PLOS ONE

Content-Based Image Retrieval Using Spatial Layout Information in Brain Tumor T1-Weighted Contrast-Enhanced MR Images

Meiyan Huang[1], Wei Yang[1], Yao Wu[1], Jun Jiang[1], Yang Gao[1], Yang Chen[2], Qianjin Feng[1]*, Wufan Chen[1], Zhentai Lu[1]

1 School of Biomedical Engineering, Southern Medical University, Guangzhou, Guangdong, China, **2** Laboratory of Image Science and Technology, Southeast University, the Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing, China

Abstract

This study aims to develop content-based image retrieval (CBIR) system for the retrieval of T1-weighted contrast-enhanced MR (CE-MR) images of brain tumors. When a tumor region is fed to the CBIR system as a query, the system attempts to retrieve tumors of the same pathological category. The bag-of-visual-words (BoVW) model with partition learning is incorporated into the system to extract informative features for representing the image contents. Furthermore, a distance metric learning algorithm called the Rank Error-based Metric Learning (REML) is proposed to reduce the semantic gap between low-level visual features and high-level semantic concepts. The effectiveness of the proposed method is evaluated on a brain T1-weighted CE-MR dataset with three types of brain tumors (i.e., meningioma, glioma, and pituitary tumor). Using the BoVW model with partition learning, the mean average precision (mAP) of retrieval increases beyond 4.6% with the learned distance metrics compared with the spatial pyramid BoVW method. The distance metric learned by REML significantly outperforms three other existing distance metric learning methods in terms of mAP. The mAP of the CBIR system is as high as 91.8% using the proposed method, and the precision can reach 93.1% when the top 10 images are returned by the system. These preliminary results demonstrate that the proposed method is effective and feasible for the retrieval of brain tumors in T1-weighted CE-MR Images.

Introduction

In the medical field, digital images are produced every day and are used by radiologists to make diagnoses. However, searching for images with the same anatomic regions or similar-appearing lesions according to their visual contents in a large image dataset is difficult. Content-based image retrieval (CBIR) is presented as a possible and promising solution to indexing images with minimal human intervention [1]. In the medical field, CBIR mainly serves for two types of application: retrieval of the same anatomic regions [2–4] and retrieval of clinically relevant lesions (e.g., lesions of the same pathological category) [5–8]. Category retrieval is only used when lesions of the same pathological category are expected to be retrieved.

The current study aims to develop a CBIR system for retrieving T1-weighted contrast-enhanced MR (CE-MR) images that contain brain tumors of the same pathological category. The goal of the proposed CBIR system is to aid diagnosis for the given query images. Concretely, the most similar tumors with the same

pathological category in the dataset are returned when a tumor is sent to the system as a query. Users can then choose the most relevant images and access their associated diagnosis information to support the diagnosis for the current case. This study focuses on three types of brain tumors with high incidence rates in clinics: gliomas, meningiomas, and pituitary tumors, percentages of which are about 45%, 15%, and 15% of all brain tumors, respectively.

Several studies have developed CBIR systems for brain tumors in MR Images. Huang et al. [9] presented a diagnostic support tool for pediatric brain diseases using the segmented lesion as a query to retrieve the most similar lesions in the database. Dube et al. [10] proposed a methodology for the image retrieval of glioblastoma multiforme (GBM) and non-GBM tumors on MR Images. Moustakas et al. [11] presented a two-tier CBIR architecture for regions of interest (ROI) of brain MR Images; this architecture was evaluated against a dataset comprising T1, T2, and PD MR scans. As opposed to other MR modalities used in the previous studies, the category retrieval of brain tumors is performed on the T1-weighted CE-MR Images in this study.
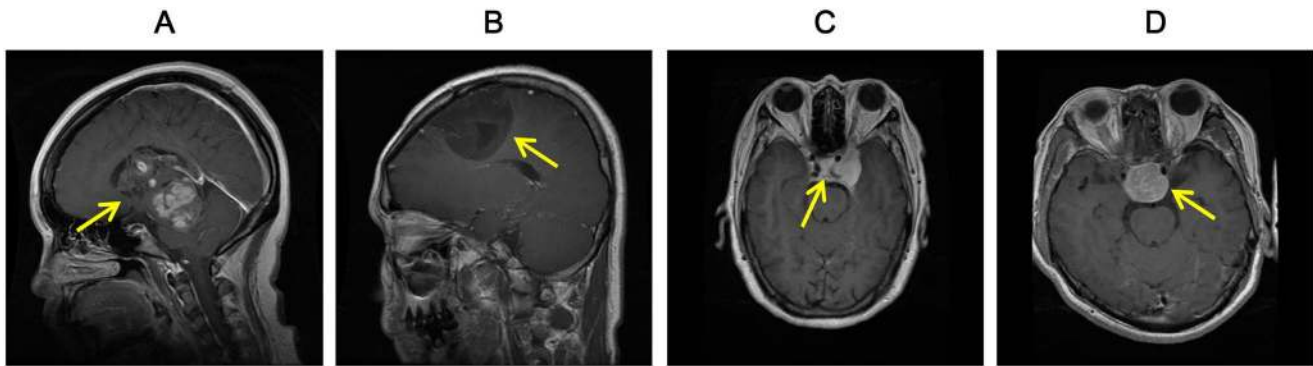
Figure 1. Four examples of brain tumors in T1-weighted CE-MR Images. The tumors are indicated by the yellow arrows in each image. (A) and (B) Gliomas in different subjects having dissimilar appearances. (C) A meningioma and (D) a pituitary tumor from different subjects showing a similar appearance.
doi:10.1371/journal.pone.0102754.g001

In general, feature extraction and distance metrics in the feature space are two crucial factors for CBIR. First, different appearances of brain tumors in MR Images demand more discriminative features. For instance, tumors with the same pathological category but from different patients or at different disease stages exhibit different appearances on the images [Figures 1(A) and 1(B)]. Conversely, lesions of different pathological categories may show visual similarity [Figures 1(C) and 1(D)]. Given that low-level visual features, such as color (intensity), texture [12–17], and shape [18,19], are not adequately discriminative to describe high-level semantic concepts [20], additional distinctive features are highly desirable [6,7]. Second, learned distance metrics, which are capable of reducing the semantic gap between visual features and semantic concepts [21], have been investigated intensively and widely used for CBIR [6,22,23].

Recent studies [5–7] have verified that image features constructed by bag-of-visual-words (BoVW) model are more discriminative than other commonly used intensity and texture features for medical images. We previously used a marginal information descriptor and region-specific BoVW to obtain structural information on tumors and their surrounding regions [6,7]. Tumor regions were partitioned into several sub-regions, and the image contents in each sub-region were represented using the BoVW-based approach. However, the tumor regions were outlined manually in these studies, which is not a trivial task and could cause intra- and inter-operator variations [24]. To promote

the feasibility of the proposed CBIR system, rectangular ROIs (Figure 2) are used in the current study instead of tumors contours for feature extraction; this method demands a more subtle extraction approach to integrate the spatial information into the BoVW-based image feature. Several other approaches have addressed this problem for object recognition tasks [25–28]. The spatial pyramid method [25], for example, is a representative of these methods. The spatial pyramid method symmetrically partitions the image into uniform cells at a series of resolution levels (i.e., 1×1, 2×2, and 4×4) and then combines all of the BoVW histograms in each cell to incorporate global and local information. With the help of spatial layout information to improve the discriminative power, the spatial pyramid method outperforms the basic BoVW model and has thus been employed in several studies [29,30]. However, compared with the symmetrical partitioning in the spatial pyramid method, the partition can be learned to make the feature more discriminative [28] when the contents of the images from a specific category are distributed with intrinsic regularity. Brain tumor T1-weghted CE-MR Images show that the layouts of the tumor, edema, and surrounding normal tissue have strong correlations [6]. For instance, meningiomas and pituitary tumors are homogeneous lesions [Figures 2(A) and 2(B)]; these two tumor types are usually not associated with edema. Meningiomas are usually adjacent to the skull, gray matter, and cerebrospinal fluid. Pituitary tumors are adjacent to the sphenoidal sinus, internal carotid arteries, and optic chiasma.
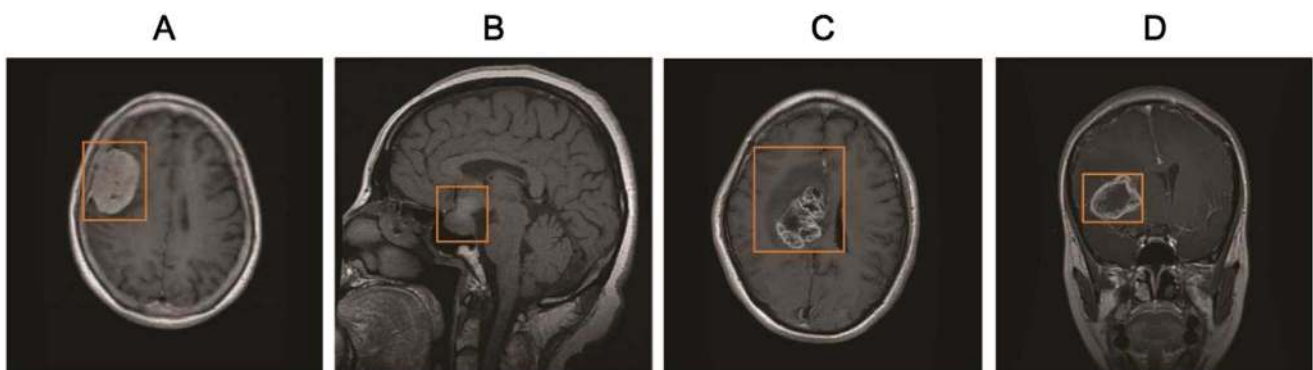

Figure 2. Illustrations of four typical brain tumors with rectangles as ROIs (orange lines) in T1-weighted CE-MR Images: (A) A meningioma located near the skull, (B) a pituitary tumor located near the sphenoidal sinus, (C) a glioma containing edema and necrosis, and (D) a glioma surrounded with edema.
doi:10.1371/journal.pone.0102754.g002

By contrast, gliomas are inhomogeneous lesions and usually surrounded by edema [Figures 2(C) and 2(D)]. Therefore, learning a spatial partition of the ROI in the current study is reasonable for obtaining discriminant features.

