and T2-FLAIR-weighted (fluid-attenuated inversion recovery) images are provided, which have been skull-stripped, linearly co-registered, and interpolated to 1-mm isotropic resolution.

The ground-truth images are segmented manually for five different classes by experienced specialists: one for necrosis, two for edema, three for non-enhancing tumor, four for enhancing tumor, and zero for everything else.

The BRATS challenge also provides an online evaluation system to upload and assess the performance of brain tumor segmentation. Dice score, positive predictive value (PPV), sensitivity, and kappa are used for the evaluation of complete tumor, tumor core, and enhancing tumor MRI images.

3 Methodology

3.1 Procedure of brain tumor segmentation

The brain tumor segmentation based on the voxel level can be considered as a classification problem. It includes image pre-processing, kernel dictionary learning, feature extraction, kernel sparse coding, and classification. However, it is not efficient enough to classify all the voxels of a full MRI volume

with size 240×240×150 in a normal personal computer using the sparse coding method. After pre-processing, we first deploy the *k*-means algorithm on T1-, T2-, and T2-FLAIR-weighted MRI images to cluster all the voxels and select the probable region R1 (label 0+1+2+3+4) that fully covers the tumor area and excludes the majority of normal tissues. Then the *k*-means clustering algorithm is performed on contrast enhanced T1- and T1C-weighted MRI images to determine the probable enhancing tumor region R4 (label 4) within R1, which is smaller than the truthful enhancing tumor region. In the end, regions R1–R4 are segmented precisely using the kernel sparse coding method based on four-modality MRI images. The schematic of the proposed method is shown in Fig. 1.

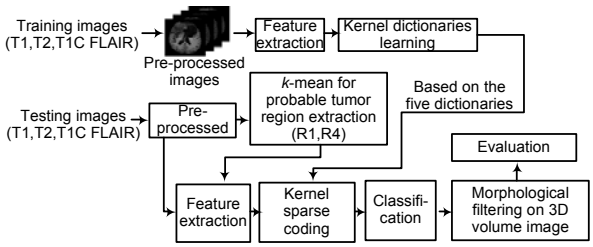

Fig. 1 Schematic of the proposed method for brain tumor segmentation

3.2 Pre-processing

Generally, pre-processing is an active part in MRI brain tumor segmentation for enhancing and correcting MRI images. In recent studies, denoising, skull stripping, and bias field correction for intensity inhomogeneity have been used commonly. The winner group (Tustison et al., 2013) in BRATS 2013 removed 1% of the highest and lowest intensities.

The pre-processing of our method includes three steps. First, we normalize MRI images to grayscale with the range of 0–255. Second, a 3D median filter with a kernel size of 3×3×3 is used to remove the noise. Third, a binary mask is created to select voxels of valid brain tissues whose intensities are larger than zero, to exclude the invalid data during matching.

3.3 Probable tumor region extraction

The intensity level of voxels in the complete tumor tissues is higher than those in the normal ones in FLAIR-weighted modalities, while it is the same as those in enhancing tumor tissues in T1C.

In this study, we perform the *k*-means clustering algorithm on T1-, T2-, and T2-FLAIR-weighted MRI images to detect the full tumor region roughly, which can cover the precise tumor area completely. This step is followed by morphological operations to fill in the holes and extend the edges within the probable area. Thereafter, the same algorithm is performed on T1 and T1C MRI images within the area obtained in the previous steps to obtain the probable enhancing tumor region, which is smaller than the truthful enhancing tumor region. Comparison of MRI images before and after pre-processing is shown in Fig. 2.

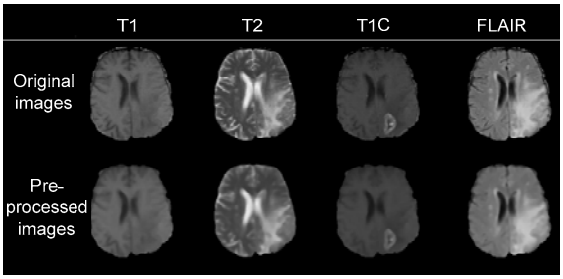

Fig. 2 Comparison of magnetic resonance imaging (MRI) images before and after pre-processing
FLAIR: fluid-attenuated inversion recovery

As shown in Fig. 3, *k*-means clustering divides the MRI data into five classes, and then each connected area of every class is marked. This can obtain *N* groups (clustering) in X1. Next, the class that has the maximum clustering center value is extracted in FLAIR-weighted MRI, which is labeled as R0 and can cover the tumor region. R1 is obtained through morphological filtering on R0. Though two classes X2 are clustered using T1- and T1C-weighted MRI

