Pulmonary Nodule Classification Using a Multiview Residual Selective Kernel Network

Herng-Hua Chang[1] · Cheng-Zhe Wu[1] · Audrey Haihong Gallogly[2]

Abstract

Lung cancer is one of the leading causes of death worldwide and early detection is crucial to reduce the mortality. A reliable computer-aided diagnosis (CAD) system can help facilitate early detection of malignant nodules. Although existing methods provide adequate classification accuracy, there is still room for further improvement. This study is dedicated to investigating a new CAD scheme for predicting the malignant likelihood of lung nodules in computed tomography (CT) images in light of a deep learning strategy. Conceived from the residual learning and selective kernel, we investigated an efficient residual selective kernel (RSK) block to handle the diversity of lung nodules with various shapes and obscure structures. Founded on this RSK block, we established a multiview RSK network (MRSKNet), to which three anatomical planes in the axial, coronal, and sagittal directions were fed. To reinforce the classification efficiency, seven handcrafted texture features with a filter-like computation strategy were explored, among which the homogeneity (HOM) feature maps are combined with the corresponding intensity CT images for concatenation input, leading to an improved network architecture. Evaluated on the public benchmark Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) challenge database with ten-fold cross validation of binary classification, our experimental results indicated high area under receiver operating characteristic (AUC) and accuracy scores. A better compromise between recall and specificity was struck using the suggested concatenation strategy comparing to many state-of-the-art approaches. The proposed pulmonary nodule classification framework exhibited great efficacy and achieved a higher AUC of 0.9711. The association of handcrafted texture features with deep learning models is promising in advancing the classification performance. The developed pulmonary nodule CAD network architecture is of potential in facilitating the diagnosis of lung cancer for further image processing applications.

Keywords Lung nodule · Image classification · Residual learning · Selective kernel · CT

Introduction

Cancer is one of the leading causes of death and a public health problem worldwide. In 2020, lung cancer cases were the second most and the number of death caused by lung cancer was the greatest. There are approximately 2.2 million people, which account for 11.4% among all cancers, suffering from lung cancer with a high mortality rate of 80% [1].

Moreover, the 5-year relative survival rate for lung cancer is only 21% practically [2]. Nevertheless, the National Lung Screening Trial has demonstrated that the mortality of lung cancer can be reduced by 20% by earlier diagnoses with low-dose computed tomography (CT) screening [2, 3]. A lung nodule is recognized as a white spot in thoracic CT scans [4]. It exhibits a round area and more solid than normal lung tissue. If it is larger than 30 mm, it is called a lung mass and has a higher probability being cancerous. Lung nodules can be divided into benign, indeterminate, and malignancy. Benign nodules usually have a smaller size and smoother contour, which are noncancerous. Contrarily, malignancies are cancerous and mostly have larger sizes and more variable structures.

Clinically, physicians and radiologists read thoracic CT scans to classify nodules slice by slice [5]. This routine is time consuming, labor expensive, and error prone.

✉ Herng-Hua Chang
herbertchang@ntu.edu.tw

[1] Computational Biomedical Engineering Laboratory (CBEL), Department of Engineering Science and Ocean Engineering, National Taiwan University, 1 Sec. 4 Roosevelt Road, Daan, Taipei 10617, Taiwan

[2] Department of Radiation Oncology, Keck Medical School, University of Southern California, Los Angeles, CA, USA

The professional experience also affects the accuracy of differentiating benign from malignant nodules [6]. With the advancement of computer technology, computer-aided diagnosis (CAD) systems have been widely employed in medical image analysis including lung cancer detection [7] and identification [8]. Generally, there are four main procedures for lung nodule classification in a CAD system [9]: (1) data collection, (2) data preprocessing, (3) feature extraction, and (4) nodule classification. Thanks to the assistance of CAD systems, the examination and classification of pulmonary nodules have been shown more accurate and robuster [8, 10].

Depending on how the image features are acquired, CAD systems can be broadly divided into two main categories: handcrafted-based and convolutional neural network (CNN)-based approaches [8, 11]. Handcrafted-based CAD systems generally adopt texture features such as the local binary pattern (LBP) [12], histogram of oriented gradient (HOG) [13], Gabor [14], gray-level co-occurrence matrix (GLCM) [14], scale invariant feature transform [15], and speeded-up robust features [16]. On the other hand, due to the rapid hardware and software advancement, deep learning strategies particularly CNN models have been extensively utilized for image classification. CNN-based CAD systems can efficiently perform feature extraction and nodule classification at the same time through automatic feature learning [17]. The success of CNN-based CAD methods motivated us the development of deep learning neural networks for lung nodule classification [18].

Specifically, we investigate a new pulmonary nodule classification framework based upon a deep learning model, which is called the multiview residual selective kernel network (MRSKNet). This pulmonary nodule classification network is constructed with a hardcore RSK block, which is conceived from the selective kernel and residual learning [19, 20]. Moreover, the input images are composed of multiple views of three-dimensional (3-D) anatomical nodules associated with their handcrafted texture features. Initially, three anatomical planes (axial, coronal, and sagittal) of pulmonary nodules are extracted from 3-D CT image volumes. Subsequently, seven different texture features are computed for each individual plane, from which the most effective feature is incorporated into the input to boost the nodule classification accuracy. We will show that the proposed pulmonary nodule classification scheme exhibits robustness across various CT images, which provides great advantages over many state-of-the-art methods. This new CAD system is potential in facilitating the burden of physicians and radiologists for rigorous pulmonary nodule classification.

The major contributions of the current work are summarized as follows:

1. An efficient CAD system for pulmonary nodule benign-malignant classification in CT images based upon a deep learning model, called MRSKNet, is investigated.
2. A central processing unit, called RSK block, is developed that takes advantage of the selective kernel and residual learning for robust decomposition of lung nodules.
3. Three anatomical planes of lung nodules in the axial, coronal, and sagittal directions along with their handcrafted texture features are independently fed in each individual branch of the network.

The remainder of this paper is organized as follows. After the "Introduction" section, we conduct a brief review of existing pulmonary nodule classification methods. We then describe the proposed MRSKNet pertinent to feature extraction and model architecture. Subsequently, we introduce the acquired image data and demonstrate experimental results along with performance analysis. Finally, we discuss the characteristics of the developed classification scheme and draw the conclusion.

Related Work

A pulmonary nodule (or mass) is a small abnormal area that forms in the lung. Different types of nodules are difficult to differentiate through human's naked eyes by their size, shape, intensity, or texture. Developing a CAD system in facilitating the classification of pulmonary nodules is a necessity and current tendency. Recently, there has been an increasing number of studies that are dedicated to the automatic differentiation between benign and malignant nodules. For feature extraction, there are two major categories. While handcrafted-based characters are acquired essentially by mathematical image processing techniques [21], CNN-based features are computed inherently from the convolutional neural network architecture [8, 22]. For the essence of classification, existing approaches can be generally divided into machine learning and deep CNN methods.

Handcrafted Features with Machine Learning Methods

Traditionally, handcrafted features combined with machine learning procedures have been simultaneously exploited for pulmonary nodule classification [23]. Handcrafted features, which provide insightful information for lung nodule analysis, usually include shape, intensity, texture, geometry, and morphology characteristics [24]. An ensemble vector consisting of distinct handcrafted features is fed into a machine learning classifier for identification. For example, Dhara et al. [25] distinguished benign from malignant nodules using the support vector machine (SVM) associated with several

shape-based and texture-based features. Orozco et al. [26] employed the wavelet transform to extract nodule traits with the SVM for nodule classification. de Sousa Costa et al. [27] utilized the mean phylogenetic distance and taxonomic diversity index as texture features, which were classified by the genetic algorithm (GA) and SVM. Firmino et al. [28] computed the HOG for nodule quality extraction followed by a rule-based classifier and the SVM to eliminate false positives. de Carvalho Filho et al. [29] adopted the phylogenetic diversity to acquire several nodule attributes and applied the GA to select the best model.

Alternatively, Sasidhar et al. [30] introduced the GLCM to calculate the texture features of nodules and employed the SVM for classification. Li et al. [31] combined the Gabor, LBP, and GLCM characters to generate an effective feature vector, which was categorized by an improved random forest (RF) classifier. Wu et al. [32] proposed a pulmonary nodule classification scheme based upon the extraction of the shape, gray level, and texture features that were classified using the RF. Rodrigues et al. [33] presented a structural co-occurrence matrix-based mechanism to differentiate nodules using three well-known classifiers: multilayer perceptron (MLP), SVM, and k-nearest neighbor (KNN). Lee et al. [34] introduced a two-step supervised learning system, which combined a GA with the random subspace method and was tested via leave-one-out validation. Farahani et al. [35] computed statistical and morphological features from nodule candidates. An ensemble of three classifiers comprising MLP, KNN, and SVM was exploited for nodule classification.

Deep Learning Models

For the past decade, deep learning approaches have been extensively utilized in image processing and computer graphics [36]. Especially, using deep CNNs for image recognition has shown outstanding achievements, e.g., AlexNet [37], VGG [38], GoogLeNet [39], ResNet [19], and so forth. Among these network architectures, GoogLeNet and ResNet are specifically designed for the analysis of large-scale data, whereas the VGG network is regarded as a more general framework. Thanks to the success of the abovementioned networks, many researchers have exploited deep CNNs for medical image processing and analysis [40, 41]. A significant advantage of using deep CNNs is that the network architecture can learn discriminative features from input images directly [42]. As such, an end-to-end model is able to execute feature extraction and nodule classification at the same time.

