Computers & Graphics 70 (2018) 288–299

Contents lists available at ScienceDirect

Computers & Graphics

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

Hybrid-feature-guided lung nodule type classification on CT images

Jingjing Yuan[a], Xinglong Liu[a], Fei Hou[b,*], Hong Qin[c], Aimin Hao[a]

[a] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, Beijing, 100191, China
[b] School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore
[c] Department of Computer Science, Stony Brook University, Stony Brook, New York, 11794, USA

ARTICLE INFO

Article history:
Received 14 June 2017
Revised 9 July 2017
Accepted 11 July 2017
Available online 21 July 2017

Keywords:
Computer tomography
Lung nodule
CNNs
Hybrid features

ABSTRACT

In this paper, we propose a novel classification method for lung nodules from CT images based on hybrid features. Towards nodules of different types, including well-circumscribed, vascularized, juxta-pleural, pleural-tail, as well as ground glass optical (GGO) and non-nodule from CT scans, our method has achieved promising classification results. The proposed method utilizes hybrid descriptors consisting of statistical features from multi-view multi-scale convolutional neural networks (CNNs) and geometrical features from Fisher vector (FV) encodings based on scale-invariant feature transform (SIFT). First, we approximate the nodule radii based on icosahedron sampling and intensity analysis. Then, we apply high frequency content measure analysis to obtain sampling views with more abundant information. After that, based on re-sampled views, we train multi-view multi-scale CNNs to extract statistical features and calculate FV encodings as geometrical features. Finally, we achieve hybrid features by merging statistical and geometrical features based on multiple kernel learning (MKL) and classify nodule types through a multi-class support vector machine. The experiments on LIDC-IDRI and ELCAP have shown that our method has achieved promising results and can be of great assistance for radiologists' diagnosis of lung cancer in clinical practice.

1. Introduction

Lung cancer has been the leading cause of cancer-related deaths in United States in 2016 [1]. There are over 150,000 deaths caused by lung and bronchus cancer, accounting for over 25% of all cancer-related deaths in the past year. A total of over 220,000 new lung and bronchus cancer cases are projected to occur in 2017. In contrast to the steady increase in survival for most cancers, advances have been slow for lung cancer, for which the overall five-year survival rate is only 18% . Among all lung cancer cases, more than one-half are diagnosed at a distant stage, decreasing the survival rate to 4%. Lung nodule is a significant sign of lung cancer. Among all the cases with nodules, approximately 20% represent lung cancer [2]. Early detection and localization of nodules can greatly improve the survival rate for lung cancer, which requires identification of nodules from malignant to innocent as an essential part for diagnosis [3,4]. However, few work has done for lung nodule classification. In this paper, we propose a lung nodule classification method based on hybrid features.

The appearances of lung nodules vary greatly. Typical shape of nodules is spherical, but can be very complicated when attached or surrounded by other anatomical structures. The malignancy of lung nodules is correlated with nodules types. Intra-parenchymal and ground glass optical (GGO) type nodules have higher chance to be lung cancer than those only connected with surrounding structures. The type of nodules depends on their corresponding intensity distribution and relative positions [5], which can be categorized as: well-circumscribed (**W**), located in the center of the lung without any connection to other tissues; vascularized (**V**), the same as **W** except for connecting to vascular structures; juxta-pleural (**J**), fully connected to pleural surface; pleural-tail (**P**), close to pleural but only connected with a thin tail; GGO (**G**), the same with **W** except for its non-solid CT value distribution.

In clinical practice, Computer Tomography (CT) can capture fine-grained details for both lung nodules and surround structures, acting as the golden standard for diagnosis. However, lung nodules are so complex that even experienced radiologists cannot confirm malignancy in some cases. Based on the coherence between nodule malignancy and type, this paper proposes an efficient nodule type classification method, which is expected to help radiologists for early cancer diagnosis in practice.

* Corresponding author.
E-mail address: houfei@ntu.edu.sg (F. Hou).

http://dx.doi.org/10.1016/j.cag.2017.07.020
0097-8493/

2. Related work

2.1. Previous work

To better help radiologists' diagnosis for lung nodules at early stages, the most straightforward way is applying enhancement towards original medical images. Image intensity-based thresholding or edge preserving smooth is most commonly used because of their ability of overcoming large image contrast between lung region and surrounding body tissues [6–8]. Gao et al. [9] propose a threshold-based approach considering anatomical information. Although simple and fast, it is extremely hard to determine the threshold value due to the correlations with image acquisition types, protocols, intensities ,and especially when there are severe lung pathologies [10].

Despite image enhancement, researchers also emphasize on designing features that can best describe nodules' characteristics. The grey-level distribution represented as a histogram [11] has been effectively used to characterize intensity variations. Filter-based feature extraction techniques are widely applied to highlight the edge and shape information for images [12,13]. Scale-invariant feature transform (SIFT) [14–16] provides a robust way invariant to image transformation, scaling and rotation. Histogram of oriented gradients (HOG) [17,18] is interpreting objects by occurrences of gradient orientation in localized portions of an image. Local binary patterns (LBP) [19,20] provides a powerful tool for local texture classification by applying multi-scale and rotation-variant property. Ciompi et al. [21] have encoded the nodule intensity distribution patterns into frequency domain and classified nodules in a bag-of-words fashion. Kim et al. [22] focus on CT imaging biomarkers such as tumor volume, tumor mass, ground-glass opacities, perfusion parameters and texture features to explore their clinical usefulness with respect to the malignancy differentiation, treatment response monitoring, and patient outcome prediction. Cirujeda et al. [23] propose an imaging biomarker of lung cancer relapse through 3-D Riesz-wavelets of morphological nodular tissue properties. Chen et al. [24] utilize the multi-task regression (MTR) scheme that leverages deep learning models of stacked denoising autoencoder (SDAE), convolutional neural networks (CNNs) and Haar-like features to analyze the relations among semantic features and achieve more accurate rating scores.

Recently, inspired by a large amount of available data and more powerful computational resources, especially parallelization ability empowered by Graphic Processing Units (GPUs), convolutional neural networks (CNNs) [25,26] have shown their abilities of outperforming the state-of-the-art in classical computer vision applications [27,28], as well as in the field of medical image analysis [29–31]. Since CNNs can be trained end-to-end with the layers automatically learning discriminative features without handcrafting design, they are best suitable for lung nodule type classification considering the complex intensity and surrounding anatomical structure distributions. Lo et al. [32] propose the CNNs-based method for pulmonary nodule detection in chest radiology images. Arnaud et al. [33] have used the multi-view CNN for pulmonary nodule false positive reduction in CT images, with a complete performance discussion over different fusion methods [30,34]. A few other studies have also extended the use of 2-D CNNs to 3-D volumetric analysis on 3-D images, i.e., CT and Magnetic Resonance (MR) images. Among all these methods, volumetric images are projected to fixed views (planes), followed by that each view is processed under 2-D CNNs and finally integrated under a multi-view fashion with the best fusion methods. Besides, Dou et al. [35] propose the 3-D CNN to process voxels from MR images for cerebral microbleeds. The 3-D CNN structure ensures contextual and spatial information well encoded and helps to achieve great results.

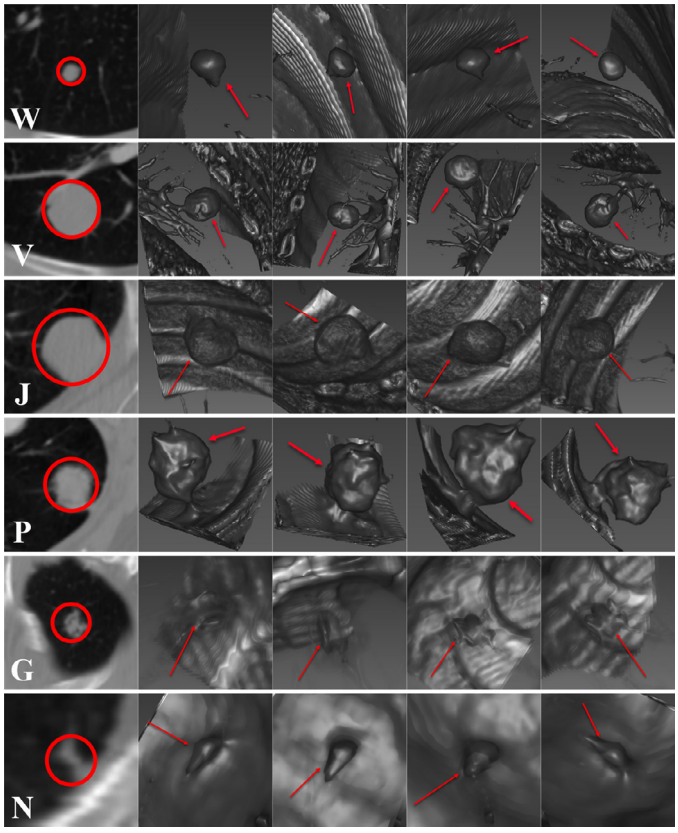

Fig. 1. Typical cases for nodule types. 1st column, original images with radii labeled in red circles. 2nd–5th columns, 3-D models (red arrow) for different nodule types. Top to bottom: well-circumscribed (**W**), vascularized (**V**), justa-pleural (**J**), pleural-tail (**P**), GGO (**G**), non-nodule (**N**). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Besides intimate connection with local image analysis, Computer Aided Diagnosis (CADx) systems have been developed and widely used to assist radiologists in diagnosis and thereby making screening more effective [36,37]. Cheng et al. [38] performs a comprehensive study on deep-learning-based CADx approaches for nodule malignancy estimation and classification.