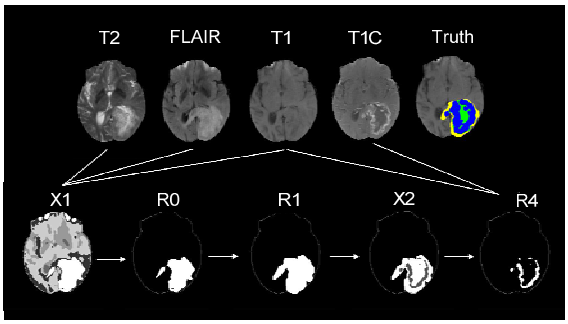

Fig. 3 Procedure of probable tumor region extraction
FLAIR: fluid-attenuated inversion recovery. Yellow region represents label 2, blue region represents label 4, and green region represents label 1+3. References to color refer to the online version of this figure

images based on the results of the first rough segmentation R1, the class that has the maximum clustering center is considered as the probable enhancing tumor area R4.

4 Kernel sparse representation with dictionary learning

4.1 Sparse coding

Sparse coding has been used widely in image understanding and signal processing such as image denoising (Hyvärinen et al., 1999; Elad and Aharon, 2006a, 2006b; Zeyde et al., 2012), compressive sensing (Bryt and Elad, 2008), image restoration (Mairal et al., 2008, 2009; Yang et al., 2009; Dong et al., 2011), and audio processing (Sivaram et al., 2010; Grosse et al., 2012), and can be interpreted as a sparse linear combination of a few basis vectors. The basic principle of sparse representation is represented as

$$\boldsymbol{y} = D \cdot \boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{y} \in \mathbb{R}^m$ is an input vector, $D \in \mathbb{R}^{m\times k}$ is a set of basis vectors which is solved as a dictionary-learning problem, and $\boldsymbol{n} \in \mathbb{R}^m$ is a vector of white noise. Sparse representation is to solve $\boldsymbol{y}\approx D\cdot\boldsymbol{x}$ and obtain sparse coding $\boldsymbol{x} \in \mathbb{R}^k$. The objective function is written as

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y} - D\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_1, \tag{2}$$

where $\|\cdot\|_1$ is the ℓ_1 norm and equals $\sum_{i=1}^{k}|\boldsymbol{x}[i]|$. It is also a convex problem called 'basis pursuit' (Chen et al., 2001).

Given a data sample $\boldsymbol{y} \in \mathbb{R}^{m\times n}$, function (2) can be rewritten as

$$\min_{\boldsymbol{D},\boldsymbol{X}} \frac{1}{2}\|\boldsymbol{Y} - D\boldsymbol{X}\|_{\mathrm{F}}^2 + \lambda\sum_{i=1}^{n}\|\boldsymbol{x}_i\|_1, \tag{3}$$

where $\boldsymbol{X} \in \mathbb{R}^{k\times n}$, and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm, which can be calculated as a convex optimization problem.

4.2 Kernel sparse coding

Sparse coding is a linear model, but not all the features are discriminative in the linear space. Therefore, the nonlinear similarity between the atoms can be taken by using kernel trick. This can be solved by mapping the data samples in the original space $\mathbb{R}$ into a higher-dimensional feature space $\psi(\cdot)$ with a nonlinear transform $\psi: \boldsymbol{y} \xrightarrow{\psi} \psi(\boldsymbol{y})$. Finally, we can obtain the kernel sparse representation for Eq. (1) as

$$\psi(\boldsymbol{y}) = \psi(D)\cdot\boldsymbol{x} + \boldsymbol{n}. \tag{4}$$

The related sparse constraint is

$$\min_{\boldsymbol{D},\boldsymbol{X}} \frac{1}{2}\|\psi(\boldsymbol{Y}) - \psi(D)\boldsymbol{X}\|_2^2 + \lambda\sum_{i=1}^{n}\|\boldsymbol{x}_i\|_1. \tag{5}$$

To make the kernel function work, the kernel function should be symmetric positive semidefinite according to Mercer's theorem (Cristianini and Shawe-Taylor, 2000). A Hilbert space with reproducing kernel $K(\cdot)$ will be deployed since the similarity between two sample vectors ($\boldsymbol{y}_1$ and $\boldsymbol{y}_2$) can be measured by inner product, noting that $K(y_i, y_j)=\psi(\boldsymbol{y}_i)^{\mathrm{T}}\psi(\boldsymbol{y}_j)$. In this study, the form of radial basis function (RBF) kernel $K(\boldsymbol{y}_i, \boldsymbol{y}_j) = \exp\left(-\gamma\|\boldsymbol{y}_i - \boldsymbol{y}_j\|_2^2\right)$ is used to obtain nonlinear similarity.

