Computer Science and Information Systems 15(3):517–531 https://doi.org/10.2298/CSIS171210020Q

Convolutional Neural Networks and Hash Learning for Feature Extraction and of Fast Retrieval of Pulmonary Nodules

Pinle Qin, Jun Chen, Kai Zhang, and Rui Chai

School of Data Science and Technology, North University of China, Taiyuan 030051, China
6833330@qq.com, {qpl, Zhangk, chairui}@nuc.edu.cn

Abstract. With a widespread use of digital imaging data in hospitals, the size of medical image repositories is increasing rapidly. This causes difficulty in managing and querying these large databases leading to the need of content based medical image retrieval (CBMIR) systems. A major challenge in CBMIR systems is the "semantic gap" that exists between the low level visual information captured by imaging devices and high level semantic information perceived by the human. Using deep convolution neural network (CNN) to construct the CBMIR system can fully characterize the high level semantic features information for medical image retrieval. The existing network mostly used for the natural images can't produce a good result directly applied to medical image. This paper used U-Net method to preprocessing under the guidance of medical knowledge. Then, multi-scale receiving field convolution module is used to extract features of the segmented images with different sizes. Finally, encoded the features and used a coarse to fine search strategy with an average search accuracy of 0.73.

Keywords: Content Based Medical Image Retrieval (CBMIR), Convolutional Neural Networks (CNN), Similarity Measure, Deep Learning.

1. Introduction

With the widespread use of computers, multi-media and storage systems, mass storage of image and multi-media content has emerged. The development of digital storage and content processing has derived the improvement of clinical and diagnostic research. Hospitals with diagnostic imaging equipment are generating large amounts of data, leading to a growing size of the medical image repository. Therefore, it necessary to develop an effective medical image retrieval system to help clinicians search these large data sets and find similar cases as well as the corresponding diagnostic records in the database.

In order to promote the production and management of such large-scale medical image data warehouse, many content-based medical image retrieval (CBMIR) methods have been proposed. M. Mizotin et al. [1] used the MRI image of early Alzheimer disease with complete case records, and proposed a Bag-of-Visual-Words method based on SIFT feature for MRI image retrieval. G. W. Jiji et al. [2] captured feature informations such as shape, texture and color from skin lesion images and proposed a content-based skin lesion image retrieval system combines with CART algorithm. G.

Quellec et al. [3] used different wavelet-basis to represent each query image, and maximize the retrieval performance of test set by learning the optimal wavelet filter of regression function. M. M. Rahman et al. [4] proposed a medical image retrieval framework based on image filtering and similarity fusion, use multi-class SVM classifiers for image prediction to filter out irrelevant images, thus reducing retrieval space of similarity measure. The researches mentioned above are basically based on the shape, color, texture or other features derived from the image itself. The performance of the retrieval system depends on these manually selected features.

With the development of the research field, despite the emergence of many new technologies, "semantic gap" has always been the most challenging issue [5]. Krizhevsky et al. [6] used the features extracted from the seventh layer of the AlexNet network for image retrieval for the first time, and tested on the ImageNet dataset. Babenko et al. [7] focused on the size of CNN feature map and improved the retrieval performance by compressing CNN feature map. Although the two tasks [6-7] performed well on the image retrieval problem, they directly performed the similarity measure in the Euclidean space, which was inefficient. Xia et al. [8] proposed a supervised hash-learning method, which learns binary Hash coding through deep learning technology for image retrieval, and has obtained a superior retrieval performance on public data sets.

In recent studies, deep learning has been widely applied to CBIR tasks. However, there are few deep learning methods to explore CBMIR tasks. A.Qayyum et al. [9] builded a CBMIR system based on CNN. Firstly, 24 categories of medical image datasets under five modalities were classified, then learned the features for image retrieval. The classification of interstitial lung disease (ILD) based on convolutional neural network is proposed in [10]. The data sets consists of 7 categories (6 ILD diseases and 1 healthy tissue), and finally got an 85.5% classification accuracy. In the literature [11], a Restricted-Boltzmann machine was used for image analysis of lung CT, which combined with generation and discernment information, proposed two different methods: one for lung classification, the other for tracheal testing. In the literature [12], a two-stage multi-instance deep learning framework is proposed for body organ recognition. In the first stage, CNN obtains the image feature information from the training set. In the second stage, the extracted information for classification task was fine-tuned, categorized images consist of 12 classes that contain both CT and MR 2D slices.