While many studies have reported detection and segmentation of pulmonary nodules, limited researches focus on nodule type classification. Farag et al. [39] have reported some basic studies in classification problem. Zhang et al. [14] designed an overlapping nodule identification procedure for nodules located at intersections among different types. After that, in light of prior work proposed by Song et al. [40], Zhang et al. [41] have proposed a multilevel patch-based context analysis for nodule classification. Although contextual information is considered, their method can only handle solid nodule types.

2.2. Our approach

Inspired by the prior works, this paper focuses on nodule type classification based on hybrid features by fusing statistical features from CNNs and geometrical features from dense SIFT vectors. The major contributions of our work include:

- A normalized spherical sampling pattern based on icosahedra, a nodule radius estimation method based on intensity analysis and a best view selection method based on high frequency content analysis.
- A multi-view multi-scale CNN to extract the most discriminative statistical features from original data automatically.

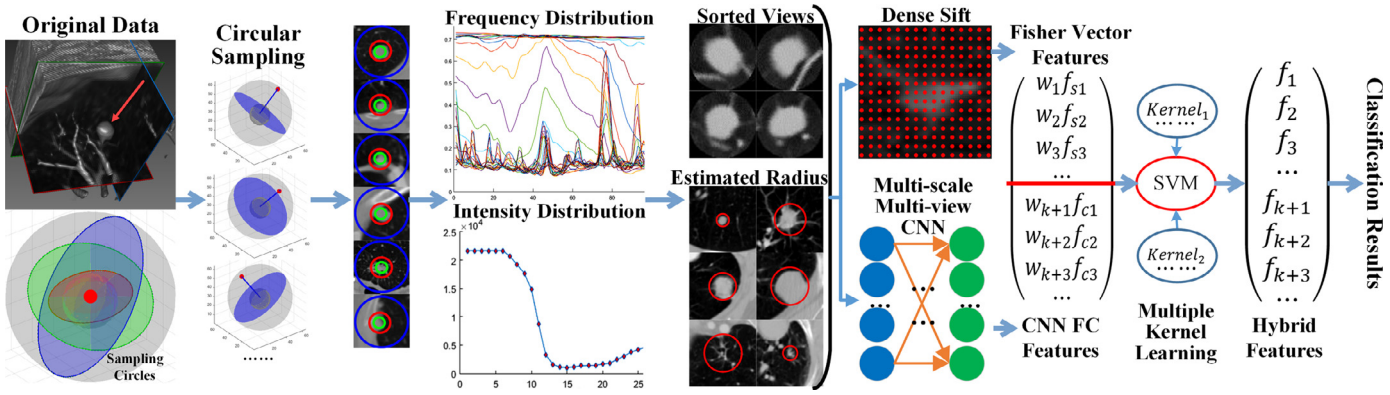

Fig. 2. Pipeline for the proposed method. First, we apply spherical sampling based on icosahedra, followed by threshold-based nodule radii estimation and high-frequency-content-based view sorting. Then, we calculate FV encodings from dense SIFT vectors and CNN features from the multi-view multi-scale CNN. Finally, we utilize MKL to transform both features into unified space by weight adjustment of different kernels and use a multi-class SVM to classify nodule types.

- An approach to combine CNN features and FV encoding features into hybrid features and use to classify nodule types accurately.

The pipeline of our method is described in Fig. 2. First, we re-sample original CT data using a linear interpolation method to make data isotropic. Then, we use icosahedra to divide the nodule volume and construct concentric circles from the inner centers to sample the original volume in different views. With this sample data, we calculate the nodule radii based on thresholding and sort the sampling views based on high frequency content analysis. After that, we build 2-D slice images from nodule volume based on the approximated radii and sorted views, which are used to both train the multi-view multi-scale CNN and construct the FV encodings after the dense SIFT calculation. Finally, we adopt multiple kernel learning to fuse the statistical CNN features with the geometrical FV features and apply the hybrid features into classification of nodule types.

The proposed method has obvious advantages towards classical nodule processing approaches. In sampling step, the icosahedron partition scheme ensures normalized sampling in original nodule volume. Meanwhile, the multi-view multi-scale CNN fashion guarantees the corresponding statistical features to capture not only nodule characteristics but also the correlations between nodules and their surrounding anatomical structures. Besides, the FV encodings from dense SIFT descriptors of different views ensure global geometrical correlations well encoded. Finally, the multiple kernel learning method helps to transform CNN features and FV encodings to unified dimensional hybrid features, which are most discriminative for nodule types.

This paper is organized as follows. Section 3 describes the data sets used to train and validate the proposed method. Section 4 details the sampling fashion, followed by the estimation method for nodule radii and sorting method for views according to importance. Section 5 describes the multi-view multi-scale CNN fashion to extract nodule statistical features and the FV encoding scheme for nodule geometrical feature extraction, followed by the multiple kernel learning approach to fuse both features into unified space as hybrid features. Finally, Section 6 validates the proposed method and applies comparisons on LIDC-IDRI and ELCAP data sets, showing the promising results.

3. Materials

3.1. LIDC-IDRI

The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) [42] consists of totally 1018 chest CT image cases, where each consists of a clinical thoracic CT scan and an associated annotation file. These CT cases are from different facilities and the imaging quality varies. The annotations are produced by four experienced thoracic radiologists from reading CT images separately, followed by non-blind communication to confirm their final diagnosis.

From the original LIDC-IDRI data sets, the count of radiologists labeling the area as nodules is called agreement level. To reduce error caused by human annotations, in our approach, only nodules at least agreement level 2 are considered. Nodule types are annotated by an experienced radiologist. Totally, we extract 1738 nodules (**W**: 905, 52.0%; **P**: 329, 19%; **V**: 219, 12.5%; **G**: 82, 6%; **J**: 203, 11.5%) and 1000 non-nodules from 744 chest CTs. The original CT is 512 × 512 pixels with in-plane spacing as 0.4–1.0 mm and slice thickness as 0.5–3 mm. We re-sample the original CT images with normalized in-plane spacing 1 mm and extract 64 × 64 × 64 mm^3 volumes centered at the nodule gravity center. In our approach, we randomly select 80% of nodules for each type as training and validation cases while the rest 20% as test. Data augmentation is applied for both training and testing cases separately to enrich varieties.

3.2. ELCAP

Early Lung Cancer Action Program (ELCAP) [43] consists of 50 low-dose documented whole lung CT scans with labeled nodules annotated by two experienced radiologists. Scans are achieved with 1.25 mm slice thickness, 0.5 mm in-plane spacing, while most nodule sizes range from 2 to 5 mm. Nodule types are annotated by an experienced radiologist.

In our approach, we select 46 cases with 421 nodules (**W**: 92, 21.8%; **V**: 49, 11.6%; **P**: 155, 36.8%; **J**: 106, 25.2%, **G**: 19, 4.6%). All ELCAP data are used as testing cases with data augmentation applied to validate the performance of the proposed method.

3.3. Data augmentation

Extracted nodules from LIDC-IDRI are greatly unbalanced for each type, which will mislead the optimization of CNNs to local minima and CNN predictions biased towards more frequent samples. Data augmentation is an efficient approach to avoid model overfit and biases by adding variances based on original data.

In our approach, both training and testing data are augmented according to two strategies. On one hand, classical methods including image rotation, scaling, flipping are randomly applied. On the other hand, augmentation is obtained with random selection based on the estimated nodule radii and sorted views.

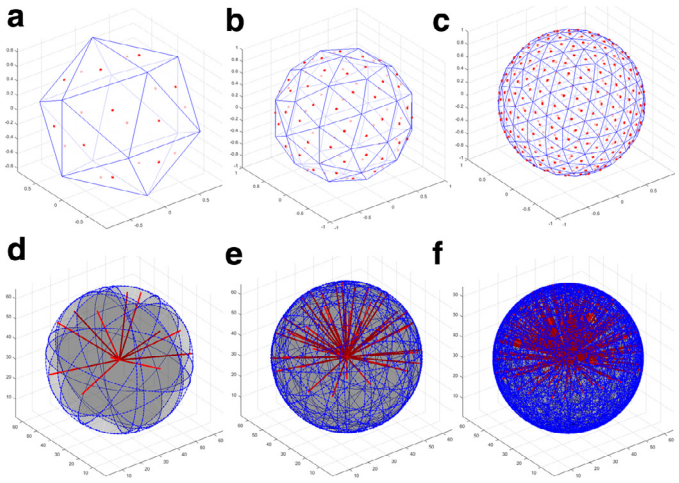

Fig. 3. Icosahedron-based sampling. Top row, icosahedron division at level 0, level 1 and level 2. Bottom row, corresponding sampling patterns with sampling planes (grey planes) and plane normals (red lines) shown. Due to symmetry, only top half sphere is considered. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

After augmentation, we achieve almost 640 training cases and 160 testing cases for each nodule type on LIDC-IDRI, while 690 testing cases in total on ELCAP.