4.3 Dictionary designation

Dictionary learning can be considered as learning sets of over-complete basis to represent data efficiently. In this subsection, the kernel dictionary learning methods based on a clustering learning algorithm (He et al., 2009; Thiagarajan et al., 2011) can be presented with constraint $\|\boldsymbol{X}_i\|_0\leq 1$. The procedure is based on the theory that every atom can be expressed by a linear combination of the other atoms in kernel space $\psi(\cdot)$, attempting to fit K 1D subspaces to the data samples. The clustering procedure can be divided into two phases, clustering assignment and clustering update. It can be solved by

$$\min_{\boldsymbol{X}} \|\psi(\boldsymbol{Y}) - \psi(D)\boldsymbol{X}\|_{\mathrm{F}}^2 \quad \text{s.t. } \|\boldsymbol{x}_i\| \leq 1. \tag{6}$$

Let $\psi(\boldsymbol{Y})=(\psi(\boldsymbol{y}_1), \psi(\boldsymbol{y}_2), \ldots, \psi(\boldsymbol{y}_n))$ be the feature vector matrix of n data samples, and $\psi(D)=(\psi(\boldsymbol{d}_1),$

$\psi(\boldsymbol{d}_2)$, …, $\psi(\boldsymbol{d}_K)$) be an $m\times K$ kernel dictionary matrix, which consists of K atoms and each atom represents the basis of every voxel in the MRI images, where $K<<n$.

To compute sparse coding $\boldsymbol{X}=(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)\in \mathbb{R}^{k\times n}$ s.t. $\|\boldsymbol{x}_i\|_0\leq 1$ (i=1, 2, …, n), dictionary D should be trained first. In clustering assignment, dictionary D is initiated by random selection from data samples $\boldsymbol{Y}$ with a large size K. Define K membership sets $\{\mathbb{C}_k\}_{k=1}^{K}$, where $\mathbb{C}_k$ contains the vectors of $\boldsymbol{Y}$ that belongs to the kth cluster. Then compute the similarity between $\psi(\boldsymbol{y}_i)$ and each atom of $\psi(D)$ as $k(\boldsymbol{y}_i, D)=\psi(\boldsymbol{y}_i)^{\mathrm{T}}\psi(D)$. If the kth atom of $\psi(D)$ results in the maximum absolute correlation, index i is placed in set $\mathbb{C}_k$. In the clustering update phase, each atom $\boldsymbol{d}_k$ in D is related to a cluster center of $\mathbb{C}_k$. Assuming that P elements are contained in $\mathbb{C}_k$, updating dictionary D can be resolved by solving

$$\max_{\boldsymbol{d}_k}\sum_{i=1}^{p}\psi(\boldsymbol{d}_k)^{\mathrm{T}}\psi(\mathbb{C}_{k(i)}), \tag{7}$$

which will result in a maximum intra-class similarity. Now we can resolve sparse coding by solving

$$\min_{\boldsymbol{X}_c}\left\|\psi(\boldsymbol{Y}_c)-\psi(\boldsymbol{D})\boldsymbol{X}_c\right\|_{\mathrm{F}}^2 \text{ s.t. } \left\|\boldsymbol{x}_i\right\|\leq 1, \tag{8}$$

where c is the number of the clustering centers.

The dictionary learning procedure is shown in Algorithm 1.

Algorithm 1 Kernel clustering dictionary learning

Input: Given a dataset $\boldsymbol{Y}=\{\boldsymbol{y}_i\in\mathbb{R}^{m\times 1}\}_{i=1}^{n}$, set the number of dictionary atoms k.

Output: dictionary $D\in\mathbb{R}^{m\times k}$.

Initialize: Select k objects from $\boldsymbol{Y}$ as dictionary atoms randomly.

Initialize dictionary $D\in\mathbb{R}^{m\times k}$ with these k atoms;
Set $\{\mathbb{C}_k\}_{k=1}^{K}$, where $\mathbb{C}_k$ contains the vectors of all data samples $\boldsymbol{Y}$ that belong to the kth cluster;
do loop
for i=1 to n
Compute similarity $k(\boldsymbol{y}_i, D)=\psi(\boldsymbol{y}_i)^{\mathrm{T}}\psi(D)$;
if the kth atom of $k(\boldsymbol{y}_i, D)$ is the maximum
then set the index i to $\mathbb{C}_k$;
end for
Update the dictionary atom $\boldsymbol{d}_k$:
$\max_{\boldsymbol{d}_k}\sum_{i=1}^{p}\psi(\boldsymbol{d}_k)^{\mathrm{T}}\psi(\mathbb{C}_{k(i)})$;
until no change in dictionary D.

4.4 Implementation for precise segmentation

The pre-processed MRI images are segmented with kernel sparse representation. The procedure is shown as follows:

Step 1: A patch of 3×3×3 around the voxel is extracted as feature vectors, combining voxel intensity and local gray intensity information.

Step 2: RBF $K(\boldsymbol{Y}_i,\boldsymbol{Y}_j)=\exp\left(-0.3\left\|\boldsymbol{Y}_i-\boldsymbol{Y}_j\right\|_2^2\right)$ is deployed on the feature vectors to map the feature vectors from the original space into a high-dimensional space.