Most of the above studies are focused on the transfer learning of a small amount of medical images. However, an interesting avenue of research, CBMIR could be the direct training of deep networks for the retrieval task itself. This paper adopts the lung CT scan as dataset, focusing on 2D slice retrieval for knowledge discovery in massive databases and offers the possibility to identify similar case histories. The CNN model is trained for extracting effective feature representations from the pixel-level information in first phase and then the learned feature representations are used for CBMIR in second. In these two phases of tasks, image feature extraction and similarity metrics are crucial. The major contributions of this work are threefold,

(a) A deep learning framework is modelled and trained on a collection of medical images.

(b) The learned features are used to present a highly efficient medical image retrieval system that works for a large collection of multimodal dataset.

(c) By modifying the network model, CNN learns the domain specific image representations and a set of hash-like (or binary coded) functions simultaneously.

The rest of paper is organized as follows: Related work is presented in Section 2, the proposed methodology is discussed in Section 3, experimental results are shown and discussed in Section 4 and a conclusion is presented in Section 5.

2. Related Work

2.1. Content Based Image Retrieval System

The common CBIR system framework is shown in Fig. 1 [13]. In CBIR, the features extracted based on image content is used for the image retrieval of large database. The feature extraction and similarity measure marked by the red box in the figure jointly determine the retrieval performance of the CBIR system.

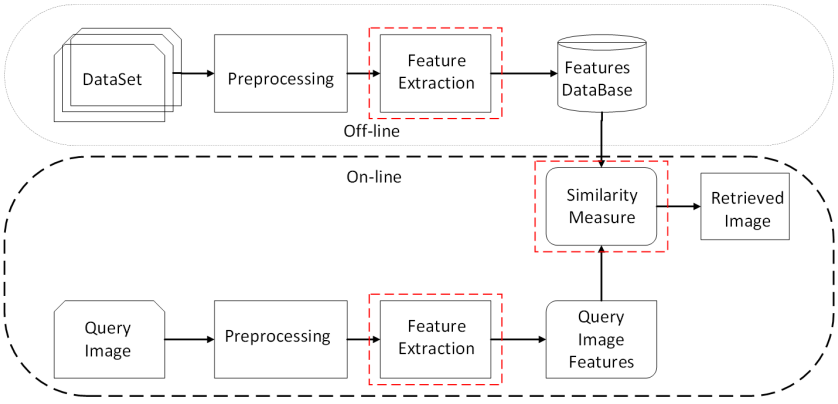

Fig. 1. The block diagram of general CBIR system

All CBIR systems are generally divided into two stages, one is off-line stage, the other is on-line stage. During the off-line processing stage, features are extracted from a large amount of images(training the systems) to build a local feature database, which is usually time-consuming, and the performance of the model depends on the quantity and quality of the training data. In the on-line retrieval stage, the image features are extracted from the query image, and calculating the distance of similarity measure with the local feature database. Then, the images with high similarity or short distance are presented to the user as search results. The preprocessing and feature extraction in two stages are the same.

2.2. Feature Extraction

In the past decades, various feature descriptors have been developed for representing images at a global level such as shape and color based features [14], texture based features [15]. Local-level feature descriptors have also been developed such as SIFT(Scale-Invariant Features Transform) [16], BoWs(Bag-of-Words) [17], etc. Although these features have been widely used in various studies for image retrieval task but are still not fully able to address the problem of semantic gap. Recently, machine learning techniques have provided a new way to reduce the semantic gap. Deep learning has given a hope for bridging this gap by learning visual features directly from images without using any hand-crafted features.

Deep learning is based on artificial neural networks, trying to mimic the way that the human brain works. It used a set of algorithms that try to model high level abstractions present in data by using a deep architecture possessing multiple processing layers, having linear and nonlinear transformation functions. In 1998, Lee et al. [18] applied CNN to hand-written digit recognition, which is the first successful CNN application that people have ever seen. Despite its initial success, CNN relies on massive data sets, effective training networks, and high-performance computers, until Krizhevsky [19] achieved breakthrough by using AlexNet at the ImageNet Challenge in 2012 that opened a new chapter in CNN. In 2015, ResidualNet [20] proposed by Microsoft won the 6th ImageNet Image Recognition Test and solved the problem of gradient diffusion in ultra-deep CNN network training, the network depth reached 152 layers and 1,000 layers were tried. Now, CNN is an alternative technology for solving computer vision problems.