Methods with 3-D CNN Models

Using 3-D raw CT images, Dai et al. [43] proposed a unique 3-D CNN called attribute-lung-nodule classification, which was derived from 3-D-DenseNet-40 for pulmonary nodule benign-malignant and image attribute categorization. Ren et al. [44] developed a manifold regularized classification deep neural network, which learned a general deep mapping schema from the original image space to a low-dimensional manifold, to analyze pulmonary nodules. Zhang et al. [9] presented a lung nodule classification framework based upon squeeze and excitation network and aggregated residual transformations, which combines the merits of ResNeXt [45] for feature reuse and SENet for feature recalibration. Liu et al. [46] introduced a multi-model ensemble learning architecture by taking advantage of a 3-D CNN, which consists of VGG [38], ResNet [19], and InceptionNet [47], for benign and malignant classification.

Methods with 2-D CNN Models

While volumetric image data provide 3-D spatial correlation of lung nodules, the network architectures are usually immense and sophisticated [48]. Accordingly, some researchers established more compact classification networks based upon two-dimensional (2-D) image data. For example, Lyu et al. [49] utilized three-level parallel ResNets with different convolution kernel sizes to extract multi-scale features of nodules, which resulted in a multi-level cross residual CNN. An et al. [50] described a two-stage CNN with the first CNN to handle ambiguous CT images followed by a simplified GoogLeNet to improve recognition accuracy. Su et al. [51] argued that a collection of 3-D views can be more informative for 3-D shape recognition and proposed a new CNN architecture that combines information from multiple views of a 3-D shape into a single and condensed shape descriptor, which offered better recognition efficiency.

Additionally, the application of anatomical planes has shown beneficial for lung nodule classification. Nibali et al. [52] employed three 2-D planar views instead of the full 3-D volume as input and adopted the ResNet architecture as the basis for exploring the effect of transfer learning, curriculum learning, and varying network depths on the classification accuracy. Sahu et al. [53] introduced a lightweight multi-section CNN model, which aggregates information from multiple cross-sections via a view pooling layer, for lung nodule categorization. Using multiple dilated convolutions instead of max-poolings to capture the scale variations, Al-Shabi et al. [54] described a gated-dilated (GD) network to classify nodules as benign or malignant. This GD architecture contains a context-aware sub-network, which inspects the input features and maneuvers the features to a suitable dilated convolution. The authors also proposed another

model called deep local–global network, which makes use of residual blocks for local feature extraction and non-local blocks for global feature extraction [55].

Handcrafted Features with Deep Learning Methods

There are some works using handcrafted features associated with deep learning networks for lung nodule investigation. For example, Xie et al. [56] suggested a multiview knowledge-based collaborative (KBC) deep model to distinguish malignant from benign nodules. The model learned nodule characteristics by decomposing a 3-D nodule into nine fixed views. For each view, a KBC submodel was constructed, where three types of appearance patches (overall appearance, heterogeneity in voxel intensity, and heterogeneity in shape) were designed to fine-tune three pre-trained ResNet-50 networks, respectively. By combining multiview knowledge-based collaborative learning, a semi-supervised adversarial classification model was further presented, which was trained with both labeled and unlabeled data to improve the accuracy of benign–malignant lung nodule classification [57]. Wang et al. [58] proposed the use of deep feature fusion of non-medical attribute training and handcrafted features to reduce the false positive error of pulmonary nodule categorization in chest radiography.

Deep Learning Features with Machine Learning Classifiers

Lastly, to analyze pulmonary nodules, several studies combined deep learning features with traditional machine learning classifiers. For example, Zhu et al. [59] presented a fully automated lung CT cancer diagnosis system called DeepLung, which involves nodule detection and classification. Specifically, a 3-D faster region with CNN (R-CNN) was designed for lung nodule detection, which conducts pixelwise multiscale learning with a U-Net-like encoder-decoder structure and 3-D dual path blocks. For nodule classification, a gradient boosting machine (GBM) [60] with 3-D dual patch network features was proposed. Zhang et al. [61] fused LBP-based texture features, HOG-based shape features, and 3-D deep dual path network (DPN) [62] features to characterize pulmonary nodules. The DPN integrates the advantages of ResNeXt [45] for feature reuse and DenseNet [63] for exploring new features. Subsequently, a GBM algorithm was employed to differentiate benign from malignant nodules.

Handcrafted feature methods usually convert an original image to a number of meaningful patterns based on mathematical formulas. However, the heterogeneity of the computed features is limited comparing with deep feature methods. While 3-D deep learning models provide abundant nodule information, they are complicated and require

excessive resources. Contrarily, 2-D CNN models are relatively efficient but it may acquire less nodule knowledge. As such, we investigate a competent 2-D deep learning architecture with multiple input branches, which strikes a good compromise between efficiency and accuracy. Moreover, handcrafted texture feature maps are incorporated into the network model to generate improved abstract features for better discrimination.

Proposed Methodology

The proposed pulmonary nodule classification framework, which refers to as MRSKNet, consists of three major phases: (1) multiple views, (2) texture features, and (3) model architecture. The code is publicly accessible through. https://github.com/yantong0116/Lung-Nodule-Classification.

Multiple Views

To acquire 3-D spatial information while reducing the complexity, Su et al. [51] presented an elegant CNN architecture, which has stimulated us the manipulation of volumetric image data for pulmonary nodule classification. Different from Su et al. [51], we extract anatomical planes from each 3-D nodule CT image volume as 2-D multiple views. Commonly, there are three frequently used orthogonal planes: axial, coronal, and sagittal. The axial plane divides the body into upper and lower portions, the coronal plane into front and back portions, and the sagittal plane into left and right portions. As volumetric CT acquisition has become routine and accessible, viewing the anatomy and pathology of organs in these three planes is beneficial when evaluating the disease.

In our approach, a solid set of volumetric CT images is constructed by stacking a series of 2-D slices for each nodule. Subsequently, we compute the 3-D centroid of the nodule in the image volume, from which the axial, coronal, and sagittal planes are extracted. Through this preprocessing procedure, we not only maintain abundant spatial information of the nodule image data but also reduce the complexity of the classification model and the corresponding training time.

Texture Features

To understand the importance of handcrafted features and to incorporate effective texture features into our classification framework, we explore three particular classes of texture features, which are the GLCM, gray-level run length matrix (GLRLM), and Tamura attributes. Suggested by the study [64], we investigate seven candidates that provide stronger distinction to reflect the characteristics of different

anatomical structures, which are entropy (ENT), homogeneity (HOM), gray level non-uniformity (GLN), run-length non-uniformity (RLN), run percentage (RP), short run emphasis (SRE), and coarseness.

GLCM Attributes

GLCM [65] is interpreted as the statistical distribution of observed intensity combinations at specified positions relative to each other in an image. Unlike the original GLCM spanning on the entire image, the philosophy underlining our perception is a filter-like manner so that the computation is restricted to a local window and duplicated through the whole image, which results in a feature map. One essential characteristic is that with respect to the pixel correlation the GLCM attribute varies rapidly in fine-texture regions and slowly in coarse texture areas. Mathematically, a GLCM is a matrix whose size is equal to the gray level number of the image intensity. The value of the matrix $C(i,j|x,y)$ is the relative frequency of two pixels separated by a distance (Δ_x, Δ_y), which is defined by d and θ, in a local window (subimage) centered at (x,y). The co-occurrence is counted based upon their intensity relationship: from intensity "i" to intensity "j." Given an $W_x \times W_y$ subimage I_s centered at (x,y) in an anatomical nodule image I_n with an intensity range $[0, L-1]$, the matrix size is $L \times L$. Among the 28 attainable texture features, we are particularly interested in ENT,

$$\text{ENT}(x,y) = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i,j|x,y)(-\ln(P(i,j|x,y))) \tag{1}$$

and HOM,

$$\text{HOM}(x,y) = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{P(i,j|x,y)}{1+(i-j)^2} \tag{2}$$

where $P(i,j|x,y)$ is the normalized matrix of $C(i,j|x,y)$.

GLRLM Attributes

GLRLM is a more objective texture analysis method so that the gray level run length is longer in a coarser texture region and vice versa [66]. Similar to GLCM, GLRLM converts an image into a particular matrix. However, a run length matrix is defined as the number of runs with a pixel of a specific gray level and its run length in a certain direction. In our approach, the GLRLM element $p(i,j|x,y)$ indicates the number j of an element with the intensity i in a particular direction in a subimage centered at (x,y). Four distinct texture features are computed and exploited in this study:

1. GLN

$$\text{GLN}(x,y) = \frac{1}{n_r(x,y)} \sum_{i=1}^{L}$$
$$\left(\sum_{j=1}^{R} p(i,j|x,y) \right)^2 = \frac{1}{n_r(x,y)} \sum_{i=1}^{L} p_g^2(i|x,y) \tag{3}$$

where $n_r(x,y)$ is the ensemble sum of $p(i,j|x,y)$ at (x,y), $p_g(i|x,y)$ is the gray level run number vector, and R is the maximum run length.

2. RLN

$$\text{RLN}(x,y) = \frac{1}{n_r(x,y)} \sum_{j=1}^{R}$$
$$\left(\sum_{i=1}^{L} p(i,j|x,y) \right)^2 = \frac{1}{n_r(x,y)} \sum_{j=1}^{R} p_r^2(j|x,y) \tag{4}$$

where $p_r(j|x,y)$ is the run length run number vector.