Step 3: A clustering method based on sparse representation described above is used to learn the five dictionaries (D_0 refers to normal tissues, D_1 refers to necrosis, D_2 refers to edema, D_3 refers to non-enhancing tumor, and D_4 refers to enhancing tumor) that are performed only once.

Step 4: Kernel sparse codes $\boldsymbol{x}_i^0,\boldsymbol{x}_i^1,\boldsymbol{x}_i^2,\boldsymbol{x}_i^3,\boldsymbol{x}_i^4$ for each voxel $\boldsymbol{y}_i$ are computed for further classification. They depend on the approximation differences among different dictionaries.

Step 5: A linear classifier is used to train and classify the voxels based on the reconstruction error:

$$\begin{aligned}\textbf{Feature}_i=\Big[&\left|\boldsymbol{y}_i-D_0\cdot\boldsymbol{x}_i^0\right|_2^2,\left|\boldsymbol{y}_i-D_1\cdot\boldsymbol{x}_i^1\right|_2^2,\\&\left|\boldsymbol{y}_i-D_2\cdot\boldsymbol{x}_i^2\right|_2^2,\left|\boldsymbol{y}_i-D_3\cdot\boldsymbol{x}_i^3\right|_2^2,\left|\boldsymbol{y}_i-D_4\cdot\boldsymbol{x}_i^4\right|_2^2\Big].\end{aligned} \tag{9}$$

The model for classification can be defined by a five-fold cross-validation dataset during the training phase.

4.5 Post-processing

There are still some voxels misclassified in the segmentation results, leaving one or more holes within the tumor region. Due to its single connected area in a tumor region, a post-processing technique is applied to fill in the area among multiple connected areas. Morphological operation is performed on the segmented image.

5 Evaluation

We evaluate the proposed brain tumor segmentation method with the BRATS 2015 training and testing datasets. To assess the performance of our segmentation, the segmentation results are uploaded to the online evaluation system, where the dice score, PPV, sensitivity, and kappa are used. Their definitions are listed as follows:

$$\mathrm{Dice}(G,S)=\frac{2|G_1\cap S_1|}{|G_1|+|S_1|}, \tag{10}$$

$$\mathrm{PPV}(G,S)=\frac{|G_1\cap S_1|}{|S_1|}, \tag{11}$$

$$\mathrm{Sensitivity}(G,S)=\frac{|G_1\cap S_1|}{|G_1|}, \tag{12}$$

$$\mathrm{Kappa}(G,S)=\frac{\dfrac{|G\oplus S|}{|G_0|+|G_1|}}{1-\dfrac{|G_0|\cdot|S_0|+|G_1|\cdot|S_1|}{(|G_0|+|G_1|)(|S_0|+|S_1|)}}-\frac{\dfrac{|G_0|\cdot|S_0|+|G_1|\cdot|S_1|}{(|G_0|+|G_1|)(|S_0|+|S_1|)}}{1-\dfrac{|G_0|\cdot|S_0|+|G_1|\cdot|S_1|}{(|G_0|+|G_1|)(|S_0|+|S_1|)}}, \tag{13}$$

where G indicates the ground-truth labels, S the predicted results, G_0 and G_1 the ground-truth labels for negative and positive, respectively, and S_0 and S_1 the predicted labels for negative and positive, respectively.

Three different tumor sub-compartments are provided for evaluation of the segmentations:

Region 1: complete tumor (labels 1+2+3+4);

Region 2: tumor core (labels 1+3+4);

Region 3: enhancing tumor (label 4).

6 Experiments and results

We conducted our experiments on the BRATS 2015 dataset. 50 HGG and 20 LGG subjects were adopted to develop our algorithm. The remaining 170 HGG subjects, 34 LGG subjects, and the testing dataset containing 110 subjects were kept for evaluation.

The proposed method of brain tumor segmentation was implemented in a personnel computer (Matlab 2015, Windows 7 64-bit system, 4 GHz I-7 CPU, 32 GB RAM).

6.1 Influence of patch size

The patch size is related to nonlinear similarities in different feature vectors and the number of dictionary atoms. Generally, the number of atoms K would be larger than that of the dimensions of feature vectors. Since the patch size of 3×3×3 has good performance and the patch sizes of 5×5×5 and 7×7×7 are not efficient or discriminative in our experiment, in the whole study we used the patch size of 3×3×3.

6.2 Influence of the number of dictionary atoms

In the dictionary learning phase, larger image patches and a larger number of dictionary atoms K lead to a higher computation cost. To study the relationship between accuracy and the number of atoms, synthetic datasets that have less variability in intensity and fewer artifacts than real images were used, while rough extraction was not included in this phase.

In the experiment, the patch size was 3×3×3. To explore the optimal value of K, a series of the number of dictionary atoms $K\in\{2, 4, 8, 16, 32, 64, 128, 256, 512\}$ were tested.

