IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 14, 2020, accepted September 27, 2020, date of publication September 30, 2020,
date of current version October 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3027812

Multi-Task Learning for Lung Nodule Classification on Chest CT

PENGHUA ZHAI, YALING TAO, HAO CHEN, TING CAI, AND JINPENG LI
HwaMei Hospital, University of Chinese Academy of Sciences, Ningbo 315010, China
Ningbo Institute of Life and Health Industry, University of Chinese Academy of Sciences, Ningbo 315010, China

Corresponding author: Jinpeng Li (lijinpeng15@mails.ucas.ac.cn)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LQ20F030013, in part by the Research Foundation of HwaMei Hospital, University of Chinese Academy of Sciences, China, under Grant 2020HMZD22, in part by the Ningbo Health Branding Subject Foundation under Grant PPXK2018-05, in part by the Zhejiang Provincial Public Service and Application Research Foundation, China, under Grant 202002N3181, and in part by the Medical Scientific Research Foundation of Zhejiang Province, China, under Grant 2021KY1028.

ABSTRACT Lung cancer is one of the leading causes of death over the world. Detecting and identifying malignant nodules on chest computed tomography (CT) plays an important role in the diagnosis and treatment of lung cancer. Computer-aided diagnosis (CAD) systems have been developed to identify lung nodules. However, the problem of a high false positive rate is still not well solved. In this paper, we propose a novel multi-task convolutional neural network (MT-CNN) framework to identify malignant nodules from benign nodules on chest CT scans. MT-CNN learns three-dimensional (3-D) lung nodule characteristics from nine two-dimensional (2-D) views, which are decomposed from different angles of each nodule. Each of 2-D MT-CNN model consists of two branches, one is the nodule classification branch (main task) and the other is the image reconstruction branch (auxiliary task). The motivation of the auxiliary task is to preserve more microscopic information in the hierarchical structure of CNN, which is beneficial to malignant nodule identification. The final classification result is obtained by integrating nine 2-D models. We test our method on the benchmark LUNA-16 and LIDC-IDRI datasets and compare it with state-of-the-art models. MT-CNN achieves the lowest false positive rate (3.2%) and highest AUC (97.3%) in LUNA-16 dataset and achieves an AUC of 95.59% in LIDC-IDRI. These results demonstrate the advantage of our method.

INDEX TERMS Multi-task learning, lung nodule classification, image reconstruction, computer-aided diagnosis, convolutional neural network.

I. INTRODUCTION

Lung cancer has the highest morbidity and mortality among all cancers over the world. About one-quarter of cancer deaths are lung cancer patients [1]. In the early stage of lung cancer, the symptoms are mild and difficult to diagnose. Patients usually have missed the best period of treatment when they are diagnosed. Early screening is an important approach to prevent lung cancer. As one of the most important early manifestations of lung cancer [2], lung nodules are radiologically visible as small structures. Currently, radiologists usually read chest computed tomography (CT) scans slice by slice to identify malignant and benign lung nodules. It is time-consuming, laborious, and subjective.

The associate editor coordinating the review of this manuscript and approving it for publication was Berdakh Abibullaev.

Computer-aided diagnosis (CAD) systems have been developed to assist radiologists in diagnosing lung nodules [3]–[5]. Generally, an automatic lung nodule diagnosis system consists of two stages: 1) candidate nodules detection and 2) false positive reduction [6]–[8]. In the first stage, the goal is to detect all suspicious nodules as much as possible. Candidate nodules inevitably contain a large number of benign nodules. Therefore, an effective supervised classifier is usually developed in the second stage to identify malignant from benign nodules.

The false positive reduction is a key step in automatic lung nodule diagnosis systems. Researchers have conducted extensive research and proposed many nodule classifiers using various nodule features extracted by visual feature descriptors [4], such as histogram of oriented gradients (HOG) [9], local binary pattern (LBP) [10] and scale-invariant feature transform (SIFT) [11]. It is also

difficult to accurately describe nodule characteristics by manual feature engineering since nodule features vary widely including shape, texture and margin.