4. Nodule spherical sampling, radii estimation and view sorting

Sampling is an important step for the proposed method. Instead of extracting features from segmented CT images, we divide nodule volumes by icosahedron pattern to analyze the characteristics of both nodules and their surrounding anatomical structures. With these sampling vectors, a threshold based approach is applied to estimate the nodule radii, followed by high frequency content analysis to sort sampling views according to nodule informatics each view holds. With estimated radii and sorted views, we first train the multi-view multi-scale CNN to use the corresponding fully connected layers to describe nodule statistical features and then calculate FV encodings to represent geometrical features between SIFT key-points. Finally, we apply the multiple kernel learning to fuse the obtained statistical and geometrical features to unified dimensions and classify nodule types according to hybrid features based on the support vector machine.

4.1. Preprocessing

The imaging qualities of original CT images are various. Besides, the inner-plane spacing and slice thickness are also different, making these data anisotropic. Towards this fact, we first linearly resample original images to unified spacing and slice thickness as 1 mm. Then, all CT values are clamped to [−1000 HU, 1000 HU] (Hounsfield Unit, HU) to remove noises and artifacts. This step is essential to reduce needless computations.

4.2. Nodule spherical sampling

There is significant difference between nodule type and classical natural image classification. It is commonly agreed that nodule type is related to not only nodule intensity distributions but also surrounding anatomical structures. Therefore, we apply the icosahedron-based sampling method which is described in Fig. 3 to capture both structures. The detailed sampling pipeline is described in Fig. 4. The algorithm for sampling is presented in Algorithm 1.

Algorithm 1: SPHERE_SAMPLING.

Input: V, nodule volume. r_{max}, maximum sampling radius. N / M / K,specified parameters.
Output: S, sampled matrix with size $N \times M \times K$ for V.

```
C ← generate N inner centers;
△r ← r_max/M;
for n = 1 → N, m = 1 → M do
    r_cur ← △r × m;
    circle ← build_circle(C_n, r_cur);
    for k = 1 → K do
        S_nmk ← linear_interp(V, circle);
return S;
```

Denote the candidate nodule volume as V and the corresponding gravity center as v_{xyz}. First, we apply the icosahedron centered at v_{xyz} with given radius r_{max} and *level*, within which the inner centers of each triangle on the divided mesh surface are calculated as $C = \{c_1, c_2, \ldots, c_N\}$. The characteristics of icosahedra assure these N centers spread uniformly, making a homogeneously sampling for nodule volume V. Then, we calculate N planes which cross v_{xyz} with plane normal vector $normal_n = (c_n - v_{xyz}), n \in \{1, 2, 3, \ldots, N\}$ as the so-called views. We use these views to slice the original CT volume and obtain corresponding 2-D

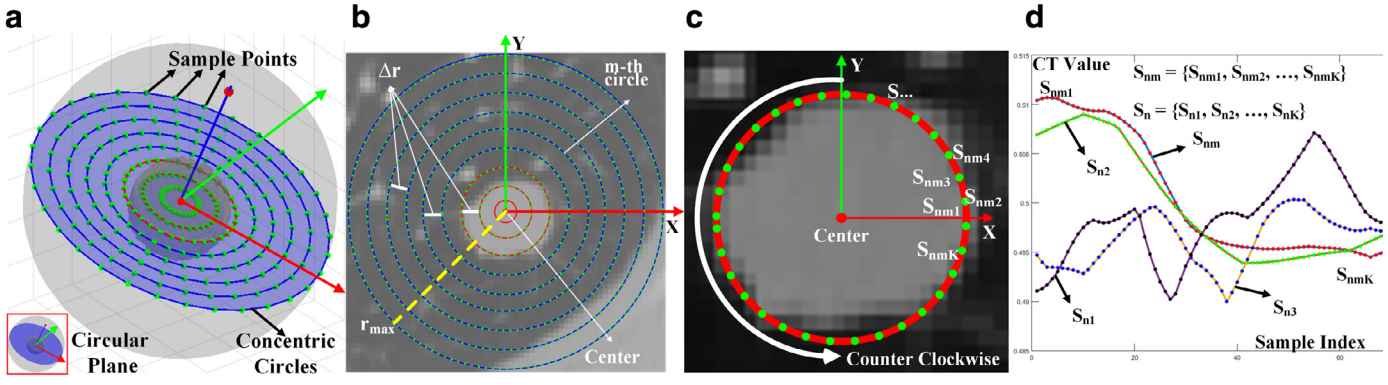

Fig. 4. Detailed scheme of spherical sampling. (a) 3-D view for the nth sampling circular plane (view) with concentric circles. The corresponding view angle is described on bottom left in red rectangle. (b) Sampling pattern for nth plane. Inner and outer nodule sampling are described using red and blue circles separately. (c) For mth circle (red circle), K points (green dots) are sampled in a counter clockwise direction. (d) Sampled values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

images, as described in Fig. 4(a). After that, we generate M concentric circles on each view (sliced 2-D image) with $r_m = \triangle r \times m, m \in \{1, 2, 3 \ldots, M\}$ and $\triangle r = r_{max}/M$, as described in Fig. 4(b). Finally, we sample K points along counter clockwise direction on each concentric circle, resulting in vectors describing CT values along corresponding circles, described as $S_{nm} = \{S_{nm1}, S_{nm2}, \ldots, S_{nmK}\}$ in Fig 4(c) and (d). Totally, sampling vectors S consist of data from N views, which is composed of M concentric circles and each circle of K data points.

4.3. Nodule radii estimation

Calculation of volume of interest (VOI) is essential for both data dimension reduction and feature accuracy increase. The intensity distribution of nodules on different sampling circles has discriminative characteristics. Intensity values change severely along nodules' contours, while remain stably low or high outside or inside nodule volume, separately. Therefore, we apply threshold method on sampling results S to estimate nodule radii. The algorithm is described in Algorithm 2.

Algorithm 2: RADIUS_ESTIMATION.

Input: S, sampled 3-D matrix with size $N \times M \times K$ for V. N / M / K, specified parameters.
Output: r_{est}, estimated radius for V.

1 $Counter \leftarrow 0$;
2 **for** $n = 1 \rightarrow N, m = 1 \rightarrow M, k = 1 \rightarrow K$ **do**
3 **if** $S_{nmk} > threshold$ **then**
4 $Counter_{nm} = Counter_{nm} + 1$;
5 $R\_Counter \leftarrow 0$;
6 **for** $m = 1 \rightarrow M, n = 1 \rightarrow N$ **do**
7 $R\_Counter_m \leftarrow R\_Counter_m + Counter_{nm}$;
8 **if** $R\_Counter_m < counter\_threshold$ **then**
9 $r_1 \leftarrow m$;
10 break;
11 $r_2 \leftarrow local\_\min(R\_Counter)$;
12 $r_{est} = \min(r_1, r_2)$;
13 return r_{est};

In our approach, with sampling results S, N circular planes generated by icosahedra and M concentric circles, we first calculate the count of K sampled values greater than given *threshold*, resulting in a counter matrix *Counter* with size $N \times M$ (Algorithm 2 Ln.2-Ln.4). Then, we accumulate the counter through all N views for each concentric circle, resulting in vector *R_Counter* with size M (Algorithm 2 Ln.6-Ln.10). Finally, the indices of first zero value and local minima in *R_Counter* are extracted as r_1 and r_2, separately. The estimated radius is calculated as the minimum between r_1 and r_2. Estimated radii for typical nodule cases are presented in Fig. 2.

4.4. View sorting

CT images are natural 3-D data, which are hard for human direct observation and diagnosis. The most practical way to have a better view for radiologists is to project 3-D voxel volumes to three 2-D fixed planes called *Axial, Sagittal* and *Coronal* planes, which is reasonable both for reducing complexness of CT images with many anatomical structures and for radiologists' better understanding. However, it should be noticed that nodules have various appearances and even severe irregular pathological areas. Therefore, *Axial, Sagittal* and *Coronal* views which are at fixed angles cannot display the most discriminative characteristics of nodules in many cases.

In our approach, we assess the quality of projection as how many irregular distributions, namely the variational changes presented in projected 2-D images. More irregular distributions captured by corresponding projected views should lead to better visualization for the appearances and characteristics of candidate nodules. The schematic view of the sorting procedure is shown in Fig. 5.

