A Survey of Dictionary Learning in Medical Image Analysis and Its Application for Glaucoma Diagnosis

Rongchang Zhao, Hong Li & Xiyao Liu

ONLINE FIRST

Archives of Computational Methods in Engineering
State of the Art Reviews

ISSN 1134-3060

Arch Computat Methods Eng
DOI 10.1007/s11831-019-09383-3

Springer

Archives of Computational Methods in Engineering
https://doi.org/10.1007/s11831-019-09383-3

ORIGINAL PAPER

A Survey of Dictionary Learning in Medical Image Analysis and Its Application for Glaucoma Diagnosis

Rongchang Zhao[1] · Hong Li[1] · Xiyao Liu[1]

Received: 18 May 2019 / Accepted: 8 December 2019

Abstract

Dictionary learning has shown its effectiveness in computer vision with the concise expression form but the powerful representation. Dictionary learning represents images with a bag of visual words (BoVW), which is a collection of atoms expressively representative for images. Recently, several task-specific dictionary learning methods have been proposed and successfully applied in medical image analysis, such as de-noising, classification, segmentation, and so on, which promotes the development of computer-aided diagnosis. In this paper, first we give a survey for dictionary learning-based medical image analysis methods including: (1) three discriminative dictionary learning frameworks, (2) CT image de-noising based on dictionary learning, and (3) histopathological image classification using sparse representation. Then, a novel method named Low-rank Shared Dictionary Learning (LRSDL), is presented to achieve accurate glaucoma diagnosis on fundus images. The LRSDL generates a shared codebook for image reconstruction and a particular one to handle the difference between the healthy and glaucomatous images. Benefit from this strategy, LRSDL not only possess distinct glaucoma-related features, but also share common patterns among all the fundus images. Experimental results show that the method effectively delivers glaucoma diagnosis with the accuracy of 92.90%. This endows dictionary learning method a great potential for glaucoma diagnosis and proves the feasibility of its application to medical image analysis.

1 Introduction

With the rapid accumulation of medical big data, building systems for automated image analysis is essential and necessary for the clinical practice and research. In the past decades, machine learning became increasingly popular in medical image analysis [1–3]. As a leading machine learning methodology, dictionary learning is good at spare image representation because it lies in the expressiveness and discrimination of the bag of visual words [4]. Most research focus on dictionary learning methods for medical image analysis, such as image de-noising [5–10], classification [11–20], segmentation [21–26], and so on.

✉ Xiyao Liu
lxyzoewx@csu.edu.cn

Rongchang Zhao
zhaorc@csu.edu.cn

Hong Li
Lihongup@csu.edu.cn

1 School of Computer Science and Engineering, Central South University, Changsha, China

Dictionary learning is a representation learning method which aims at finding a sparse representation of the input data in the form of a linear combination of basic elements. The elements are called as atoms or visual words which forming a dictionary. In the classical dictionary approach, visual words are generated from all training images and collected into a dictionary. Based on the generation of the representative dictionary, a global representation per image is extracted as the histogram of visual words and provides a unique feature for image classification, de-noising, segmentation, and so on. Existing applications of dictionary learning methods in medical image analysis are mainly concentrated in: (1) CT image de-noising, where dictionary can effectively distinguish the noise-free image and noisy part from sparse representation; (2) Histopathological image classification, where dictionary can capture its diverse histology features to make accurate diagnose.

While the discriminative power of dictionary learning is nowadays demonstrated, one of the fundamentally important topic in computer vision and medical image analysis is finding the sparse representation for dictionary learning model. Nowadays, most researches are devoted to improving the performance of dictionary learning model by enhancing the

Published online: 13 December 2019

representation ability of the sparse representation, especially the discrimination power of the learned dictionary. However, there are still challenging due to: (1) inevitably introduced noise in the process of medical image acquisition, transmission and recording because of the limitations of medical imaging system, blurs or masks the important information of the images; (2) lack of effective method to design discriminative class-specific dictionaries; (3) the high variability and extreme inhomogeneity of medical images make it difficult to capture available features.

In this paper, we give a survey for dictionary learning in medical image analysis and its application for glaucoma diagnosis. Firstly, we introduce and compare three of the most widely used discriminative dictionary learning models, including discriminative K-SVD (D-KSVD), label consistent K-SVD (LC-KSVD) and Fisher discrimination dictionary learning (FDDL); Secondly, we summarize the applications of dictionary learning methods in medical image analysis, including applications in CT image de-noising and histopathological image classification; Lastly, we utilize the low-rank shared dictionary learning (LRSDL) to achieve glaucoma automatic diagnosis on fundus images for the first time.

2 Review of Dictionary Learning Methods

In recent years, multitudes of researchers have been devoted to the study of designing a discriminative dictionary to improve the performance of dictionary learning models. In other words, dictionary atoms should have different characteristics according to the requirements of specific application problems. In this section, we will introduce three discriminative dictionary learning methods proposed recently, including Discriminative K-SVD, Label Consistent K-SVD and Fisher Discrimination Dictionary Learning. We focus on their improvements of objective function and the relationship among them.