With the development of deep learning, convolutional neural network (CNN) has achieved remarkable success in image processing [12], [13]. Compared with traditional computer vision technology, deep learning can automatically learn the image feature representation and can dig the potential features of the data as much as possible. Deep learning can also more easily adapt to different fields and applications. Deep learning avoids the subjectivity of feature engineering and can significantly reduce the workload of radiologists. Therefore, researchers have applied CNN to automatic lung nodule diagnosis systems. Ciompi *et al.* [14] constructed an ensemble classifier to automatically recognize pulmonary peri-fissural nodules. They classified nodules based on multiple 2-D views of the nodules. In order to describe nodule morphology in 2-D views, the method used the output of a pre-trained CNN known as OverFeat. This method achieved an AUC of 86.8%, which was close to radiologists. However, 2-D CNN is unable to detect 3-D spatial information of nodules. Lung nodule diagnosis from volumetric CT scans is essentially a 3-D object recognition problem. Zhao *et al.* [15] proposed a new agile CNN framework to conquer the challenges of a small-scale medical image database and the small size of the nodules based on the hybrid of LeNet and AlexNet. This method achieved an AUC of 87.7%.

Considering the 3-D characteristic of lung nodules, researchers have also tried to identify nodules by 3-D CNN due to the 3-D characteristics of nodules [16]–[18]. Dou *et al.* [19] developed a 3-D fully convolutional network (FCN) to screen candidates and designed a hybrid-loss 3-D residual network to distinguish malignant from benign nodules. Experimental results on the public large-scale LUNA-16 dataset demonstrated superior performance of this method. Zhu *et al.* [20] constructed a gradient boosting machine (GBM) with 3-D dual path network (DPN) features for nodule classification, and this method achieved an accuracy of 90.44% on LIDC-IDRI. Ardila *et al.* [21] designed a binary classifier based on the 3-D Inception network. The model contains two paths: one is used for unrestricted 3-D global image analysis, and the other is used for specific identification of local structures containing potential lung nodules. Finally, the two output values are combined to calculate the probability of lung cancer.

Although 3-D CNN has achieved good results, the sensitivities of the models have encountered a bottleneck. In this paper, we propose a novel multi-task CNN (MT-CNN) to automatically identify malignant from benign nodules. We decompose each nodule to nine 2-D views from different angles based on existing researches [18], [22]. In multi-task learning, sharing parameters between different tasks can provide better generalization performance for each task [23]. Therefore, we speculate image self-reconstructing auxiliary task enables the model to preserve more microscopic information.

Our main contributions are summarized as follows: (1) In order to take full advantage of microscopic information, we design a MT-CNN. MT-CNN includes two branches, i.e., nodule classification main branch and image reconstruction auxiliary branch. We also design a loss penalty to balance the weights of the two branches. (2) We integrate multiple 2-D models for lung nodule classification. The proposed architecture can learn the 3-D spatial information of nodules with the fusion of multiple 2-D models. Compared with 3-D models, the proposed architecture not only reduces computational complexity but also learns more discriminative features from diagonal and flat of nodules. (3) We validate our proposed strategies on the public benchmark LIDC-IDRI and LUNA-16 dataset. The results obtained by 2-D MT-CNN with fewer parameters are comparable to state-of-the-art methods.

In the following, we describe dataset and our method in Section II and Section III, respectively. Section IV reports the experimental results. We discuss some key issues in Section V. Conclusions are drawn in Section VI.

II. DATASET

We train and evaluate the proposed method on LIDC-IDRI [24] and LUNA-16 [25] dataset. LIDC-IDRI contains 1,018 CT scans. For each CT scan, two-stage image annotation is performed by four experienced radiologists. In the first stage, each radiologist independently reviews CT scans and marks lesions belonging to one of the three categories (''nodule $\geq$ 3.0 *mm*,'' ''nodule $<$ 3.0 *mm*,'' and ''non-nodule $\geq$ 3.0 *mm*''). In the second stage, each radiologist independently reviews the anonymized marks from the three other radiologists to render a final opinion.