As shown in Fig. 4, the accuracy grows noticeably with the increase of dictionary atomic number K until K reaches 128, where the accuracy grows scarcely thereafter. Therefore, K=128 was chosen as the optimal value in our experiments.

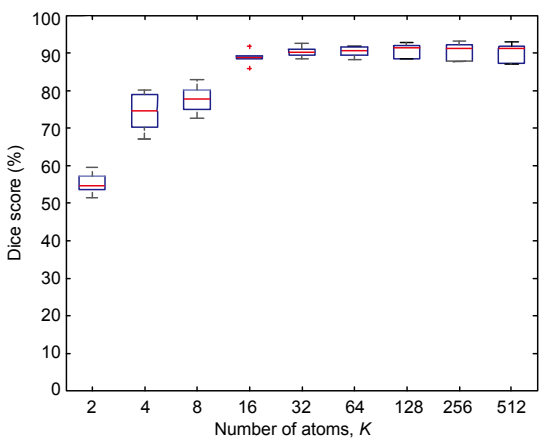

Fig. 4 Relationship between the number of dictionary atoms *K* and segmentation accuracy

Results were obtained using a patch size of 3×3×3 for a complete tumor region

7 Segmentation results

Table 1 summarizes the performance of the proposed method with various evaluation metrics using the BRATS 2015 training datasets.

Table 2 summarizes the performance of our method and the published rankings of the BRATS competition with the state-of-the-art methods.

The segmentation results using kernel sparse coding are shown in Fig. 5, where some slices of four subjects including the original MRI images, the ground truth, and corresponding segmentation results are shown.

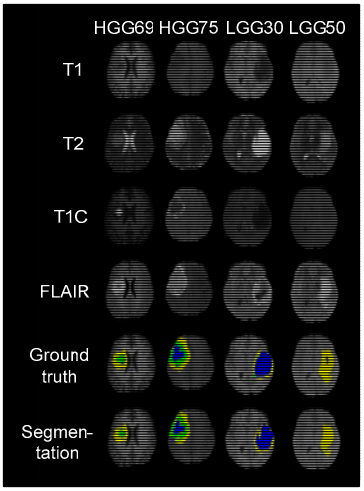

Fig. 5 Subjects from the brain tumor segmentation (BRATS) 2015 training dataset
FLAIR: fluid-attenuated inversion recovery. The color coding is: yellow region refers to edema, green region refers to necrosis and non-enhancing tumor, and blue region refers to enhancing tumor. References to color refer to the online version of this figure

8 Discussions

In this paper, we have proposed a brain tumor segmentation method based on kernel sparse coding. This method includes MRI image dataset preprocessing, feature extraction, dictionary learning, image sparse representation via sparse coding, and linear discrimination. Compared with other segmentation methods, such as region growing and deep learning (Rathi and Palani, 2015), the proposed method is competitive to the other groups in the (BRATS) challenge. Besides, to finish a fully automatic brain tumor segmentation in a personal computer, our method needs only about 40 s. It can achieve high segmentation accuracy with less time.

Generally, it is difficult for the conventional segmentation methods to achieve good results in brain tumor segmentation (Liu et al., 2014; Menze et al., 2015). The state-of-the-art methods are usually based on the pixel level, and can also be considered as a clustering or classification problem (Salman Al-Shaikhli et al., 2015).

Because of the spatial volume effect in MRI images, the performance of pixel-level methods based on information of 2D images is limited. Some researchers have used the volumetric (3D) image

Table 1 Average results of 204 subjects of the brain tumor segmentation (BRATS) 2015 training datasets for validation, including 170 HGG and 34 LGG

Grading	Dice			PPV			Sensitivity			Kappa
	Complete	Core	Enhancing	Complete	Core	Enhancing	Complete	Core	Enhancing	
HGG	0.85	0.75	0.82	0.96	0.73	0.82	0.74	0.80	0.83	1.00
LGG	0.79	0.63	0.71	0.84	0.59	0.78	0.81	0.77	0.68	1.00

HGG: high-grade glioma; LGG: low-grade glioma; PPV: positive predictive value

Table 2 Ranking of the brain tumor segmentation (BRATS) challenge with state-of-the-art methods

User	Dice			PPV			Sensitivity		
	Complete	Core	Enhancing	Complete	Core	Enhancing	Complete	Core	Enhancing
Oskar	0.84	0.66	0.39	0.84	0.70	0.47	0.85	0.72	0.43
Sergio	0.87	0.73	0.68	0.89	0.74	0.72	0.86	0.77	0.70
S. Reza	0.81	0.66	0.71	0.95	0.82	0.78	0.73	0.61	0.75
Dongjin	0.86	0.79	0.59	0.88	0.84	0.60	0.86	0.81	0.63
Axel	0.85	0.74	0.68	0.85	0.74	0.62	0.85	0.78	0.77
Proposed method	0.83	0.69	0.58	0.84	0.76	0.60	0.82	0.80	0.65

PPV: positive predictive value

information in segmentation. Fletcher-Heath et al. (2001) and Wu et al. (2012) used 3D connected components to build the tumor shape. The conditional random fields provide a way to integrate spatial information into the clustering or classification process (Held et al., 1997; Ruan and Bloyet, 2000; Ahmadvand and Daliri, 2015).