2.1 Discriminative K-SVD

When sparse representation [27] was initially applied to image processing problems, the dictionary was formed by all the training images. The major drawback connected with this approach is that the representational power of the dictionary especially lies in the quality and quantity of the training image. In order to overcome the shortcoming, K-SVD algorithm [28] was widely employed for learning a smaller-sized dictionary from the given training images while maintaining the representational power of the dictionary, which finds a solution for the following problem:

$$\begin{aligned} <D, X> = \underset{D,X}{\arg\min} \parallel Y - DX \parallel_2 \\ \textit{subject to} \parallel x_i \parallel_0 \leq T \end{aligned} \tag{1}$$

where Y is the set of input samples, D is the dictionary, $X = [x_1, \ldots, x_N]$ is the coefficient matrix, and T determines the sparse degree of the representation coefficient. It's obvious from the expression that the K-SVD only focuses on the representational power of the dictionary without considering its capability for discrimination.

Zhang et al. [29] proposed Discriminative K-SVD (D-KSVD) to obtain a dictionary that has both representational power and good discrimination capability based on the K-SVD model for general sparse representations. It added a simple linear regression term as the penalty term and its objective function is as follow:

$$\begin{aligned} <D, W, X> = \underset{D,W,X}{\arg\min} \parallel Y - DX \parallel_2 \\ + \gamma \parallel H - WX \parallel_2 + \beta \parallel W \parallel_2 \\ \textit{subject to} \parallel x_i \parallel_0 \leq T \end{aligned} \tag{2}$$

where H is the label of the training samples, W is the parameter of the classifier, and γ and β are scalars controlling the relative contribution of the corresponding terms. The first regularization term determines representation ability and the second term indicates discrimination power.

2.2 Label Consistent K-SVD

Jiang et al. [30] proposed Label Consistent K-SVD (LC-KSVD) to learn a discriminative dictionary for sparse coding. Besides the reconstruction error and the classification error mentioned in D-KSVD, they introduced a new label consistency constraint called "discriminative sparse-code error" into the objective function. The newly added term can connect label information with each dictionary atom to enforce discriminability in sparse codes during the dictionary learning process. Its objective function can be defined as follows:

$$\begin{aligned} <D, W, A, X> = \underset{D,W,A,X}{\arg\min} \parallel Y - DX \parallel_2 \\ + \gamma \parallel H - WX \parallel_2 + \beta \parallel Q - AX \parallel_2 \\ \textit{subject to} \parallel x_i \parallel_0 \leq T \end{aligned} \tag{3}$$

where the same notation and terms used in D-KSVD have the same meaning, and $Q = [q_1, q_2, \ldots, q_N] \in \mathbb{R}^{K \times N}$ are the "discriminative" sparse codes of input samples $Y$for classification. And $q_i = [q_i^1, q_i^2, \ldots, q_i^K]^t = [0, \ldots, 1, 1, \ldots, 0]^t \in \mathbb{R}^K$ is the class indicator of the input sample y_i, where the nonzero value indicates that the input sample has the same label as the dictionary atoms. The term $\parallel Q - AX \parallel_2$ represents the discriminative sparse-code error, which enforces that the transformed sparse codes AX approximate the discriminative sparse codes Q. It forces the samples from the same class to have very similar sparse representations, which results in good classification performance even using a simple linear classifier.

2.3 Fisher Discrimination Dictionary Learning

Yang et al. [31] proposed a novel Fisher discrimination-based dictionary learning (FDDL) scheme. The Fisher discrimination criterion was imposed on the coding coefficients to improve its discrimination. Meanwhile, they learned a structured dictionary which associated the dictionary atoms with the corresponding class labels. It made the reconstruction error discriminative. In order to achieve the above target functions, they designed the following objective function:

$$
\begin{aligned}
<D, X> = & \arg\min_{D,X} L(Y, D, X) \\
& + \alpha \parallel X \parallel_1 + \beta f(X)
\end{aligned} \tag{4}
$$

where $Y = [Y_1, Y_2, \ldots, Y_c]$ is the set of training samples, and Y_i is the subset of the training samples from class i, and c is the total number of classes. Both the dictionary $D = [D_1, D_2, ..., D_c]$ and the coding coefficients $X = [X_1, X_2, ..., X_c]$ are discriminative, where D_i is the class-specified sub-dictionary associated with category and X_i is the sub-matrix containing the coding coefficients of Y_i over D. The first item determines the discrimination of the learned dictionary; the second term is the sparsity constraint; $f(X)$ indicates the discrimination ability of the coefficient matrix X.

The first term in Eq. 4 is usually called as the discriminative fidelity term. It can be formulated as follow:

$$
\begin{aligned}
L(Y, D, X) = & \parallel Y_i - DX_i \parallel_F^2 + \parallel Y_i - D_i X_i^i \parallel_F^2 \\
& + \sum_{j \neq i} \parallel D_j X_i^j \parallel_F^2
\end{aligned} \tag{5}
$$

By minimizing the above formula, each class-specific sub-dictionary in the whole structured dictionary is good at representing the training samples from the associated class but poor at representing other classes.

The third term $f(X)$ which called as Fisher-based discriminative coefficient term, is designed as follow:

$$
f(X) = \sum_{i=1}^{c} \parallel X_i - M_i \parallel_F^2 - \parallel M_i - M \parallel_F^2 + \parallel X \parallel_F^2 \tag{6}
$$

where M_i and M are the mean vector of X_i and X respectively. Through minimizing the function, the coding coefficient is discriminative since it can minimize the intra-class differences and maximize the inter-class differences of X at the same time.