In LIDC-IDRI, the slice thickness of CT images varies from 0.6 *mm* to 5.0 *mm* with a median of 2.0 *mm*. In our study, the thick-slice scans larger than 2.5 *mm* are not used, as these are not recommended anymore due to it can easily lead to the omission of some small nodules [18], [26], [27]. Following the procedures used in previous studies, we select the nodules which were annotated by at least three radiologists for this study, and calculated the median malignancy level of each nodule as a final malignant score. Therefore, the number of nodules in our dataset with different malignancy scores is shown in Fig. 1. According to the malignancy score, we annotated a nodule whose *score* $>$ 3 as benign, a nodule whose *score* $=$ 3 as uncertain, and a nodule whose *score* $<$ 3 as malignant. After that, our dataset contains 1,109 candidate nodules, including 369 benign, 405 uncertain and 335 malignancy nodules. There are 266 benign, 175 malignant and 239 uncertain nodules not included in our dataset. To reduce the impact of uncertain evaluation, we excluded all uncertain lung nodules from our dataset.

LUNA-16 dataset is a subset of LIDC-IDRI, and it includes 888 CT scans. The candidate locations are computed using three existing candidate detection algorithms [28]–[30]. As lesions can be detected by multiple candidates, those that are located $<=$5 *mm* are merged. Using this method,

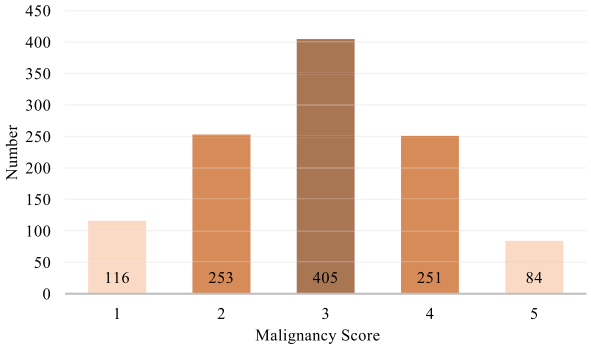

FIGURE 1. The number of nodules with different malignancy scores in LIDC-IDRI dataset.

1,120 out of 1,186 nodules are detected with 551,065 candidates. For convenience, the corresponding class label (0 for non-nodule and 1 for nodule) for each candidate is provided on the list. It has to be noted that there can be multiple candidates per nodule. Therefore, LUNA-16 is an extremely unbalanced dataset, where the number of non-nodules is about 500 times the number of nodules.

III. METHOD

In order to achieve automatic benign-malignant lung nodules classification on chest CT, we propose a MT-CNN architecture, as shown in Fig. 2. We first extract nine 2-D views from a candidate nodule cube from different projection angles. Then, for each view, we construct a 2-D MT-CNN model, which consists of a nodule classification branch and an image reconstruction branch. Finally, the classification result is obtained by the weighted fusion of the prediction results of nine 2-D MT-CNN models.

A. VIEW EXTRACTION

Since chest CT scans have a variable spatial resolution, we resample resolution to a unified value of $1.0 \times 1.0 \times 1.0\ mm^3$ using the spline interpolation [31]. The pixel intensity range is rescaled from $(-1000, 400)$ to $(0, 1)$. In order to reduce the search space and improve the performance of the proposed model, we also perform lung parenchyma segmentation before extracting views.

For each nodule, we first extract a cube of $64 \times 64 \times 64$, which contains the candidate nodule, as shown in Fig. 2(a). To improve the generation ability and prediction performance, every cube embraces a complete nodule, and the nodule is always located at the center of the cube. For each cube, we extract nine 2-D views from different angles, as shown in Fig. 2(b).

