Computers and Electrical Engineering 90 (2021) 106992

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Fine-grained correlation analysis for medical image retrieval☆,☆☆

Xiaoqin Wang [a], Rushi Lan [a,*], Huadeng Wang [b], Zhenbing Liu [b], Xiaonan Luo [c]

[a] Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin 541004, China
[b] School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China
[c] National Local Joint Engineering Research Center of Satellite Navigation and Location Service, Guilin University of Electronic Technology, Guilin 541004, China

ARTICLE INFO

Keywords:
Feature fusion
K-means clustering
Medical image retrieval
Canonical correlation analysis

ABSTRACT

Feature fusion in medical image retrieval remains a challenging task because of the high-dimensional data and massive amount of irrelevant information in images. To solve these issues, we propose a novel feature fusion method, called fine-grained correlation analysis (FGCA), for medical image retrieval. First, we analyze the problem that there are many irrelevant local regions in a category. To solve this problem, an image is partitioned into some fine-grained samples. Then, the fine-grained samples with similar characteristics are tagged with the same label by the k-means clustering algorithm. Finally, we investigate how the correlation relationship extracted from the fine-grained samples helps fuse different features and obtain the more discriminative and less redundant information for medical image retrieval. Experiments on three medical image datasets show that our proposed FGCA approach works better than the conventional methods.

1. Introduction

For the last several years, medical imaging technology, such as magnetic resonance imaging (MRI), computed tomography (CT), photon emission tomography (PET) and X-ray, has produced massive numbers of medical images. Selecting some similar images from a large dataset is particularly difficult for doctors. It is a time-consuming and labor-intensive task. Medical image retrieval is a useful technique that provides some medical images similar to a given query image. It plays an important role in medical image analysis and raises the necessity of constructing intelligent retrieval systems. An effective feature representation is the key for an effective medical image retrieval system. Currently, some medical image retrieval systems with a single feature have been used. However, these systems have the limitation of retrieving medical images using the single feature, which cannot comprehensively and effectively describe the characteristics of a medical image. Therefore, it is necessary to design a feature representation that combines different characteristics of images for medical image retrieval systems.

There are several representative medical image retrieval algorithms that are based on a single feature. The first one is the local binary pattern (LBP) [1,2]. It encodes the sign of the differences between the center pixel and its surrounding neighbors to represent images, and it has been widely used as a texture analysis technique, such as in face recognition [3] and biomedical image retrieval

☆ This work was partially supported by the National Natural Science Foundation of China (Nos. 61702129, 61772149, 61762028, U1701267, and 61936002), the Guangxi Science and Technology Project (Nos. 2019GXNSFFA245014, 2018GXNSFAA294127, AD18281079, and AD18216004), Guangxi Key Laboratory of Image and Graphic Intelligent Processing, China (No. 2001), and by the Project of Cultivating Excellent Dissertations for Graduate of GUET, China (No. 17YJPYSS15).

☆☆ This paper is for CAEE special section VSI-aicv4. Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. Yujie Li.

* Corresponding author.
E-mail address: rslan2016@163.com (R. Lan).

https://doi.org/10.1016/j.compeleceng.2021.106992
Received 16 January 2020; Received in revised form 23 August 2020; Accepted 18 January 2021
Available online 5 February 2021

(local ternary cooccurrence patterns (LTCoP)) [4]. LTCoP calculates the gray values of the center pixel and its surrounding neighbors to represent a biomedical image. Another texture feature extraction method is the local wavelet pattern (LWP) [5]. The histogram of compressed scattering coefficients (HCSC) [6] uses the scattering transform to obtain the representations of a medical image and is proposed for medical image retrieval. Although the above approaches have extracted effective representations for pattern recognition, they cannot reflect the different characteristics of patterns to provide a more discriminative feature for medical image retrieval.

Feature fusion is a technique that aims to optimize combined features. Different features extracted from the same image always reflect different characteristics, such as shape, texture and color. Feature fusion helps improve the pattern recognition performance by keeping relevant information and eliminating redundant information [7]. There are two simple feature fusion methods. One is called serial feature fusion [8] and is based on a union vector. The other is based on a complex vector and called parallel feature fusion [9]. Ref. [10] verifies that serial feature fusion and parallel feature fusion can improve the classification accuracy [11]. However, it is difficult for these two methods to reveal plenty of inherent correlations between two sets of variables [12]. To remedy this problem, Sun et al. proposed a new feature fusion method using canonical correlation analysis (CCA) [13]. CCA is a powerful statistical technique that reveals the linear correlation relationship between two multidimensional variables [14]. The CCA-based feature fusion method has been widely applied in face recognition. Ref. [15] also proposed an integrated scattering feature, called scattering transform canonical correlation analysis (ST-CCA), for medical image retrieval. It integrated two types of compressed scattering data via canonical correlation analysis and took account of more comprehensive characteristics of medical images. However, CCA-based feature fusion only considers the correlation between two sets features, and it ignores the similarity of the samples within the same category, which is very important in pattern recognition. It is also an unsupervised feature fusion method. To take advantage of label information and transform CCA from an unsupervised method into a supervised method, Ref. [16] proposed a variant of CCA called generalized canonical correlation analysis (GCCA). It does not only consider the correlation between two sets of variables, but it also considers the within-class information of training samples. It gets better face recognition rates than CCA-based feature fusion by maximizing the correlation between two sets of features and minimizing the within-class scatter matrix. However, CCA and GCCA have difficulties calculating the projection matrix with high-dimensional data, and the local information of an image is not considered when extracting the correlation relationship. Therefore, both of them are not suitable for medical image retrieval.