The purpose of the three dictionary learning method mentioned above is to obtain a discriminative dictionary. Their relationships are shown in Fig. 1. It's obvious that the three methods incorporate discriminative terms into objective function to get a discriminative dictionary.

3 Applications in Medical Image Analysis

Dictionary learning methods have been successfully exploited in several medical image analysis tasks, including image de-noising and a variety of classification works. In this paper, we focus on the CT image de-noising and histopathological image classification. Finally, on behalf of demonstrating the potential of dictionary learning in the field of medical image analysis, we apply the dictionary learning method to the glaucoma diagnosis task.

3.1 CT Image De-noising

In the past decade, Computed Tomography (CT) has become an important tool in clinical diagnosis, but it also brings high radiation doses to patients during X-ray CT examinations. Therefore, in clinical practice, decreasing the X-ray tube current is generally utilized to decrease radiation dose, so as to reduce the harm it causes to human. Nevertheless, low-dose CT (LDCT) is bound to provide degraded images due to increased noise and artifacts.

Noise can be removed efficiently by sparse representation model. Given a noisy image, sparse representation model assumes that the noise-free image can be sparsely represented via a redundant dictionary, whereas the dictionary cannot represent the noisy parts sparsely. Then the noise can be removed efficiently according to this difference.

Based on the idea mentioned above, many dictionary learning methods have been applied to improve the performance of sparse representation model for medical image de-noising. Yu et al. [6] first proposed a method of CT image

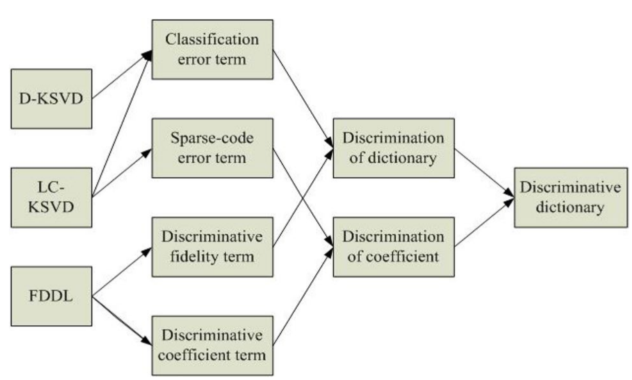

Fig. 1 Relationship between D-KSVD, LC-KSVD, and FDDL

de-noising based on sparse representation using a global dictionary. They applied the K-SVD algorithm to train a global dictionary used to approximate the image's local feature. If different images had similar features, they could use one dictionary to de-noise these images although these images were different. When it was used to reconstruct the image, they added gauss weight, which improved the performance of de-noising. Bai et al. [14] proposed a method of de-noising based on clustering integration and sparse dictionary learning, then applied it to medical image de-noising. In this approach, the kernel regression and clustering integration were combined to improve clustering performance. The sparse dictionary learning algorithm was then used to learn the dictionaries of each cluster. Since the method used both self-similarity and sparse prior knowledge, training dictionary was stable and adaptable. To prove the performance of their method, they took several experiments on two CT images and two MR images. Experimental results showed that the method had good performance and good stability between running time and efficiency comparing with several current excellent algorithms.

As for artifacts, it is found in [7] that high-contrast artifacts can also obtain prominent sparse representation. Therefore, it's hard to suppress it effectively via ordinary dictionary learning processing. Shi et al. [8] introduced an artifact suppressed dictionary learning algorithm (ASDL), which exploited scale and orientation information of artifacts to construct discriminative dictionaries for artifact suppression. The algorithm contained two main steps. On the first stage, three novel discriminative dictionaries were devised to characterize high-frequency artifact components in different orientations. Then, the residual noise and artifacts were suppressed by the general dictionary learning processing.

3.2 Histopathological Image Classification

Histopathological images carry rich structural information, making them important to the diagnosis of many diseases. Simultaneously, due to the diversity of histology features and rich geometric structure, it's a great challenge for doctors to make an analysis of histopathological images and diagnose diseases. In recent years, computer-aided diagnosis (CAD) technology has developed rapidly and been applied widely, which has been a great assist to doctors. Recently, sparse representation has played an important role in image classification tasks and gradually become one of the hottest research fields. Liu et al. [12] firstly exploited sparse representation for CAD classification. Subsequently, the sparse representation framework has been frequently applied to the medical domain.

Srinivas et al. [13, 19] firstly applied the sparse representation to histopathological image representation and classification. Considering the importance of the relevance of color information for histopathological image classification tasks, they proposed a simultaneous sparsity model (SHIRC) for multiple color channels of histopathological images. The authors exploited the correlations among the RGB channels of the color images in a sparse linear model setting under suitable channel-wise constraints. Subsequently, inspired by that pathologists often diagnose a tissue image based on local objects and the presence or absence of these local objects in an image matters much more than their absolute spatial location, they infused the simultaneous sparsity model with a robust locally adaptive flavor (LA-SHIRC) to address the issue of correspondence of local image objects located at different spatial locations. In fact, they replaced the whole image with local desired object blocks, then obtained the class decision of each test image block through SHIRC model, and combined the block decisions to identify the global image at last.