3. RP

$$\text{RP}(x,y) = \frac{n_r(x,y)}{n_p(x,y)} \tag{5}$$

where $n_p(x,y)$ indicates the number of all elements being considered in the subimage centered at (x,y).

4. SRE

$$\text{SRE}(x,y) = \frac{1}{n_r(x,y)} \sum_{i=1}^{L} \sum_{j=1}^{R} \frac{p(i,j|x,y)}{j^2} = \frac{1}{n_r(x,y)} \sum_{j=1}^{R} \frac{P_r(j|x,y)}{j^2} \tag{6}$$

Tamura Attributes

Tamura features [67] were developed according to the human visual and psychological perception. There are six Tamura texture features and the first three features are associated with human vision, which are more distinctive than the last three attributes. We are particularly interested in the coarseness (COA) feature, which is utilized to measure the grain size of image attributes. A larger COA value represents a ruder image. For more details, please refer to the papers [64, 67]. Similar to the GLCM and GLRLM scenarios, a local COA computation manner in a subimage is exploited.

Model Architecture

The core block (RSK block) of our proposed network architecture is first introduced, followed by describing the automatic lung nodule classification model.

RSK Block

Lung nodules generally have different scales and various structures. A large filter is too difficult to thoroughly extract

subtle features for small nodules. On the other hand, a small filter's window is too narrow to capture the overall information of large nodules. To handle the extraction of multiscale features from diverse nodules, an RSK block with two different scales of filters is proposed.

Our RSK block is derived from the "SK convolution" in SKNet [20]. Rather than using a normal 3×3 convolution filter, the SK convolution has been employed in a bottleneck residual block of ResNet [19]. We integrate the design concepts of the residual learning and selective kernel network to develop the RSK block, which consists of three major operations: split, fuse, and select, as illustrated in Fig. 1. In the basic architecture, the split phase contains two parallel convolution operations with different kernel sizes. One more branch of the identity mapping is added to the network so that there are three branches in the beginning of the RSK block: (1) Conv3, (2) Conv5, and (3) identity mapping. The purpose of including the identity mapping is to provide the original information for straightness processing. Accordingly, the architecture can learn the weights of all features and adjust the weights by the corresponding features at last. It enables neurons to adaptively accommodate their receptive field (RF) sizes based upon multiple scales of input information during the inference.

In the first phase of split, the input image (or feature map) is denoted as $X \in R^{H \times W \times C}$, where H, W, and C are the height, width, and number of channels in X, respectively. After two parallel convolution (Conv3 and Conv5) procedures and one identity mapping operation, the input quantity X is transformed into $U_i \in R^{H' \times W' \times C'}$, where H', W', and C' represent the height, width, and number of channels after the transformations, and the subscript i indicates the corresponding branch. The convolution is similar to the residual bottleneck block, where the Conv3 procedure comprises a 1×1 convolution, a 3×3 convolution, and a 1×1 convolution,
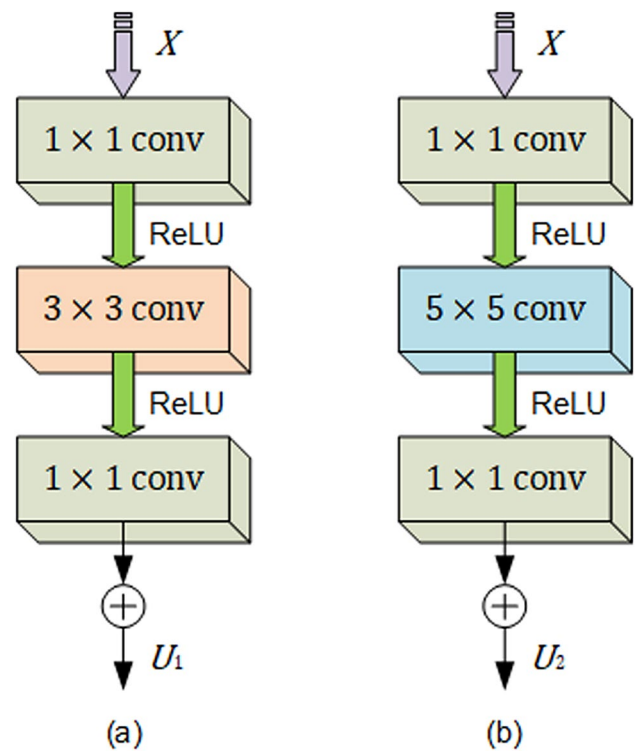
Fig. 2 Bottleneck computation of the proposed convolution procedures. **a** Conv3; **b** Conv5

interconnected by the ReLU activation function, as demonstrated in Fig. 2(a). While the first two convolutions are followed by the same batch normalization and ReLU function, the last convolution is followed by the batch normalization only. The channel number in the first 1×1 convolution is halved to reduce the computational complexity and increased to the number of the output channels C' in the last 1×1 convolution. Similarly, the Conv5 procedure is constructed by

Fig. 1 Illustration of the proposed RSK block

$\oplus$: element-wise summation　　　$\otimes$: element-wise product
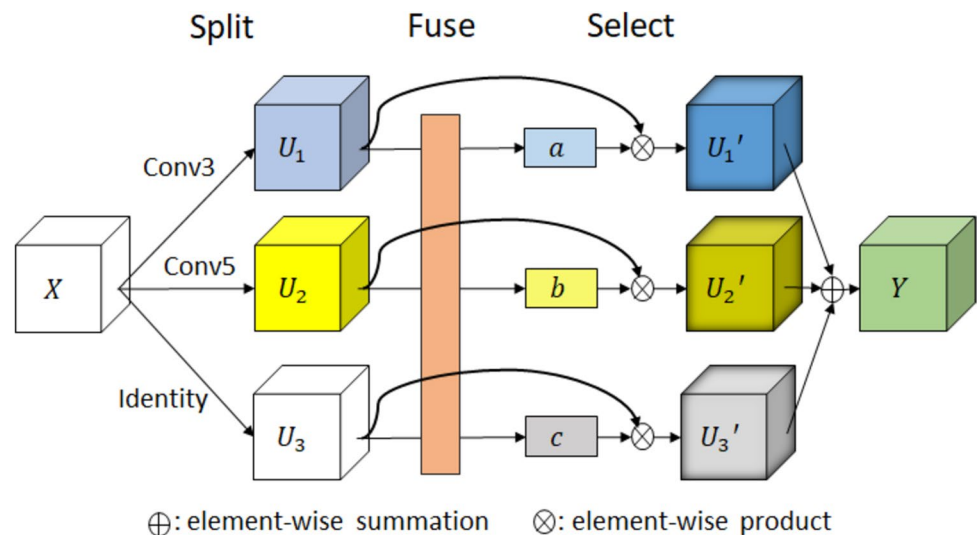
changing the kernel size of the middle convolution to 5×5, as shown in Fig. 2(b). For the identity mapping process, if the input and output channels are different, a 1×1 convolution followed by the batch normalization is employed.

Figure 3 depicts the detailed architecture of the second phase fuse in the RSK block. The design is to empower neurons to adaptively regulate their RF sizes by using dynamic gates to consolidate the information from the three upstream branches, which contain different feature maps, into the next layer. To achieve this, the ReLU activation function is first applied to the feature maps, which are fused via the element-wise summation with

$$\widetilde{U}(h, w, c) = \sum_{i=0}^{N} U_i(h, w, c) \tag{7}$$

where N represents the number of branches, which is 3 in our scenario. The global average pooling (GAP) is then employed to acquire the global information and flatten the feature map to generate the channel-wise statistical score:

$$S(c) = \text{GAP}\left(\widetilde{U}(h, w, c)\right) = \frac{1}{H\prime \times W\prime} \sum_{h=1}^{H\prime} \sum_{w=1}^{W\prime} \widetilde{U}(h, w, c) \tag{8}$$

where $S \in R^{C\prime}$ is the output after the GAP operation. Subsequently, a fully connected (FC) layer is executed to produce a compact feature vector,

$$Z(d) = \text{FC}(S(c)) \tag{9}$$

where $Z \in R^{d \times 1}$ is the output after the FC1 layer and d denotes the adjustable number of channels. A reduction ratio r is introduced to determine the value of d with

$$d = C\prime / r \tag{10}$$

where $C\prime$ is the number of the output channels. A second FC layer, FC2, further divides Z into three branches for subsequent processing.

The last phase select exploits the soft attention strategy across different channels to adaptively choose multiscale features. Specifically, the Softmax function is administered to the channel-wise digits to acquire the soft attention weights a, b, and c corresponding to the three feature maps using

$$a = \frac{e^{A_s Z}}{e^{A_s Z} + e^{B_s Z} + e^{C_s Z}} \tag{11}$$

$$b = \frac{e^{B_s Z}}{e^{A_s Z} + e^{B_s Z} + e^{C_s Z}} \tag{12}$$

$$c = \frac{e^{C_s Z}}{e^{A_s Z} + e^{B_s Z} + e^{C_s Z}} \tag{13}$$

where $A_s, B_s, C_s \in R^{C\prime \times d}$ are the outputs from the FC2 layer and $a, b, c \in R^{C\prime \times 1}$ are the soft attention weights for the corresponding feature maps $U_i, i = 1, 2, 3$. To perform element-wise multiplication, we reshape the dimension of the soft attention weights $a, b, c \in R^{H\prime \times W\prime \times C\prime}$, which is identical to the size of the feature maps U_i. The resultant feature map Y is computed using the element-wise summation of the soft attention weights multiplied by the three corresponding feature maps:

$$Y = U_1\prime + U_2\prime + U_3\prime = a \otimes U_1 + b \otimes U_2 + \otimes U_3 \tag{14}$$

where $\otimes$ is the element-wise multiplication operator and $Y \in R^{H\prime \times W\prime \times C\prime}$ is the final product of the RSK block.