Thiagarajan et al. (2014) developed a *k*-line clustering algorithm for sparse coding in brain tumor segmentation combining location and intensity features. Salman Al-Shaikhli et al. (2015) adopted the *k*-means-generating spare vector dictionary (K-SVD) method for brain tumor classification, using topological and texture features to learn the dictionary. Kong et al. (2016) presented a denoising method based on sparse representation for 3D diffusion tensor imaging by learning adaptive dictionary with the context redundancy between neighbor slices, and evaluated it on both simulated and real datasets. Nasir et al. (2014) used sparse coding methods to not only classify brain tumor into eight different categories but also detect their existence. Tong et al. (2013) proposed a fixed discriminative dictionary learning for segmentation (F-DDLS) strategy to learn a dictionary and perform segmentation online using discriminative dictionary learning and sparse coding techniques.

In terms of segmentation performance, the following factors should be taken into account in brain tumor segmentation: The pre-processing techniques are important in the preparation of MRI images, because they can reduce noise and correct non-uniformity within each MRI image. In our experiments, we used a 3D median filter on the input four-modality MRI images to enhance the segmentation results. The samples for training dictionary also have an impact on accuracy. The samples with little or no noise could achieve a high accuracy. Conversely, selecting samples randomly from multiple MRI image datasets would achieve unbiased samples. This is an alternative trick to obtain a strong dictionary.

The choice of features has a great influence on the segmentation results. Some studies have shown that texture and other high-order statistical features can improve the performance of segmentation (Sachdeva et al., 2013; Sompong and Wongthanavasu, 2014). In our future study, the proposed method can also be extended by using high-order statistical features instead of performing the dictionary learning algorithm on the voxels.

There is also a potential linear correction among different modalities of MRI images. Some researchers (Dvořák and Menze, 2015; Juan-Albarracin et al., 2015) focused on the fusion of multiple modalities in T1-, T2-, T1C-, and T2-FLAIR-weighted MRI images, aiming at improving the expression ability of tumor characteristics.

The dictionary learning method using intensity of the voxel and its neighbor information has improved the results by kernel tricks, considering the nonlinear similarity, whereas the simple linear discrimination has difficulty in classifying the voxels between healthy and pathological tissues in some special cases. Other classifiers can also be used, such as extremely randomized trees and support vector machine (SVM).

It is a promising method to improve the segmentation performance by enhancing the expression ability of effective features in the future. We can extend the proposed method to consider not only the relationship between multimodal MRI images but also the correlation in labels.

9 Conclusions

Fully automatic brain tumor segmentation can improve the diagnosis efficiency and reduce the job intensity of physicians. It is an important application of computer-aided diagnosis (CAD). The segmentation based on the pixel level has been considered as a classification problem. Thus, the feature choice, expression, and mapping can influence the classification capabilities. The proposed method extracts the nonlinear features of MRI images through kernel dictionary learning and sparse coding performed on the feature vectors. In the end, morphological processing has been used to fill in the area among multiple connected components to improve the segmentation quality. Experiment results suggested that the proposed method has better capacity and higher segmentation accuracy with low computation cost.

Acknowledgements

The authors appreciate Bjoern MENZE et al. for sharing the data.

References

Ahmadvand A, Daliri MR, 2015. Improving the runtime of MRF based method for MRI brain segmentation. *Appl Math Comput*, 256:808-818. https://doi.org/10.1016/j.amc.2015.01.053

Atkins MS, Mackiewich BT, 1998. Fully automatic segmentation of the brain in MRI. *IEEE Trans Med Imag*, 17(1):98-107. https://doi.org/10.1109/42.668699

Bryt O, Elad M, 2008. Compression of facial images using the K-SVD algorithm. *J Vis Commun Imag Represent*, 19(4): 270-282. https://doi.org/10.1016/j.jvcir.2008.03.001

Chen SS, Donoho DL, Saunders MA, 2001. Atomic decomposition by basis pursuit. *SIAM Rev*, 43(1):129-159. https://doi.org/10.1137/S003614450037906X

Chong VFH, Zhou JY, Khoo JBK, et al., 2004. Tongue carcinoma: tumor volume measurement. *Int J Radiat Oncol Biol Phys*, 59(1):59-66. https://doi.org/10.1016/j.ijrobp.2003.09.089

Cristianini N, Shawe-Taylor J, 2000. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, p.189.