Whether it's SHIRC or LA-SHIRC, they both used all the training samples as atoms to construct dictionaries. However, recent work has shown that adaptive dictionaries significantly outperform ones constructed by simply stacking training samples together. To optimize the above frameworks, a novel method of Histopathological Image Classification based on Discriminative Feature-Oriented Dictionary Learning (DFDL) was proposed in [11, 18]. The main challenge in the histopathological analysis is the geometric richness of tissue images which results in difficulty to obtain reliable discriminative features for classification. To overcome the challenge, [11] proposed an automatic feature discovery framework via learning class-specific dictionaries and presented a low-complexity method for classification. It introduced the label information of training samples and directly learned health dictionary and diseased dictionary. The framework worked by minimizing intra-class differences, while simultaneously emphasizing inter-class differences in the dictionary learning stage.

Because of the geometric richness of histopathologic image and the various types of cells, the morphological and geometric changes of cells may be large in the homogeneous images, while there are certain similarities in non-homogeneous images, which results in a distance between image features belonging to the same class may be greater than the different category. Therefore, the healthy and diseased dictionaries learned by DFDL had high similarity. To make up for this shortcoming, Tang et al. [17] proposed a Discriminative feature-oriented dictionary learning method with the Fisher criterion (FCD-FDL) for histopathological image classification. In order to improve the difference between health dictionary and diseased dictionary, they used Fisher criterion to directly constrain the intra-class distance and the inter-class distance of the learning dictionary (minimize the intra-class distance and maximize the inter-class distance), rather than constrained sparse representation coefficient and obtained a more discriminative healthy dictionary and diseased dictionary. At the same time, better performance of the sparse representation was obtained by minimizing the reconstruction error based on the same class samples over the learned dictionaries and maximizing reconstruction error based on different class samples. Finally, the classifier was constructed by employing the reconstruction error vector of test samples.

As described earlier, local image regions of the histopathologic image from different classes may share common features. To solve the problem furtherly, Li et al. [20] proposed a new framework named analysis-synthesis model learning with shared features (ALSF) algorithm for histopathologic image classification. This framework could represent both similarities and differences in histo-pathological images from distinct classes more accurately due to the learning of a low rank shared dictionary and a shared analysis operator. And it combined analysis dictionary (determines the sparse code from the images) and synthesis dictionary (yields images by multiplying with the sparse code) to learn the classifier and the feature extractor at the same time, which reduced the computation load efficiently.

Comparing with the above three methods, as shown in Table 1, it shows that with the improvement of dictionary discrimination ability, the classification performance of the dictionary learning framework also improves. The results indicate the potential of dictionary learning method in the application of medical image classification tasks.

4 Low-rank Shared Dictionary Learning for Glaucoma Diagnosis

To demonstrate the availability of dictionary learning in medical image analysis, we apply it to glaucoma diagnosis task. Considering the high variability among glaucoma individuals, we propose to utilize low-rank shared dictionary learning(LRSDL) [32] method to availably capture discriminative features and obtain more reasonable representation. Finally, the method achieves accuracy of 92.90% on RIM-ONER2 datasets.

4.1 Computer-Aided Glaucoma Diagnosis

Glaucoma is an irreversible chronic blinding fundus disease [33]. The optic nerve injury caused by glaucoma cannot be recovered and cured. Therefore, early diagnosis and timely intervention are the most effective ways to prevent and treat glaucoma [34]. In recent years, with the successful application of computer-aided diagnosis (CAD) technology in clinical practice, the glaucoma recognition method based on computer vision technology has developed rapidly, and plays an indispensable role in the early diagnosis and automatic screening of glaucoma [35–37]. At present, the research of glaucoma recognition method based on computer-aided analysis mainly focuses on the detection and segmentation of optic disc and optic cup, and then the cup/disc ratio (CDR) is used as the classification criterion [38–40]. Although these methods have achieved good results, they still face the following challenges:

(1) The description of fundus features of glaucoma is inaccurate and incomprehensive. The fundus characteristics of different types of glaucoma at different periods are various. Especially, some pathological characteristics are manifested outside the optic disc, such as atrophy around the optic disc and retinal nerve fiber layer defect. Therefore, it's not comprehensive to identify glaucoma only according to CDR.
(2) Existing identification methods cannot cope with the high variability and extreme inhomogeneity of optic disc structure from fundus image across subjects.

To settle the above challenges, we extract the color, Gabor and Histogram of Gradient(HOG) features of the images to comprehensively describe the fundus characteristics. More specifically, we adopt LRSDL, which generates a shared dictionary and particular dictionary simultaneously, to deal

Table 1 Classification results comparison of different methods on ADL dataset

Class	Kidney		Lung		Spleen		Method
	Healthy	Diseased	Healthy	Diseased	Healthy	Diseased	
Healthy	0.8110	0.1890	0.7500	0.2500	0.6500	0.3500	SHIRC
	0.8723	0.1277	0.9234	0.0766	0.8999	0.1001	DFDL
	0.8809	0.1191	0.9509	0.0491	0.9064	0.0936	FCDFDL
	0.8550	0.1450	–	–	–	–	ALSF
Diseased	0.1946	0.8054	0.1500	0.8500	0.1167	0.8833	SHIRC
	0.1405	0.8595	0.0576	0.9424	0.0599	0.9401	DFDL
	0.1311	0.8689	0.0375	0.9625	0.0409	0.9591	FCDFDL
	0.1300	0.8700	–	–	–	–	ALSF

with high variability and extreme heterogeneity among individuals.