MRSKNet

A reliable classifier should recognize different objects based upon both semantic and comprehensive features. A small filter is able to resolve more intricate attributes but it is unable to catch semantic information. Contrarily, a large filter is capable of capturing semantic knowledge but it would lose detailed features. As such, the proposed deep learning architecture adaptively extracts multiscale features and vigorously integrates different RF sizes into a robust classifier to well detect the characteristics of pulmonary nodules.

Our MRSKNet is conceived from the integration of the multiview CNN [51], residual learning [52], and SKNet [20]. The intention is to consider the anatomical planes of medical image volumes as the multiple views of 3-D images for the categorization of pulmonary nodules. Not only do the multiview input preserve the spatial information of volumetric nodule images, but they also decrease the complexity of input image data, leading to the acceleration of model training. Figure 4 illustrates the network architecture of our proposed MRSKNet that exhibits three branches correlating with the three different anatomical planes as described earlier. For each branch, there are one intensity nodule CT image and one corresponding handcrafted feature image for concatenation input.
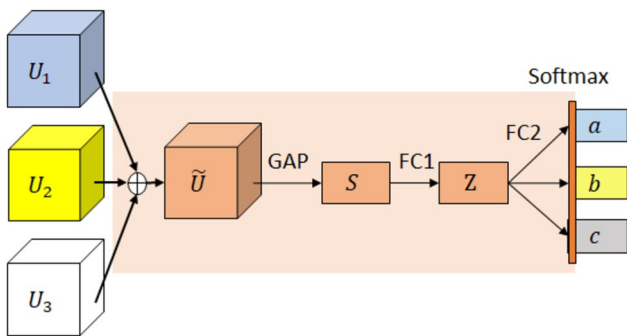

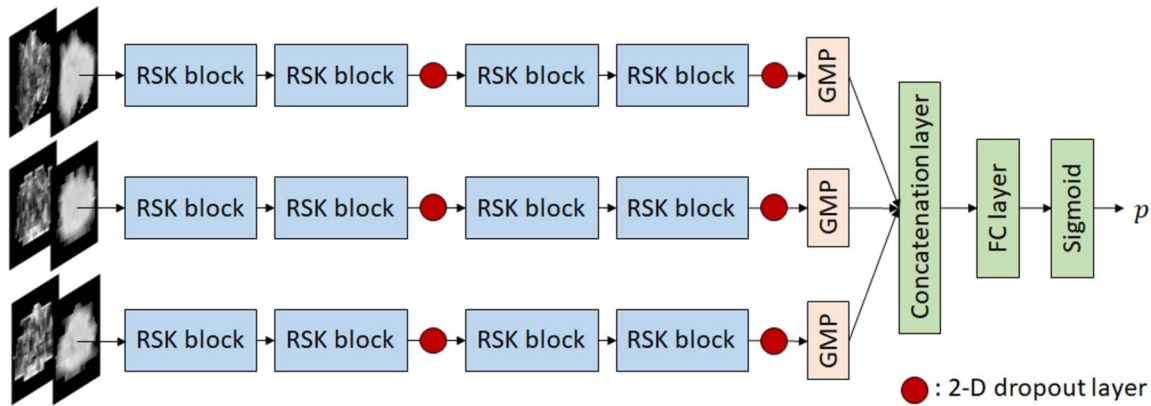**Fig. 3** Illustration of the fuse process in the RSK block

Fig. 4 Architecture illustration of the proposed MRSKNet for pulmonary nodule classification

In each branch, the main structure contains four sequential RSK blocks. Two 2-D dropout layers are deliberately included after the second and fourth RSK blocks to avoid overfitting. As adjacent pixels within feature maps are strongly correlated, the 2-D dropout layers, which randomly zero out the entire channels, help to promote independence between the feature maps. Afterward, the GMP operation is employed to flatten these features followed by the concatenation layer to integrate the three branches before entering the last FC layer. Through the Sigmoid function, the system eventually predicts the likelihood of the malignancy *p*.

Experimental Results

Data Acquisition

To realize the capability of the proposed pulmonary nodule classification framework, the public benchmark Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset [68], which contains 1018 instances and 1010 patients, was utilized. Specifically, the image dataset was acquired from The Cancer Imaging Archive (TCIA) [69] and the LIDC nodule size report from the Vision and Image Analysis Group of Cornell University [70] was exploited. Each instance provides a series of thoracic images and one XML file, which stores annotation and related information such as the UID number, nodule position, and likelihood of malignancy judged by four radiologists. Those nodules that were found by at least three radiologists were adopted for the experiments. The nodules were labeled the malignancy level from 1 to 5 by every radiologist and the median of all malignancy levels was regarded as the final score for each instance. A nodule with a score less than 3 was considered benign, whereas a nodule with a score larger than 3 was considered malignant. Nodules with a score equal to 3 were discarded as they were undetermined.

In short, 698 subjects with 1386 nodules, which include 447 benign nodules, 430 malignant nodules, and 509 indeterminate nodules, were qualified. For binary classification, only 877 benign and malignant nodules were employed.

Data Preprocessing

In order to extract the nodule from an entire pulmonary CT image for the purpose of categorization, the raw image data were preprocessed. We first readjusted the intensity of the acquired CT images by setting the out-of-scan pixels from -2000 to 0 and rescaling the pixel values to the HU so that the converted image I_c was rescaled to a new range of $[-1024, 3071]$. As the acquired pulmonary images exhibited various resolutions, I_c was resampled to a specified isotropic resolution of $1 \times 1 \times 1 \text{mm}^3$ per voxel with the cubic spline interpolation and renormalized to [0, 1] for better training and recognition. Subsequently, the normalized image was multiplied by the nodule mask, which was generated according to the nodule locations in the corresponding XML file, to extract the nodule image as illustrated in Fig. 5. To capture the essence of the nodule, a smaller image with 32×32 was utilized. Lastly, the whole nodule was stored in a $32 \times 32 \times 32$ image volume with respect to the maximum cross-section plane. Three anatomical planes in the axial, coronal, and sagittal directions, with each stored in a 32×32 image, were produced according to the centroid of the nodule and denoted as I_n for feeding the pulmonary nodule classification network.

Data Augmentation and Feature Computation

Since deep learning-based networks require abundant data for effective training, the diversity and quantity of image data plays an important role in achieving excellent performance. However, as in our scenario, one major drawback of applying deep learning-based networks to medical images is
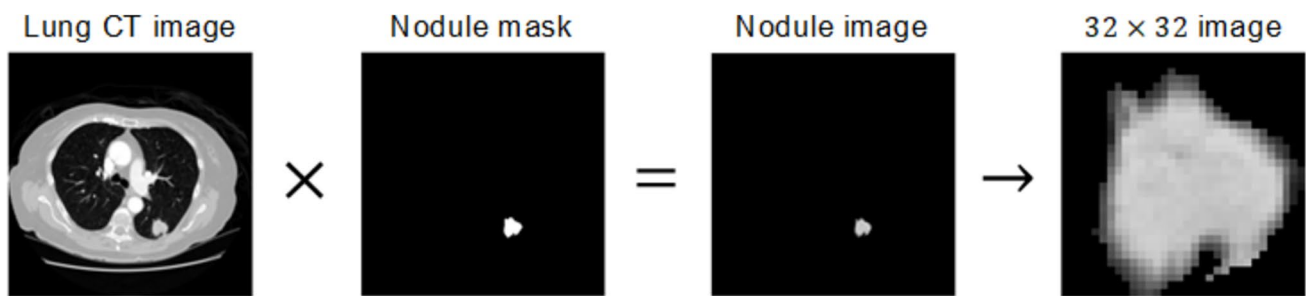
Fig. 5 Pulmonary nodule extraction process

the limited amount of accessible data. After screening and preprocessing of the pulmonary CT images in the LIDC-IDRI database, there were only 877 nodules for the experiments. To increase the diversity and quantity of the image data for training, a series of augmentation processes were executed. Each image was first generated rotated images in seven angles (i.e., 45°, 90°, 135°, 180°, 225°, 270°, and 315°), which resulted in eight time the amount of original image data. All original images were then horizontally flipped followed by the same seven angle rotation operations. Eventually, the amount of total pulmonary nodules was increased by sixteen times. For each nodule image, seven corresponding texture feature images of ENT, HOM, GLN, RLN, RP, SRE, and COA were computed for concatenation input as illustrated in Fig. 6. As they are correlated with the nodule CT image according to the mathematical formulas, no prior medical knowledge is incorporated into the classification model.

Implementation Details

Each branch of the proposed MRSKNet started with a $32 \times 32 \times 2$ concatenated pulmonary nodule image. While the reduction ratio r in Eq. (10) was a constant of 8, the values of $H\prime \times W\prime \times C\prime$ were set to $16 \times 16 \times 32$ and $8 \times 8 \times 64$ for the first two RSK blocks and last two RSK blocks, respectively. Table 1 presents the essential parameter setting of the proposed model for pulmonary nodule classification. As the input image size was small (32×32), we constructed a relatively miniature network, which provides higher efficiency than the classical ResNet and SKNet for our specific task. Our MRSKNet was implemented with Pytorch, which is one of the most popular Python environments for deep learning-based development. All experiments were executed on an Intel® Core i7-10700F CPU @ 2.90 GHz ×16 workstation running 64-bit Linux Ubuntu 18.04.5 LTS with 128 GB RAM. The machine was equipped with a NVIDIA GeForce RTX 3090 GPU of 24 GB RAM for parallel computation.