The contributions of the current study can be summarized as follows:

(1) A partition learning algorithm is proposed based on the concept that the best partition will lead to the largest difference among the BoVW histograms of the sub-regions. This idea allows the region with variable appearance to be partitioned and image contents in each partitioned sub-region to be consistent; thus, combinational histograms of sub-regions can bring more discriminative information. We present a novel objective function of the partition learning method, provide an optimization approach, and evaluate the effectiveness of the method in the CBIR of brain tumor T1-weighted CE-MR Images.

(2) A distance metric learning approach, called Rank Error-based Metric Learning (REML), is introduced to reduce the semantic gap between high-level semantic concepts and low-level visual features in the proposed CBIR system. A novel objective function that integrates rank error is proposed, and a stochastic gradient descent-based optimization strategy is presented to find the optimal solution of the objective function. REML can project image features to a low-dimensional feature space, where the learned distance is expected to reflect the differences between the semantic concepts.

Materials and Methods

2.1 Ethics Statement

This study was approved by the Ethics Committees of Nanfang Hospital and General Hospital, Tianjin Medical University. Patient records/information was anonymized and de-identified prior to analysis.

2.2 Image Data

At present, only a certain number of slices of brain T1-weighted CE-MR with a large slice gap, not 3-D volumes, are usually acquired and available for clinical practice in China. Considering the difficulty to build a 3D model using such sparse data, we opted to build a 2D image-based CBIR system for clinical applications. In our implementation, we selected slices from 3D images primarily based on the visible size of tumors. The slice containing the largest tumor size and its adjacent 1 to 10 slices were selected as the typical slices for constructing our dataset. Slices in different views (transverse, coronal, and sagittal) were processed together in the proposed CBIR system. For instance, when one slice in a specific view was fed into the system as a query, slices containing tumors with the same category may be expected to be returned even when they are in different views. We expect that the retrieval method would only work in realistic situations when a certain number of 2D slices in different views were acquired for clinical practice.

In this study, the brain T1-weighted CE-MR dataset was acquired from Nanfang Hospital, Guangzhou, China, and General Hospital, Tianjin Medical University, China, from 2005 to 2010. This dataset consisted of 3064 slices from 233 patients. Each slice is 512×512 pixels, and the pixel size is 0.49 mm×0.49 mm. Three types of brain tumors, namely, meningiomas, gliomas, and pituitary tumors, are apparent in the dataset (Table 1). For each patient, three experienced radiologists initially consulted the patient pathology report to obtain the pathology type and then labeled the images. Each radiologist dealt with all images independently. Afterward, the radiologists discussed and reached a consensus regarding the label of every tumor in each image. In the current study, two images containing tumors of the same category were defined to be relevant (similar); otherwise, they were considered irrelevant (dissimilar).

2.3 BoVW Model

The BoVW model has been widely used in image classification and CBIR [4,31–33]. It involves four steps. First, patches represented by the local descriptors are sampled in each ROI in an image of the given image dataset. Second, a dictionary is learned by a clustering algorithm, and each of the cluster centers is a visual word. Third, local descriptors of a ROI in a new image are quantized to the learned dictionary by a coding method. Finally, a BoVW histogram is constructed to represent the ROI by a pooling method.

In the BoVW framework, selections of the patch sampling method and the local descriptor are two basic tasks. As introduced in our previous study [7], raw patches sampled in a dense fashion are applicable for capturing discriminative information in brain tumor T1-weghted CE-MR Images. Therefore, in this study, we retain the use of raw patches as the local descriptor and sample these patches in a dense manner. After the patch sampling step, a dictionary can be learned from the sampled patches by a clustering algorithm. In this study, we used the k-means clustering algorithm

Table 1. Summary of the image dataset.

Tumor category	Number of patients	Number of slices	View	Number of slices
Meningiomas	82	708	Transverse	209
			Sagittal	231
			Coronal	268
Gliomas	89	1426	Transverse	494
			Sagittal	495
			Coronal	437
Pituitary tumors	62	930	Transverse	291
			Sagittal	320
			Coronal	319

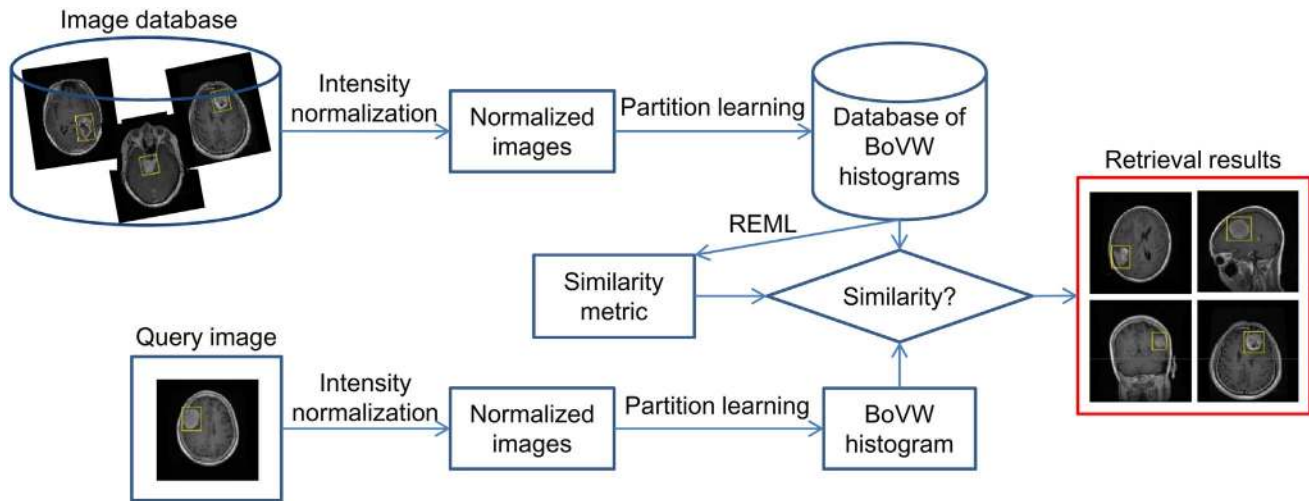doi:10.1371/journal.pone.0102754.t001

Figure 3. Flowchart of the proposed CBIR system.
doi:10.1371/journal.pone.0102754.g003

to learn V visual words in a dictionary D because of its simplicity and effectiveness.

For a ROI in a new image, the local descriptor of each pixel in this ROI can be mapped to D, and a code for a local descriptor is generated; this process is called coding. A basic coding method in the BoVW approach is done by assigning an image patch to one visual word, also called hard-assignment. In the hard-assignment method, a patch is mapped to a visual word that is most similar to the current patch. If hard-assignment is used for coding, each code has only one non-zero element. This coding method may ignore all ambiguity regarding the meaning of a patch. Unlike hard-assignment, soft-assignment assigns a degree of similarity to each of the visual words, and a small group of elements in a code can be non-zero. This coding method can help in modeling the inherent uncertainty of the image patch, while considering the continuous nature of the image patches [34]. Aside from soft-assignment, sparse coding, such as locality-constrained linear coding [35], is another alternative to hard-assignment.

To represent a ROI in an image, all coding results over m spatial regions of the ROI are aggregated, and an mV-dimensional vector $\boldsymbol{x} = (x_{11}, \cdots, x_{1V}, \cdots, x_{ij}, \cdots, x_{m1}, \cdots, x_{mV})$ is obtained. This process is called pooling. Each dimension of the pooled feature x_{ij} is achieved by taking the jth visual word of all codes $c_{\cdot j}$ in the specific spatial region R_i, and performing a predefined operator on the set of visual words. Therefore, x_{ij} can be defined as follows:

$$x_{ij} = op\left(X_{R_i}^{c \cdot j}\right),$$

where $X_{R_i}^{c \cdot j}$ is the set of visual words of codes $c_{\cdot j}$ in the spatial region R_i. Generally, $op(\cdot)$ is the operator for computing the statistics of the visual words under the p-norm; thus, x_{ij} can be calculated by:

$$x_{ij} = \frac{1}{|R_i|}\left(\sum\nolimits_{\alpha_{ij} \in X_{R_i}^{c \cdot j}} \alpha_{ij}^p\right)^{1/p},$$

when $p = 1$, $op(\cdot)$ is the average pooling, and when $p \to \infty$, $op(\cdot)$ is the max pooling.