4.2 Low-rand Shared Dictionary Learning for Glaucoma Diagnosis

In practice, ophthalmologists mainly diagnose glaucoma by judging whether the fundus image from the patient has the pathological characteristics of glaucoma. In general, only a few small areas of the patient's fundus image contain pathological features. In other words, different fundus images from healthy and glaucoma individuals not only possess distinct class-specific features but also share common patterns. Based on this theory, we apply the Low-rank Shared Dictionary Learning (LRSDL) method [32] to glaucoma classification.

LRSDL model mainly contains two process: dictionary learning and classification. During the learning process, LRSDL introduces a shared dictionary D_0 to the model based on FDDL method. With the addition of D_0, a sample Y_i will be well represented by the collaboration of the particular dictionary D_i and the shared dictionary D_0. Formally, the total dictionary learned by the model is expressed as $D' = [DD_0]$, and let $X' = [X^T, (X^0)^T]^T$ and $X'_i = [(X_i)^T, (X_i^0)^T]^T$. The discriminative fidelity term $L(Y, D, X)$ in (5) is redefined as:

$$L'(Y, D', X') = \| Y_i - D_i X_i^i - D_0 X_i^0 \|_F^2 + \| Y_i - D' X'_i \|_F^2 + \sum_{j \neq i} \| D_j X_i^j \|_F^2 \tag{7}$$

where the shared dictionary is added to the reconstruction error term, but is excluded from the class-dictionary learning term. The Fisher-based discriminative coefficient term $f(X)$ is described as:

$$f'(X') = f(X) + \| X^0 - M^0 \|_F^2 \tag{8}$$

where M^0 is the mean vector of X^0 columns, and the additional item forces the coefficients of all training samples represented via the shared dictionary to be similar, so that the contribution of the shared dictionary to the reconstruction of every sample is approximately the same. Finally, the objective function of the LRSDL is:

$$<D', X'> = \arg\min_{D,X} L'(Y, D', X') + \alpha \| X' \|_1 + \beta f'(X') + \gamma \| D_0 \|_* \tag{9}$$

where the term $\| D_0 \|_*$ can constrain $rank(D_0)$ to be small to avoid the shared dictionary from containing the element of discriminative dictionary. Both the additional terms are crucial to improving the performance of dictionary learning model.

After the dictionary construction, we first acquire the coefficient vector x' for a new test sample y relied on the total dictionary D'. Formally, x' is calculated as follow:

$$x' = \arg\min_{x'} \| y - D'x' \|_2^2 + \alpha \| x' \|_1 + \beta \| x^0 - m^0 \|_2^2 \tag{10}$$

Then, we extract the contribution of the shared dictionary before classification. That's to say, the class label of y is determined by the particular dictionary. Formally, the classification scheme can be described as:

$$\arg\min_{1 \leq i \leq C} (\omega \| y' - D_i x^i \|_2^2 + (1 - \omega) \| x - m_i \|_2^2) \tag{11}$$

where $y' = y - D_0 x^0$, and $\omega \in [0, 1]$ is a preset weight for balancing the contribution of the two terms.

The whole implementation framework of our experiment is demonstrated as Fig. 2. It's shown that the total framework is consisted of three distinguished parts: feature extraction, dictionary construction and glaucoma diagnosis. In feature extraction stage, in order to depict the changes of the optic disc for glaucoma accurately and comprehensively, this experiment comprehensively utilized color distribution, Gabor filtering and Histogram of Gradient (HOG) methods to describe the fundus image features from different aspects such as color, structure and texture. The features extracted from fundus images are shown as Fig. 3. The three features are combined at last to comprehensively describe fundus images in the form of vectors. In the dictionary construction stage, the obtained eigenvectors are fed to LRSDL model to achieve dictionary learning. Considering that different fundus images not only possess distinct class-specific features but also usually share common patterns with each other, LRSDL model gets a particular and shared dictionary at the same time in the dictionary learning stage. This way the fundus image can be described more reasonably. In the classification stage, the test samples are represented by the combined dictionary and its sparse coefficient is obtained. Finally, the classification criterion adds the intra-class distance of the coefficient on the basis of reconstruction error. Specially, the reconstruction error is calculated with shared dictionary excluded.