B. 2-D MT-CNN MODEL

The architecture of the proposed MT-CNN is shown in Fig. 2(c). For each view, we construct a 2-D MT-CNN model, which consists of three parts (see Fig. 3.), i.e., parameter sharing network (ShareNet), nodule classification network (ClsNet) and image reconstruction network (ImgNet). ShareNet contains three convolutional layers and three max-pooling layers. The images are first fed into three 2-D convolution layers with 24, 32, and 64 filters of size 5×5, 3×3, and 3×3, respectively. ClsNet branch contains three fully connected layers with 256, 16 and 2 neurons, respectively. ImgNet contains three deconvolutional layers and three upsampling layers, and it is structurally symmetrical to ShareNet. The ShareNet and ImgNet branch form an autoencoder [32], in which the ShareNet is an encoder and the ImgNet branch is a decoder.

For each 2-D MT-CNN, the input is a 2-D nodule view (i.e., an image of 64×64). We get a malignant score by sigmoid activation function from ClsNet branch. The reconstructed image is also obtained by sigmoid activation function. We use rectified linear units (ReLU) activation function [33] except for the output layers. For the reconstructed image, it has same size and similar texture feature as the input image.

We update weights during the error back propagation process by jointly optimizing ClsNet branch and ImgNet branch. We express the final loss for each 2-D MT-CNN model as:

$$\mathcal{L}_{mt} = \mathcal{L}_{cls} + \alpha \mathcal{L}_{img}, \tag{1}$$

where $\mathcal{L}_{cls}$ and $\mathcal{L}_{img}$ are denoted the loss function of ClsNet and ImgNet, respectively. α controls the tradeoff between ClsNet branch and ImgNet branch, and $\alpha \in (0, 1)$. If $\alpha = 0$, the proposed method becomes a 2-D CNN (ST-CNN) model that is only used for lung nodule classification.

We train ShareNet by jointly these two branches. Therefore, the network can force the image feature representation learned in ShareNet to have both good nodule classification capability and strong image reconstruction capability. We also design a tradeoff α to balance the weight between ImgNet and ClsNet to prevent the model from focusing on ImgNet and reducing the nodule classification ability.

We apply the cross-entropy loss function to assess the model and optimize it to achieve the best nodule classification performance. For ImgNet, we calculate cross-entropy loss for each pixel of an image, and appoint the mean value as the loss of ImgNet branch. The cross-entropy loss function can be formulated as

$$\mathcal{L}(y_i, \hat{y}_i) = y_i log\hat{y}_i + (1 - y_i)log(1 - \hat{y}_i), \tag{2}$$

where y_i is the real label of the i-th sample. If the nodule is malignant, $y_i = 1$, otherwise, $y_i = 0$. $\hat{y}_i$ is the probability of predicting the i-th sample as a positive sample. The larger the prediction probability, the higher the possibility of the nodule is a malignant nodule.

Finally, the loss of 2-D MT-CNN model is formulated as

$$\mathcal{L}_{mt} = \mathcal{L}(y_i, \hat{y}_i) + \alpha \frac{1}{N} \sum_{i=0}^{N} \mathcal{L}(z_i, \hat{z}_i), \tag{3}$$

where N is the number of pixels per input image, and z_i represents the i-th pixel of the input image and $\hat{z}_i$ is the corresponding predicted value of the pixel. In our work, we process per input image to 64×64, therefore, the $N = 64 \times 64$.