In the basic BoVW model, the whole ROI is considered as a spatial region, and all codes in the spatial region are pooled together; thus, the spatial information is ignored.

2.4 Proposed Method

The proposed CBIR system mainly contains two components. (1) Offline feature extraction and (2) online image retrieval. In (1), a bounding box is first set to cover the tumor region in each image in the database as the ROI. The proposed partition learning method is used to partition the ROI into a set of sub-regions
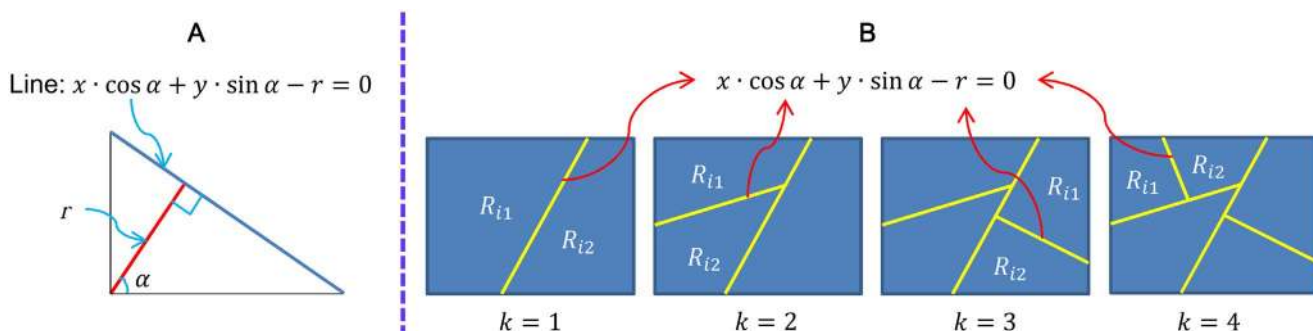

Figure 4. (A) Normal form of a linear equation. (B) Formation of the spatial partition patterns by successive partition of regions.
doi:10.1371/journal.pone.0102754.g004

Table 2. Algorithm 1.

Input:
Bounding boxes of brain tumor T1-weghted CE-MR Images $\varphi = \{I_i\}_{i=1}^{N}$; Maximum iteration number t_{max}.
Output:
$\{\alpha_j, r_j\}_{j=1}^{k}$ (Final solution of the partition learning method).
■ Initialize $\{\alpha_j, r_j\}_{j=1}^{k} = \{\pi/4, 50\}_{j=1}^{k}$ and the step sizes $\delta_\alpha = 0.01$ and $\delta_r = 1$.
■ For $j = 1, \cdots, k$
Find the region with maximal area among k regions.
For $t = 1, \ldots, t_{max}$
■ Initialize the parameters:
$\alpha_1 = \alpha_j + \delta_\alpha, \alpha_2 = \alpha_j - \delta_\alpha, r_1 = r_j + \delta_r, r_2 = r_j - \delta_r$
■ Calculate $h(\alpha_1, r), h(\alpha_2, r), h(\alpha, r_1)$ and $h(\alpha, r_2)$, respectively, using:
$h(\alpha, r) = \sum_{i=1}^{N} (D_{KL}(H_{i1} \| H_{i2}) + D_{KL}(H_{i2} \| H_{i1}) - \beta \| S_{i1} - S_{i2} \|)$
■ Optimize Eq. (1) using:
$f(\alpha^*, r^*) = \arg\max_{\alpha, r} h(\alpha, r)$
■ Update $a_j \leftarrow a^*$ and $r_j \leftarrow r^*$

according to the largest difference among the BoVW histograms of these sub-regions. Thus, the BoVW histograms that best capture the discriminative information can be built by concatenating the histograms in the partitioned sub-regions and saved into the database. According to the BoVW histograms extracted from the ROIs with known categories, an optimized similarity metric is learned for subsequent online image retrieval using the proposed REML method. In (2), a bounding box is also needed to cover the tumor region in a query image. The BoVW histogram is built online using the learned partitioned sub-regions and fed to the CBIR system. By calculating the similarity between the BoVW histograms of the query image and each image in the database according to the learned similarity metric, the CBIR system returns the images that are relatively similar to the query image. The flowchart of the proposed CBIR system is shown in Figure 3.

2.5 Intensity Normalization

In this study, we use the raw intensity as local patch descriptor to construct the BoVW histogram, which heavily relies on pixel intensities. However, image intensities in MR Images do not have a fixed meaning and vary considerably within or between subjects. Therefore, an intensity normalization step is needed to preprocess the MR Images. First, images were fed to an implementation of N3 bias field correction algorithm [36] to deal with the intensity heterogeneity. The intensities were then normalized as follows: intensity values at the 1% and 99% quantiles are computed for the brain region and then these two values are used to scale the intensities to [0, 1] through the min–max method.

2.6 Partition learning

As described in Section 1, while the spatial pyramid method outperforms the basic BoVW model, spatial partitions are fixed without considering the empirical or theoretical knowledge of the objects in the image. Therefore, a novel method is proposed to learn the optimal partitions of the spatial regions for pooling in this section.

Given a ROI in a brain tumor T1-weghted CE-MR Image, a series of lines are used to separate this ROI to sub-regions

[Figure 4(B)]. The learned partitions are decided by finding the largest difference among the BoVW histograms of the sub-regions. Based on this idea, we use the normal form of a linear equation to represent a line to partition an image [Figure 4(A)]. In addition, the Kullback–Leibler divergence is used to measure the difference between the BoVW histograms of two sub-regions; this divergence has been proven to be effective in discriminating two different classes [37]. Therefore, the objective function of the partition learning can be defined as follows:
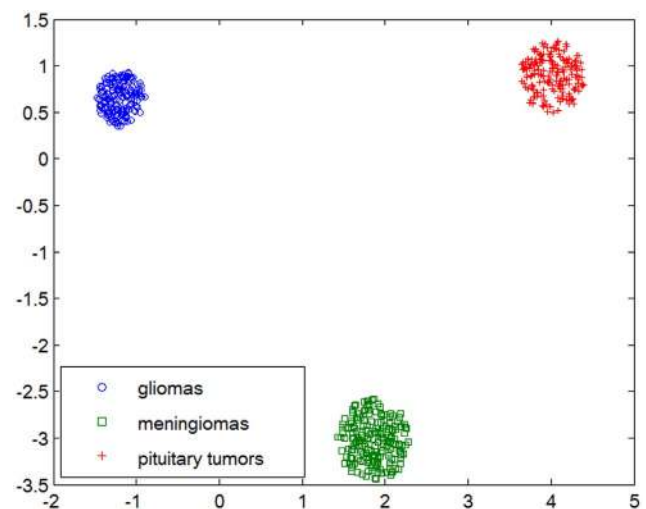

Figure 5. BoVW histograms of brain tumor T1-weghted CE-MR Images are projected into 2-dimensional feature space by REML. In the training dataset, three types of brain tumors are linearly separable in the transformed space learned by REML.

$$f(\alpha,r) = \tag{1}$$

$$\operatorname*{arg\,max}_{\alpha,r} \sum_{i=1}^{N} \left(D_{KL}(H_{i1}\|H_{i2}) + D_{KL}(H_{i2}\|H_{i1}) - \beta\|S_{i1} - S_{i2}\| \right)$$

where α and r are two parameters in a linear equation [Figure 4(A)]; $D_{KL}(H_{i1}\|H_{i2}) + D_{KL}(H_{i2}\|H_{i1})$ is the symmetrized Kullback–Leibler divergence; N is the image number in the dataset; β is a parameter that controls the trade-off between the maximized divergence values and the area constraint; H_{i1} and H_{i2} are the BoVW histograms in regions R_{i1} and R_{i2}, respectively; and S_{i1} and S_{i2} represent the areas of regions R_{i1} and R_{i2}, respectively. Moreover, the constraint of the areas of the two sub-regions in Eq. (1) is employed to ensure similar areas in the two sub-regions because the partition is meaningless when one partitioned area has a much smaller area than the other partitioned area. For an extreme example, when one partitioned area only contains one pixel and the other partitioned area contains significantly more pixels, the divergence values may be large; however, this partition cannot capture useful information from the image.

To optimize the objective function, we first define the ROIs of the images. For each brain tumor T1-weghted CE-MR Image, outlining the tumor by hand is not practical. Therefore, we locate a bounding box to cover the tumor region in each image. For a tumor region in an image, the bounding box is generated manually by specifying the opposite corner of a rectangle. The relative locations between the bounding box and the tumor region can be different for each image in our experiment. Starting with the bounding box as the spatial partition with one region, we partition the region further into two parts with a line. A disjoint partition set denoted as $P = \bigcup_{k}\{P_k\}$ is obtained after k successive partitions. When the number of partitions is equal to k, the number of the separated regions is $k+1$. In particular, the bounding box is the case where $k=0$. For each partition, we aim to find a line to separate a given region into two sub-regions [Figure 4 (B)] that can maximize the difference between these two sub-regions.