For medical image retrieval, it is necessary to capture the heterogeneous nature of images for predictive modeling. It is worth noting that many medical images share similar structural information or local information, and the heterogeneous natural information and the similar local structural information can be defined as the fine-grained characters (samples) [17]. In this paper, we investigate how the correlation analysis based on fine-grained samples helps fuse different features and obtain the more discriminative and less redundant information of an image, and we propose a novel feature fusion framework called fine-grained correlation analysis (FGCA) for medical image retrieval. Our contributions are summarized as follows:

- Extracting some fine-grained samples of a medical image, analyzing the correlation between different fine-grained samples within the same category, and generating the pseudolabel [18] for each fine-grained sample;
- Proposing a novel feature fusion framework called FGCA that extracts the discriminative feature representation from different combined features for medical image retrieval;
- Evaluating the proposed FGCA framework on three medical image datasets.

The rest of this paper is organized as follows. In Section 2, we briefly introduce the preliminary knowledge. The proposed FGCA framework for medical image retrieval is presented in Section 3. The experiments and results on different medical image datasets are shown in Section 4. The conclusion is finally drawn in Section 5.

2. Preliminary

2.1. Canonical correlation analysis

CCA [14] establishes the correlation criterion function between two sets of variables and then analyzes the linear relationship. Suppose there are two zero-mean matrices $X = [x_1, x_2, \dots, x_N]$ and $Y = [y_1, y_2, \dots, y_N]$, and $X \in R^p$, $Y \in R^q$, N is the number of variables, and p and q are the dimensions of the attributes of each variable for X and Y, respectively. Let $S_{xx} \in R^{p\times p}$ and $S_{yy} \in R^{q\times q}$ denote the within-set scatter matrices of X and Y, respectively, as follows:

$$S_{xx} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - m_x\right)\left(x_i - m_x\right)^T,$$

$$S_{yy} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - m_y\right)\left(y_i - m_y\right)^T,$$

where $x_i \in X$ and $y_i \in Y$, and m_x and m_y are the mean vectors of X and Y, respectively. Let S_{xy} denote the between-set scatter matrix of X and Y as follows:

$$S_{xy} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - m_x\right)\left(y_i - m_y\right)^T.$$

The aim of CCA is to find the transformation pairs α and β by maximizing the correlation criterion function of X and Y, which is defined as follows:

$$\max_{\alpha,\beta} J(\alpha,\beta) = \frac{\alpha^T S_{xy}\beta}{\sqrt{\alpha^T S_{xx}\alpha \cdot \beta^T S_{yy}\beta}} \quad s.t. \quad \alpha^T S_{xx}\alpha = \beta^T S_{yy}\beta = 1. \tag{1}$$

Lagrange multiplier method is performed for Eq. (1), and the question is converted to solve the two generalized eigenproblems $M_{xy} = S_{xx}^{-1}S_{xy}S_{yy}^{-1}S_{yx}$ and $M_{yx} = S_{yy}^{-1}S_{yx}S_{xx}^{-1}S_{xy}$. α and β are obtained by solving the eigenvalues and eigenvectors of M_{xy} and M_{yx} as follows:

$$M_{xy}\alpha = \lambda^2\alpha, \tag{2}$$

$$M_{yx}\beta = \lambda^2\beta. \tag{3}$$

The number of nonzero eigenvalues in Eq. (2) and Eq. (3) is $d \le r$, where r is the rank of S_{xy}, and they will be sorted in decreasing order, $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_d$. The transformation matrices A and B are consisted of the sorted eigenvectors corresponding to the nonzero eigenvalues, i.e., $A = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ and $B = (\beta_1, \beta_2, \ldots, \beta_d)$. The linear transformations of X^* and Y^* are shown as follows:

$$X^* = X^T(\alpha_1, \alpha_2, \ldots, \alpha_d) = X^T A,$$

$$Y^* = Y^T(\beta_1, \beta_2, \ldots, \beta_d) = Y^T B,$$

In Ref. [13], the fusion strategies are defined as $Z_1 = (X^*, Y^*)$ and $Z_2 = X^* + Y^*$, where Z_1 is linear feature fusion strategy I (LFFS-I), and Z_2 is linear feature fusion strategy II (LFFS-II).