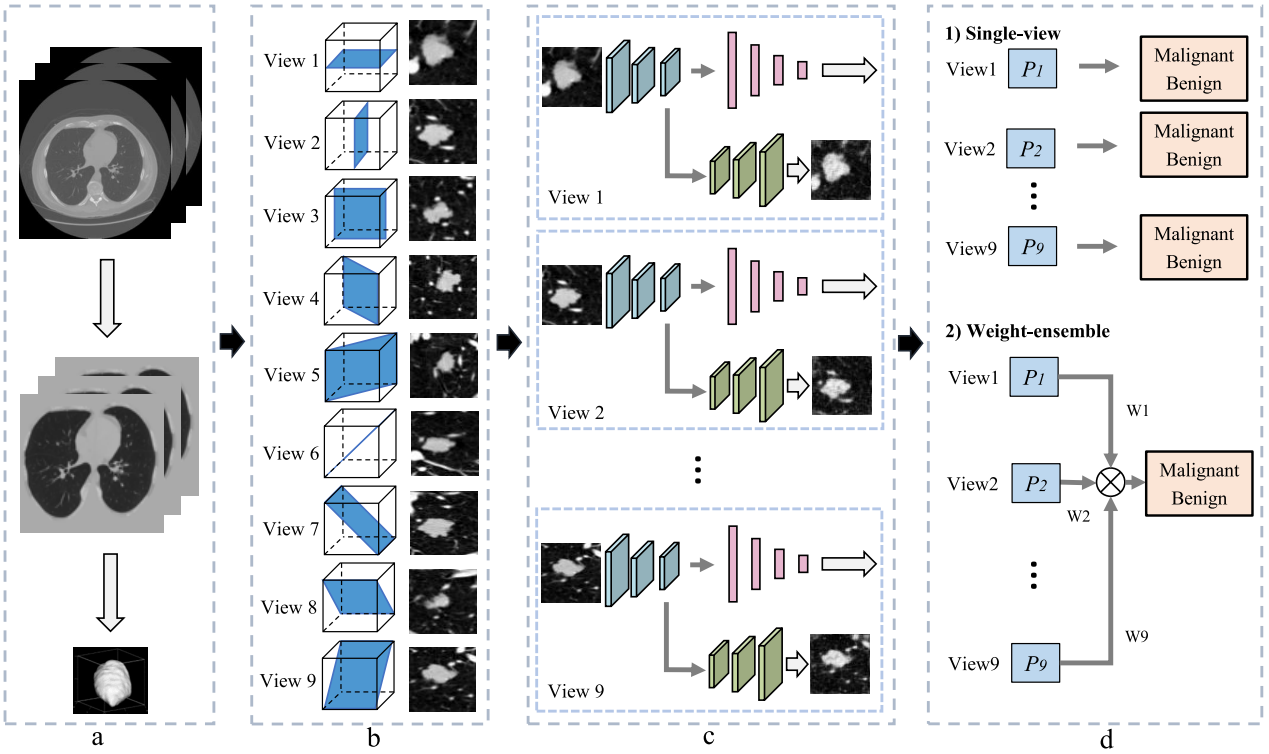

FIGURE 2. The framework of MT-CNN. We design a multi-task convolutional neural network to identify malignant and benign nodules. (a) lung parenchyma segmentation and candidate nodule extraction. (b) extracting 2-D views from nine different angles of a candidate nodule cube. (c) 2-D MT-CNN. (d) 3-D MT-CNN.