Based on the extracted nodule VOI, we re-sample nodule volume V again with r_{est} as the maximum radius to obtain more unified nodule distribution samples. Inspired by aforementioned analysis and unified sampling patterns by icosahedra, the problem of view sorting is transformed into how to sort N sampling views from icosahedron-based division according to their corresponding importance. In fact, according to our spherical sampling pattern, views with more abundant information should present more variational changes in their corresponding sampling vectors S_n, which can be straightforward solved in frequency domain since variations of signals in spatial domain definitely imply presence of high frequency components in frequency domain. Therefore, we apply the high frequency content analysis [44] as quantization of frequency changes in sampling data S_n for nth view, defined as,

$$freq_n = \sum_{m=1}^{M} D_C(S_{nm}), \tag{1}$$

where $D_C(S_{nm})$ is the complex domain difference between target and observed Short Time Fourier Transform (STFT) and defined as,

$$D_c(S_{nm}) = \frac{1}{K} \sum_{k=1}^{K} |\widehat{X}_k - X_k|, \tag{2}$$

where X_k is the STFT value of S_{nmk} and $\widehat{X}_k$ is the polar form of X_k, K is sampling count for each concentric circle.

Based on sampled intensity S, we first calculate the STFT and frequency for the K data points in S_{nm} corresponding to the mth concentric circle on nth view. Then, we accumulate all M frequencies for the nth view as $freq_n, n \in \{1 \ldots N\}$, which is an quantization standard and indicator for the corresponding view. Finally, all N views are sorted by $freq_n$ in descending order, indicating from most important to less.

Typical results for view sorting procedure are shown in Fig. 7. For each case, the top row shows the original slices on axial plane; the middle row presents slices sorted by *freq*; the bottom row shows 3-D views for corresponding nodules. It can be obviously observed that nodule characteristics in sorted slices are better expressed than those on fixed axial planes, which can also be confirmed in 3-D views.

5. Hybrid feature extraction

5.1. Multi-view multi-scale CNN feature extraction

Multi-view CNNs have been proven efficient in 3-D natural object classification [45], which is most suitable for nodule classification since CT images are naturally 3-D. However, it is also common sense that nodule types are related to not only nodule intensity distribution but also relative positions with surrounding anatomical structures. Therefore, inspired by the multi-view approach, with estimated nodule radius r_{est} and sorted views, we build 2-D images from slicing original volumes to capture characteristics of both nodules and the surround anatomical structures. The process for building slices is presented in Algorithm 3.

After building input images according to calculated scales and views, we build the proposed CNN, whose structure is described in Fig. 6. We have conducted the CNN as an end-to-end classifier with first 224×224 pixels input, then classical convolutional and

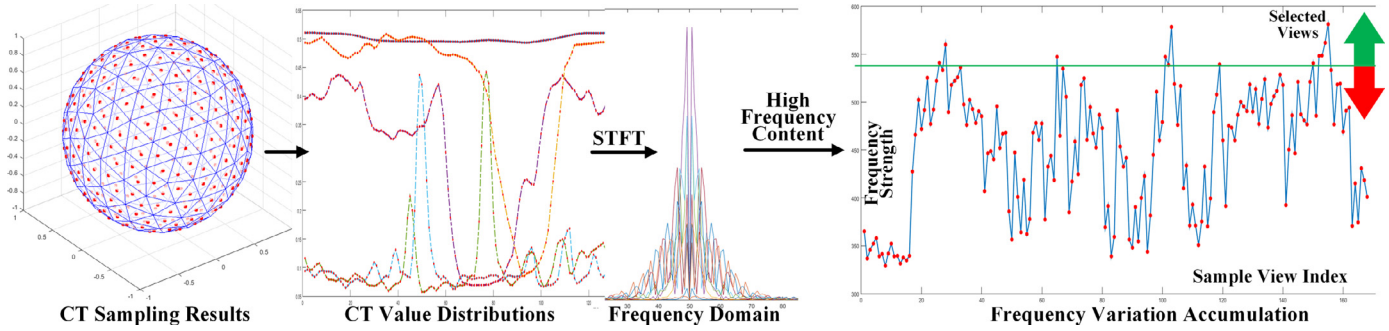

Fig. 5. View sorting based on high frequency content analysis. Short Time Fourier Transform (STFT) is applied to transform sampling data from spatial domain to frequency domain. High Frequency Content Analysis is utilized as quantization of variations.

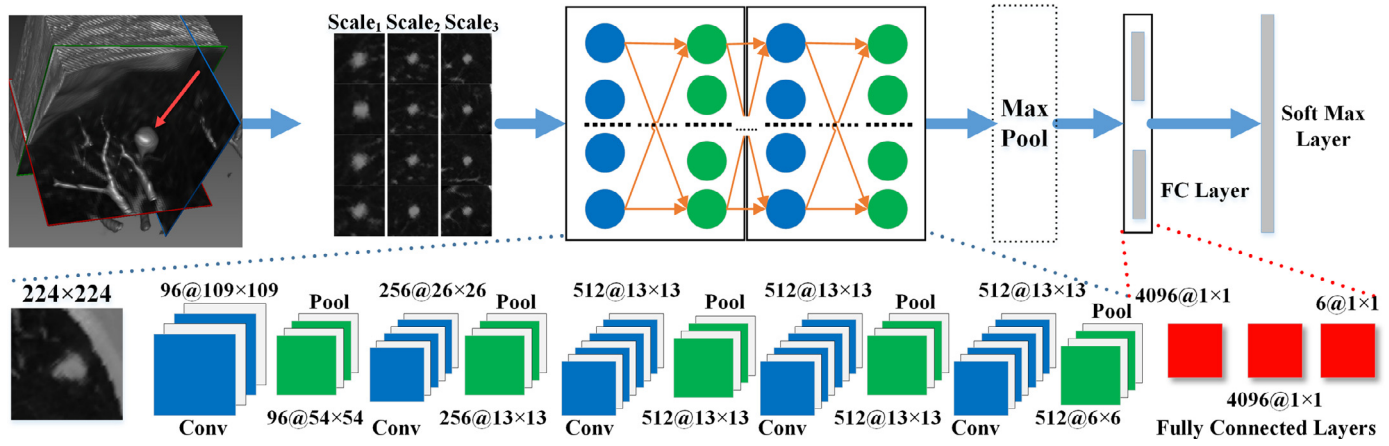

Fig. 6. The proposed CNN architecture. The pre-trained model is without max-pooling layer to train an independent CNN without affection of views. This model is then fine-tuned by combining all views together through the max-pooling.

Algorithm 3: BUILD_SLICES.

Input: V, candidate volume. *Scales*, given multi-scales. r_{max}, maximum sampling radius.
Output: *Slices*, generated nodule slices.

$S \leftarrow$ SPHERE_SAMPLING(V, r_{max});
$r_{est} \leftarrow$ RADIUS_ESTIMATION(S);
Slices $\leftarrow \Phi$;
for *scale in Scales* **do**
  $S \leftarrow$ SPHERE_SAMPLING(V, $r_{est} \times scale$);
  Views $\leftarrow$ VIEW_SORT(S);
  slice $\leftarrow$ sample V with *Views* and current *scale*;
  Slices $\leftarrow$ *Slices* $\cup$ *slice*
return *Slices*;

pooling layers, finally three fully connected layers, within which the first two are used to extract nodule features and the last is used as input for the soft-max classifier.

The initialization of weights for CNN layers also can definitely affect the final performance of CNN model. Random initialization can lead the optimization of CNNs into local minima in some cases. Therefore, in our approach, we adopt the pre-trained VGGNet [28] model as base to initialize the weights in pre-train stage.

After initialization, we apply a two-stage training scheme to obtain the CNN model. First, for each training case, we adopt three scales and four selected views, leading to twelve images and use each image to train the CNN, separately. Then, based on the pre-trained CNN model, each twelve images corresponding to the same nodule are grouped as a batch to fine-tune the CNN model with a max-pooling layer. Finally, we extract the CNN fully connected layer responses as statistical features for each nodule candidate. The algorithm for CNN feature extraction is described in Algorithm 4.

Algorithm 4: CNN_FEATURES.

Input: V, candidate volume. *Scales*, given multi-scales. r_{max}, maximum sampling radius.
Output: f_c, CNN extracted statistical features.