2.2. Generalized canonical correlation analysis

Generalized canonical correlation analysis (GCCA) was proposed for face recognition [16]. It converts CCA from an unsupervised model to a supervised model. GCCA utilizes label information to analyze the correlation relationship between two sets of variables, which allows GCCA to achieve better performance than CCA-based feature fusion. The objective of GCCA is that the two sets of variables have the maximum correlation when the projected variables remain in the minimum within-class scatter matrix. Therefore, the difference between GCCA and CCA is the within-class scatter matrices S_{gx} and S_{gy}, which are defined as follows:

$$S_{gx} = \sum_{i=1}^{C}\sum_{j=1}^{c_i}(x_{ij} - m_i^x)(x_{ij} - m_i^x)^T,$$
$$S_{gy} = \sum_{i=1}^{C}\sum_{j=1}^{c_i}(y_{ij} - m_i^y)(y_{ij} - m_i^y)^T,$$

where C is the number of classes, and c_i is the number of samples in class i. $x_{ij} \in X$ and $y_{ij} \in Y$ denote the jth sample in class i. m_i^x and m_i^y are the mean vectors of X and Y in class i, respectively.

3. Proposed feature fusion framework

3.1. Motivation

CCA-based and GCCA-based feature fusion algorithms have been effectively applied to face images. However, the unique characteristics of medical images bring about new challenges to the above algorithms. One is that medical images always are high-dimensional data, which make it difficult for CCA and GCCA to perform eigenvalue decomposition. Another is that medical images have considerable amounts of local irrelevant information within in the same category, which is neglected by the CCA and GCCA methods. Fig. 1 shows the correlation relationship between fine-grained samples, and (a) illustrates the correlation relationship of different fine-grained samples X^*X^{*T}, which are transformed by GCCA using two medical images from the same category. In this figure, the white represents that fine-grained sample i and fine-grained sample j have the highest correlation, and the dark elements represent those with lower correlation. The implication of Fig. 1(a) is that there is considerable irrelevant fine-grained information in the same category. Fig. 1(b) shows the distribution of fine-grained samples X^*. The orange points are obtained from the same medical image, and the blue points are obtained from another medical image. The results clearly show that it is difficult to distinguish the two images based on the GCCA method. To deal with these problems, we propose a novel feature fusion framework for medical image retrieval.

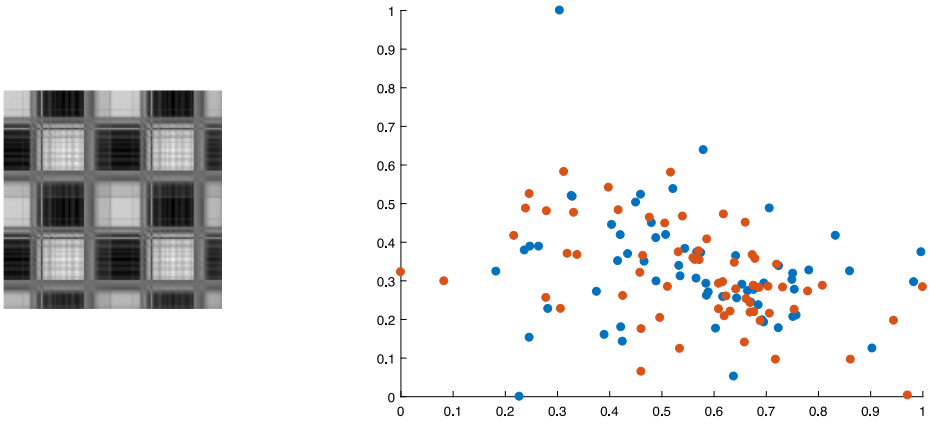


Fig. 1. The correlation relationship between fine-grained samples. (a) Visualization of the covariance matrix between fine-grained samples X^*X^{*T}, where the white represents the highest correlation, and the dark represents the elements with lower values. (b) The distribution of fine-grained samples X^*, which are obtained from two medical images within the same category.

3.2. Fine-grained correlation analysis

To investigate the correlations between fine-grained samples and solve the issues described in Section 3.1, we propose a novel fine-grained correlation analysis (FGCA) framework for medical image retrieval. Take two matrices $X \in R^{N\times p}$ and $Y \in R^{N\times q}$, where N is the number of medical images, and p and q are the feature dimensions of X and Y, respectively. The fine-grained samples of X and Y are obtained by partitioning each image into m subsamples, and the subsamples of X and Y have the dimensions of p' and q', respectively. Then, the fine-grained training samples are defined as $X' \in R^{M\times p'}$ and $Y' \in R^{M\times q'}$, where $M = m\times N$ is the number of fine-grained samples. To remove the local irrelevant fine-grained samples in the same category and gather fine-grained samples with similar attributes into the same category, the k-means clustering algorithm is performed to train the fine-grained samples. Then, the pseudolabel of each fine-grained sample is defined as follows:

$$l_{ij} = \begin{cases} 1 & if\ jth\ fine\text{-}grained\ sample\ belongs\ to\ the\ class\ i \\ 0 & others. \end{cases}$$