Dong WS, Zhang L, Shi GM, 2011. Centralized sparse representation for image restoration. Proc IEEE Int Conf on Computer Vision, p.1259-1266. https://doi.org/10.1109/ICCV.2011.6126377

Duarte-Carvajalino JM, Sapiro G, 2009. Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization. *IEEE Trans Image Process*, 18(7):1395-1408. https://doi.org/10.1109/TIP.2009.2022459

Dvořák P, Menze B, 2015. Structured prediction with convolutional neural networks for multimodal brain tumor segmentation. Proc Multimodal Brain Tumor Image Segmentation Challenge, p.13-24.

Elad M, Aharon M, 2006a. Image denoising via learned dictionaries and sparse representation. Proc IEEE Computer Society Conf on Computer Vision and Pattern Recognition, p.895-900. https://doi.org/10.1109/CVPR.2006.142

Elad M, Aharon M, 2006b. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans Image Process*, 15(12):3736-3745. https://doi.org/10.1109/TIP.2006.881969

Fletcher-Heath LM, Hall LO, Goldgof DB, et al., 2001. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artif Intell Med*, 21(1-3): 43-63. https://doi.org/10.1016/S0933-3657(00)00073-7

Gibbs P, Buckley DL, Blackband SJ, et al., 1996. Tumour volume determination from MR images by morphological segmentation. *Phys Med Biol*, 41(11):2437-2446. https://doi.org/0.1088/0031-9155/41/11/014

Grosse R, Raina R, Kwong H, et al., 2012. Shift-invariance sparse coding for audio classification. http://arxiv.org/abs/1206.5241

He ZS, Cichocki A, Li YQ, et al., 2009. K-hyperline clustering learning for sparse component analysis. *Signal Process*, 89(6):1011-1022. https://doi.org/10.1016/j.sigpro.2008.12.005

Held K, Kops ER, Krause BJ, et al., 1997. Markov random field segmentation of brain MR images. *IEEE Trans Med Imag*, 16(6):878-886. https://doi.org/10.1109/42.650883

Hyvärinen A, Hoyer P, Oja E, 1999. Image denoising by sparse code shrinkage. Proc Intelligent Signal Processing, p.1-31.

Juan-Albarracin J, Fuster-Garcia E, Manjon JV, et al., 2015. Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. *PLoS ONE*, 10(5):e0125143. https://doi.org/10.1371/journal.pone.0125143

Juergens KU, Seifarth H, Range F, et al., 2008. Automated threshold-based 3D segmentation versus short-axis planimetry for assessment of global left ventricular function with dual-source MDCT. *Am J Roentgenol*, 190(2): 308-314. https://doi.org/10.2214/AJR.07.2283

Kistler M, Bonaretti S, Pfahrer M, et al., 2013. The virtual skeleton database: an open access repository for biomedical research and collaboration. *J Med Int Res*, 15(11): e245. https://doi.org/10.2196/jmir.2930

Kong YY, Li YJ, Wu JS, et al., 2016. Noise reduction of diffusion tensor images by sparse representation and dictionary learning. *BioMed Eng*, 15:5. https://doi.org/10.1186/s12938-015-0116-3

Liu J, Li M, Wang JX, et al., 2014. A survey of MRI-based brain tumor segmentation methods. *Tsinghua Sci Technol*, 19(6):578-595. https://doi.org/10.1109/TST.2014.6961028

Mairal J, Elad M, Sapiro G, 2008. Sparse representation for color image restoration. *IEEE Trans Imag Process*, 17(1):53-69. https://doi.org/10.1109/TIP.2007.911828

Mairal J, Bach F, Ponce J, et al., 2009. Non-local sparse models for image restoration. Proc 12th Int Conf on Computer Vision, p.2272-2279. https://doi.org/10.1109/ICCV.2009.5459452

Menze BH, Jakab A, Bauer S, et al., 2015. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imag*, 34(10):1993-2024. https://doi.org/10.1109/TMI.2014.2377694

Mittelhäußer G, Kruggel F, 1995. Fast segmentation of brain magnetic resonance tomograms. Proc 1st Int Conf on Computer Vision, Virtual Reality and Robotics in Medicine, p.237-241. https://doi.org/10.1007/978-3-540-49197-2_27

Nasir M, Baig A, Khanum A, 2014. Brain tumor classification in MRI scans using sparse representation. In: Elmoataz A, Lezoray O, Nouboud F, et al. (Eds.), Image and Signal Processing. Springer, Cham, p.629-637. https://doi.org/10.1007/978-3-319-07998-1_72

Olabarriaga SD, Smeulders AWM, 2001. Interaction in the segmentation of medical images: a survey. *Med Imag Anal*, 5(2):127-142.

https://doi.org/10.1016/S1361-8415(00)00041-4

Prastawa M, Bullitt E, Ho S, et al., 2004. A brain tumor segmentation framework based on outlier detection. *Med Imag Anal*, 8(3):275-283. https://doi.org/10.1016/j.media.2004.06.007