2.3. Similarity Measure

Image retrieval is a computer vision technique that gives a way for searching relevant images in large databases. Recently, with the advent of deep learning for hashing, we are able to perform effective end-to-end learning of binary representations directly from input images. Among them, one of the more interesting concerns is the simultaneous feature learning and hashing. However, very deep versions of these networks accuracy gets saturated and often degrades [21]. In addition, the continuous relaxation of hash codes to train deep networks to be able to learn with more viable continuous optimization methods (gradient-descent based methods) could potentially lead to uncontrolled quantization and distance approximation errors during binarization.

To address the short-comings of the existing approaches, this paper attempts to solve the problems by the following methods: 1) Design a novel deep hash function learning framework using deep residual networks for representation learning; 2) During binarization of the hash code, we utilize multiple hashing related losses and regularizations to control the quantization error and to encourage the hash codes to be maximally independent of each other.

3. Multi-scale Receptive Field Convolutional Neural Network

Due to high-resolution CT examination showed that there are many different sizes of nodules in the lung. In the face of large-scale pulmonary nodule image data, first considered the imaging characterization used basic medical knowledge to preprocessing; secondly, for different sizes of nodules, designed convolution module with different receptive field to prevent small nodule features are filtered, and in order to avoid gradient dispersion or too many network layers lead to feature loss, the residual module is introduced to improve; Finally, the last three fully connected layers of the network are utilized to extract the features of retrieval tasks to improve the accuracy and efficiency.

3.1. Pulmonary Nodule Image Preprocessing

According to the pathological definition, pulmonary nodules are tumors that grow in lung tissue. In order to effectively reduce the search space for pulmonary nodules, it is necessary to identify the boundary of two lungs, that is, to extract the entire lung tissue from other tissues. However, due to the heterogeneity of the lobes and the structures similar to the lung tissue density (such as arteries, veins, bronchi, and bronchioles), segmentation of the lung parenchyma presents challenges.

CNN based image segmentation methods have been providing state-of-the-art performance in the last few years. U-net [22] is widely used because of its short training time and low training parameters. As opposed to classification tasks, the architecture of segmentation tasks requires both convolutional encoding and decoding units. The encoding unit is used to encode input images into a larger number of maps with lower dimensionality. The decoding unit is used to perform up-convolution (de-convolution) operations to produce segmentation maps with the same dimensionality as the original input image.

3.2. The InRes-Net Model Architecture

The first stage of the CBMIR framework proposed in this paper is feature extraction. Because the size of nodules is between 3mm and 30mm, the designed network is required to extract features with different receptive fields. Inception-resnet v2 [23] is very good at training features in different receptive fields. By simplifying the model and migrating its ideas to AlexNet, the network (InRes-Net) shown in Fig. 2 was built.

Space compression module is shown in Fig. 3, the inV2 module in Fig. 2, is MaxPooling and a series of different convolution operations compress the dimensions of the input data over different paths. The leftmost input module represents Nxkxk features, number of output channels does not change after MaxPooling path, but the feature size is reduced to k/2 x k/2. Conv represents the convolution layer, and the latter parameters denote the convolution kernel size and step size. The three paths all reduce the feature map to k/2 x k/2, the number of output channels is N/4, N/4 and N/2. The concat operation splices the output of the four paths to form an intermediate feature with 2N channels and a size of k/2 x k/2.

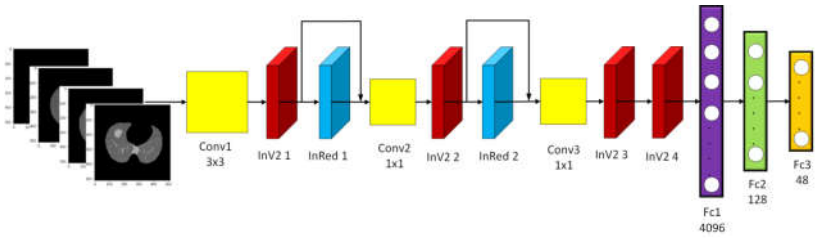

Fig. 2. InRes-Net network structure

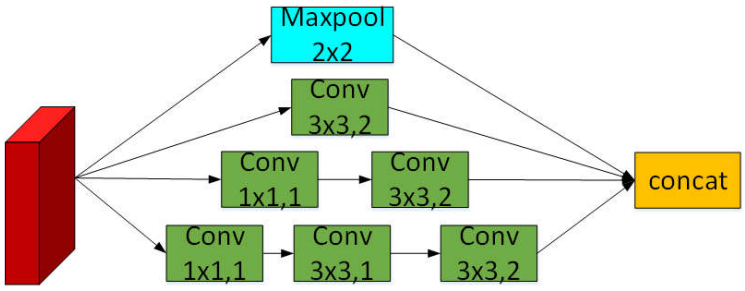

Fig. 3. space compression module (inV2)

The inRed module, shown in Fig. 4, allows the model to be trained on different receptive fields. Finally, the input and concat output are quickly connected to form a residual module.