After k-means clustering, each fine-grained sample has only one pseudolabel. Based on the pseudolabels, the respective within-class scatter matrices $S_{fx'}$ and $S_{fy'}$ of X' and Y' are defined as follows:

$$S_{fx'} = \sum_{i=1}^{K}\sum_{j=1}^{M} l_{ij}^{x'}(x'_{ij} - m_i^{x'})(x'_{ij} - m_i^{x'})^T,$$
$$S_{fy'} = \sum_{i=1}^{K}\sum_{j=1}^{M} l_{ij}^{y'}(y'_{ij} - m_i^{y'})(y'_{ij} - m_i^{y'})^T,$$

where K is the number of clusters, and $l_{ij}^{x'}$ and $l_{ij}^{y'}$ are the pseudolabels of X' and Y', respectively. x'_{ij} and y'_{ij} are the jth samples of the ith class of X' and Y', respectively. $m_i^{x'}$ and $m_i^{y'}$ are the average vectors of class i of X' and Y', respectively. The between-set scatter matrix of X' and Y' is presented as follows:

$$S_{fx'y'} = \sum_{i=1}^{M}(x'_i - m^{x'})(y'_i - m^{y'})^T,$$

where $r' = rank\left(S_{fx'y'}\right)$ and $S_{fx'y'} = S_{fy'x'}^T$. $m^{x'}$ and $m^{y'}$ are the global averages of X' and Y', respectively. The objective function of FGCA is written as follows:

$$J(w_{x'}, w_{y'}) = \frac{w_{x'}^T S_{fx'y'} w_{y'}}{\sqrt{w_{x'}^T S_{fx'} w_{x'} \cdot w_{y'}^T S_{fy'} w_{y'}}},$$
$$s.t.\ w_{x'}^T S_{fx'} w_{x'} = w_{y'}^T S_{fy'} w_{y'} = 1. \tag{4}$$

where $w_{x'}$ and $w_{y'}$ are the projective pairs. FGCA finds $w_{x'}$ and $w_{y'}$ by maximizing Eq. (4) with restrictions. The Lagrange multiplier method is used to solve the optimization problem, and the projective pairs $w_{x'}$ and $w_{y'}$ are composed by the first $d' \le r'$ (r' is the rank of $S_{fx'y'}$) pair eigenvectors after sorting the eigenvalues from large to small, which are shown as $W_{x'} = \left[w_{x'}^1, w_{x'}^2, \ldots, w_{x'}^{d'}\right]$ and

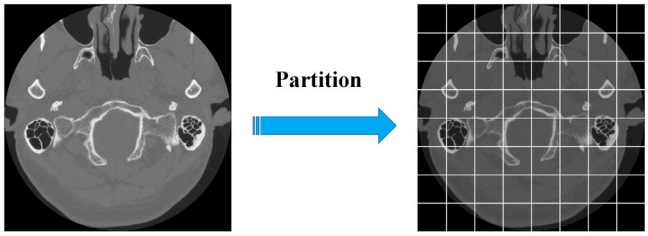


Fig. 2. An example process of extracting fine-grained samples: (a) a CT image sized 512 × 512; and (b) 64 fine-grained samples, and each fine-grained sample is sized 64 × 64.

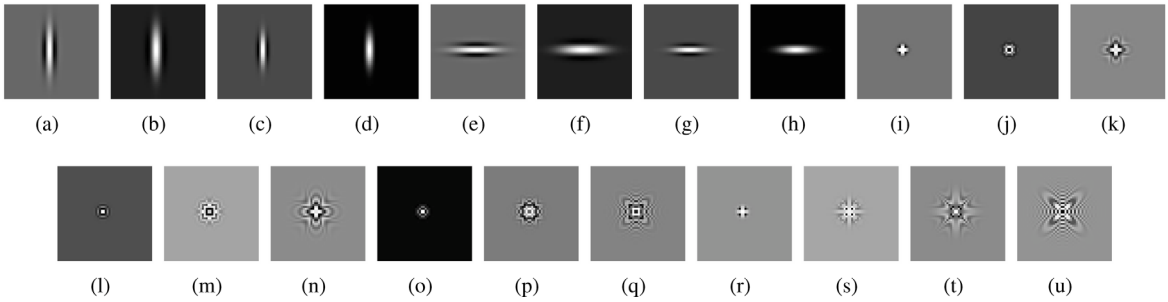


Fig. 3. Filters: The images of (a)~(h) are Gabor filters, and the images of (i)~(u) are Schmid filters.

$W_{y'} = \left[w_{y'}^1, w_{y'}^2, \ldots, w_{y'}^{d'}\right]$. Finally, the transformed feature sets of X'^* and Y'^* are defined as follows:

$$X'^* = X' \left[w_{x'}^1, w_{x'}^2, \ldots, w_{x'}^{d'}\right] = X'W_{x'},$$

$$Y'^* = Y' \left[w_{y'}^1, w_{y'}^2, \ldots, w_{y'}^{d'}\right] = Y'W_{y'}.$$

The feature fusion strategies of our proposed FGCA method are the same as those in Section 2. In this paper, we use the LFFS-II to evaluate the performances of our proposed FGCA algorithm, that is:

$$Z'^* = X'^* + Y'^*,$$

where Z'^* is the fused feature of a fine-grained sample. We concentrate the feature vectors of all fine-grained samples of the image to obtain the final feature vector $\overline{Z}$.