In the current study, the coordinate descent-like iterations and greedy forward selection methods are used to approximately optimize Eq. (1). The parameters α and r in a linear equation are treated as two sets of variables in which alternating coordinate descent-like iterations are performed to find the best partition for a given region. To increase the number of partitions, greedy forward selection is used to compute the next best partition by selecting the parameters that increase the objective function the most. The kth partition occurs in the region with the maximal area among the k regions. The complete partition learning procedure is listed in Table 2.

Compared with the spatial pyramid method that pre-sets a fixed partition for all bounding boxes, the proposed partition learning method generates a set of spatial partitions based on maximizing the difference between the two sub-regions. This idea is similar to coarse segmentation, which isolates the objects from their background according to their dissimilarities [38]. In addition, instead of using a pyramid, only the final partition set P_k generated by k partitions is used as we expect the partition to adjust its resolution depending on the spatial distribution of discriminative information for the images. Although this idea is straightforward, it contributes to a considerable improvement in retrieval accuracy in the experiment. Moreover, the proposed partition learning method can use empirical or theoretical knowledge of the brain tumors in T1-weghted CE-MR Images to discover descriptive partitions and better present the spatial configuration of the semantic components of the brain tumors in T1-weghted CE-MR Images.

2.7 Distance Metric Learning

When a query image is presented, the similarity between the feature vectors (BoVW histograms in this study) of the query image and the images in the dataset is measured. Image representation by visual features usually results in information loss, that is, the performance of a CBIR system suffers from the gap between the low-level visual features and the semantic concepts. Furthermore, a CBIR system cannot perform well when common distance metrics, such as the Euclidean distance or χ^2 distance, are used to measure the similarity of the features because BoVW histograms may not be located in the Euclidean space. Different visual words have different contributions to image similarity, whereas Euclidean distance provides the same weight to all visual words. Moreover, visual words in BoVW histograms may be relative. Therefore, a distance metric learning method is used to find a linear transformation that projects the image features into a new and meaningful feature space to reduce this semantic gap. Previous

Table 3. Algorithm 2.

Input:
Testing query set $\{x_q\}$ and their associated retrieved ranking list $\{(x_i, T_i)\}_q$; Parameter γ; Maximum iteration number t_{max}.
Output:
L^* (Final solution of REML).
Initialize the projection $L_0 = L_{LFDA}$.
Estimate the step size of gradient: $a = 0.1 \Big/ E\left(\left\|\frac{\partial J(L)}{\partial L}\|L=L_0\right\|_2\right)$.
For $t=1,\ldots,t_{max}$
■ Calculate a stochastic gradient in Eq. (11) on a randomly sampled query with respect to L_{t-1}
$\nabla L_{t-1} = \frac{\partial \ell_q(L)}{\partial L_{t-1}} + 2\gamma L_{t-1}$.
■ Update L:
$L_t = L_{t-1} - \frac{a}{1 + 2t/t_{ymax}}\nabla L_{t-1}$.

Table 4. Parameter settings in the experiments.

Parameter	Description	Setting
w	Patch size ($w \times w$)	7
V	Dictionary size	1000
k	Number of partition	9
β	Parameter defined in Eq. (1)	$1/\tau_k$ (τ_k: maximal area among the k regions)
t_{max}	Maximum iteration number defined in Algorithms 1 and 2	1000
σ	Scaling parameter defined in Eq. (7)	1
γ	Regularization weight defined in Eq. (10)	0.0001
d	Dimensionality of transformed feature space of REML, MPP, and LFDA	3
d_{CFML}	Dimensionality of transformed feature space of CFML	2
-	Coding method described in Section 2.1	soft-assignment
p	p-norm described in Section 2.1	∞

doi:10.1371/journal.pone.0102754.t004

studies show that well-designed distance metrics can perform better than the Euclidean distance on the CBIR system [39,40].

The squared Mahalanobis distance is used to compute the similarity between the feature vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$:

$$d_{ij} = \left\| \boldsymbol{L}(\boldsymbol{x}_i - \boldsymbol{x}_j) \right\|_2^2 = (\boldsymbol{x}_i - \boldsymbol{x}_j)^T \boldsymbol{L}^T \boldsymbol{L}(\boldsymbol{x}_i - \boldsymbol{x}_j), \qquad (2)$$

where $\boldsymbol{L}$ is a transformation matrix of size $d \times D$, where d and D are the dimensionalities of the transformed and the original feature space, respectively. In brief, the distance metric learning problem can be formulated as

$$\boldsymbol{L}^* = \underset{\boldsymbol{L} \in \mathbb{R}^{d \times D}}{\arg\min} J(\boldsymbol{L}) \qquad (3)$$

where $J(\boldsymbol{L})$ is an objective function. An optimal projection $\boldsymbol{L}^*$ can be obtained to project the image features to a new feature space by minimizing this objective function, making the squared Mahalanobis distance between data with the same labels closer and that with different labels farther (Figure 5).
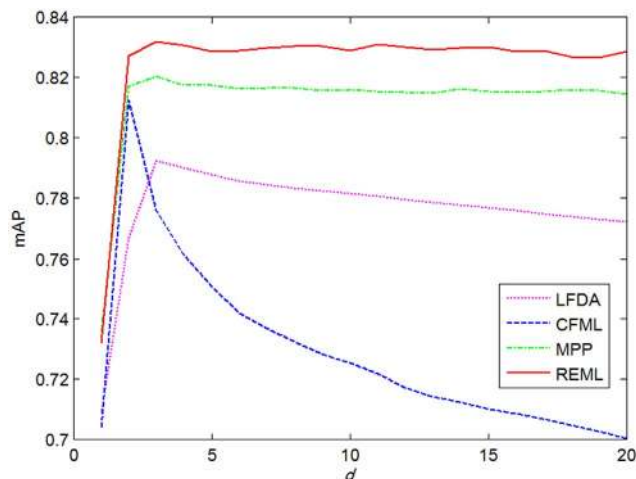
Figure 6. mAPs of different d values in LFDA, CFML, MPP, and REML.
doi:10.1371/journal.pone.0102754.g006

A number of distance metric learning algorithms have been proposed to find an optimal $\boldsymbol{L}$ for minimizing an objective function; examples of these algorithms include kernel-based distance metric learning [22], Xing's method [41], local fisher discriminant analysis (LFDA) [42], large margin nearest neighbor (LMNN) [43], and close-form metric learning (CFML) [44]. Most of these algorithms are designed to achieve high classification accuracy. However, the evaluation measures for CBIR systems are very different from those for classification. In our previous study [6], a distance metric learning method called maximum mean average precision projection (MPP) was proposed to optimize retrieval measures directly. We also verified that the MPP method yields good performance compared with other approaches [6]. In this study, we propose a novel metric learning method based on the minimization of rank error to achieve the projection $\boldsymbol{L}$.

2.7.1 Objective function. In this study, two images with tumors of the same category are considered relevant; otherwise, they are considered irrelevant. Given a query image I_q with feature vector x_q, all images in the dataset are retrieved and sorted in a ranking list according to increasing order of the squared Mahalanobis distance $\{d_j\}_{j=1}^{N}$, where d_j is the distance between x_q and x_j, and N is the image number in the dataset. Meanwhile, $T_j \in \{0,1\}$ (1 for relevant and 0 otherwise) is used to represent the relevance between I_q and an image I_j in the dataset. For a set of retrieved images $I_i(i \in S)$ with $T_i = 1$ and a set of retrieved images $I_j(j \in D)$ with $T_j = 0$, where S is the set of relevant pairs and D is the set of irrelevant pairs, d_i is made smaller so that $d_i < \min_{j \in D} d_j$.

Therefore, the loss function can be defined based on the rank error of the relevance:

$$RE(\boldsymbol{L}) = \frac{1}{n_S} \sum_{i \in S} 1\left\{ d_i > \min_{j \in D} d_j \right\} \qquad (4)$$

where $1\{\cdot\}$ is the indicator function. The distance gap of $\boldsymbol{x}_i$ is defined as

$$\Delta d_i = d_i - \min_{j \in D} d_j = d_i + \max_{j \in D}(-d_j) \qquad (5)$$
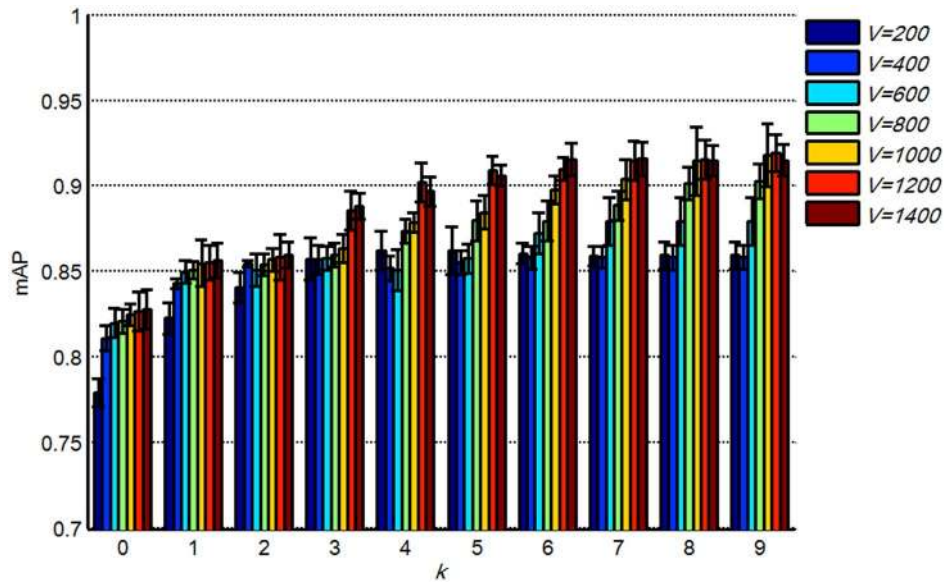
Thus, Eq. (4) can be rewritten as