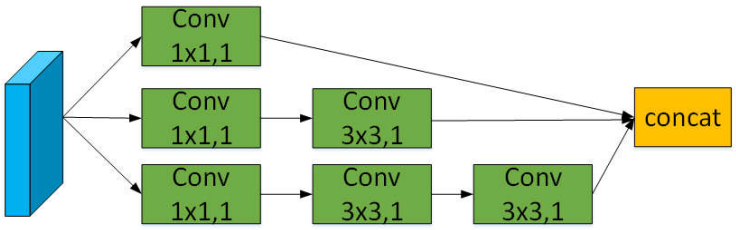

Fig. 4. inRed module

The proposed framework includes a spatial compression module, a residual module, and a fully connected hash layer for hash generation. It provides a unique advantage for the convolution neural network to hashing. (1) Training deeper networks: The expression ability of deep network increases with the increase of the network but when

the network reaches a certain depth, adding extra layers leads to higher error rate of training and verification. Residual network is used to solve this problem by adding a quick connection added to the convolution output; (2) Easy to optimize: One of the major problems in the network training process is the disappearance of the gradient, introducing an activation unit and normalizing the input batch can improve this situation.

3.3. Hash Coding and Retrieval

We assume that the final outputs of the classification layer FC3 rely on a set of *h* hidden attributes with each attribute *on* or *off*. In other points of view, images inducing similar binary activations would have the same label. To fulfill this idea, we embed the latent layer *H* between FC1 and FC3 as shown in the middle row of Fig. 2. The latent layer *H* is a fully connected layer, and its neuron activities are regulated by the succeeding layer FC3 that encodes semantics and achieves classification. The proposed latent layer *H* not only provides an abstraction of the rich features from FC1, but also bridges the mid-level features and the high-level semantics. In our design, the neurons in the latent layer *H* are activated by sigmoid functions so the activations are approximated to $\{0,1\}$.

We adopt a coarse-to-fine search strategy for rapid and accurate image retrieval. We firstly retrieve a set of candidates with similar high-level semantics, that is, with similar hidden binary activations from the latent layer. Then, to further filter the images with similar appearance, similarity ranking is performed based on the deepest mid-level image representations.

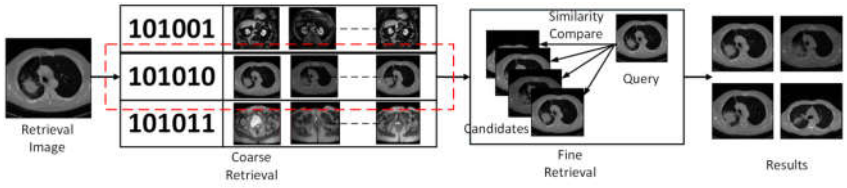

Fig. 5. Image retrieval

Coarse Retrieval. For images, the output of FC2 is expressed as Out(H), then each of binary codes (H is the number of nodes of FC2) is obtained by the sigmoid function. The binary code of the output is described by h:

$$H^j = \begin{cases} 1 & Out^j(H) \geq 0.5 \\ 0 & others \end{cases} \quad (1)$$

Let $\tau = \{I_1, I_2, ..., I_n\}$ represent a dataset consisting of n medical images, each image corresponding to the binary code expressed as $\tau_H = \{H_1, H_2, ..., H_n\}$, where $H_i \in (\{0,1\})^h$. Given a query image I_q, the corresponding binary code is H_q, if the Hamming distance between H_q and $H_i \in \tau_H$ is less than the threshold, we get an m-sized candidate pool $P = \{I_1^c, I_2^c, ..., I_m^c\}$.