Slices $\leftarrow$ BUILD_SLICES(V, *Scales*, r_{max});
CNN $\leftarrow$ *initialize from based model*;
for *slice in Slices* **do**
  CNN $\leftarrow$ *pre_train*(*slice*);
CNN $\leftarrow$ *fine_tune*(*max_pool*(*Slices*));
$f_c \leftarrow$CNN fully connected layer responses;
return f_c;

5.2. Fisher vector feature extraction

Scale-invariant feature transform (SIFT) [46] is a stable and robust local feature extraction method based on extrema in spatial domain, which is broadly applied in many computer vision areas. SIFT vectors focus on local features, which cannot express correlations between key-points due to lack of global information. Meanwhile, Fisher vector (FV) [47] obtains model by applying Gaussian Mixture Model (GMM) to input samples, resulting in unified representations which can encode both local features and global correlations between data points. It is straightforward to use FV to

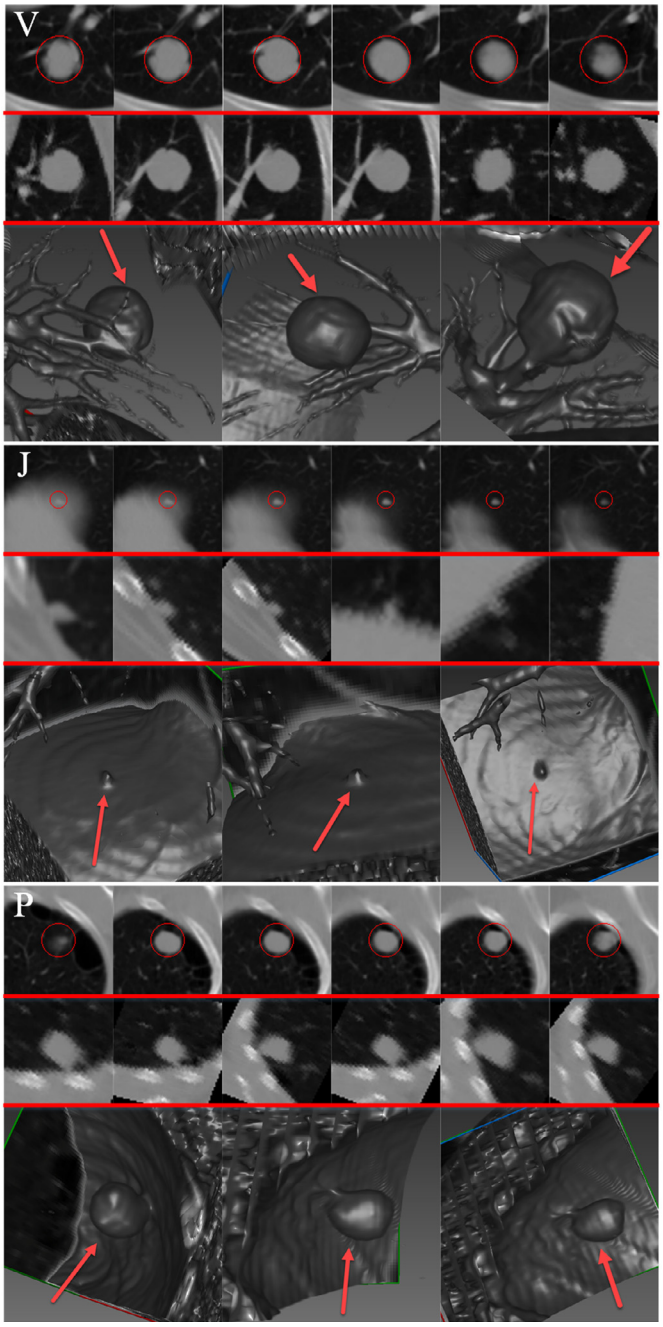

Fig. 7. View sorting results for typical cases. For each case, top row, original slices on axial plane; middle row, view sorting slices; bottom row, 3-D views for nodules.

encode SIFT features to obtain unified descriptors. However, it is also obvious that medical images are different from natural images due to their low lumen and low contrast, making it hard to locate the classical SIFT key-points. Therefore, we apply the dense SIFT method [48] to calculate uniform SIFT vectors. Besides, we also apply the improved FV method [49] with normalization and non-linear additive kernel to obtain encodings.

The pipeline of FV feature extraction is shown in Fig. 8. In our approach, we first build image batches according to estimated nodule radii and view order, and calculate dense SIFT features for each slice corresponding to the same nodule. Then, we stack all SIFT vectors and apply principle component analysis (PCA), followed by a whitening operation for dimensionality reduction, noise removal and feature correlation decoupling. Finally, we utilize the Gaussian Mixture Model (GMM) [50] and encode SIFT vectors grouped by nodule to build corresponding unified FV presentations. The extraction algorithm is described in Algorithm 5,

Algorithm 5: FISHER_FEATURES.

Input: V, candidate volume. *Scales*, given multi-scales.r_{max}, maximum sampling radius.
Output: f_s, Fisher geometrical features.

```
Slices ← BUILD_SLICES(V, Scales, r_max);
Sifts ← {};
for slice in Slices do
    sift ← dense_sift(slice);
    Sifts ← Sifts ∪ sift;
PCA_Sifts ← PCA(Sifts);
mean, covar, prior ← GMM(PCA_Sifts);
f_s ← fisher_encoding(PCA_Sifts, mean, covar, prior);
return f_s;
```

5.3. Feature fusion based on multiple kernel learning

The achieved CNN features represent patterns of intensity distribution, regarded as statistical features. Meanwhile, FV features process inner differences and distances between sample data points, regarded as geometrical features. Both features are measured by different dimensions and simply combing them with stacking is unreasonable. In fact, the unified feature vectors combing both features can be described as,

$$F(x_i, x_j) = \sum_{k \in K} w_k \cdot f_k(x_i, x_j), \tag{3}$$

where f is the original features and w is the weights to transform the stacked features for unified non-dimensional vectors.

The schematic description for achieving hybrid features is described in Fig. 9. In our approach, we apply the multiple kernel learning (MKL) [51] method for heterogeneous information fusion. The key step lies in learning corresponding weights w from original feature space. Therefore, we first project features with linear combination of multiple kernels and utilize a multi-label support vector machine (SVM) to assess the weights. Finally, we obtain the hybrid features as f and train SVM model to classify corresponding nodule types.

6. Experiments

In our approach, we apply the proposed method on LIDC-IDRI for training, while validate the model on both LIDC-IDRI and ELCAP data sets. All cases are extracted as $64 \times 64 \times 64\,\text{mm}^3$ volume. For LIDC-IDRI, we totally select 1738 nodules and 1000 non-nodules from 744 chest CTs, which are then annotated by an experienced radiologist to label corresponding nodule types. 80% of LIDC-IDRI candidates are used for training and validating, while 20% are used for testing. For ELCAP, we treat all 421 nodules as testing cases. Data augmentation is applied on both training and testing cases, separately. After augmentation, we obtain almost 640 training and 160 testing cases for each nodule type on LIDC-IDRI and achieve 680 testing cases on ELCAP.

6.1. Parameter selection

The parameter *number of views* has been discussed in [45]. Meanwhile, we adopt three scales and four views, resulting in twelve 2-D slices for each nodule. Following that, we are using three scales which actually represent corresponding multiple of

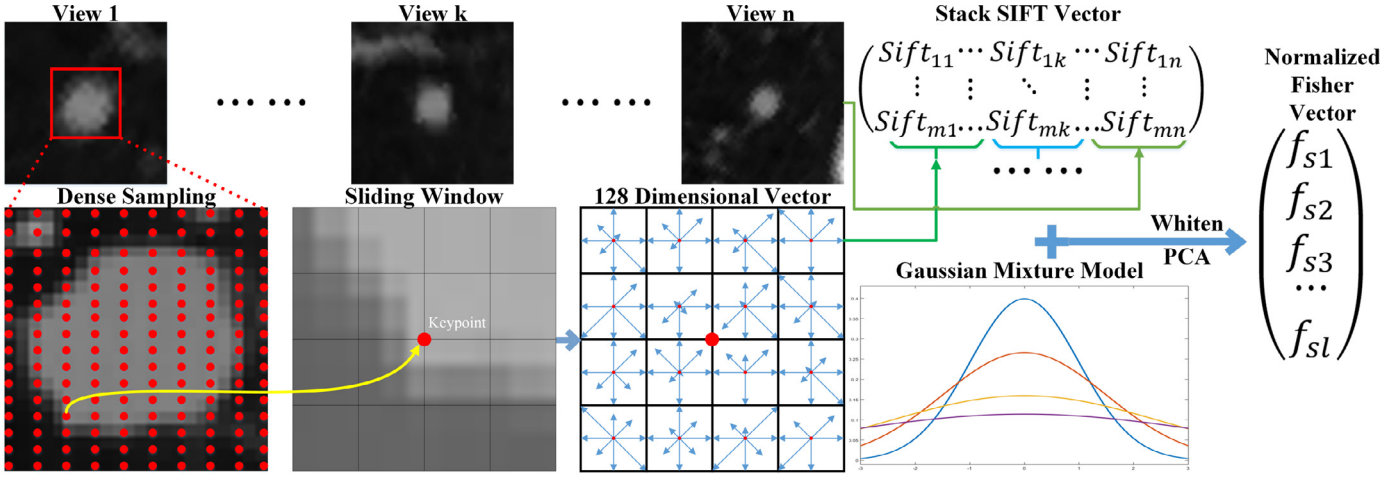

Fig. 8. The proposed pipeline for Fisher vector feature extraction. Firstly, image slice patches are built based on estimated nodule radii and sorted views. Then, dense SIFT vectors are calculated and stacked, followed by PCA and whitening operations to obtain the final FV encodings.