3.3. Feature extraction via FGCA

In this section, we present the proposed FGCA framework for medical image retrieval in detail. To obtain fine-grained samples, the region-based segmentation method is used to partition each medical image. The fine-grained samples of a medical image are nonoverlapping and have unique characteristics. An example of this process is depicted in Fig. 2, which shows that each medical image sized 512 × 512 is partitioned into 64 fine-grained samples, and each fine-grained sample is sized 64 × 64.

To effectively perform feature fusion, each fine-grained sample has four kinds of features extracted. (1) Gabor feature X_g: The Gabor feature is extracted by the Gabor filter [19]. It describes the local structure information of the spatial frequency, spatial position and direction selectivity and extracts texture features on different scales and in different directions. (2) Schmid feature X_s: The Schmid feature is obtained by the Schmid filter. It is worth noting that the Schmid filter is an isotropic ring filter. Ref. [20] used the both Gabor feature and Schmid feature to perform medical image retrieval. Fig. 3 shows the filters used in this paper. The first 8 are Gabor filters, and the rest are Schmid filters. For each fine-grained sample, we first perform the filtering operation, and then the average of all the filter samples is obtained. Finally, the sum of the vertical projection is the filtered feature. For a fine-grained sample sized 64 × 64, the dimensions of Gabor feature X_g and Schmid feature X_s are 1 × 64. (3) PCA feature X_p: The PCA feature is extracted using principal component analysis (PCA). For each fine-grained sample, the top 95% of the principal components are used to form the PCA feature. (4) TDPCA feature X_t: The TDPCA feature is extracted using two-dimensional principal component analysis (TDPCA). Compared to PCA, TDPCA performs the transform based on 2D image matrices [21]. In this paper, the first two-dimensional projective vectors are used to form the TDPCA feature. The combined features are composed by combining any two of the above features, and an example is shown in Table 1.

Table 1
The combined features used in this paper.

Notations	Definitions
$X_g - X_s$	Combined features of X_g and X_s
$X_g - X_p$	Combined features of X_g and X_p
$X_g - X_t$	Combined features of X_g and X_t
$X_s - X_p$	Combined features of X_s and X_p
$X_s - X_t$	Combined features of X_s and X_t
$X_p - X_t$	Combined features of X_p and X_t

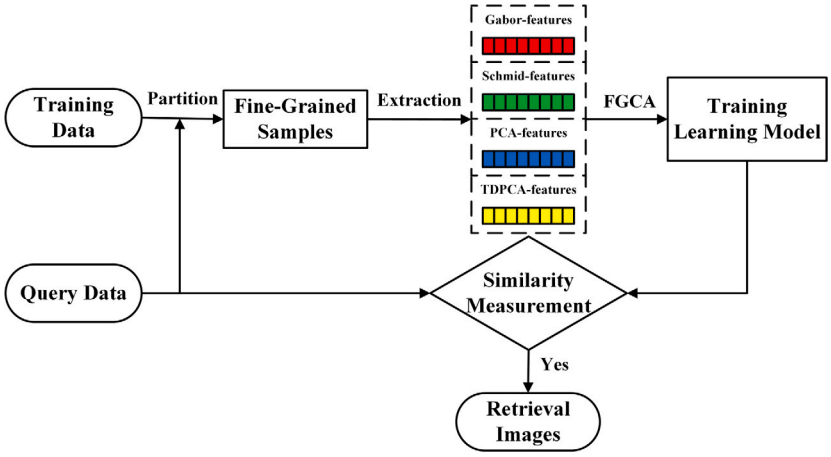


Fig. 4. Overview of our proposed fine-grained correlation analysis framework for medical image retrieval.

Fig. 4 depicts the proposed FGCA framework for medical image retrieval. First, the training images and a query image are partitioned into 64 fine-grained samples, and each fine-grained sample is sized 64×64. Then, four kinds of features, i.e., the Gabor feature, the Schmid feature, the PCA feature and the TDPCA feature, are extracted from each fine-grained sample. The different training learning models are obtained by our proposed FGCA algorithm based on different combined features. Finally, the retrieval images are selected by the similarity measurement.

4. Experiments and results

4.1. Experimental setting

In this paper, we conduct experiments on three publicly available CT image datasets, namely, EXACT09-CT [22], TCIA-CT [23] and NEMA-CT [5], to evaluate our proposed FGCA. All the CT images are sized 512×512. EXACT09-CT dataset contains 675 images from 19 different categories. TCIA-CT dataset includes 604 images from 8 different categories, and NEMA-CT dataset consists of 315 images from 9 different categories. Fig. 5 shows the sample images that are randomly chosen from the three datasets. In this paper, the training data are composed of two images randomly selected from each category. All images in each dataset are matched with the query one.