Figure 7. mAPs of the proposed method with different _V_ and _k_; the bars show the means and standard deviations of mAPs.
doi:10.1371/journal.pone.0102754.g007

$$RE(\boldsymbol{L}) = \frac{1}{n_S} \sum_{i \in S} 1\{\Delta d_i > 0\} \qquad (6)$$

Given the max function in Eq. (5), the distance gap is noncontinuous and nondifferentiable. The max function can be approximated using the log-sum-exp method [45]:

$$\max_{j}(-d_j) \approx \frac{1}{\sigma} \log \sum_{j} \exp(-d_j/\sigma) \qquad (7)$$

where σ is the scaling parameter and set to 1 in this study. The indicator function in Eq. (6) is also noncontinuous and nondifferentiable; thus, the logistic loss function can be used to approximate the indicator function [46] and is defined as

$$l(v) = \log(1 + \exp(v)). \qquad (8)$$

The loss function can be approximated, and can become continuous and differentiable (denoted as $\ell_q(\boldsymbol{L})$):

$$\ell_q(\boldsymbol{L}) = \frac{1}{n_S} \sum_{i \in S} \log(1 + \exp(\Delta d_i))$$
$$= \frac{1}{n_S} \sum_{i \in S} \log\left(1 + \exp(d_i) \sum_{j \in D} \exp(-d_j)\right) \qquad (9)$$

If $-d_i$ is represented as a ranking function, $\ell_q(\boldsymbol{L})$ can be transformed to the loss function of "IR Push" [46]. To the best of our knowledge, the approximation of rank error has not been used in distance metric learning.

Based on the loss function, the proposed objective function is defined as follows:

$$J(\boldsymbol{L}) = \sum_{q} \ell_q(\boldsymbol{L}) + \gamma tr(\boldsymbol{L}^T \boldsymbol{L}). \qquad (10)$$

On the right hand side of Eq. (10), the second term is the regularization term, where $tr(\boldsymbol{L}^T \boldsymbol{L})$ is the trace of matrix $\boldsymbol{L}^T \boldsymbol{L}$. This term ensures the sparsity of the matrix $\boldsymbol{L}^T \boldsymbol{L}$ and makes the solution more reliable [47]. Parameter γ is used to control the trade-off between the empirical loss and the regularization term, which can be set empirically or estimated by cross-validation.

2.7.2 Optimization method of the objective function. To achieve the optimized projection $\boldsymbol{L}$, the objective function defined in Eq. (10) should be minimized. The gradient-based optimization method is used to find the optimal solutions; thus the following issues should be considered.

(1) The computation of $J(\boldsymbol{L})$ or its gradient over all query pairs is time-consuming and impractical on a large training dataset.

Figure 8. mAPs of the proposed method with different _w_ and _k_; the bars show the means and standard deviations of mAPs.
doi:10.1371/journal.pone.0102754.g008

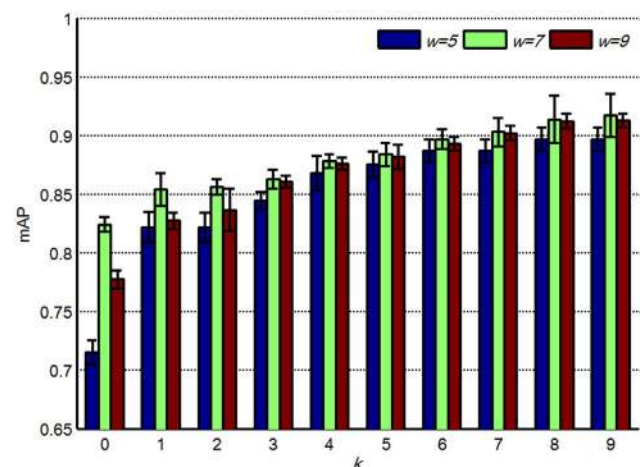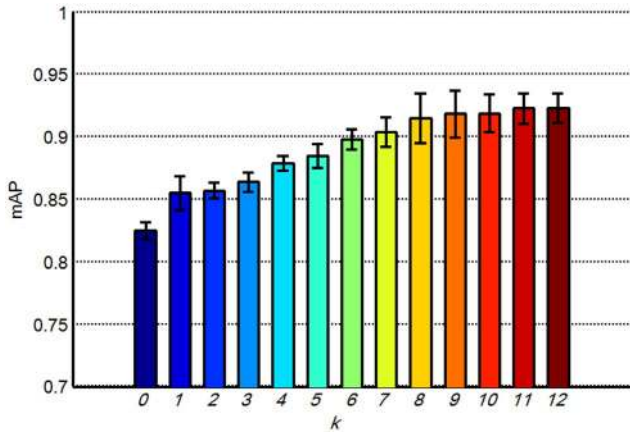**Figure 9. mAPs of the proposed method with different k; the bars show the means and standard deviations of mAPs.**
doi:10.1371/journal.pone.0102754.g009

Therefore, the stochastic gradient descent algorithm is applied in this study for optimizing $J(L)$; this algorithm exhibits reasonable performance in the experiments. The gradient of $J(L)$ in Eq. (10) with respect to L is

$$g = \frac{\partial J(L)}{\partial L} = \sum_q \frac{\partial \ell_q(L)}{\partial L} + 2\gamma L. \tag{11}$$

Define $\beta_i = \exp(d_i) \sum_{j \in D} \exp(-d_j)$, then Eq. (9) can be rewritten as

$$\ell_q(L) = \frac{1}{n_S} \sum_{i \in S} \log(1 + \beta_i). \tag{12}$$

Now, the gradient of $\ell_q(L)$ with respect to L can be defined as

$$\frac{\partial \ell_q(L)}{\partial L} = \frac{1}{n_S} \sum_{i \in S} \frac{1}{1 + \beta_i} \frac{\partial \beta_i}{\partial L}, \tag{13}$$

where

$$\frac{\partial \beta_i}{\partial L} = \beta_i \left(\frac{\partial d_i}{\partial L} - \sum_{j \in D} \frac{\partial d_j}{\partial L} \right). \tag{14}$$

Define $C_{ij} = (x_i - x_j)(x_i - x_j)^T$, then Eq. (3) can be rewritten as

$$d_{ij} = \left\| L(x_i - x_j) \right\|_2^2 = (x_i - x_j)^T L^T L (x_i - x_j) = tr(L^T L C_{ij}). \tag{15}$$

Therefore, the gradient of d_{ij} with respect to L is

$$\frac{\partial d_{ij}}{\partial L} = 2L C_{ij}. \tag{16}$$

(2) Given the non-convexity of $J(L)$, the final solution of L is highly dependent on the initial estimation L_0. The initial

estimation L_0 is expected to be very close to the global optimal solution. Some other metric learning algorithms can be applied to obtain L_0. As reported in our previous study [6], LFDA [42] can be used to obtain an appropriate linear transformation with relatively high retrieval performance. Thus, LFDA is also used to estimate the initial L_0 in this study. The LFDA projection matrix L_{LFDA} can be defined as:

$$L_{LFDA} = \arg\max_{L \in \mathbb{R}^{d \times D}} \left[tr\left(\left(L^T \tilde{S}^{(w)} L \right)^{-1} L^T \tilde{S}^{(b)} L \right) \right],$$

where

$$\tilde{S}^{(w)} = \frac{1}{2} \sum_{i,j=1}^{n} \tilde{W}_{i,j}^{(w)} (x_i - x_j)(x_i - x_j)^T,$$

$$\tilde{S}^{(b)} = \frac{1}{2} \sum_{i,j=1}^{n} \tilde{W}_{i,j}^{(b)} (x_i - x_j)(x_i - x_j)^T,$$

and $tr()$ denotes matrix trace. $\tilde{S}^{(w)}$ and $\tilde{S}^{(b)}$ are the local within-class scatter matrix and the local between-class scatter matrix, respectively. $\tilde{W}_{i,j}^{(w)}$ and $\tilde{W}_{i,j}^{(b)}$ can be defined as

$$\tilde{W}_{i,j}^{(w)} = \begin{cases} A_{i,j}/n_l & if \quad y_i = y_j = l, \\ 0 & if \quad y_i \neq y_j, \end{cases}$$

$$\tilde{W}_{i,j}^{(b)} = \begin{cases} A_{i,j}(1/n - 1/n_l) & if \quad y_i = y_j = l, \\ 1/n & if \quad y_i \neq y_j, \end{cases}$$

where $A_{i,j}$ is an affinity matrix, $A_{i,j}$ is large if x_i and x_j are 'close' and $A_{i,j}$ is small if x_i and x_j are 'far apart', y_i is the label of x_i, n_l is the sample number of class l, and n is the total sample number.