For the feature extraction, the values of L in Eq. (1) and R in Eq. (3) were all set to 32. The MRSKNet was trained with a batch size of 256 for 50 epochs. The initial learning rate was set to 1×10^{-4} for faster gradient descent and declined to 1×10^{-5} after a half of the epochs for preventing overshooting. The Adam optimizer was utilized for finding the best system parameters with the default hyperparameters β_1 and β_2 equal to 0.9 and 0.999, respectively. All network weights were initialized with the Kaiming uniform distribution $\mu\left(-\sqrt{k}, \sqrt{k}\right)$ with $k = 1/\tau$, where τ is the number of input channels multiplied by the kernel size in a convolution layer or the number of the input channels in a FC layer. The employment of the Kaiming function was beneficial for our network, which contains the ReLU activation function to avoid the gradient vanishing and exploding problems. Finally, the binary cross entropy loss function was adopted for this pulmonary nodule classification scheme, based upon which best model within 50 epochs was acquired.

Evaluation Protocols

Ten-fold cross validation was exploited to evaluate the proposed MRSKNet, i.e., the ratio of the training to validation images was set to 9:1. Because there were 447 benign nodules and 430 malignant nodules, the amount of benign nodules accounted for approximately 51% in each fold with random selection. Table 2 details the numbers of benign and malignant nodules for training and validation after data augmentation in each fold. The average evaluation scores of the ten-fold experiments were regarded as the ultimate performance of our pulmonary nodule classification framework.

In addition to the assessment of the receiver operating characteristic (ROC) and area under ROC (AUC), a number of popular evaluation metrics including accuracy, recall, specificity, and precision were employed, which are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

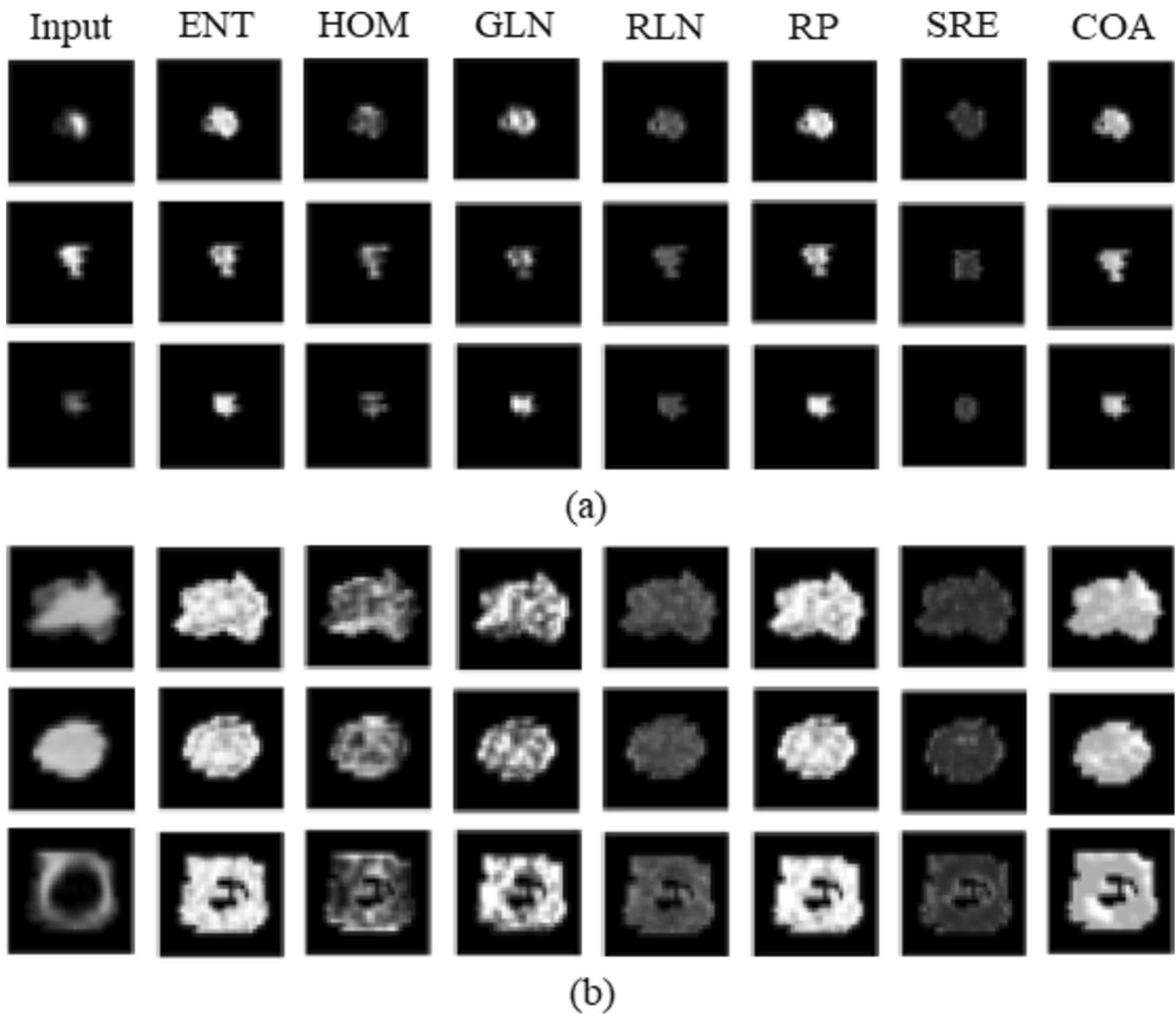| Input	ENT	HOM	GLN	RLN	RP	SRE	COA

(a)

(b)

Fig. 6 Illustration of handcrafted texture feature images for pulmonary nodule classification. **a** Benign nodules. **b** Malignant nodules

$$\text{Recall} = \frac{TP}{TP + FN} \quad (16)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (17)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (18)$$

**Table 1** Essential parameter setting of the proposed MRSKNet

Layer	Output	Property
RSK block	$16 \times 16 \times 32$	Stride $= 2, r = 8$
RSK block	$16 \times 16 \times 32$	Stride $= 1, r = 8$
Dropout		Dropout rate $= 0.2$
RSK block	$8 \times 8 \times 64$	Stride $= 2, r = 8$
RSK block	$8 \times 8 \times 64$	Stride $= 1, r = 8$
Dropout		Dropout rate $= 0.2$
Ultimate	1×1	1×1 GMP, 1D FC layer, Sigmoid

Table 2 Image numbers for ten-fold cross validation on the LIDC-IDRI database

Fold	Phase	Benign	Malignant	Total
1–7	*Training	6432	6192	12,624
	Validation	45	43	88
8–10	*Training	6448	6192	12,640
	Validation	44	43	87

*Training with data augmentation

Fig. 7 ROC curves based upon ten-fold cross validation on the LIDC-IDRI database. **a** Input with the original CT images only. **b** Input with additional HOM texture feature map concatenation

(a)