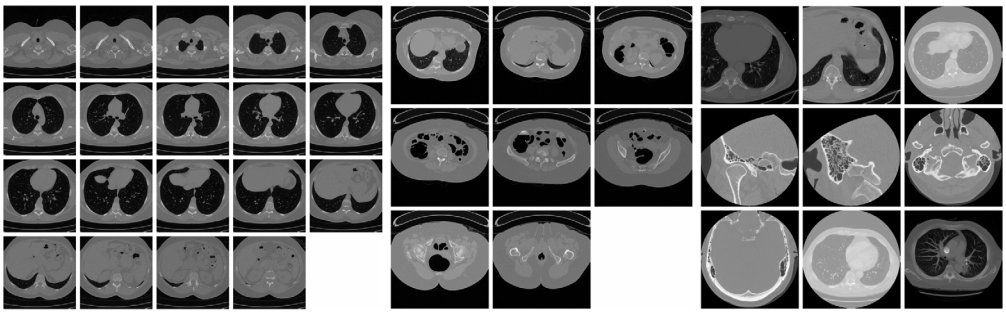
To quantitatively evaluate the retrieval performance of the FGCA, three standard metrics are used, namely, the average retrieval precision (ARP), the average retrieval recall (ARR) and the $F\text{-}score$ [24]. The similarity measurement is calculated by the distance FI as follows:

$$FI(\overline{Z}_1, \overline{Z}_2) = \sum_{i=1}^{D} min(\overline{Z}_1(i), \overline{Z}_2(i)),$$

where $\overline{Z}_1$ and $\overline{Z}_2$ are two feature vectors, and D is the dimension. The image features corresponding to the first 10 FIs are used to calculate the ARP, ARR and $F\text{-}score$. In this paper, the parameter k of the k-means clustering algorithm is $[10 : 2 : 500]$; the ratio of the projection vectors of CCA, GCCA and FGCA is set as $[1 : 1 : 20]$; and the best performance is shown in this paper. The comparative algorithms include the single feature methods LBP [1], LTCoP [4], LWP [5], HCSC [6] and ST-CCA [15] and the feature fusion methods CCA [13] and GCCA [16].

4.2. Comparison of the feature fusion methods

In this section, we compare the quality of different combined features using three feature fusion methods, namely, the CCA-based feature fusion method, the GCCA-based feature fusion method and our proposed FGCA feature fusion method. The different



(a) (b) (c)

Fig. 5. Sample images of the used datasets. Each image corresponds to one category. (a) 19 images from the EXACT09-CT dataset. (b) 8 images from the TCIA-CT dataset. (c) 9 images selected from the NEMA-CT dataset.

Table 2
Quantitative comparison of the different feature fusion methods for the six combined features of the three CT image datasets.

Combined features	CCA [13]			GCCA [16]			FGCA		
	ARP	ARR	$F-score$	ARP	ARR	$F-score$	ARP	ARR	$F-score$
$X_g - X_s$	0.9581	0.2575	0.3990	0.9611	0.2584	0.4005	0.9635	0.2590	0.4014
$X_g - X_p$	0.9578	0.2571	0.3986	0.9602	0.2580	0.3999	0.9623	0.2586	0.4008
$X_g - X_t$	0.9605	0.2579	0.3999	0.9603	0.2580	0.3999	0.9629	0.2588	0.4011
$X_s - X_p$	0.9589	0.2576	0.3992	0.9597	0.2580	0.3999	0.9633	0.2590	0.4014
$X_s - X_t$	0.9593	0.2574	0.3991	0.9600	0.2581	0.4000	**0.9647**	**0.2591**	**0.4017**
$X_p - X_t$	0.9595	0.2577	0.3995	0.9600	0.2579	0.3998	0.9616	0.2583	0.4005

Table 3
Quantitative comparison of the five kinds of single feature methods and the FGCA method on the EXACT09-CT dataset.

Performance	Methods					
	LBP [1]	LTCoP [4]	LWP [5]	HCSC [6]	ST-CCA [15]	FGCA
ARP	0.6503	0.7348	0.8300	0.9150	0.9172	**0.9312**
ARR	0.1951	0.2216	0.2487	0.2883	0.2880	**0.2935**
$F-score$	0.3002	0.3405	0.3827	0.4384	0.4383	**0.4464**

combined features of X_g-X_s, X_g-X_p, X_g-X_t, X_s-X_p, X_s-X_t and X_p-X_t are used to investigate the effectiveness of different feature fusion methods. Table 2 shows the average ARP, average ARR and average F-$score$, which are based on the different feature fusion methods for the six combined features using three medical image datasets. We can observe that the performances of our proposed FGCA feature fusion method are better than those of CCA and GCCA, especially in the combined feature X_s-X_t. FGCA obtains $[0.56\%, 0.66\%, 0.66\%]$ and $[0.49\%, 0.39\%, 0.43\%]$ relative improvements in terms of the ARP, ARR and F-$score$ for the CCA and GCCA for the combined feature X_s-X_t, respectively. These results demonstrate that our proposed FGCA feature fusion method eliminates the redundant information and extracts more discriminative features from fine-grained samples than CCA-based and GCCA-based feature fusion methods. The combined feature X_s-X_t is used in the next three experiments for our proposed FGCA method.