Fine Retrieval. Given a query image I_q and candidate pool P, an image of top-K is retrieved and sorted in candidate pool P using the features extracted from FC1. Let V_q represent eigenvector of the query image q, V_i^P represent the image I_i^c in the candidate pool P and the Euclidean distance on the corresponding eigenvector of the image i in I_q and P Euclidean distance.

$$dist_i = \left\| V_q - V_i^P \right\| \tag{2}$$

The smaller the Euclidean distance, the more similar the images are. Each is sorted by similarity, so an image of top-K is retrieved.

3.4. Training

The model was trained using Stochastic Gradient Descent (SGD) with backpropagation. It was optimized with a very low learning rate of 0.0001 with maximum of 30 epochs of SGD. The Negative Log Likelihood (NLL) was used as an objective function or criterion in SGD training. The weights of all layers were initialized using Gaussian distribution having mean of zero and standard deviation of 0.01. FC2 fully connected layers were initialized like LSH [24] which uses random projections for constructing the hashing bits. The learning rate was kept constant for all iterations of stochastic gradient i.e., 0.0001 as for dataset used, decaying the learning rate, increases the training time as compared to keeping it constant.

4. Experimental Results

In this paper, a popular and widely-used deep learning tool Tensorflow has been used for developing and training the proposed deep learning framework. The simulations have been performed on Inter (R) Xeon (R) CPU E5-2683 v3 Laptop with Ubuntu 14.04 having NVIDIA M40 GPU with clock speed of 2.40 GHz with a RAM of 16.00 GB. The proposed method has been evaluated in terms of retrieval results.

4.1. Dataset and Preprocessing

For data-driven learning, having a large annotation dataset is crucial. Unlike natural images (eg. ImageNet), the primary problem with medical images is collecting difficulties, followed by the high cost of annotation, so there is no large-scale medical annotation images comparable to ImageNet's data volume. Fortunately, a few open challenges provide a considerable amount of medical image data sets. Such as Lung Node Analysis 2016 (LUNA 2016) [25], Data Science Bowl 2017 [26] and Tianchi Medical AI Contest [27].

This paper uses the publicly available LIDC/IDRI database, 888 CT scans are included, excluded scans with a slice thickness greater than 2.5 mm. The advantage of

the LIDC/IDRI dataset compared to other datasets is that the CT scans are annotated by 4 experienced radiologists. The comment file is a csv file. Take one of them as shown in Table 1.

Table 1. seriesuid is the slice number in each CT scan; coordX, coordY and coordZ correspond to x, y and z in the world coordinate system respectively; diameter_mm represents the diameter of the pulmonary nodule in millimeters (mm).

seriesuid	coordX	coordY	coordZ	diameter_mm
SeriesInstanceUID	-100.56	67.26	-231.81	6.44

In the dataset, the slice is 512x512 pixels. CT scans of different patients may come from different machines, the amount of slices per CT scan and slice thickness is also different. Thus, we need to preprocessing the dataset and resampling it so that each pixel represents a 1x1x1 cube.

After resampling, there are significant differences in the pulmonary parenchyma from the surrounding organs through the contrast test in Fig. 6(b). Select the empirical threshold -400HU to generate lung parenchyma masks in Fig. 6(c).

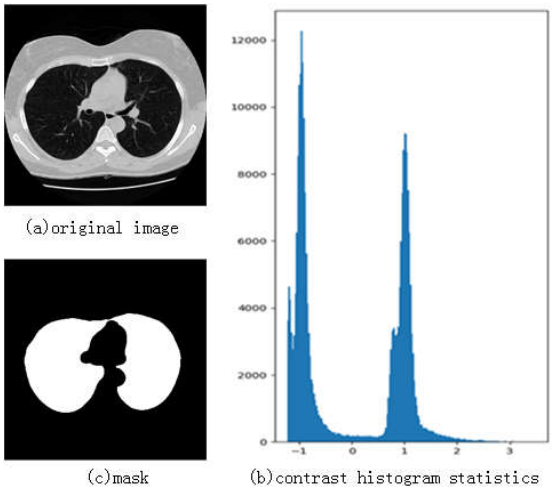

Fig. 6. Generate mask

The method in Fig. 6 is low efficiency. Take into consideration the mask as the label of corresponding image, training U-net in a pair-wised manner which completed adaptive image segmentation as shown Fig. 7. After all preprocessing work, InRes-Net gets a set of clean input data, which can converge quickly during training and also reduce the search space during testing.