4.3 Experiments and Results

The experiment is conducted on a publicly available glaucoma image dataset, named RIM-ONE R2, which includes 455 images from different individuals, 255 from healthy individuals and 200 from patients with glaucoma at different stages. In the experiment, 150 images are randomly selected from healthy individuals and glaucoma patients respectively as training datasets, and the rest are used as test samples. As

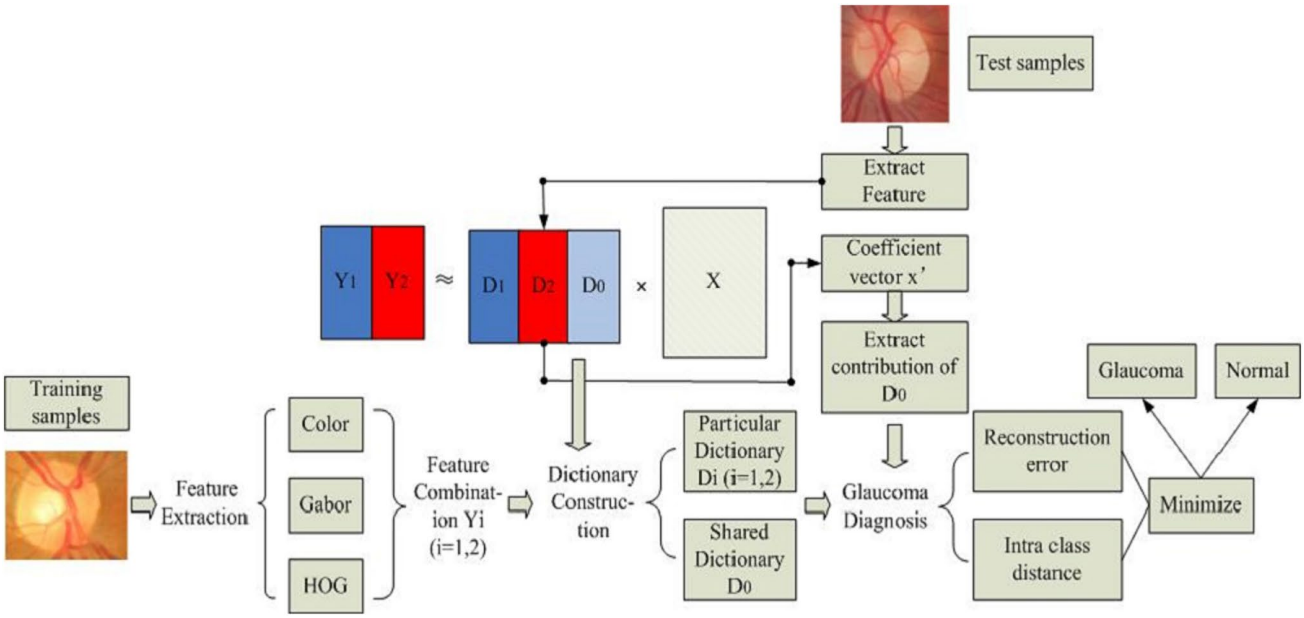

Fig. 2 The overall framework of the experiment

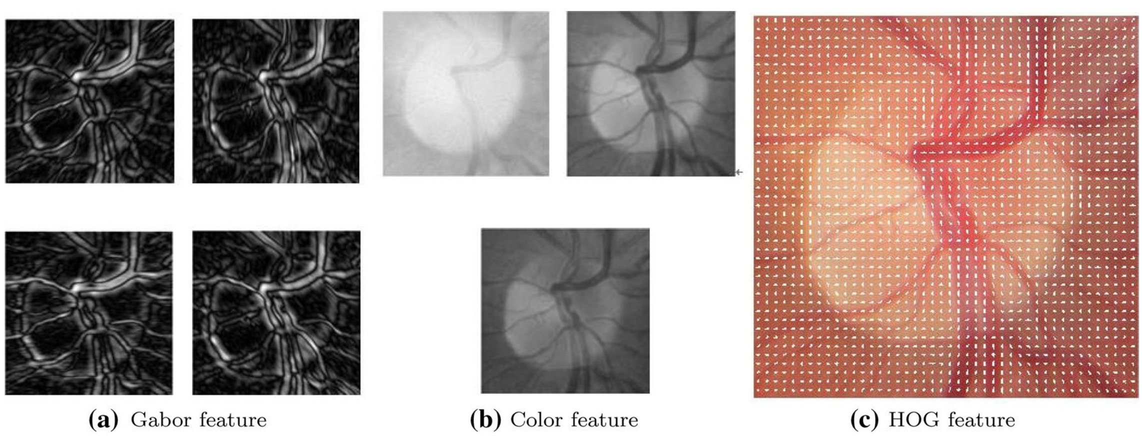

Fig. 3 Feature extraction on fundus images. **a** Multi-direction gabor feature highlight the texture information of the image; **b** extract the information of the three channels of RGB respectively; **c** HOG feature focus on capturing the edge information of blood vessels

for dictionary construction, the shared dictionary contains 10 dictionary atoms, each particular dictionary contains 60 words, and finally, the whole dictionary contained 130 words.

For evaluation, we employ three evaluation criteria to measure the performance of our method, including: $Sensitivity = TP/(TP + FN)$, $Specificity = TN/(TN + FP)$, and $Accuracy = (TP + TN)/(TP + TN + FN + FP)$, where TP and TN denote the number of true positives and true negatives, respectively, and FP and FN denote the number of false positives and false negatives, respectively. Sensitivity refers to the proportion of the samples actually suffering from glaucoma that are correctly recognized, which is equivalent to the ability of the algorithm to recognize glaucoma. Specificity represents the ability of the algorithm to recognize healthy samples. Accuracy is the proportion of all test samples correctly classified.