4.3. Experiments on the EXACT09-CT dataset

In this experiment, the EXACT09-CT dataset is used to verify the effectiveness of our proposed FGCA approach. Two images per category are randomly chosen to form the training set, and all the images are used as the testing set. Each image is partitioned into 64 fine-grained samples, and the number of fine-grained samples is $2 \times 19 \times 64 = 2432$. Table 3 displays the retrieval results of the compared single feature methods and our proposed approach. The ARP, ARR and F-$score$ of FGCA are presented in this table using the combined feature X_s-X_t. We can observe that our proposed method can achieve the best performance compared to the performances of the other methods. The proposed method outperforms ST-CCA by 1.53%, 1.91% and 1.85%, and HCSC by 1.77%, 1.80% and 1.82% in terms of the ARP, ARR and F-$score$, respectively.

4.4. Experiments on the TCIA-CT dataset

In this section, we use the TCIA-CT dataset to illustrate our proposed FGCA framework. Two images per category are randomly chosen to form the training set while all images are used for the testing set. Then, the number of fine-grained training samples is $8 \times 2 \times 64 = 1024$. Table 4 shows the performances of existing medical image retrieval methods and our proposed FGCA method. This table shows the retrieval results of our proposed FGCA method based on the combined feature X_s-X_t. Table 4 clearly shows

Table 4
Quantitative comparison of the five kinds of single feature methods and the FGCA method on the TCIA-CT dataset.

Performance	Methods					
	LBP [1]	LTCoP [4]	LWP [5]	HCSC [6]	ST-CCA [15]	FGCA
ARP	0.6691	0.7440	0.8840	0.9512	0.9618	**0.9717**
ARR	0.0974	0.1092	0.1309	0.1452	0.1464	**0.1480**
$F-score$	0.1700	0.1904	0.2280	0.2520	0.2542	**0.2569**

Table 5
Quantitative comparison of the five kinds of single feature methods and the FGCA method on the NEMA-CT dataset.

Performance	Methods					
	LBP [1]	LTCoP [4]	LWP [5]	HCSC [6]	ST-CCA [15]	FGCA
ARP	0.9055	0.9215	0.9532	0.9833	0.9644	**0.9911**
ARR	0.2933	0.3031	0.3133	**0.3364**	0.3270	0.3359
$F-score$	0.4431	0.4562	0.4716	0.5013	0.4884	**0.5018**

that our proposed FGCA method gets the best performance compared to the performances of the other single feature methods including LBP, LTCoP, LWP, HCSC and ST-CCA. In particular, the ARP of the FGCA method exceeds 0.97, and the ARPs of the other methods are below 0.9618. The FGCA method outperforms ST-CCA by 1.03%, 1.09% and 1.06% in terms of the ARP, ARR and F-$score$, respectively. The FGCA method outperforms HCSC by 2.16%, 1.93% and 1.94% in terms of the ARP, ARR and F-$score$, respectively. In addition, the FGCA method significantly improves the retrieval performances compared to those of texture-based methods, i.e., LBP, LTCoP and LWP.

4.5. Experiments on the NEMA-CT dataset

To further verify the performance of FGCA, the NEMA-CT dataset is employed here. In this experiment, two images are randomly selected from each category to form the training set, and each image is partitioned into 64 fine-grained samples. The different performance metrics are calculated using the first 10 images. The quantitative comparison results between five kinds of single feature medical image retrieval methods and our proposed FGCA method are shown in Table 5. The performances of the FGCA method are assessed using the combined feature X_s-X_t. The FGCA method outperforms the ST-CCA by 2.77%, 2.72% and 2.74% in terms of the ARP, ARR and F-$score$, respectively. Although the ARR of FGCA is slightly smaller than that of HCSC, the ARP and F-$score$ of FGCA are higher than of HCSC. Our proposed FGCA method outperforms LWP, which has the best performances among the other texture feature methods, by 3.98%, 7.21% and 6.40% in terms of the ARP, ARR and F-$score$, respectively.

5. Conclusion

In this paper, we proposed a novel feature fusion approach, called fine-grained correlation analysis (FGCA), for medical image retrieval. It aims to solve the high-dimensional problem of medical images and to find more discriminative fusion features. The correlation analysis based on the fine-grained simples shows that FGCA is able to eliminate redundant information and irrelevant information. In the proposed FGCA method, each image is partitioned into a set of fine-grained samples, and then the pseudolabels are generated for the fine-grained samples by the k-means clustering algorithm. The fine-grained samples retained the same identical pseudolabel have similar characteristics. The correlation of these fine-grained samples is beneficial for extracting the more discriminative projection matrices. The experiments on three benchmark datasets demonstrated the effectiveness of our proposed approach in medical image retrieval.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.compeleceng.2021.106992.