(b)
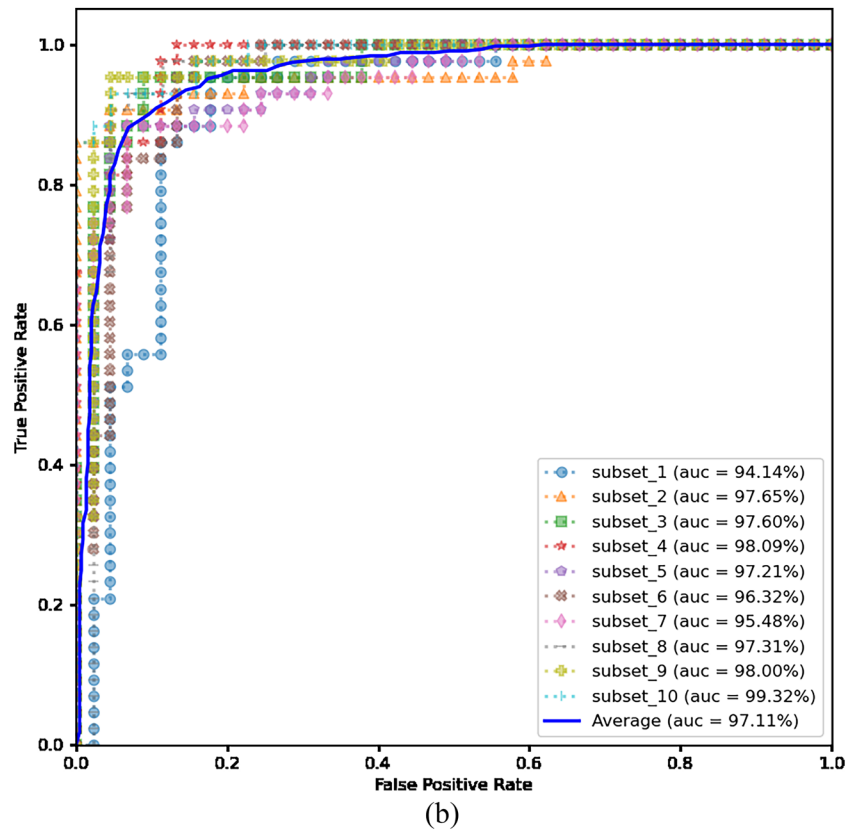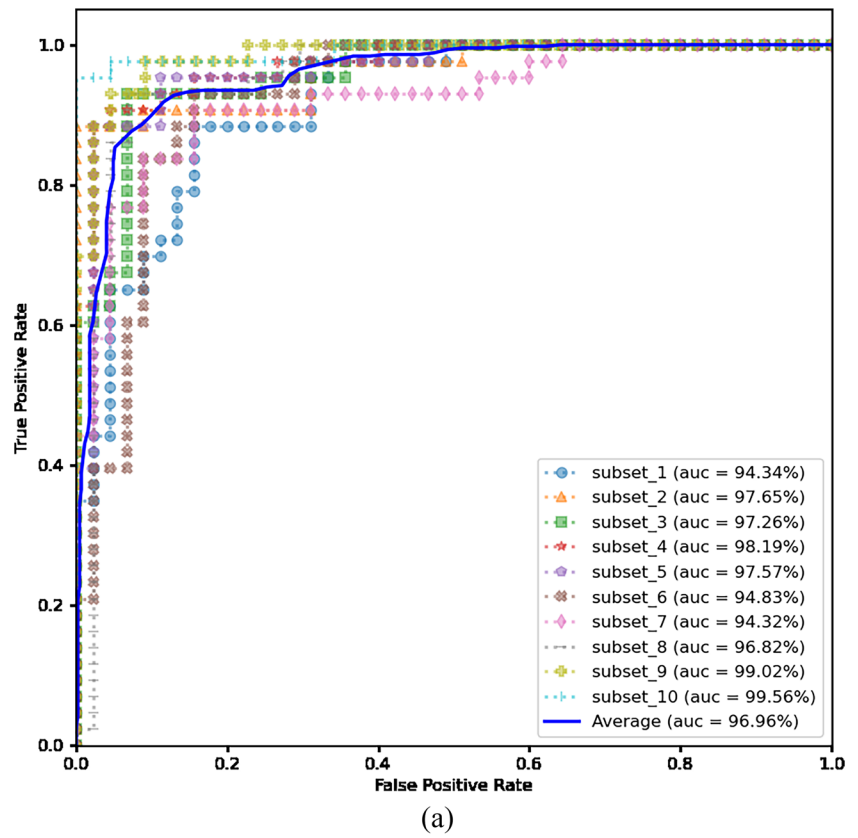
where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives. Among the described evaluation metrics, AUC, accuracy, and recall provides more distinctive judgement for our nodule classification experiments.

Performance Evaluation

Three 32×32 nodule CT images in the axial, coronal, and sagittal directions were delivered to the proposed MRSKNet for categorization. To reinforce the performance with diversified input, seven texture feature maps derived from the CT images were computed as illustrated in Fig. 6. Each texture feature map, say HOM, was combined with its original intensity image as the intake $(32 \times 32 \times 2)$ of the nodule classification model as depicted in Fig. 4. For the scenario of the sole intensity CT image input, the parameters of the MRSKNet were identical to Table 1, except the dropout rate, which was set to 0.1 instead of 0.2. All in all, there were eight different input variations: one single CT image and seven integration images in each branch of the network.

Figure 7(a) depicts the ROC curves of using the original nodule CT image planes in three different views for the input in light of ten-fold cross validation, where the average AUC attained 0.9696. Table 3 summarizes the performance measures with respect to five various evaluation metrics for the eight different input scenarios. The experiments feeding only the intensity CT images produced the highest average precision score of 0.9395, whereas the experiments of the concatenation scenarios with the ENT texture feature maps generated the greatest average specificity score of 0.9365. It was the combination of the intensity images and the HOM texture feature maps that achieved the best average AUC, accuracy, and recall scores, which were 0.9711, 0.9366, and 0.9556, respectively. An apparent improvement was the 2.10% increment of recall using the additional HOM feature maps, which reduced FN and increased the identified number of malignant nodules. Clinically speaking, this is beneficial for early detection of lung cancer. The ROC curves

Table 3 Performance evaluation of different input settings based upon ten-fold cross validation on the LIDC-IDRI database

Input	AUC	Accuracy	Recall	Specificity	Precision
CT	0.9696	0.9349	0.9346	0.9356	**0.9395**
+ENT	0.9669	0.9242	0.9139	**0.9365**	0.9314
+HOM	**0.9711**	**0.9366**	**0.9556**	0.9177	0.9222
+GLN	0.9675	0.9159	0.9209	0.9144	0.9114
+RLN	0.9691	0.9324	0.9498	0.9150	0.9200
+RP	0.9661	0.9254	0.9305	0.9187	0.9228
+SRE	0.9696	0.9349	0.9405	0.9246	0.9253
+COA	0.9692	0.9323	0.9320	0.9333	0.9320

Table 4 Performance analysis using different RSK block and layer channel settings on the LIDC-IDRI database

RSK block number	Layer channel	AUC	Accuracy	Recall
2	(32, 32)	0.9679	0.9269	0.9297
4	(32, 32), (64, 64)	**0.9696**	**0.9349**	0.9346
6	(32, 32), (64, 64), (128, 128)	0.9602	0.9288	**0.9356**

of the experiments with the additional HOM feature maps based upon ten-fold cross validation are shown in Fig. 7(b).

Ablation Study

To better understand the characteristics of our proposed MRSKNet for pulmonary nodule classification, various parameter and structure settings were investigated. We first diversified the numbers of the RSK block and channels per layer while maintaining other parameter settings in Table 1. When the quantity of the RSK block was increased, the channel number was grown from 32 to 128 as presented in Table 4, where the setting of four RSK blocks gained the best overall performance. Subsequently, four distinct dropout rates were studied to realize their sensitivity to our classification model. Table 5 reports the performance evaluation scores with respect to varied dropout rates, where the left values indicate the scenario with the original CT image only and the right values with the additional HOM texture feature map concatenation.

As presented in Eq. (10), the reduction ratio r plays an essential role in determining the output channel number after the FC1 layer. The influence of the reduction ratio with different values on the classification efficiency was also examined as presented in Table 6, where the setting with $r = 8$ accomplished the best performance in both input scenarios. Experiments using a reduced model with only one single branch were conducted. The nodule classification results using the axial plane (AP), coronal plane (CP), or sagittal plane (SP) as input are summarized in Table 7, which indicated worse performance than the use of three branches simultaneously. To further demonstrate the advantage of our MRSKNet, two baseline network models were

Table 5 Sensitivity analysis using different dropout rates on the LIDC-IDRI database

Dropout rate	AUC	Accuracy	Recall
0	0.9615 | 0.9657	0.9327 | 0.9242	0.9334 | 0.9523
0.1	**0.9696** | 0.9693	**0.9349** | 0.9350	**0.9346** | 0.9596
0.2	0.9683 | **0.9711**	0.9293 | 0.9366	0.9230 | 0.9556
0.3	0.9661 | 0.9701	0.9220 | **0.9396**	0.9122 | 0.9292

CT \| CT+HOM

Table 6 Sensitivity analysis using different reduction ratio values on the LIDC-IDRI database

Ratio r	AUC	Accuracy	Recall
2	0.9694 | 0.9653	0.9247 | 0.9167	0.9119 | 0.8963
4	0.9681 | 0.9695	0.9243 | 0.9299	0.9210 | 0.9338
8	**0.9696** | 0.9711	**0.9349** | **0.9366**	**0.9346** | **0.9556**
16	0.9686 | **0.9729**	0.9259 | 0.9290	0.9266 | 0.9319

CT \| CT + HOM

exploited for comparison. One model consists of four bottleneck residual blocks [19], which is referred to as Basic ResNet. The other comprises four bottleneck blocks with the SK convolution [20], which is called Basic SKNet. For fair comparison, these two baseline models have the same network architecture as the proposed MRSKNet except that the RSK block is replaced with the bottleneck residual block and the SK block, respectively. Table 7 suggests that our proposed MRSKNet outperformed these two networks with the highest scores according to the three achievement measures.

Performance Comparison

For completeness, the proposed pulmonary nodule classification framework based upon MRSKNet was compared to a number of state-of-the-art methods, which adopted the same database for performance validation. Table 8 compares our MRSKNet with the contemporary techniques that exploited deep learning strategies for pulmonary nodule classification on the LIDC-IDRI database. It was obvious that the greatest scores of AUC with 0.9711, recall with 0.9556, and precision with 0.9222 were achieved by the proposed network model. Our accuracy score (0.9366) was the second largest among all competitive methods. One of our advantages was the high recall rate, which was 3.35% larger than the second best method, thanks to the introduction of the HOM feature map concatenation. We struck a better compromise between recall and specificity than other models. Overall, our MRSKNet produced more accurate pulmonary nodule classification judgement comparing with the state-of-the-art methods.