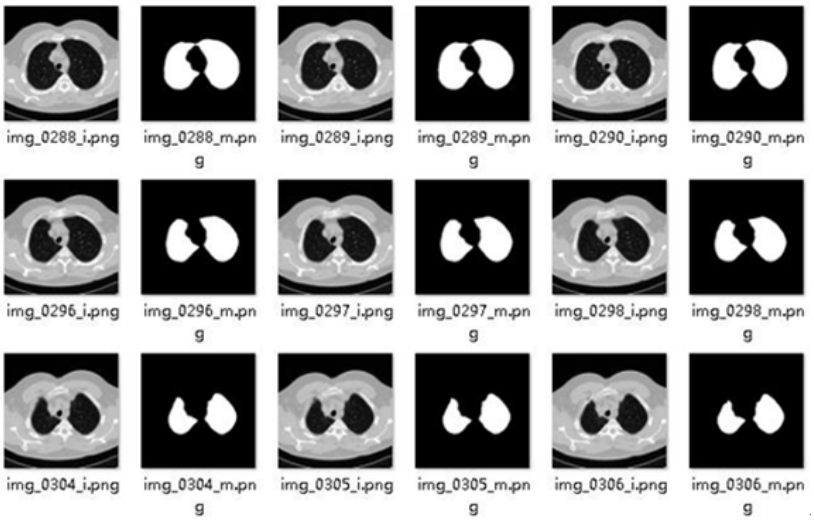

Fig. 7. segmentation results. name like 'img_xxxx_i.png' in left is original image, the right is corresponding mask.

4.2. Retrieval Performance and Analysis

The performance of the proposed framework for CBMIR has been tested using most frequently used performance measure for CBMIR systems i.e., Precision and Recall. The mathematical expression for precision and recall are,

$$precision = \frac{The\ number\ of\ related\ images\ retrieved}{The\ number\ of\ images\ retrieved} \tag{3}$$

$$recall = \frac{The\ number\ of\ related\ images\ retrieved}{The\ total\ number\ of\ related\ images\ in\ the\ dataset} \tag{4}$$

Feature representations from all three fully connected layers of the trained model have been used for retrieval of medical images. The precision vs recall plots for feature representations extracted from FC1, FC2 and FC3 are depicted in Fig. 8. It shows that precision is high for feature representations of FC1+FC3 as compared to feature representations of layer FC1, FC2 and FC3. Hence, coarse search eliminating the irrelevant images the search area is reduced in large database for similarity measurement.

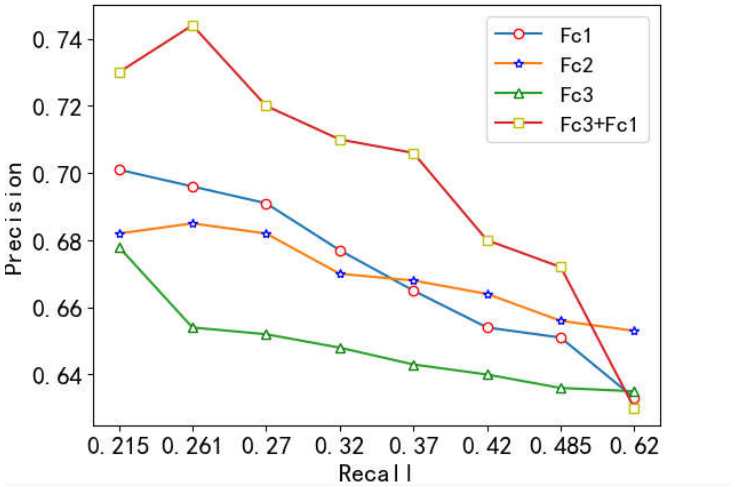

Fig. 8. Precision vs Recall for CBMIR.

The retrieval performance has also been evaluated using mean average precision. Fig. 9 and 10 show the retrieved images, of query image from normal and nodule classes respectively. The retrieval results are shown in a ranked order Top-10, where the most relevant image found after feature comparison is presented first. The retrieved results demonstrate the interclass variance.