We conduct ten times experiments randomly on the dataset, and the best results reaches accuracy of 0.9290, sensitivity of 0.9600, and specificity of 0.9143. The experimental results demonstrate that our method is capable of distinguishing glaucoma. To prove the competitiveness of our method, we compare our method with several existing state-of-the-art glaucoma screening baselines. As shown in Table 2, our method obtains higher accuracy of 0.9290 based on RIM-ONE R2 dataset, which increases the accuracy by

Table 2 Performance comparison between LRSDL and existing methods (No. represents number of image from dataset; Acc., Sen. and Spec. represent the accuracy, sensitivity and specificity respectively.)

Method	Dataset (No.)	Acc.	Sen.	Spec.
RF [41]	RIM-ONE R2 (455)	0.8242	0.8706	0.7650
	Drishti-GS (101)	0.7843	0.9571	0.7073
DENet [42]	SCES (1676)	0.8429	0.8478	0.8380
	SINDI (5783)	0.7495	0.7876	0.7115
Our LRSDL	RIM-ONE R2 (455)	0.9290	0.9600	0.9143

10.48% compared with [41], and our algorithm can get better results on fewer datasets compared with [42]. This mainly profit from the strong expressive power and discrimination power of the dictionary learned by LRSDL model.

5 Conclusion

Dictionary learning methods have emerged as one of the most successful methods for medical image analysis. In recent years, numerous studies have proved their practical value. This paper summarizes some main concepts and methods of dictionary learning and reviews some of the published research on the application of dictionary learning methods to de-noising and classification for medical images. Most of these studies have obtained very good results. Especially, we apply low-rank shared dictionary learning to achieve accurate glaucoma diagnosis. The practice of dictionary learning method proves its prospect in the field of medical image analysis.

Funding This study was funded by NSFC (Grant Numbers 61702558, 61602527), Hunan Natural Science Foundation (Grant Numbers 2017JJ3411, 2017JJ3416), Primary Research & Developement Plan of Hunan Province (Grant Number 2017WK2074), National Key Research and Development Program of China (Grant Number 2017YFC0840104), China Scholarship Council (Grant Number 201806375006).

Compliance with Ethical Standards

Conflict of interest The author declares that he has no conflict of interest.

References

1. Khan A, Doucette JA, Cohen R, Lizotte DJ (2012) Integrating machine learning into a medical decision support system to address the problem of missing patient data. In: 2012 11th international conference on machine learning and applications, Boca Raton, FL, pp 454–457
2. Seixas JL, Barbon S, Mantovani RG (2015) Pattern recognition of lower member skin ulcers in medical images with machine learning algorithms. In: 2015 IEEE 28th international symposium on computer-based medical systems, Sao Carlos, pp 50–53
3. Mothkur R, Km P (2018) Machine learning will transfigure medical sector: a survey. In: 2018 international conference on current trends towards converging technologies (ICCTCT), Coimbatore, pp 1–8
4. Song L, Peng J (2012) Dictionary learning research based on sparse representation. In: 2012 international conference on computer science and service system, Nanjing, pp 14–17
5. Zhang CZ, Zhao L (2011) Image de-noising based on learned dictionary. In: 2011 International conference on multimedia technology, Hangzhou, pp 3101–3104
6. Yu F, Chen Y, Luo L (2013) CT image denoising based on sparse representation using global dictionary. In: 2013 ICME International conference on complex medical engineering, Beijing, pp 408–411
7. Chen Y, Yin X, Shi L et al (2013) Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing. Phys Med Biol 58:5803–5820
8. Shi L, Chen Y, Shu H, Luo L, Toumoulin C, Coatrieux J (2014) Low-dose CT image processing using artifact suppressed dictionary learning. In: 2014 IEEE 11th international symposium on biomedical imaging (ISBI), Beijing 2014, pp 1127–1130
9. Shan B, Hao W, Zhao R (2012) Infrared image de-noising based on K-SVD over-complete dictionaries learning. In: 2012 5th international congress on image and signal processin, Chongqing 2012, pp 316–320
10. Ji J, Ren F, Ji H, Yao Y, Hou G (2018) Generalised non-locally centralised image de-noising using sparse dictionary. IET Image Process 12(7):1072–1078
11. Vu TH, Mousavi HS, Monga V, Rao G, Rao UKA (2016) Histopathological Image Classification Using Discriminative Feature-Oriented Dictionary Learning. IEEE Trans Med Imaging 35(3):738–751
12. Liu M et al (2011) Sparse classification for computer aided diagnosis using learned dictionaries. Proc Med Image Comput Comput Assist Interv (MICCAI) 6893:41–48
13. Srinivas U, Mousavi HS, Monga V, Hattel A, Jayarao B (2014) Simultaneous sparsity model for histopathological image representation and classification. IEEE Trans Med Imag 33(5):1163–1179
14. Bai J, Song S, Fan T et al (2018) Medical image denoising based on sparse dictionary learning and cluster ensemble. Soft Comput. https://doi.org/10.1007/s00500-017-2853-7
15. Sun X, Nasrabadi NM, Tran TD (2015) Task-driven dictionary learning for hyperspectral image classification with structured sparsity constraints. IEEE Trans Geosci Remote Sens 53(8):4457–4471
16. Diamant I, Klang E, Amitai M, Konen E, Goldberger J, Greenspan H (2017) Task-driven dictionary learning based on mutual information for medical image classification. IEEE Trans Biomed Eng 64(6):1380–1392
17. Tang HZ, Li X, Zhang XG, Zhang DB, Wang X, Mao LZ (2018) Discriminative feature-oriented dictionary learning method with Fisher criterion for histopathological image classification. Acta Autom Sin 44(10):1842–1853
18. Vu TH, Mousavi HS, Monga V, Rao UKA, Rao G (2015) DFDL: Discriminative feature-oriented dictionary learning for histopathological image classification. In: IEEE 12th international symposium on biomedical imaging (ISBI), New York, NY, pp 990-994
19. Srinivas U, Mousavi HS, Jeon C, Monga V, Hattel A (2013) SHIRC: A simultaneous sparsity model for histopathological image representation and classification. In: Proceedings of