Rathi VPGP, Palani S, 2015. Brain tumor detection and classification using deep learning classifier on MRI images. *Res J Appl Sci Eng Technol*, 10(2):177-187. https://doi.org/10.19026/rjaset.10.2570

Rousson M, Lenglet C, Deriche R, et al., 2004. Level set and region based surface propagation for diffusion tensor MRI segmentation. In: Sonka M, Kakadiaris IA, Kybic J (Eds.), Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis. Springer Berlin Heidelberg, p.123-134. https://doi.org/10.1007/978-3-540-27816-0_11

Ruan S, Bloyet D, 2000. MRF models and multifractal analysis for MRI segmentation. Proc 5th Int Conf on Signal Processing, p.1259-1262. https://doi.org/10.1109/ICOSP.2000.891775

Sachdeva J, Kumar V, Gupta I, et al., 2013. Segmentation, feature extraction, and multiclass brain tumor classification. *J Dig Imag*, 26(6):1141-1150. https://doi.org/10.1007/s10278-013-9600-0

Salman Al-Shaikhli SD, Yang MY, Rosenhahn B, 2015. Brain tumor classification and segmentation using sparse coding and dictionary learning. *BioMed Tech (Berl)*, 61(4): 413-429. https://doi.org/10.1515/bmt-2015-0071

Salman YM, Assal MA, Badawi AM, et al., 2005. Validation techniques for quantitative brain tumors measurements. Proc 27th Annual Int Conf of the Engineering in Medicine and Biology Society, p.7048-7051. https://doi.org/10.1109/IEMBS.2005.1616129

Shanthi KJ, Kumar MS, 2007. Skull stripping and automatic segmentation of brain MRI using seed growth and threshold techniques. Proc Int Conf on Intelligent and Advanced Systems, p.422-426. https://doi.org/10.1109/ICIAS.2007.4658421

Sivaram GSVS, Nemala SK, Elhilali M, et al., 2010. Sparse coding for speech recognition. Proc IEEE Int Conf on Acoustics, Speech and Signal Processing, p.4346-4349. https://doi.org/10.1109/ICASSP.2010.5495649

Sompong C, Wongthanavasu S, 2014. MRI brain tumor segmentation using GLCM cellular automata-based texture feature. Proc Int Computer Science and Engineering Conf, p.192-197. https://doi.org/10.1109/ICSEC.2014.6978193

Taheri S, Ong SH, Chong VFH, 2010. Level-set segmentation of brain tumors using a threshold-based speed function. *Imag Vis Comput*, 28(1):26-37. https://doi.org/10.1016/j.imavis.2009.04.005

Thiagarajan JJ, Ramamurthy KN, Spanias A, 2011. Optimality and stability of the K-hyperline clustering algorithm. *Patt Recogn Lett*, 32(9):1299-1304. https://doi.org/10.1016/j.patrec.2011.03.005

Thiagarajan JJ, Ramamurthy KN, Rajan D, et al., 2014. Kernel sparse models for automated tumor segmentation. *Int J Artif Intell Tools*, 23(3):1460004. https://doi.org/10.1142/S0218213014600045

Tong T, Wolz R, Coupé P, et al., 2013. Segmentation of MR images via discriminative dictionary learning and sparse coding: application to hippocampus labeling. *NeuroImage*, 76:11-23. https://doi.org/10.1016/j.neuroimage.2013.02.069

Tustison N, Wintermark M, Durst C, et al., 2013. ANTs and arboles. Proc NCI-MICCAI BRATS, p.47-50.

Wang ZZ, Vemuri BC, 2004. Tensor field segmentation using region based active contour model. In: Pajdla T, Matas J (Eds.), Computer Vision-ECCV. Springer Berlin Heidelberg, p.304-315. https://doi.org/10.1007/978-3-540-24673-2_25

Wong K, 2005. Medical image segmentation: methods and applications in functional imaging. In: Suri JS, Wilson DL, Laxminarayan S (Eds.), Handbook of Biomedical Image Analysis, Volume II: Segmentation Models Part B. Springer, Boston, US, p.111-182. https://doi.org/10.1007/b104806

Wu P, Xie K, Zheng Y, et al., 2012. Brain tumors classification based on 3D shape. In: Jin D, Lin S (Eds.), Advances in Future Computer and Control Systems. Springer, Berlin, p.277-283. https://doi.org/10.1007/978-3-642-29390-0_45

Yang JC, Yu K, Gong YH, et al., 2009. Linear spatial pyramid matching using sparse coding for image classification. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.1794-1801. https://doi.org/10.1109/CVPR.2009.5206757

Zeyde R, Elad M, Protter M, et al., 2012. On single image scale-up using sparse-representations. In: Boissonnat J, Chenin P, Cohen A, et al. (Eds.), Curves and Surfaces. Springer, Berlin, p.711-730. https://doi.org/10.1007/978-3-642-27413-8_47