A projection matrix L_{LFDA} was first obtained using the LFDA method. The matrix L_{LFDA} was then fed to the proposed REML method as the initial solution L_0 for further optimization. The complete REML procedure is listed in Table 3.
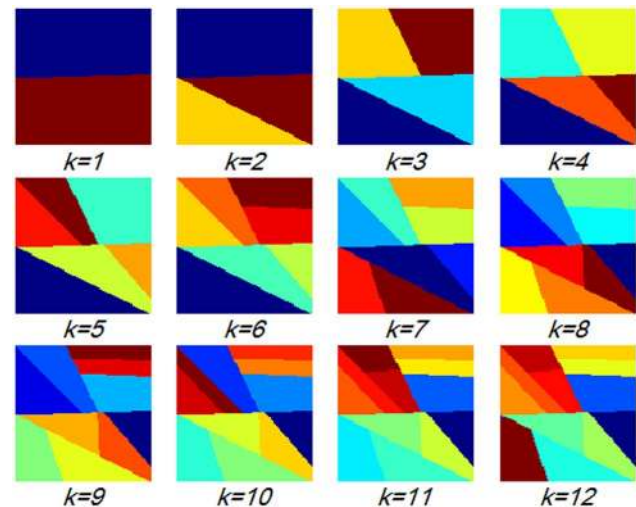

Figure 10. Partition patterns with different k when $V = 1000$ and $w = 7$.
doi:10.1371/journal.pone.0102754.g010

Experimental Results

3.1 Retrieval Evaluation Measures

The retrieval evaluation measures for the proposed CBIR system are presented in this section. For a given number of retrieved images, the Precision and Recall are defined as:

$$\text{Precision} = \sum_{j=1}^{K} T_j \Big/ K,$$

$$\text{Recall} = \sum_{j=1}^{K} T_j \Big/ \sum_{j=1}^{N} T_j,$$

where $K = 1, \cdots, N$ is the number of retrieved images and N is the total image number in the dataset. The precision-recall pairs for varying numbers of retrieved images are usually plotted in a precision-recall curve to evaluate the retrieval system. In the ranking list, precision at the top n retrieved images (Prec@n in short) is denoted by:

$$\text{Prec}@n = \frac{1}{n} \sum_{j=1}^{N} T_j 1\{\pi(\boldsymbol{x}_j) \leq n\},$$

where $\pi(\boldsymbol{x}_j)$ represents the position or rank of the retrieved image $\boldsymbol{x}_j$ in the ranking list. The average precision (AP) is the average of the precisions at the positions where a relevant image exists:

$$\text{AP} = \frac{1}{\sum_{j=1}^{N} T_j} \sum_{j=1}^{N} T_j \times \text{Prec}@j$$

Finally, mAP is the mean AP over all of the queries:

$$\text{mAP} = \frac{1}{M} \sum_{q=1}^{M} \text{AP}_q$$

where M is the number of queries.

mAP and Prec@10 were used as the major quantitative measures to evaluate the performance of the algorithms. Furthermore, precision-recall curves were drawn for some cases to enable intuitive comparison.

3.2 Experimental Settings

In the experiments, the performance of the category retrievals of the brain tumors was evaluated through five-fold cross-validation method. The 233 patients were randomly partitioned into five subsets. Partitioning ensured that the slices from one patient do not exist in the training and test dataset at the same time. All experiments were repeated five times, and the final results were reported as the mean and standard deviation of the results from
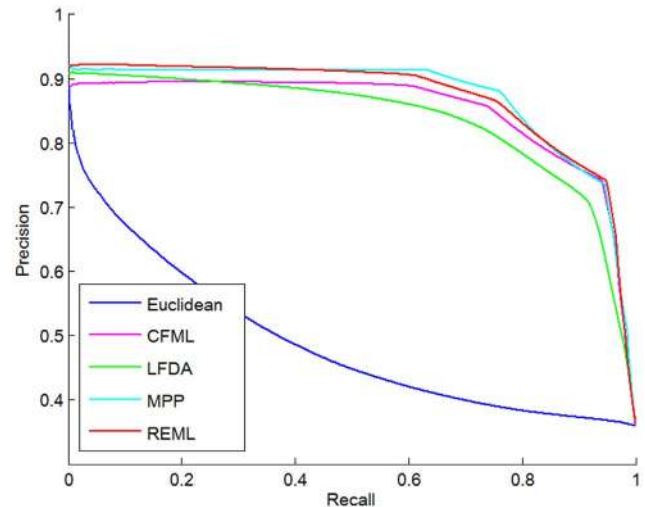

Figure 11. Precision-recall curves using different distance metric learning methods.
doi:10.1371/journal.pone.0102754.g011

the individual runs. In each run, four subsets were used as a training dataset, and the remaining one was used as a test dataset. For the given training and test datasets, each image in the test dataset was adopted as a query to retrieve the training dataset to report the performance.

The parameter setting of all the algorithms tested in this study was carefully considered to obtain optimum performance during the experiments. Table 4 provides a summary of the parameter settings in the experiments.
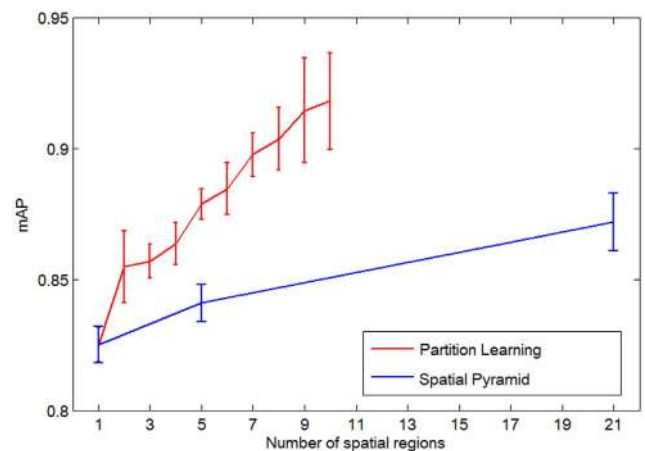
Figure 12. mAPs of the partition learning and spatial pyramid methods with different numbers of spatial regions.
doi:10.1371/journal.pone.0102754.g012

Table 5. Summary of the retrieval performance of different distance metrics (mean $\pm$ std %).

	Euclidean	CFML	LFDA	MPP	REML
mAP	46.5±0.1	86.6±1.0	89.2±1.9	90.1±2.47	91.8±1.8
Prec@10	79.4±0.2	89.3±0.2	90.9±0.2	91.7±0.1	93.1±0.1

doi:10.1371/journal.pone.0102754.t005

3.3 Experiments

3.4.1 Parameter Optimization. The reduced dimensionality d of the transformed feature space is an important parameter for L. We investigated the retrieval performance at different d values of LFDA, CFML, MPP, and the proposed REML to determine the optimal d. In addition, parameters V, w, and k were set to 1000, 7, and 0, respectively. The highest mAP values were achieved when d was set to 3 for LFDA, MPP, and REML, as shown in Figure 6. At the same time, the best performance of CFML was obtained when d was fixed to 2. Furthermore, REML and MPP were more robust than LFDA and CFML when d varied. The mAP values of REML and MPP were maintained at an approximately constant level, whereas the mAP values of LFDA and CFML significantly decreased when d was greater than 3 and 2, respectively. Therefore, the d values of LFDA, CFML, MPP, and REML were set to 3, 2, 3, and 3, respectively, in the subsequent experiments.

The dictionary size V is a crucial parameter in the partition learning method. We investigated the retrieval performance at different V with the distance metric learned by REML to determine the optimal V. Parameters d and w were set to 3 and 7, respectively. In addition, k ranged from 0 to 9. Figure 7 shows that the retrieval performance was improved by increasing V with different k values. Large dictionaries contain numerous discriminative data, resulting in a higher retrieval performance compared with small dictionaries. However, in addition to high memory and computational costs, retrieval accuracy increased slowly at $V>$ 1000 (paired t-test $p = 0.0095$ when V ranged between 800 and 1000; paired t-test $p = 0.049$ when V ranged between 1000 and 1200). Therefore, to balance the trade-off between memory and computation costs and accuracy, the dictionary size V was fixed to 1000 in subsequent experiments.

The patch size w is another important parameter in the partition learning method. Three different patch sizes, 5, 7, and 9, with the distance metric learned by REML were used to assess the effect of patch size on the retrieval performance. Parameters d and V were set to 3 and 1000, respectively. In addition, k ranged from 0 to 9. Figure 8 shows that retrieval performance was improved by increasing w from 5 to 7 with different k values. In general, larger patches lead to more discriminative information for identifying different objects. However, in high-dimensional feature spaces associated with larger patches, a larger dictionary is necessary to construct the basis of the space, and more patches are required to train the model. As shown in Figure 8, the retrieval accuracy of $w = 7$ is higher than those of $w = 5$ (paired t-test $p = 0.0068$) and $w = 9$ (paired t-test $p = 0.0081$). Therefore, a moderate-sized patch is more applicable in the proposed method than patches of other sizes. Thus, w was set to 7 in subsequent experiments.