the IEEE international symposium biomedical imaging, pp 1118–1121
20. Li X, MongaV , Rao UKA (2018) Analysis-synthesis model learning with shared features: a new framework for histopathological image classification. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), Washington, DC, pp 203–206
21. Roy S et al (2015) Subject-specific sparse dictionary learning for atlas-based brain MRI segmentation. IEEE J Biomed Health Inform 19(5):1598–1609
22. Shi Y, Wang D, Liu Z (2015) Hippocampus segmentation in MR images using sparse patch representation and discriminative dictionary learning. In: IET international conference on biomedical image and signal processing (ICBISP 2015), Beijing, pp 1–5
23. Guo Y, Zhan Y, Gao Y, Jiang J, Shen D (2013) MR prostate segmentation via distributed discriminative dictionary (DDD) learning. In: 2013 IEEE 10th international symposium on biomedical imaging, San Francisco, CA, pp 868–871
24. Dong Y, Hansen PC, Kjer HM (2018) Joint CT reconstruction and segmentation with discriminative dictionary learning. IEEE Trans Comput Imag 4(4):528–536
25. Jin P, Ye DH, Bouman CA (2015) Joint metal artifact reduction and segmentation of CT images using dictionary-based image prior and continuous-relaxed potts model. In :2015 IEEE international conference on image processing (ICIP), Quebec City, QC, pp 798–802
26. Onofrey JA, Staib LH, Papademetris X (2019) Segmenting the brain surface from CT images with artifacts using locally oriented appearance and dictionary learning. IEEE Trans Med Imag 38(2):596–607
27. Wright J, Yang A, Ganesh A, Sastry S, Ma Y (2009) Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell 31:210–227
28. Aharon M, Elad M, Bruckstein A (2006) K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Signal Process 54(11):4311
29. Zhang Q, Li B (2010) Discriminative K-SVD for dictionary learning in face recognition. In: 2010 IEEE computer society conference on computer vision and pattern recognition, San Francisco, CA, pp 2691–2698
30. Jiang Z, Lin Z, Davis LS (2013) Label consistent K-SVD: learning a discriminative dictionary for recognition. IEEE Trans Pattern Anal Mach Intell 35(11):2651–2664
31. Yang M, Zhang L, Feng X, Zhang D (2011) Fisher discrimination dictionary learning for sparse representation. In: 2011 International conference on computer vision, Barcelona, pp 543–550
32. Vu TH, Monga V (2017) Fast low-rank shared dictionary learning for image classification. IEEE Trans Image Process 26(11):5160–5175
33. Roodhooft J (2002) Leading causes of blindness worldwide. Bull Soc Belge Ophtalmol 283:19–25
34. Michelson G, Hornegger J, Wärntges S, Lausen B (2008) The papilla as screeningparameter for early diagnosis of glaucoma. Deutsches Arzteblatt Int 105:334–35
35. Haleem MS, Han L, Hemert JV, Li B (2013a) Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: a review. Comput Med Imag Graph 37:581–596
36. Khalil T, Khalid S, Syed AM (2014) Review of machine learning techniques for glaucoma detection and prediction. In: Science and information conference, London, pp 438–442
37. Nawaldgi S (2016) Review of automated glaucoma detection techniques. In: 2016 international conference on wireless communications, signal processing and networking (WiSPNET), Chennai, pp 1435–1438
38. Yin F et al (2012) Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis. In: 2012 25th IEEE international symposium on computer-based medical systems (CBMS), Rome, pp 1–6
39. Cheng J, Yin F, Wong DWK, Tao D, Liu J (2015) Sparse dissimilarity-constrained coding for glaucoma screening. IEEE Trans Biomed Eng 62(5):1395–1403
40. Poshtyar A, Shanbehzadeh J, Ahmadieh H (2013) Automatic measurement of cup to disc ratio for diagnosis of glaucoma on retinal fundus images. In: 2013 6th international conference on biomedical engineering and informatics, Hangzhou, pp 24–27
41. Zhao R, Chen Z, Duan X, Chen Q, Liu K, Zhu C (2017) Automated glaucoma detection based on multi-channel features from color fundus images. J Comput Aided Des Comput Graph 29:998–1006
42. Fu H, Cheng J, Xu Y, Zhang C, Wong DWK, Liu J, Cao X (2018) Disc-aware ensemble network for glaucoma screening from fundus image. IEEE Trans Med Imag 37(11):2493–2501