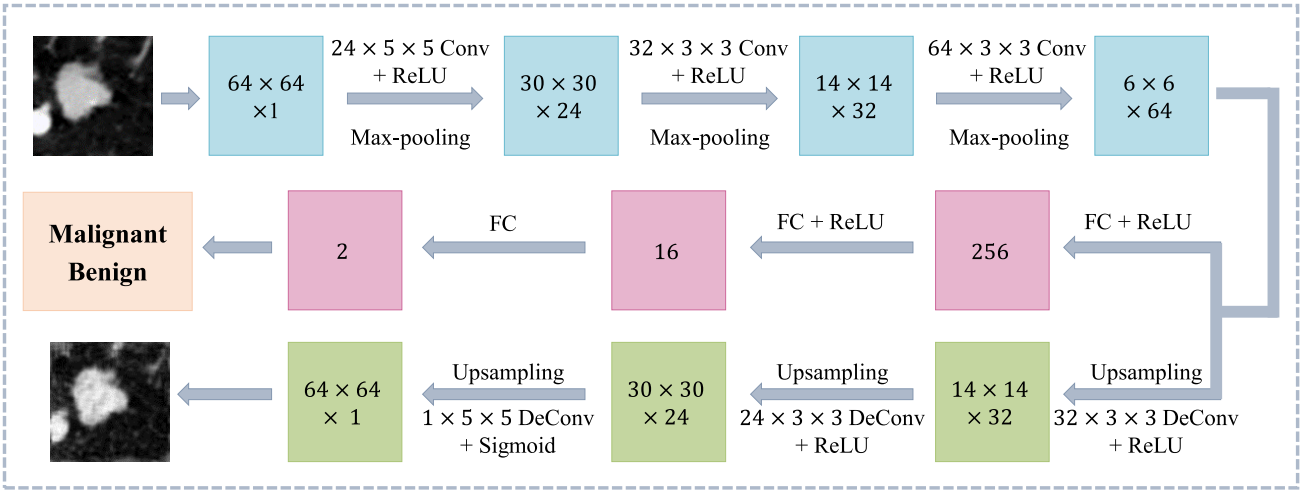

FIGURE 3. The architecture of 2-D MT-CNN model of each view for lung nodule classification. Conv: convolutional layer. FC: fully connected layer. The architecture of the 2-D MT-CNN model of each view for lung nodule classification: (a) In the first row, it is the parameter sharing network with three convolutional layers. The input image size is 64 × 64 with one channel. The filter size is 5 × 5, 3 × 3 and 3 × 3, respectively. In the middle row, we show the nodule classification branch with three fully connected layers. In the bottom, it is the image reconstruction network with three deconvolutional layers, and it is structurally symmetrical to the parameter sharing network.

C. FUSION APPROACH

The 2-D MT-CNN contains nine 2-D MT-CNN models, which is shown in Fig. 2(d). The final prediction results are obtained by fusing these nine 2-D models. The MT-CNN can learn the 3-D characteristics by fusing the outputs of multiple 2-D MT-CNN. We get the final results by computing the

weight average prediction scores of the nine 2-D MT-CNN, which is formulated as

$$\mathcal{P} = \sum_{v=0}^{\mathcal{V}} \omega_v p_v, \tag{4}$$

where p_v represents the prediction probability of the v-th model. ω_v is the weight of v-th model. Due to the problem of data imbalance, we use AUC to estimate the weight. In the training process, we train model for each view separately without fusion. In the testing process, we get the final result by fusing the nine views. The AUC value is the result obtained for each view in the training phase. $\mathcal{V}$ is the total number of views, in our paper, $\mathcal{V} = 9$. $\mathcal{P}$ is the final prediction probability of a nodule obtained by fusing all view results with weight.

D. TRAINING

To train the proposed network, we adopt the Adam learning algorithm [34]. The training epochs and batch size are set as 150 and 64, respectively. Also, we apply early stopping to monitor the training process with a patience of ten. In order to avoid overfitting, dropout with the probability of 0.5 is implemented on the first and second fully connected layer in ClsNet. The initial learning rate is 1e-4, and turns to 1e-5 after half of epochs, and finally 5e-6 for the last one-quarter of the epochs. For parameter initialization, the sets of Kaiming's [35] is used for the weights, and bias is set to 0. The initialization method that particularly considers the rectifier nonlinearities, and it enables us to train rectified models directly from scratch. The training processing is shown in algorithm 1.

Algorithm 1 Multi-Task CNN for Lung Nodule Classification

Input: dataset $D_v = \{(\boldsymbol{X}, \boldsymbol{y})_i\}_{i=1}^{N}$, hyper-parameter α.;
Output: P, the prediction probability of a nodule is a malignant nodule.;

```
Training:
for v = 1 to 9:
    Train the multi-task CNN for view v by minimizing equation 3.
    Calculate AUC_v and P_v for view v.
end for
Validation:
W_v = AUC_v / sum(AUC_v)_{v=1}^{9}
Calculate P for each lung nodule by equation 4.
Repeat training and validation to search parameter α by maximizing equation 4.
return P;
```

Since data augmentation alleviates the overfitting by adding variants to the dataset [36], as well as the limitation of the data, we generated four augmented data using random image translation, rotation, and flip. The translation step was selected from $[6, 6]$ pixels, and the rotation angle was randomly selected from $\{90°, 180°, 270°\}$. For flip, we also randomly flip the image horizontally or vertically. For LIDC-IDRI, we apply these transformations to positive and negative samples, and for LUNA-16 dataset, we only apply these transformations to positive samples.