References

[1] Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 2002;971–87.
[2] Zhu Z, You X, Chen CP, Tao D, Ou W, Jiang X, et al. An adaptive hybrid pattern for noise-robust texture analysis. Pattern Recognit 2015;48:2592–608.
[3] Ge S, Li C, Zhao S, Zeng D. Occluded face recognition in the wild by identity-diversity inpainting. IEEE Trans Circuits Syst Video Technol 2020.
[4] Murala S, Wu QJ. Local mesh patterns versus local binary patterns: biomedical image indexing and retrieval. IEEE J Biomed Health Inf 2013;18:929–38.
[5] Dubey SR, Singh SK, Singh RK. Local wavelet pattern: a new feature descriptor for image retrieval in medical ct databases. IEEE Trans Image Process 2015;24:5892–903.
[6] Lan R, Zhou Y. Medical image retrieval via histogram of compressed scattering coefficients. IEEE J Biomed Health Inf 2016;21:1338–46.
[7] Jaswal G, Kaul A, Nath R. Multiple feature fusion for unconstrained palm print authentication. Comput Electr Eng 2018;72:53–78.
[8] Liu C, Wechsler H. A shape-and texture-based enhanced fisher classifier for face recognition. IEEE Trans Image Process 2001;10:598–608.

[9] Yang J, Yang J-Y, Gao J-Z. Handwritten character recognition based on parallel feature combination and generalized kl expansion. J Softw 2003a;14:490–5.
[10] Yang J, Yang J-y, Zhang D, Lu J-f. Feature fusion: parallel strategy vs. serial strategy. Pattern Recognit 2003b;36:1369–81.
[11] Lan R, Zhou Y, Liu Z, Luo X. Prior knowledge-based probabilistic collaborative representation for visual recognition. IEEE Trans Cybern 2018.
[12] Ou W, Long F, Tan Y, Yu S, Wang P. Co-regularized multiview nonnegative matrix factorization with correlation constraint for representation learning. Multimedia Tools Appl 2018;77:12955–78.
[13] Sun Q-S, Zeng S-G, Liu Y, Heng P-A, Xia D-S. A new method of feature fusion and its application in image recognition. Pattern Recognit 2005;38:2437–48.
[14] Hotelling H. Relations between two sets of variates. In: Breakthroughs in statistics. Springer; 1992, p. 162–90.
[15] Lan R, Wang H, Zhong S, Liu Z, Luo X. An integrated scattering feature with application to medical image retrieval. Comput Electr Eng 2018;69:669–75.
[16] Sun Q-S, Liu Z-d, Heng P-A, Xia D-S. A theorem on the generalized canonical projective vectors. Pattern Recognit 2005;38:449–52.
[17] Larsson M, Stenborg E, Toft C, Hammarstrand L, Sattler T, Kahl F. Fine-grained segmentation networks: Self-supervised segmentation for improved long-term visual localization. In: Proceedings of the IEEE international conference on computer vision. 2019. p. 31–41.
[18] Tang J, Hu X, Gao H, Liu H. Discriminant analysis for unsupervised feature selection. In: Proceedings of the 2014 SIAM international conference on data mining. SIAM; 2014, p. 938–46.
[19] Mehrotra R, Namuduri KR, Ranganathan N. Gabor filter-based edge detection. Pattern Recognit 1992;25:1479–94.
[20] Lan R, Zhong S, Liu Z, Shi Z, Luo X. A simple texture feature for retrieval of medical images. Multimedia Tools Appl 2018;77:10853–66.
[21] Yang J, Zhang D, Frangi AF, Yang J-y. Two-dimensional pca: a new approach to appearance-based face representation and recognition. IEEE Trans Pattern Anal Mach Intell 2004;26:131–7.
[22] Lo P, Van Ginneken B, Reinhardt JM, Yavarna T, De Jong PA, Irving B, et al. Extraction of airways from ct (exact'09). IEEE Trans Med Imaging 2012;31:2093–107.
[23] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging archive (tcia): maintaining and operating a public information repository. J Digit Imaging 2013;26:1045–57.
[24] Li Z, Zhang X, Müller H, Zhang S. Large-scale retrieval for medical image analytics: A comprehensive review. Med Image Anal 2018;43:66–84.

Xiaoqin Wang received the B.S. degree in information and computing science from Guilin University of Technology, and the M.S. degree in computer science and technology from Guilin University of Electronic Technology, Guilin, China. Her main research interests include image processing, data mining and machine learning.

Rushi Lan is an Associate Professor with School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China. His research interests include image classification, image denoising, and natural language processing.

Huadeng Wang is an Associate Professor of School of Computer Science and Information Security, Guilin University of Electronic Technology, China. His current research interests include medical data mining, machine learning, and computer graphics.

Zhenbing Liu is a Professor and master supervisor in School of Computer Science and Information Security, Guilin University of Electronic Technology, China. His main research interests include image processing, machine learning and pattern recognition.

Xiaonan Luo is currently a Professor with School of Computer Science and Information Security at Guilin University of Electronic Technology, Guilin, China. His current research interests include computer vision, image processing, computer graphics and CAD.