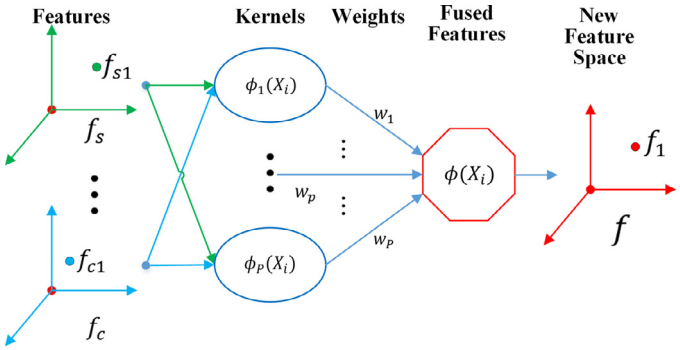

Fig. 9. The multiple kernel learning approach for feature fusion. By applying multiple kernel learning, different weights are calculated to fuse features from different spaces into hybrid features in the new feature space.

calculated rest to determine the sampling area, considering following circumstances: scale 2.5 is used to make sure more nodule details are captured; scale 4.0 is used to capture more surrounding details around nodules.

In preprocessing, all images are re-sliced to an inner plane space as 1 mm, since the maximum resolution of original CT images is around 1 mm; CT values within [−1000,1000] HU are linearly normalized into [0,1]. In sampling stage, we are using a level-2 icosahedron mesh, but only 168 ($N=168$) inner centers are used due to symmetry; 32 ($M=32$) concentric circles are built for each circular plane since r_{max} is defined as 32 and each circle represents 1 mm; 128 ($K=128$) sampling points are evenly distributed on each concentric circle; maximum sampling radius (r_{max}) is set to 32 mm based on the fact that extracted nodules are all 64 × 64 × 64 mm^3 volumes. In feature extraction stage, we apply two fully connected layers with 4096-D as CNN features. Meanwhile, we apply the 18 × 18 dense SIFT, followed by a PCA operation to reduce the feature dimension to 32-D, resulting in 4096-D FV features. All default parameters are presented in Table 1.

6.2. Results

Classification results on LIDC-IDRI for typical cases of each type are shown in Figs. 10 and 11. Both built slices for training the CNN (top row in each case) and 3-D appearances (bottom row in each case) of corresponding nodules (pointed by arrows) are shown. The characteristics of each nodule type can be easily observed in 3-D view.

Table 1
Parameter selection for the proposed method.

Parameter	Value	Description
Scales	2.5, 3.0, 4.0	Sampling scales
N	168	Number of circular planes
M	32	Number of concentric circles
K	128	Number of sampling points
r_{max}	32 mm	Maximum sampling radius
threshold	0.15	Minimum accepted CT value

6.3. Performance

Totally, we apply the proposed CNN and FV model on LIDC-IDRI with almost 3860 training and 1000 testing cases (640 and 160 cases for each type, separately). Data augmentation is applied on both training and testing data, separately. As described in Table 3, classification rate for each type is presented in the corresponding confusion matrix. Totally, we achieve an overall classification rate as 93.1% (911 out of 979) through all types of candidates. We can also work out that the proposed method recognize almost all **W** type nodules accurately, while hits a lower score for **P** type nodules. We believe this mainly lies in the hardness for **P** type nodule radii estimation.

6.4. Comparison

Although many researches have been done in nodule detection and malignancy estimation, there are still few focusing on nodule accurate type classification. Patch-based classification by Zhang et al. [41] which is validated on ELCAP [43] focuses on the same field with us. Thus, we applied the proposed method on ELCAP to achieve more comparable results. The confusion matrices for both the proposed and patch-based method are presented in Table 2.

It should be noted that patch-based method [41] can only handle **W**, **P**, **V** and **J** nodules and acquire an overall classification rate as 89%. Besides, we validate our proposed method for all nodule types with an overall classification rate as 93.0% (4 types, 384 out of 413) and 93.9% (all 6 types, 647 out of 689). It can be informed that despite extending classification to six classes, the proposed method has also outperformed PB method in both single-class and overall classification accuracy except **J** type nodules. From Table 2, we can conclude that classification errors of the proposed method for **J** type nodules mainly lie in **N** type, which PB method cannot handle. Therefore, if we exclude the **N** type and **G** type nodules,

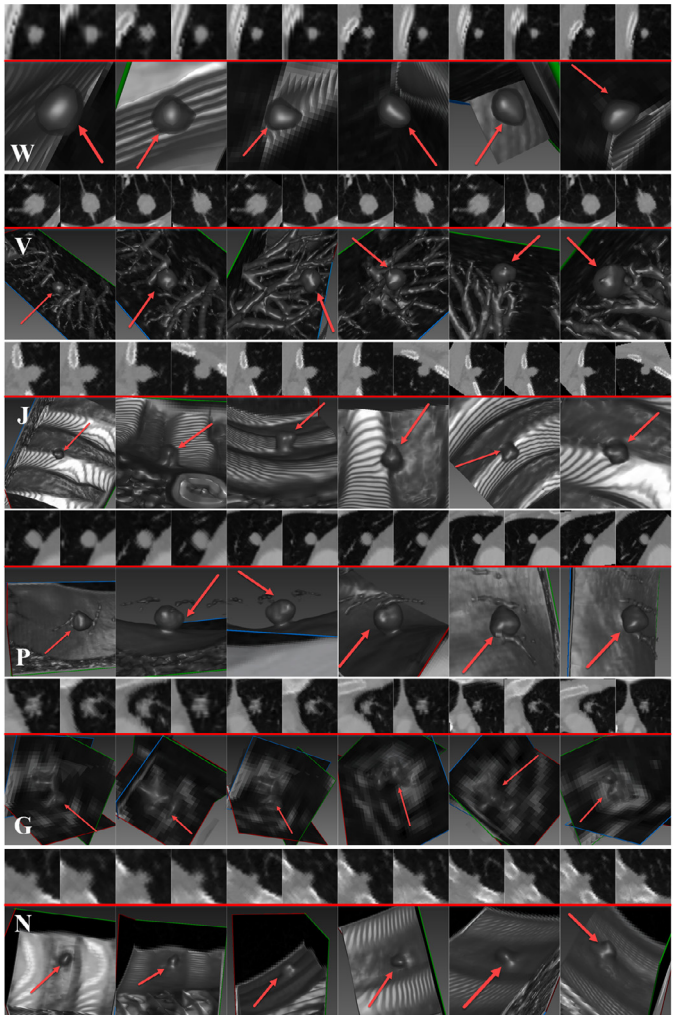

Fig. 10. Classification results for typical cases of each type. Corresponding nodule types are labeled. Arrows are used to point out nodule positions. Nodule type characteristics can be easily recognized in 3-D views.

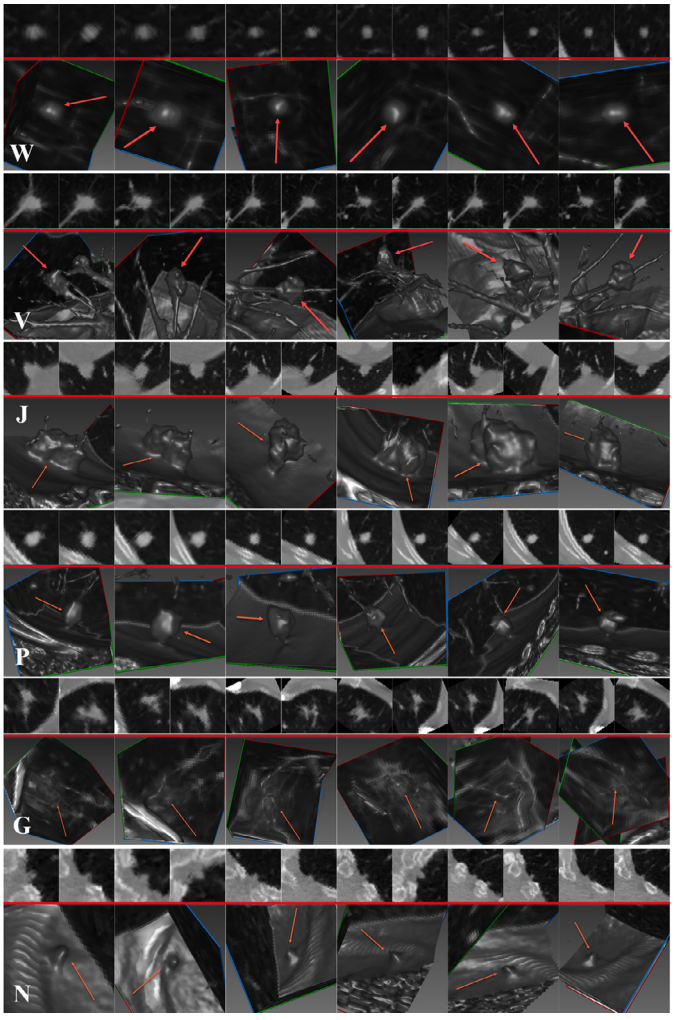

Fig. 11. Classification results for another set of typical cases.

our proposed method has outperformed the PB method in all four classes (**W**, **J**, **P**, **V**).