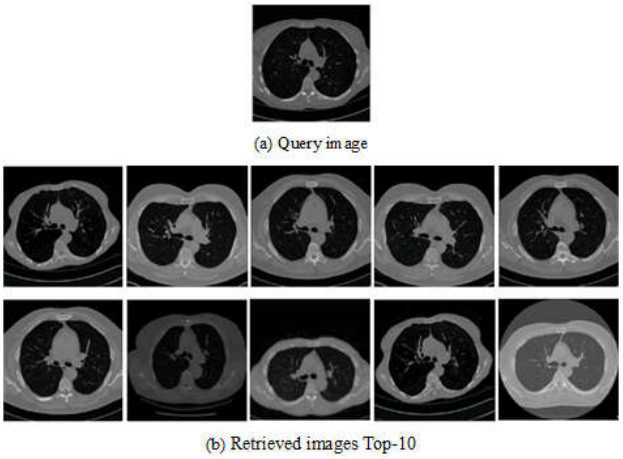

Fig. 9. Results of normal image retrieval.

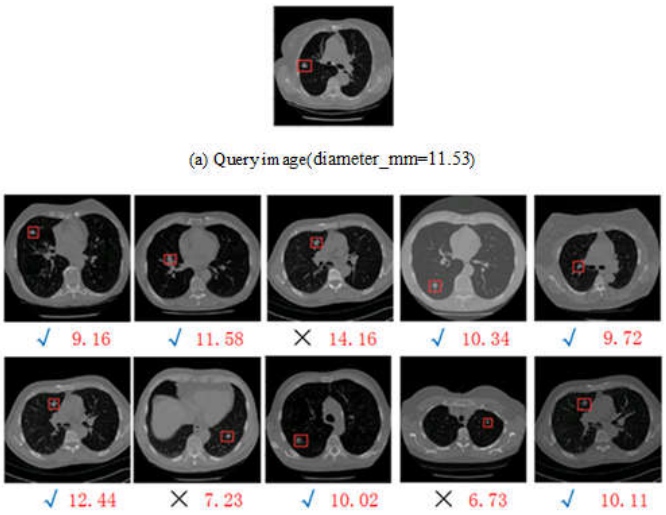

Fig. 10. Retrieved images Top-10. The blue check marks indicate the query and retrieved image share the same label; the black crosses indicate otherwise, the red text is diameter_mm.

To verify that the FC3 feature can be directly used for image classification, the medical images are divided into various classes based on diameter of the pulmonary nodule. 1186 CT slices with nodule are labeled according to the size of the nodule in Table 2, and 1000 normal(diameter=0) CT slices. The data from each class was divided randomly into training and testing set using 70% and 30% images for training and testing set. The output of last fully connected layer (FC3) has been changing to a Softmax function having eight outputs, which produce a probability distributions for each class label. The classification performance is summarized in Table 3.

Table 2. data sets for image classification. diameter_mm is the size of nodules, the range of diameter is chosen to ensure the dataset is uniform. quantity is the different amount of nodules. class is the label of nodules

diameter_mm	=0	0~5	5~6	6~8	8~13	13~18	18~23	>23
quantity	1000	270	232	276	234	100	55	19
class	0	1	2	3	4	5	6	7

As Shown on Table 3, our proposed system demonstrates the efficacy of the method in the classification task. The retrieval results are improved since the images retrieved only belong to the predicted class by the classification framework.

Specially, Fig. 10 shows the Top-10 images retrieved by different features. As can be seen in Table 3, due to imbalanced class distributions, AlexNet classifies the images with diversity, normal data retrieve more images with the same label. Although the data

set has reached 80GB, for each CT scan, the nodule exists only several of slices for each CT scan, which is reluctant to training CNN, we can collect more data to solve the problem.

Table 3. FC3 feature for image classification.

class	0	1	2	3	4	5	6	7
score(mAP)	84.6	77.5	76.5	73.8	76.7	65.7	52.1	52.0

Specially, Fig. 10 shows the Top-10 images retrieved by different features. As can be seen in Table 3, due to imbalanced class distributions, AlexNet classifies the images with diversity, normal data retrieve more images with the same label. Although the data set has reached 80GB, for each CT scan, the nodule exists only several of slices for each CT scan, which is reluctant to training CNN, we can collect more data to solve the problem.

5. Conclusions

We present a simple yet effective deep learning framework to create the hash-like binary codes for fast image retrieval. We add a latent-attribute layer in the deep CNN to simultaneously learn domain specific image representations and a set of hash-like functions. Our method relies on a large amount of medical images and direct training of InRes-Net for the retrieval task. Experimental results show that, with only a simple modification of the deep CNN, our method achieves 0.73 retrieval precision on the LUNA2016 datasets. We intend to further improve the retrieval performance by using a larger dataset and adapt the network for 3D volumetric applications by defining further classes that incorporate the different geometric views of a 3D slice.