The number of partition k is the third parameter in the partition learning method that should be determined carefully. We varied k from 0 to 12 to investigate the variation of the retrieval performance with different k values by using the distance metric learned by REML. In addition, parameters V, w, and d were set to

1000, 7, and 3, respectively. Figure 9 shows the mAP values of different k. Increasing k was proven to be useful up to 9, as shown in Figure 9. Moreover, increasing k beyond 9 increased computational costs with no obvious improvement in the retrieval performance. Figure 10 shows the partition patterns with different k values. The larger k leads to finer partitions and obtains more discriminative information of the objects. However, when the images are too finely partitioned, the individual bins of the histograms may yield too few matches. This result may be why increasing k beyond 9 does not result in evident improvement in the retrieval performance. Figure 10 also shows that the spatial partitions of the brain tumor images are irregular, which may be attributed to the complex appearance of the brain tumors in the T1-weghted CE-MR Images (Figure 2).

3.3.2 Evaluation of the Effectiveness of REML. The BoVW histogram, which was constructed using the optimal parameters described above, was used as the feature input to learn different distance metrics. The retrieval performance of different distance metrics is summarized in Table 5 and Figure 11. As listed in Table 5, the proposed REML outperforms the Euclidean distance, CFML, LFDA, and MPP methods in terms of both mAP and Prec@10. Compared with the MPP, which had the second best performance, REML resulted in significant improvement ($p = 0.03$, paired t-test) in mAP (1.7% improvement from 90.1% to 91.8%). Compared with the Euclidean distance, CFML, LFDA, and MPP methods, the precision-recall curves in Figure 11 reveal that REML is the most effective distance metric learning method for improving the performance of the proposed CBIR system.

3.3.3 Evaluation of the Effectiveness of Partition learning. The partition learning method was compared with the spatial pyramid method introduced in a previous study [25] to validate the performance of the proposed method. To provide a fair comparison, we fixed $V = 1000$, $w = 7$, and $d = 3$, and used the distance matric learned by REML on both the partition learning method and the spatial pyramid method. In addition, k was set to 9 in the partition learning method, whereas the spatial level was set to 3 in the spatial pyramid method. Therefore, the number of spatial regions for pooling was 10 for the partition learning method and 21 for the spatial pyramid method ($1+4+16 = 21$), respectively. Furthermore, the partitions were performed on the rectangle ROIs for both the partition learning method and the spatial pyramid method. The comparison results are illustrated in Figure 12. In this experiment, the spatial pyramid method achieved a mean mAP value of 0.825 ± 0.007 at pyramid level 0 (bounding box i.e., 1×1), 0.841 ± 0.007 at pyramid levels 0 and 1 (1×1 and 2×2), and 0.872 ± 0.011 at levels 0, 1, and 2 (1×1, 2×2 and 4×4). The performance decreased for spatial pyramid method when we go higher than level 2. As shown in Figure 12, the performance of the partition learning method with spatial regions equal to 10 ($k = 9$) outperformed the best performance (spatial regions = 21) of the spatial pyramid method. This result supports our claim that quantizing the image space by symmetric divisions is inappropriate to describe the spatial layout information and can

Table 6. Retrieval performance of the proposed method for different categories of brain tumor (mean $\pm$ std%).

Tumor category	mAP	Prec@10	Prec@20
Meningioma	81.7±2.7	88.9±2.9	87.9±2.9
Glioma	97.4±1.0	95.7±2.0	95.9±1.9
Pituitary tumor	93.2±1.7	94.5±2.1	94.2±2.0

doi:10.1371/journal.pone.0102754.t006

Table 7. Comparison of the proposed method with two relevant methods in terms of mAP and Prec@10 (%).

Methods	Yang et al. [6]	Huang et al. [7]	Proposed
mAP	87.3	91.0	**91.8**
Prec@10	89.3	91.7	**93.1**

be improved by the partition learning method. In addition, the computation and memory costs of the proposed method are much lower than those of the spatial pyramid method (feature dimensionality is $1000 \times 9 = 9000$ in the proposed method vs. $1000 \times 21 = 21000$ in the spatial pyramid method). These experimental results indicate that the partition learning method is more flexible than the spatial pyramid method and can better capture the discriminative information of the brain tumors in T1-weighted CE-MR Images.

3.3.4 Retrieval Performance of the Proposed Method. The optimal parameters described above were used to construct the BoVW histogram and learn the distance metric for the proposed method. The retrieval performance of the proposed method for the different categories of brain tumors was evaluated, and the results are listed in Table 6. The retrieval performance of gliomas is higher than that of meningiomas and pituitary tumors, which supports the fact that the extracted features contain adequate discriminative information to distinguish gliomas from other tumor types ($p<0.01$, two sample t-test). The inferior performance of meningiomas may be due to the unbalanced distribution of samples in the dataset and their extremely similar appearance to pituitary tumors.

3.3.5 Comparison with Relevant Methods. To demonstrate the performance of the proposed method, it was compared with the performance of two other methods for brain tumor retrieval in T1-weighted CE-MR Images. Yang et al. [6] proposed a feature extraction method by capturing tumor margin information and designed a distance metric learning method by maximizing the mAP to retrieve brain tumors in the T1-weghted CE-MR Images. Huang et al. [7] used a fixed separation method combined with the BoVW model to capture the information of the intensity values in the tumor region and the intensity variation of the tumor-surrounding region for brain tumor retrieval. The dataset used in these two methods was identical to that used in the proposed method. For the proposed method, the optimal parameters described above were used to construct the BoVW histogram and learn the distance metric. Table 7 shows the comparison results of three methods. The retrieval results of the two compared methods are reported in the corresponding papers. The best result in each evaluation measure is shown in boldface in Table 7. The proposed method yielded the highest mAP and Prec@10 among the three methods, as listed in Table 7. This result demonstrates that the proposed method is effective and robust in retrieving brain tumor in T1-weighted CE-MR Images.

3.3.6 Computation Cost. A PC with Intel P4 3.0 GHz processor and 16 GB RAM was used as the workstation in this study. In the offline feature extraction step, learning an optimal partition with $V = 1000$, $w = 7$, and $k = 9$ from a training dataset with 2450 images consumed 1 h. Furthermore, learning an optimized similarity metric using the extracted features and the proposed REML method consumed 2 min. In the online image retrieval step, the processing time was approximately 5 s.

Discussion and Conclusion

Effective image features are crucial to produce satisfactory retrieval results. A partition learning algorithm is proposed to be integrated into the basic BoVW framework to extract discriminative features in this study. The coordinate descent-like iterations and greedy forward selection methods were used to optimize the objective function of the partition learning method. Therefore, the final solution of the partition learning method may vary for different initial conditions. In addition, the solution may offer a local optimal solution instead of a global optimal solution, which is a limitation in the partition learning method. However, the partition learning method still performed well during the experiments and achieved higher mAP values than other relevant methods (Table 7).

The major limitation of REML is that the learned distance metric is globally linear. In future work, we aim to learn multiple local projections for different regions in the feature space, similar to the multi-metric LMNN [43]. Moreover, the objective function of REML is optimized by the stochastic gradient descent algorithm; thus, the final solutions of L vary for different initial solutions and different query pairs randomly selected. An optional solution is to stabilize the REML solution by aggregation. Another competitive solution is to use the virtual convex cost function defined by the LambdaRank [48] algorithm and directly optimize the cost function with respect to the matrix M, where $M = L^T L$ and M is a positive semidefinite matrix.

Although the proposed CBIR system is based on 2D CE-MR Images, the proposed system can be extended to 3D cases. As introduced in Section 2.2, 3D MR Images are difficult to obtain in the routine clinics of hospitals in China. Before extending the proposed method to 3D cases, the normalization of space resolutions in different directions among different subjects must be solved. Furthermore, 3D patches can be extracted to construct BoVW histograms and represent 3D regions. Region partitioning in 3D space is a more difficult task than region partitioning in 2D space because of the increasing complexity of the former. Instead of calculating the parameters in a line, parameters in a plane must be solved for region partition in 3D space.

Another aspect of this work that can be improved is the image intensity standardization step. The proposed method uses the N3 algorithm and min–max method to remove bias field artifacts and normalize image intensity, respectively. However, retrieval can benefit from sophisticated intensity standardization steps, such as inhomogeneity correction, noise removal filtering, and intensity standardization [49,50]. Better intensity matching corresponds to more discriminative features.

The development of the CBIR system for brain tumors in MR Images has been widely investigated. Thus, placing our results in context helps to reveal the significance of our results. However, direct comparisons of quantitative retrieval results across publications are difficult and not always fair because of inconsistent tumor category and imaging protocol. Therefore, the proposed method was only compared with two brain tumor retrieval methods [6,7],

which were all performed using the same dataset. As listed in Table 7, mAP and Prec@10 of the proposed method are higher than those of Yang et al. [6] and Huang et al. [7].