Meanwhile, the most classical comparable scheme is Bag-of-Visual-Words representation (BOW) [12] towards our proposed method. Therefore, in our approach, we have compared the proposed method in the following ways,

- CNN_S, the classical pure CNN fashion without max-pooling layer, using the soft-max layer as the classifier. The network architecture is the same as described in Fig. 6. Multi-view multi-scale scheme is also applied.
- CNN_S+Pooling, same with CNN_S, but with max-pooling layer.
- CNN_F, the classical CNN fashion without max-pooling. Independent images are used to train the CNN without any group or batch information. The descriptors are arranged as 4096-D vectors from fully connected layer responses for each input image.
- CNN_F+Pooling, the multi-view multi-scale CNN+Pooling approach. Images are grouped into batches according to corresponding nodules and processed into unified features via max-pooling layer. The descriptors are organized as 4096-D vectors from CNN fully connected layer responses for each corresponding nodule.
- FV+SIFT. We utilize comparison with FV encodings [47] under SIFT descriptors. Totally, GMM models are applied to determine 32 cluster centers, while SIFT vectors are reduced to 32-D by PCA and whitening, leading to 4096-D FV encodings finally.

Table 2
Confusion matrix for comparison with PB [41] on ELCAP.

Type/method		G	W	N	P	V	J
G	PB	N/A	N/A	N/A	N/A	N/A	N/A
	The proposed	**0.793**	0	0.138	0	0	0.069
W	PB	N/A	0.89	N/A	0.06	0.04	0.01
	The proposed	0	**0.961**	0.039	0	0	0
N	PB	N/A	N/A	N/A	N/A	N/A	N/A
	The proposed	0	0.028	**0.972**	0	0	0
P	PB	N/A	0.03	N/A	0.91	0.03	0.03
	The proposed	0	0	0.063	**0.938**	0	0
V	PB	N/A	0.05	N/A	0.06	0.86	0.03
	The proposed	0	0	0.068	0	**0.932**	0
J	PB	N/A	0.03	N/A	0.06	0.04	**0.87**
	The proposed	0	0	0.169	0	0	0.831

Prop = The proposed method.
PB = Patch-based method [40].
G = Ground glass optic.
N = Non-nodule.
P = Pleural-tail.
W = Well-circumscribed.
V = Vascularized.
J = Juxta-pleural.

Table 3
Confusion matrix for comparison. Classical CNN_S, CNN_S+Pooling, CNN_F, CNN_F+Pooling, FV+SIFT [47] and BOW+SIFT [12] are considered.

Type/method		G	W	N	P	V	J
G	CNNS	0.809	0.045	0.101	0	0.045	0
	CNNS+Pooling	0.854	0.023	0.079	0.011	0.034	0
	CNNF	0.596	0.067	0.101	0.045	0.124	0.067
	CNNF+Pooling	0.809	0.034	0.112	0	0.045	0
	FV+SIFT	0.615	0.021	0.077	0.070	0.154	0.063
	BOW+SIFT	0.524	0.063	0.063	0.091	0.168	0.091
	The proposed	**0.899**	0.011	0.079	0	0.011	0
W	CNNS	0.022	0.801	0.018	0.074	0.055	0.030
	CNNS+Pooling	0.011	0.867	0.011	0.066	0.033	0.011
	CNNF	0.052	0.683	0.041	0.074	0.125	0.026
	CNNF+Pooling	0.011	0.963	0.015	0	0.011	0
	FV+SIFT	0.022	0.735	0.040	0.036	0.157	0.009
	BOW+SIFT	0.036	0.664	0.027	0.054	0.206	0.013
	The proposed	0	**0.993**	0.004	0.004	0	0
N	CNNS	0.015	0.045	0.865	0.020	0.020	0.035
	CNNS+Pooling	0.010	0.040	0.885	0.010	0.020	0.035
	CNNF	0.035	0.070	0.735	0.045	0.030	0.085
	CNNF+Pooling	0.005	0.035	0.885	0.010	0.030	0.035
	FV+SIFT	0.129	0.041	0.676	0.071	0.006	0.076
	BOW+SIFT	0.153	0.029	0.571	0.071	0.018	0.159
	The proposed	0.010	0.035	**0.915**	0.005	0.025	0.010
P	CNNS	0.027	0.054	0.116	0.705	0.027	0.071
	CNNS+Pooling	0	0.063	0.089	0.786	0	0.063
	CNNF	0.036	0.116	0.107	0.607	0.009	0.125
	CNNF+Pooling	0.018	0.045	0.152	0.759	0.009	0.018
	FV+SIFT	0.067	0.044	0.081	0.659	0.015	0.133
	BOW+SIFT	0.089	0.111	0.126	0.474	0.030	0.170
	The proposed	0	0.018	0.116	**0.866**	0	0
V	CNNS	0.064	0.113	0.043	0	0.780	0
	CNNS+Pooling	0.028	0.078	0.050	0.028	0.816	0
	CNNF	0.057	0.142	0.064	0.043	0.688	0.007
	CNNF+Pooling	0.007	0.085	0.078	0	0.830	0
	FV+SIFT	0.063	0.162	0.035	0.077	0.606	0.056
	BOW+SIFT	0.063	0.232	0.007	0.092	0.507	0.099
	The proposed	0.014	0.035	0.043	0	**0.908**	0
J	CNNS	0.012	0.018	0.096	0.036	0.006	0.831
	CNNS+Pooling	0.012	0	0.042	0.042	0	0.904
	CNNF	0.018	0.036	0.102	0.133	0.006	0.705
	CNNF+Pooling	0.012	0.030	0.066	0.054	0.006	0.831
	FV+SIFT	0.048	0.030	0.139	0.157	0.030	0.596
	BOW+SIFT	0.048	0.048	0.181	0.108	0.024	0.590
	The proposed	0.018	0.006	0.048	0	0	**0.928**

G = Ground glass optic.
N = Non-nodule.
P = Pleural-tail.
W = Well-circumscribed.
V = Vascularized.
J = Juxta-pleural.

- BOW+SIFT. We compare the proposed method with the most classical BOW [12] scheme using SIFT features. We apply the K-Means method to obtain the vocabulary with size 4096, leading to 4096-D BOW encodings for each nodule.

All comparisons except for CNN_S and CNN_S+Pooling are validated with a multi-class SVM as a classifier. The confusion matrices for all comparisons are presented in Table 3.

Totally, for CNN_S, CNN_S+Pooling, CNN_F, CNN_F+Pooling, FV+SIFT and BOW+SIFT, the overall classification rates are 80.6, 85.9, 68.1, 86.8, 65.5 and 56.6%, separately. We can conclude that although BOW method is efficient for natural image classification and retrieval, it cannot compete in medical image fields. FV encoding outperforms the classical BOW method but is either not acceptable. The classical CNN fashion (CNN_S, CNN_S+Pooling, CNN_F, CNN_F+Pooling) has been proven efficient in classification, especially when multi-view multi-scale scheme is applied (CNN_S, CNN_S+Pooling, CNN_F+Pooling). The application of max-pooling layer also helps to raise classification accuracy. From Table 3, we can conclude that the proposed CNN feature extraction method outperforms the previous methods and achieve significant improvements. The multi-view multi-scale scheme assures the proposed CNN capture more distinctive features. At last, by fusing FV encodings and CNN features, we have improved the accuracy compared with both CNN-only and FV-only. The advantages by fusion mainly lie in classification for **G** (9%), **V** (8%), **P** (11%) and **J** (10%) types. This is because FV encodings are based on dense SIFT vectors which are irrelevant with nodule radii. Therefore, FV encodings improve the final hybrid features by fulfilling CNN missing features caused by radii wrong estimation.

7. Conclusions

In this paper, we have devised a hybrid-feature-based lung nodule type classification method. We utilize the normalized spherical sampling method based on icosahedra. Meanwhile, we apply threshold method to estimate nodule radii and high frequency content analysis to sort views planes. After that, we train the multi-view multi-scale CNN and calculate the Fisher vector encodings. Finally, by fusing the CNN features and FV features

through MKL, we obtain the hybrid features which are most descriptive for nodule characteristics. The experiments on LIDC-IDRI and ELCAP have shown we achieve an overall classification rate of 93.1% (911 out of 979) and 93.9% (647 out of 689) separately, which is promising in clinical practice. We hope our method could be of great assistance for radiologists for diagnosis of lung cancer. Future work will lie in two aspects. One is a scale invariant and more robust radius estimation method towards very tiny nodules (radius $\leq$ 3 mm) and juxta-pleural nodules. The other is automatically labeling out types, positions and sizes for nodules not centered in images with few human interactions.

Acknowledgments

This research is supported in part byNational Natural Science Foundation of China (Grant Nos. 61532002, 61672149, 61672077), National Science Foundation of USA (Grant Nos. IIS-0949467, IIS-1047715, and IIS-1049448).