References

1. Mizotin M, Benois-Pineau J, Allard M, et al.: Feature-based brain MRI retrieval for Alzheimer disease diagnosis[C]// IEEE International Conference on Image Processing. IEEE, 2012:1241-1244.
2. Jiji G W, Savariraj J D R P. Content-based image retrieval in dermatology using intelligent technique[J]. Image Processing Iet, 2015,9(4):306-317.
3. Quellec G, Lamard M, Cazuguel G, et al. Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval[J]. IEEE Transactions on Image Processing, 2011, 21(4):1613-1623.
4. Rahman M M, Antani S K, Thoma G R. A learning-based similarity fusion and filtering approach for biomedical image retrieval using SVM classification and relevance feedback.[J]. Information Technology in Biomedicine IEEE Transactions on, 2011, 15(4):640-646.
5. Smeulders A W M, Worring M, Santini S, et al. Content-Based Image Retrieval at the End of the Early Years[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(12):1349-1380.

6. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
7. Babenko A, Slesarev A, Chigorin A, et al. Neural Codes for Image Retrieval[J]. 2014, 8689:584-599.
8. R.Xia, Y.Pan, H.Lai, C.Liu, and S.Yan. Supervised hashing for image retrieval via image representation learning. In Proc.AAAI, 2014. 1,2,6,7.
9. Qayyum A, Anwar S M, Awais M, et al. Medical Image Retrieval using Deep Convolutional Neural Network[J]. 2017.
10. Anthimopoulos M, Christodoulidis S, Ebner L, et al. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network[J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1207.
11. Van T G, De B M. Combining Generative and Discriminative Representation Learning for Lung CT Analysis with Convolutional Restricted Boltzmann Machines[J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1262-1272.
12. Yan Z, Zhan Y, Peng Z, et al. Multi-instance Deep Learning: Discover Discriminative Local Anatomies for Bodypart Recognition[J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1-1.
13. Qayyum A, Anwar S M, Awais M, et al. Medical image retrieval using deep convolutional neural network[J]. Neurocomputing, 2017.
14. Jain A K, Vailaya A. Image retrieval using color and shape[J]. Pattern Recognition, 1996, 29(8):1233-1244.
15. Bengio Y, Courville A, Vincent P. Unsupervised Feature Learning and Deep Learning: A Review and New Perspectives[J]. 2012.
16. Lowe D G. Object Recognition from Local Scale-Invariant Features[C]// The Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 2001:1150.
17. Yang J, Jiang Y G, Hauptmann A G, et al. Evaluating bag-of-visual-words representations in scene classification[C]// ACM Sigmm International Workshop on Multimedia Information Retrieval, Mir 2007, Augsburg, Bavaria, Germany, September. DBLP, 2007:197-206.
18. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
19. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
20. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015:770-778.
21. Conjeti S, Roy A G, Katouzian A, et al. Deep Residual Hashing[J]. 2016.
22. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[M]// Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. 2015:234-241.
23. Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[J]. 2016.
24. Gionis A, Indyk P, Motwani R. Similarity Search in High Dimensions via Hashing[C]// International Conference on Very Large Data Bases. Morgan Kaufmann Publishers Inc. 2000:518--529.
25. https://luna16.grand-challenge.org/data/
26. https://www.kaggle.com/c/data-science-bowl-2017
27. https://tianchi.aliyun.com/competition/introduction.htm?raceId=231601

Pinle Qin received the PhD degree in computer application technology from Dalian University of Technology (DLUT), Dalian, Liaoning, P.R. China, in 2008. He is

currently an associate professor with the School of Data Science and Technology, North University of China (NUC). His current research interests include computer vision, medical image processing and deep learning. Contact him at 6833330@qq.com

Jun Chen received his undergraduate degree in computer application technology from North University of China (NUC), Taiyuan, Shanxi, in 2014. Currently, he is pursing Master degree from NUC, and his areas of interest are digital image processing, medical image processing and computer vision.

Kai Zhang received his undergraduate degree in computer application technology from North University of China (NUC), Taiyuan, Shanxi, in 2016.

Rui Chai received the PhD degree in computer application technology from Beijing University of Technology (BUT), Beijing, P.R. China, in 2016. Currently, his research interests include image processing, medical image processing and deep learning. Contact him at chairui@ nuc.edu.cn

Received: December 10, 2017; Accepted: August 15, 2018.