In conclusion, this study presented a CBIR system for 2D brain tumor T1-weighted CE-MR Images. A partition learning method is integrated into the basic BoVW framework to extract discriminative features, which is very useful to the category retrieval task of brain T1-weighted CE-MR Images. A distance metric learning method is also proposed to reduce the sematic gap in the proposed CBIR system. The effectiveness of the partition learning and REML methods is demonstrated on a large T1-weighted CE-MR Image dataset of brain tumors.

Author Contributions

Conceived and designed the experiments: MH QF. Performed the experiments: MH WY YW. Analyzed the data: MH JJ. Contributed reagents/materials/analysis tools: YG WC ZL. Wrote the paper: MH WY YC QF.

References

1. Muller H, Michoux N, Bandon D, Geissbuhler A (2004) A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. Int J Med Inform 73: 1–23.
2. Rahman MM, Desai BC, Bhattacharya P (2008) Medical image retrieval with probabilistic multi-class support vector machine classifiers and adaptive similarity fusion. Computerized Medical Imaging and Graphics 32: 95–108.
3. Iakovidis DK, Pelekis N, Kotsifakos EE, Kopanakis I, Karanikas H, et al. (2009) A pattern similarity scheme for medical image retrieval. IEEE Transactions on Information Technology in Biomedicine 13: 442–450.
4. Avni U, Greenspan H, Konen E, Sharon M, Goldberger J (2011) X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words. IEEE Transactions on Medical Imaging 30: 733–746.
5. Yang W, Lu Z, Yu M, Huang M, Feng Q, et al. (2012) Content-Based Retrieval of Focal Liver Lesions Using Bag-of-Visual-Words Representations of Single- and Multiphase Contrast-Enhanced CT Images. Journal of Digital Imaging 25: 708–719.
6. Yang W, Feng Q, Yu M, Lu Z, Gao Y, et al. (2012) Content-based retrieval of brain tumor in contrast-enhanced MRI images using tumor margin information and learned distance metric. Medical Physics 39: 6929–6942.
7. Huang M, Yang W, Yu M, Lu Z, Feng Q, et al. (2012) Retrieval of brain tumors with region-specific bag-of-visual-words representations in contrast-enhanced MRI images. Comput Math Methods Med 2012: 280538.
8. Oliveira JD, Machado A, Chavez G, Lopes AP, Deserno T, et al. (2010) MammoSys: A content-based image retrieval system using breast density patterns. Computer Methods and Programs in Biomedicine 99: 289–297.
9. Huang HK, Nielsen JF, Nelson MD, Liu LF (2005) Image-matching as a medical diagnostic support tool (DST) for brain diseases in children. Computerized Medical Imaging and Graphics 29: 195–202.
10. Dube S, El-Saden S, Cloughesy TF, Sinha U. Content based image retrieval for MR image studies of brain tumors; 2006 Aug. 30 2006-Sept. 3 2006; New York. pp. 3337–3340.
11. Moustakas J, Marias K, Dimitriadis S, Orphanoudakis SC (2005) A two-level CBIR platform with application to brain MRI retrieval. IEEE International Conference on Multimedia and Expo 2005.
12. Antel SB, Collins DL, Bernasconi N, Andermann F, Shinghal R, et al. (2003) Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis. NeuroImage 19: 1748–1759.
13. Chaplot S, Patnaik LM, Jagannathan NR (2006) Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomedical Signal Processing and Control 1: 86–92.
14. Nazari MR, Fatemizadeh E (2010) A CBIR system for human brain magnetic resonance image indexing. International Journal of Computer Applications 7: 33–37.
15. Klöppel S, Abdulkadir A, Hadjidemetriou S, Issleib S, Frings L, et al. (2011) A comparison of different automated methods for the detection of white matter lesions in MRI data. NeuroImage 57: 416–422.
16. Lahmiri S, Boukadoum M. Brain MRI classification using an ensemble system and LH and HL wavelet sub-bands features; 2011 11–15 April Paris, France. pp. 1–7.
17. Zhang YD, Dong ZC, Wu LN, Wang SH (2011) A hybrid method for MRI brain image classification. Expert Systems with Applications 38: 10049–10053.
18. Robinson GP, Tagare HD, Duncan JS, Jaffe CC (1996) Medical image collection indexing: shape-based retrieval using kd-trees. Computerized Medical Imaging and Graphics 20: 209–217.
19. Petrakis EGM, Faloutsos C (1997) Similarity searching in medical image databases. IEEE Transactions on Knowledge and Data Engineering 9: 435–447.
20. Chu WW, Hsu CC, Cardenas AF, Taira RK (1998) Knowledge-based image retrieval with spatial and temporal constructs. IEEE Transactions on Knowledge and Data Engineering 10: 872–888.
21. Guan H, Antani S, Long LR, Thoma GR. Bridging the semantic gap using ranking SVM for image retrieval; 2009.
22. Chang C, Yeung DY (2007) Kernel-based distance metric learning for content-based image retrieval. Image and Vision Computing 25: 695–703.
23. Yang L, Jin R, Mummert L, Sukthankar R, Goode A, et al. (2010) A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. IEEE Trans Pattern Anal Mach Intell 32: 30–44.

24. Jiang J, Wu Y, Huang M, Yang W, Chen WF, et al. (2013) 3D brain tumor segmentation in multimodal MR images based on learning population- and patient-specific feature sets. Comput Med Imaging Graph 37: 512–521.
25. Lazebnik S, Schmid C, Ponce J (2006) Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. Computer Vision and Pattern Recognition. pp. 2169–2178.
26. Harada T, Ushiku Y, Yamashita Y, Kuniyoshi Y (2011) Discriminative Spatial Pyramid. IEEE Conference on Computer Vision and Pattern Recognition. pp. 1617–1624.
27. Sharma G, Jurie F. Learning discriminative spatial representation for image classification; 2011. pp. 6.1–6.11.
28. Jiang Y, Yuan J, Yu G. Randomized Spatial Partition for Scene Recognition; 2012. pp. 730–743.
29. Yang Y, Newsam S. Spatial pyramid co-occurrence for image classification; 2011. pp. 1465–1472.
30. Wu J, Rehg JM (2011) Centrist: A visual descriptor for scene categorization. IEEE Transactions on Pattern Analysis and Machine Intelligence 33: 1489–1501.
31. Jégou H, Douze M, Schmid C (2010) Improving bag-of-features for large scale image search. International Journal of Computer Vision 87: 316–336.
32. Bosch A, Muñoz X, Oliver A, Martí J. Modeling and classifying breast tissue density in mammograms; 2006; New York, USA. pp. 1552–1558.
33. Li FF, Perona P. A bayesian hierarchical model for learning natural scene categories; 2005. pp. 524–531.
34. van Gemert JC, Veenman CJ, Smeulders AWM, Geusebroek J-M (2010) Visual Word Ambiguity. IEEE Transactions on Pattern Analysis and Machine Intelligence 32: 1271–1283.
35. Wang JJ, Yang JC, Yu K, Lv FJ, Huang T, et al. (2010) Locality-constrained Linear Coding for Image Classification. Computer Vision and Pattern Recognition. pp. 3360–3367.
36. Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 17: 87–97.
37. Liu C, Shum HY. Kullback-Leibler Boosting; 2003. pp. 587–594.
38. Baqci U, Chen X, Udupa JK (2012) Hierarchical Scale-Based Multi-Object Recognition of 3D Anatomical Structures. IEEE Transactions on Medical Imaging 31: 777–789.
39. Frome A, Singer Y, Malik J (2006) Image retrieval and classification using local distance functions. Neural Information Processing Systems. pp. 417–424.
40. Chang C, Yeung DY (2007) Kernel-based distance metric learning for content-based image retrieval. Image and Vision Computing 25: 695–703.
41. Xing EP, Ng AY, Jordan MI, Russell SJ. Distance metric learning with application to clustering with side-information; 2002. pp. 505–512.
42. Masashi S (2007) Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis. Journal of Machine Learning Research 8: 1027–1061.
43. Weinberger KQ, Saul LK (2009) Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research 10: 207–244.
44. Alipanahi B, Biggs M, Ghodsi A. Distance metric learning vs. Fisher discriminant analysis; 2008 July 13–17; Chicago, Illinois. pp. 598–603.
45. Chen M, Liew SC, Shao Z, Kai C (2013) Markov Approximation for Combinatorial Network Optimization. IEEE Transactions on Information Theory 59: 6301–6327.
46. Rudin C (2009) The P-Norm Push: A Simple Convex Ranking Algorithm that Concentrates at the Top of the List. The Journal of Machine Learning Research 10: 2233–2271.
47. Huang K, Ying Y, Campbell C (2011) Generalized sparse metric learning with relative comparisons. Knowledge and Information Systems 28: 25–45.
48. Burges C, Ragno R, Le Q (2006) Learning to rank with nonsmooth cost functions. Neural Information Processing Systems. pp. 193–200.
49. Nyul LG, Udupa JK (1999) On Standardizing the MR Image Intensity Scale. Magnetic Resonance in Medicine 42: 1072–1081.
50. Bagci U, Udupa JK, Bai L (2010) The role of intensity standardization in medical image registration. Pattern Recognition Letters 31: 315–323.