References

[1] Siegel R, Miller K, Jemal A. Cancer statistics, 2017. CA-CANCER J Clin 2017;67(1):7–30.
[2] Erasmus J, Connolly J, McAdams H, Roggli V. Solitary pulmonary nodules: part I. morphologic evaluation for differentiation of benign and malignant lesions. Radiographics 2000;20(1):43–58.
[3] Wu D, Lu L, Bi J, Shinagawa Y, Boyer K, Krishnan A, et al. Stratified learning of local anatomical context for lung nodules in ct images. In: Proceedings of the computer vision and pattern recognition; 2010. p. 2791–8.
[4] Farag A, Elhabian S, Graham J, Farag A, Falk R. Toward precise pulmonary nodule descriptors for nodule type classification. In: Proceedings of the medical image computing and computer-assisted intervention; 2010. p. 626–33.
[5] Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical ct images. IEEE Trans Med Imag 2003;22(10):1259–74.
[6] Yao J, Bliton J, Summers R. Automatic segmentation and measurement of pleural effusions on CT. IEEE Trans Biomed Eng 2013;60(7):1834–40.
[7] Song Y, Cai W, Zhou Y, Feng D. Feature-based image patch approximation for lung tissue classification. IEEE Trans Med Imag 2013;32(4):797–808.
[8] Mansoor A, Bagci U, Mollura D. Near-optimal keypoint sampling for fast pathological lung segmentation. In: Proceedings of the conference of the IEEE engineering in medicine and biology society; 2014. p. 6032–5.
[9] Gao Q, Wang S, Zhao D, Liu J. Accurate lung segmentation for X-ray CT images. In: Proceedings of the international conference on computing, networking and communications, 2; 2007. p. 275–9.
[10] van Rikxoort E, de Hoop B, Viergever M, Prokop M, van Ginneken B. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Med Phys 2009;36(7):2934–47.
[11] Kumar A, Kim J, Cai W, Fulham M, Feng D. Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. J Digit Imag 2013;26(6):1025–39.
[12] Galaro J, Judkins A, Ellison D, Baccon J, Madabhushi A. An integrated texton and bag of words classifier for identifying anaplastic medulloblastomas. In: Proceedings of the conference of the IEEE engineering in medicine and biology society; 2011. p. 3443–6.
[13] Khurd P, Bahlmann C, Maday P, Kamen A, Gibbs-Strauss S, Genega E, et al. Computer-aided gleason grading of prostate cancer histopathological images using texton forests. In: Proceedings of the international symposium on biomedical imaging; 2010. p. 636–9.
[14] Zhang F, Cai W, Song Y, Lee MZ, Shan S, Dagan D. Overlapping node discovery for improving classification of lung nodules. In: Proceedings of the conference of the IEEE engineering in medicine and biology society; 2013. p. 5461–4.
[15] Zhang F, Song Y, Cai W, Zhou Y, Shan S, Feng D. Context curves for classification of lung nodule images. In: Proceedings of the digital image computing: techniques and applications; 2013. p. 1–7.
[16] Zhang F, Song Y, Cai W, Zhou Y, Fulham M, Eberl S, et al. A ranking-based lung nodule image classification method using unlabeled image knowledge. In: Proceedings of the international symposium on biomedical imaging; 2014. p. 1356–9.
[17] Unay D, Ekin A. Dementia diagnosis using similar and dissimilar retrieval items. In: Proceedings of the international symposium on biomedical imaging; 2011. p. 1889–92.
[18] Song Y, Cai W, Kim J, Feng DD. A multistage discriminative model for tumor and lymph node detection in thoracic images. IEEE Trans Med Imag 2012;31(5):1061–75.
[19] Jacobs C, Sánchez C, Saur S, Twellmann T, de Jong P, van Ginneken B. Computer-aided detection of ground glass nodules in thoracic CT images using shape, intensity and context features. In: Proceedings of the medical image computing and computer-assisted intervention, part III; 2011. p. 207–14.
[20] Sorensen L, Shaker S, de Bruijne M. Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Trans Med Imag 2010;29(2):559–69.
[21] Ciompi F, Jacobs C, Scholten E, Wille M, de Jong P, Prokop M, et al. Bag-of-frequencies: a descriptor of pulmonary nodules in computed tomography images. IEEE Trans Med Imag 2015;34(4):962–73.
[22] Kim H, Park CM, Goo JM, Wildberger JE, Kauczor H-U. Quantitative computed tomography imaging biomarkers in the diagnosis and management of lung cancer. Investig Radiol 2015;50(9):571—583.
[23] Cirujeda P, Cid YD, Müller H, Rubin D, Aguilera TA, Loo BW, et al. A 3-d Riesz–covariance texture model for prediction of nodule recurrence in lung CT. IEEE Trans Med Imag 2016;35(12):2620–30.
[24] Chen S, Ni D, Qin J, Lei B, Wang T, Cheng J-Z. Bridging computational features toward multiple semantic features with multi-task regression: a study of CT pulmonary nodules. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. in: Proceedings of the medical image computing and computer-assisted intervention, part II; 2016. p. 53–60.
[25] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–324.
[26] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44.
[27] Chatfield K, Simonyan K, Vedaldi A, Zisserman A. Return of the devil in the details: delving deep into convolutional nets. CoRR 2014. abs/1405.3531.
[28] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. CoRR 2014a. abs/1409.1556.
[29] Rongjian L, Zhang W, Suk H-I, Wang L, Li J, Shen D, et al. Deep learning based imaging data completion for improved brain disease diagnosis. In: Proceedings of the medical image computing and computer-assisted intervention, part III; 2014. p. 305–12.
[30] Roth HR, Lu L, Seff A, Cherry K, Hoffman J, Wang S, et al. A new 2.5d representation for lymph node detection using random sets of deep convolutional neural network observations. In: Proceedings of the medical image computing and computer-assisted intervention, part I; 2014. p. 520–7.
[31] Brosch T, Yoo Y, Li DK, Traboulsee A, Tam R. Modeling the variability in brain morphology and lesion distribution in multiple sclerosis by deep learning. In: Proceedings of the medical image computing and computer-assisted intervention, part II; 2014. p. 462–9.
[32] Lo S, Chan H, Lin J, Li H, Freedman M, Mun S. Artificial convolution neural network for medical image pattern recognition. Neural Netw 1995;8(7):1201–14.
[33] Setio A, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel S, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imag 2016;35(5):1160–9.
[34] Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: Proceedings of the medical image computing and computer-assisted intervention, part II; 2013. p. 246–53.
[35] Dou Q, Chen H, Yu L, Zhao L, Qin J, Wang D, et al. Automatic detection of cerebral microbleeds from mr images via 3d convolutional neural networks. IEEE Trans Med Imag 2016;35(5):1182–95.
[36] Jacobs C, van Rikxoort E, Twellmann T, Scholten ET, de Jong PA, Kuhnigk JM, et al. Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. Med Imag Anal 2014;18(2):374–84.
[37] Liu X, Hou F, Qin H, Hao A. A cade system for nodule detection in thoracic ct images based on artificial neural network. Sci Chin Inf Sci 2017;60(7):072106.
[38] Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in CT scans. Sci Rep 2016;6:24454.
[39] Farag A, Ali A, Graham J, Farag A, Elshazly S, Falk R. Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest. In: Proceedings of the international symposium on biomedical imaging; 2011. p. 169–72.
[40] Song Y, Cai W, Wang Y, Feng DD. Location classification of lung nodules with optimized graph construction. In: Proceedings of the international symposium on biomedical imaging; 2012. p. 1439–42.
[41] Zhang F, Song Y, Cai W, Lee M, Zhou Y, Huang H, et al. Lung nodule classification with multilevel patch-based context analysis. IEEE Trans Biomed Eng 2014;61(4):1155–66.
[42] Armato III Samueland McLennan G, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on ct scans. Med Phys 2011;38(2):915–31.
[43] Vison, Group I.A. I-ELCAP: international early lung cancer action program. 2013. http://www.via.cornell.edu/databases/lungdb.html.
[44] Brossier P, Bello J, Plumbley M. Real-time temporal segmentation of note objects in music signals. In: Proceedings of the international computer music conference 2004.
[45] Su H, Maji S, Kalogerakis E, Learned-Miller E. Multi-view convolutional neural networks for 3d shape recognition. In: Proceedings of the international conference on computer vision; 2015. p. 945–53.
[46] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis 2004;60(2):91–110.
[47] Perronnin F, Dance C. Fisher kernels on visual vocabularies for image categorization. In: Proceedings of the computer vision and pattern recognition; 2007. p. 1–8.
[48] Yang J, Yu K, Gong Y, Huang T. Linear spatial pyramid matching using sparse coding for image classification. In: Proceedings of the computer vision and pattern recognition; 2009. p. 1794–801.

[49] Perronnin F, Sánchez J, Mensink T. Improving the Fisher kernel for large-scale image classification. In: Proceedings of the European conference on computer vision, part IV; 2010. p. 143–56.
[50] Huang Y, Englehart KB, Hudgins B, Chan ADC. A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses. IEEE Trans Biomed Eng 2005;52(11):1801–11.
[51] Gönen M, Alpaydı n E. Multiple kernel learning algorithms. J Mach Learn Res 2011;12:2211–68.