Table 7 Performance comparison between different classification networks under identical configuration on the LIDC-IDRI database

Model	AUC	Accuracy	Recall
AP-MRSKNet	0.9566	0.9223	0.9302
CP-MRSKNet	0.9441	0.9104	0.9289
SP-MRSKNet	0.9357	0.9046	0.9212
Basic ResNet	0.9689	0.9227	0.9276
Basic SKNet	0.9689	0.9317	0.9374
MRSKNet	**0.9711**	**0.9366**	**0.9556**

Discussion

Stimulated by the success of deep CNN models in medical image processing applications, we have proposed a new pulmonary nodule classification system in CT images, which is called MRSKNet. Rather than employing the entire image volume for input, the strategy of feeding the network with three different views in the axial, coronal, and sagittal planes was leveraged in this study to achieve cost-effective design. Although volumetric image data can provide more contexture information than planar images, more complicated network architectures with larger and deeper layers are usually necessitated. Moreover, it is highly possible that redundant spatial correlation links appear in the colossal network structure, which hampers the learning efficacy. As can be observed from Table 8 that those methods [9, 43, 44, 46] with 3-D image input did not present contexture advantages over other approaches according to the accuracy measures.

Multiscale resolution has been extensively exploited to extract contexture information at various scales to enrich the processing mechanism. To take advantage of multiscale manipulation, the core network (RSK block) of our MRSKNet was designed with three different analytic pathways. In addition to various tracks of scope resolution, the RSK block comprises three distinct phases of split, fuse, and select to integrate the multiscale contexture into a collective unit for subsequent processing as illustrated in Fig. 1. Since the dimension of most nodules was between 3 and 26 mm, our construction with two different convolution entities of Conv3 and Conv5 was sufficed for capturing the pulmonary nodule characteristics in 32×32 images for categorization. However, if the input image size is increased, say 64×64 or 128×128, larger convolution windows, e.g., Conv7, with longer bottleneck structures may be required to effectively reflect the contexture at different scales.

While deep learning schemes automatically provide a wide variety of perceptive features, the majority are beyond the common interpretation of the human beings and the uniqueness of each individual feature map is deteriorating as the network advances deeper. As described earlier, handcrafted texture features are capable of extracting the essence of image structures for CAD. To understand the impact of critical handcrafted texture features associated with deep learning processing on pulmonary nodule classification, we investigated seven different texture features. Specifically, we combined the intensity CT images with the corresponding texture feature maps as the input of the MRSKNet in an attempt to boost the classification performance. Interestingly, the CT images along with the HOM feature maps outperformed other input scenarios as presented in Table 3. This may be due to the fact that the HOM map computes the nodule homogeneity attribute that exhibits more intensity variations and complements the original intensity image for

Table 8 Performance comparison between MRSKNet and contemporary methods based upon deep learning classification on the LIDC-IDRI database

Study	Year	Nodule number (benign, malignant)	AUC	Accuracy	Recall	Specificity	Precision
Nibali et al	2017	831 (421, 410)	0.9459	0.8990	0.9107	0.8864	0.8935
*Zhu et al	2018	1004 (450, 554)	-	0.9044	-	-	-
*Dai et al	2018	1011	0.9690	0.9147	0.9126	0.9167	-
Xia et al	2019a	1945 (1301, 644)	0.9570	0.9160	0.8652	0.9400	-
Xia et al	2019b	1945 (1301, 644)	0.9581	0.9253	0.8494	**0.9628**	-
*Fu	2019	1186 (650, 536)	-	0.8993	0.8334	0.9105	-
Al-Shabi et al	2019b	848 (442, 406)	0.9514	0.9257	0.9221	-	0.9185
Al-Shabi et al	2019a	848 (442, 406)	0.9562	0.8864	0.8866	-	0.8738
Zhang et al	2019	1004 (450, 554)	0.9687	**0.9378**	-	-	-
*Zhang et al	2020	1004 (450, 554)	0.9563	0.9167	-	-	-
*Ren et al	2020	1226 (795, 431)	-	0.9000	0.8100	0.9500	-
*Liu et al	2020	1268 (863, 405)	0.9390	0.9060	0.8370	0.9390	-
Lyu et al	2020	-	0.9705	0.9219	0.9210	0.9150	-
MRSKNet	2022	877 (447, 430)	**0.9711**	0.9366	**0.9556**	0.9177	**0.9222**

*Adopted 3-D image data as input

-Unavailable

better discrimination between benignancy and malignancy, as illustrated in Fig. 6. The incorporation of the HOM feature can also be regarded as perceiving pulmonary nodules with a more advanced imaging modality other than CT. Nevertheless, only seven out of possibly hundred texture features were studied for this particular classification task. Exhaustive searching of appropriate image texture features regarding the concatenation amount and computation manner, and other potential image processing applications is beyond the scope of the current work, which is worth investigating in the future.

In order to justify the appropriateness of the major components in the proposed MRSKNet, we systematically performed ablation studies. As summarized in Table 4, the design of two sets of two consecutive RSK blocks followed by a dropout layer in each branch achieved the best ensemble performance. This implies that deeper network architectures are prone to deterioration than network models with a moderate size. In Tables 5 and 6, as the greatest scores in each individual evaluation measure were accomplished by different settings, the utilized dropout rates and reduction values were primarily determined according to the AUC metric, which, in our opinion, is the most critical judgement indicator for this binary classification task. To demonstrate the effectiveness of the proposed RSK block, two competitive networks with the same architecture as the MRSKNet but replaced with different core blocks are compared in Table 7. It was obvious that the best evaluation scores were all produced by our proposed framework.

One practical issue is that different nodule numbers of benignancy and malignancy were adopted by distinct studies on the LIDC-IDRI database as presented in Table 8. Diverse

folds of cross validation were employed and many works generated inadequate evaluation measures, which hampered more perceptive comparison. Nevertheless, not only did we provide sufficient assessment scores, but our proposed scheme also surpassed the competing methods in overall AUC, accuracy, and precision evaluation. Without other available and reliable datasets for further cross-dataset improvement and evaluation, our classification ability is restricted on the LIDC-IDRI database as many competing methods in Table 8. Lastly, we extracted the pulmonary nodule from the original CT image in a 32×32 subimage for efficient computation, which could exclude important perinodular information and reduce the benign-malignant classification accuracy.

Conclusions

In summary, we have investigated a new CAD framework, which is named MRSKNet, for the task of pulmonary nodule classification in distinguishing malignancy from benignancy in CT images. We took advantage of the residual learning and selective kernel to accommodate the diversity of lung nodules with various shapes and obscure characters. The strategy of the multiview input with three anatomical planes was exploited to reinforce the model efficacy. To improve the performance, the intensity CT images along with the HOM feature maps was concatenated for feeding the network. Experimental results on the LIDC-IDRI challenge database validated the capability of our nodule classification scheme, which achieved high-throughput performance evaluation. Larger AUC, recall, and precision scores were generated by our system in comparison with the competing

methods. We believe that the established categorization network architecture is of potential in many lung CT image classification applications. Future research directions include the use of a larger nodule subimage, more hand-crafted texture feature incorporation, and the evaluation on additional nodule classification image datasets.

Data Availability The inclusion of datasets is future work, which will be collected from any possible public domain.

Declarations

Ethics Approval This paper does not contain any studies with human participants or animals performed by any of the authors.

Consent to Participate This paper does not contain patient data.

Competing Interest The authors declare no competing interests.

References

1. Ferlay J, et al.: Cancer statistics for the year 2020: An overview. International Journal of Cancer 149:778-789, 2021
2. Siegel RL, Miller KD, Fuchs HE, Jemal A: Cancer Statistics, 2021. CA: A Cancer Journal for Clinicians 71:7–33, 2021
3. Lu S, et al.: Iterative reconstruction of low-dose CT based on differential sparse. Biomedical Signal Processing and Control 79:104204, 2023
4. Jennifer S, Sharmila S: A Neutrosophic Set Approach on Chest X-rays for Automatic Lung Infection Detection. Information Technology and Control 52:37-52, 2023
5. Poap D, Wozniak M, Damaševičius R, Wei W: Chest radiographs segmentation by the use of nature-inspired algorithm for lung disease detection. Proc. 2018 IEEE Symposium Series on Computational Intelligence (SSCI): City, 18–21 Nov. 2018
6. Jaszcz A, Połap D, Damaševičius R: Lung X-Ray Image Segmentation Using Heuristic Red Fox Optimization Algorithm. Scientific Programming 2022:4494139, 2022
7. Ma H, et al.: Automatic pulmonary ground-glass opacity nodules detection and classification based on 3D neural network. Medical Physics 49:2555-2569, 2022
8. Thakur SK, Singh DP, Choudhary J: Lung cancer identification: a review on detection and classification. Cancer and Metastasis Reviews 39:989-998, 2020
9. Zhang G, Yang Z, Gong L, Jiang S, Wang L, Zhang H: Classification of lung nodules based on CT images using squeeze-and-excitation network and aggregated residual transformations. La radiologia medica 125:374-383, 2020
10. Saba T: Automated lung nodule detection and classification based on multiple classifiers voting. Microscopy Research and Technique 82:1601-1609, 2019
11. Muzammil M, Ali I, Haq IU, Khaliq AA, Abdullah S: Pulmonary Nodule Classification Using Feature and Ensemble Learning-Based Fusion Techniques. IEEE Access 9:113415-113427, 2021
12. Farag AA, Ali A, Elshazly S, Farag AA: Feature fusion for lung nodule classification. International Journal of Computer Assisted Radiology and Surgery 12:1809-1818, 2017
13. Chen S, et al.: Automatic Scoring of Multiple Semantic Attributes With Multi-Task Feature Leverage: A Study on Pulmonary Nodules in CT Images. IEEE Transactions on Medical Imaging 36:802-814, 2017
14. Wei G, Cao H, Ma H, Qi S, Qian W, Ma Z: Content-based image retrieval for Lung Nodule Classification Using Texture Features and Learned Distance Metric. Journal of Medical Systems 42:13, 2017
15. Zhang F, et al.: Lung Nodule Classification With Multilevel Patch-Based Context Analysis. IEEE Transactions on Biomedical Engineering 61:1155-1166, 2014
16. Mao K, Deng Z: Lung Nodule Image Classification Based on Local Difference Pattern and Combined Classifier. Computational and Mathematical Methods in Medicine 2016:1091279, 2016
17. Cicero M, et al.: Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs. Investigative Radiology 52:281-287, 2017
18. Khan MA, et al.: VGG19 Network Assisted Joint Segmentation and Classification of Lung Nodules in CT Images. Diagnostics 11:2208, 2021
19. He K, Zhang X, Ren S, Sun J: Deep residual learning for image recognition. Proc. Proceedings of the IEEE conference on computer vision and pattern recognition: City
20. Li X, Wang W, Hu X, Yang J: Selective kernel networks. Proc. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition: City
21. Liu M, et al.: Three-Dimensional Modeling of Heart Soft Tissue Motion. Applied Sciences 13:2493, 2023
22. Liu X, Yang L, Chen J, Yu S, Li K: Region-to-boundary deep learning model with multi-scale feature fusion for medical image segmentation. Biomedical Signal Processing and Control 71:103165, 2022
23. Uthoff J, et al.: Machine learning approach for distinguishing malignant and benign lung nodules utilizing standardized perinodular parenchymal features from CT. Medical Physics 46:3207-3216, 2019
24. Zheng B, Yang D, Zhu Y, Liu Y, Hu J, Bai C: 3D gray density coding feature for benign-malignant pulmonary nodule classification on chest CT. Medical Physics 48:7826-7836, 2021
25. Dhara AK, Mukhopadhyay S, Dutta A, Garg M, Khandelwal N: A Combination of Shape and Texture Features for Classification of Pulmonary Nodules in Lung CT Images. Journal of Digital Imaging 29:466-475, 2016
26. Madero Orozco H, Vergara Villegas OO, Cruz Sánchez VG, Ochoa Domínguez HdJ, Nandayapa Alfaro MdJ: Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. BioMedical Engineering OnLine 14:9, 2015
27. de Sousa Costa RW, da Silva GLF, de Carvalho Filho AO, Silva AC, de Paiva AC, Gattass M: Classification of malignant and benign lung nodules using taxonomic diversity index and phylogenetic distance. Med Biol Eng Comput 56:2125-2136, 2018
28. Firmino M, Angelo G, Morais H, Dantas MR, Valentim R: Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. BioMedical Engineering OnLine 15:2, 2016
29. de Carvalho Filho AO, Silva AC, Cardoso de Paiva A, Nunes RA, Gattass M: Computer-Aided Diagnosis of Lung Nodules in Computed Tomography by Using Phylogenetic Diversity, Genetic Algorithm, and SVM. Journal of Digital Imaging 30:812-822, 2017
30. Sasidhar B, Geetha G, Khodanpur B, Babu DR: Automatic classification of lung nodules into benign or malignant using SVM classifier. Proc. Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications: City

31. Li X-X, Li B, Tian L-F, Zhang L: Automatic benign and malignant classification of pulmonary nodules in thoracic computed tomography based on RF algorithm. IET Image Processing 12:1253-1264, 2018

32. Wu W, Hu H, Gong J, Li X, Huang G, Nie S: Malignant-benign classification of pulmonary nodules based on random forest aided by clustering analysis. Physics in Medicine & Biology 64:035017, 2019

33. Rodrigues MB, et al.: Health of things algorithms for malignancy level classification of lung nodules6:18592–18601, 2018

34. Lee MC, et al.: Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction. Artificial Intelligence in Medicine 50:43-53, 2010

35. Farahani FV, Ahmadi A, Zarandi MHF: Hybrid intelligent approach for diagnosis of the lung nodule from CT images using spatial kernelized fuzzy c-means and ensemble learning. Mathematics and Computers in Simulation 149:48-68, 2018

36. Suganyadevi S, Seethalakshmi V, Balasamy K: A review on deep learning in medical image analysis. International Journal of Multimedia Information Retrieval 11:19-38, 2022

37. Krizhevsky A, Sutskever I, Hinton GE: ImageNet classification with deep convolutional neural networks, 2012

38. Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. arXiv:14091556, 2014

39. Szegedy C, et al.: Going deeper with convolutions. Proc. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): City

40. Li R, Xiao C, Huang Y, Hassan H, Huang B: Deep Learning Applications in Computed Tomography Images for Pulmonary Nodule Detection and Diagnosis: A Review. Diagnostics 12:298, 2022

41. Murugesan M, Kaliannan K, Balraj S, Singaram K, Kaliannan T, Albert JR: A Hybrid deep learning model for effective segmentation and classification of lung nodules from CT images. Journal of Intelligent & Fuzzy Systems 42:2667-2679, 2022

42. Lu S, et al.: Soft Tissue Feature Tracking Based on Deep Matching Network. Computer Modeling in Engineering \& Sciences 136:363--379, 2023

43. Dai Y, Yan S, Zheng B, Song C: Incorporating automatically learned pulmonary nodule attributes into a convolutional neural network to improve accuracy of benign-malignant nodule classification. Physics in Medicine & Biology 63:245004, 2018

44. Ren Y, et al.: A manifold learning regularization approach to enhance 3D CT image-based lung nodule classification. International Journal of Computer Assisted Radiology and Surgery 15:287-295, 2020

45. Xie S, Girshick R, Dollár P, Tu Z, He K: Aggregated Residual Transformations for Deep Neural Networks. Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): City, 21–26 July 2017

46. Liu H, et al.: Multi-model Ensemble Learning Architecture Based on 3D CNN for Lung Nodule Malignancy Suspiciousness Classification. Journal of Digital Imaging 33:1242-1256, 2020

47. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA: Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Proc. AAAI Conference on Artificial Intelligence: City

48. Zuo W, Zhou F, He Y, Li X: Automatic classification of lung nodule candidates based on a novel 3D convolution network and knowledge transferred from a 2D network. Medical Physics 46:5499-5513, 2019

49. Lyu J, Bi X, Ling SH: Multi-Level Cross Residual Network for Lung Nodule Classification. Sensors 20:2837, 2020

50. An Y, Hu T, Wang J, Lyu J, Banerjee S, Ling SH: Lung Nodule Classification using A Novel Two-stage Convolutional Neural Networks Structure'. Proc. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): City

51. Su H, Maji S, Kalogerakis E, Learned-Miller E: Multi-view convolutional neural networks for 3d shape recognition. Proc. Proceedings of the IEEE international conference on computer vision: City

52. Nibali A, He Z, Wollersheim D: Pulmonary nodule classification with deep residual networks. International Journal of Computer Assisted Radiology and Surgery 12:1799-1808, 2017

53. Sahu P, Yu D, Dasari M, Hou F, Qin H: A Lightweight Multi-Section CNN for Lung Nodule Classification and Malignancy Estimation. IEEE Journal of Biomedical and Health Informatics 23:960-968, 2019

54. Al-Shabi M, Lee HK, Tan M: Gated-Dilated Networks for Lung Nodule Classification in CT Scans. IEEE Access 7:178827-178838, 2019

55. Al-Shabi M, Lan BL, Chan WY, Ng K-H, Tan M: Lung nodule classification using deep Local–Global networks. International Journal of Computer Assisted Radiology and Surgery 14:1815-1819, 2019

56. Xie Y, et al.: Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT. IEEE Transactions on Medical Imaging 38:991-1004, 2019

57. Xie Y, Zhang J, Xia Y: Semi-supervised adversarial model for benign–malignant lung nodule classification on chest CT. Medical Image Analysis 57:237-248, 2019

58. Wang C, Elazab A, Wu J, Hu Q: Lung nodule classification using deep feature fusion in chest radiography. Computerized Medical Imaging and Graphics 57:10-18, 2017

59. Zhu W, Liu C, Fan W, Xie X: Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. Proc. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV): City

60. Friedman JH: Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics 29:1189-1232, 2001

61. Zhang G, Yang Z, Gong L, Jiang S, Wang L: Classification of benign and malignant lung nodules from CT images based on hybrid features. Physics in Medicine & Biology 64:125011, 2019

62. Chen Y, Li J, Xiao H, Jin X, Yan S, Feng J: Dual path networks, Long Beach, California, USA: Curran Associates Inc., 2017

63. Huang G, Liu Z, Maaten LVD, Weinberger KQ: Densely Connected Convolutional Networks. Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): City, 21–26 July 2017

64. Chang H, Hsieh C: Brain segmentation in MR images using a texture-based classifier associated with mathematical morphology. Proc. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): City

65. Haralick RM, Shanmugam K, Dinstein I: Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics SMC-3:610–621, 1973

66. Galloway MM: Texture analysis using gray level run lengths. Computer Graphics and Image Processing 4:172-179, 1975

67. Tamura H, Mori S, Yamawaki T: Textural Features Corresponding to Visual Perception. IEEE Transactions on Systems, Man, and Cybernetics 8:460-473, 1978

68. Armato III SG, et al.: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. Medical Physics 38:915-931, 2011

69. Clark K, et al.: The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. Journal of Digital Imaging 26:1045-1057, 2013

70. Reeves AP, Biancardi AM: The Lung Image Database Consortium (LIDC) Nodule Size Report. Release: 2011–10–27–2 At: /lidc/, 2011