E. EVALUATION

To yield more reliable results, we apply ten times 10-fold cross-validation to the benchmark LIDC-IDRI. The area under the receiver operating characteristic curve (AUC), sensitivity, and specificity are used to evaluate our proposed method. Sensitivity represents the proportion of malignant nodules that are correctly predicted. Specificity measures the proportion of benign nodules that are correctly detected. AUC is a comprehensive indicator describing the classification results of benign and malignant nodules. In evaluating this method, we do not use any data augmentation.

IV. EXPERIMENTS

A. COMPARISONS IN BENIGN-MALIGNANT CLASSIFICATION

We compare the proposed approach with existing lung nodule classification approaches in the recent five years on LUNA-16 and LIDC-IDRI, and results are summarized in Table 1. LUNA-16 dataset evaluated in our method is highly unbalanced, the ratio of positive and negative samples is approximately 1: 500, but for LIDC-IDRI, the ratio is approximately 1:1. In order to fairly compare all the methods, we should evaluate all methods in the same environment. Therefore, we repeated all the experiments under exact conditions. We rewrite the code of all the methods except for that of Xie *et al.* [22], where the public codes are available. For all the methods, we performed 10-times 10-fold cross-validation under the same dataset partition. As we can see, our method has the highest AUC (95.59%) in the LIDC-IDRI dataset, and has the highest AUC (97.30%) and specificity (96.80%) in the LUNA-16 dataset.

For method Xie *et al.* [22], they also extract nine views from a nodule cube. And for each view, they construct a knowledge-based collaborative (KBC) model, where three types of image patches are designed to fine-tune three pre-trained ResNet-50 network that characterize the nodule characteristics, including overall appearance, voxel and shape heterogeneity. Finally, the nine KBC models are jointly used to classify lung nodules. They constructed twenty-seven KBC models and achieve the AUC of 94.04% with a sensitivity of 89.67%. Our method only uses nine models and achieves an AUC of 95.59% with a sensitivity of 87.74%.

For method Dou (2017) [44], they propose a novel method employing three-dimensional (3-D) CNN for false positive reduction. The proposed framework has been extensively validated in the LUNA-16 challenge held in conjunction with ISBI 2016, where they achieved the highest competition performance metric (CPM) score in the false positive reduction track. Compared with their method, our MT-CNN has higher

TABLE 1. Performance of Ours Proposed Model and Other Lung Nodule Classification Methods.

Methods	Dataset	Sensitivity (%)	Specificity (%)	AUC (%)
Kumar et al., 2015 [37] (Autoencoder + decision tree)	LIDC-IDRI	78.25 ± 0.01	79.15 ± 0.03	78.70 ± 0.01
Han et al., 2015 [38] (3D GLCM feature + SVM)	LIDC-IDRI	78.57 ± 0.05	88.89 ± 0.03	83.73 ± 0.02
Shewaye et al., 2016 [39] (heterogeneous feature + classifier)	LIDC-IDRI	86.11 ± 0.02	89.28 ± 0.02	87.70 ± 0.01
Yan et al., 2016 [40] (3D CNNs)	LIDC-IDRI	$\mathbf{90.37} \pm 0.00$	88.27 ± 0.01	90.48 ± 0.01
Dey et al., 2018 [41] (multi-output + 3D DenseNet)	LIDC-IDRI	85.71 ± 0.03	93.10 ± 0.03	89.40 ± 0.03
Xie et al., 2018 [22] (knowledge-based collaborative + ResNet)	LIDC-IDRI	89.30 ± 0.02	$\mathbf{95.56} \pm 0.02$	94.43 ± 0.02
Xie et al., 2019 [42] (semi-supervised adversarial model)	LIDC-IDRI	89.67 ± 0.02	94.24 ± 0.03	94.04 ± 0.02
Shen et al., 2019 [43] (deep hierarchical CNN)	LIDC-IDRI	81.54 ± 0.02	94.35 ± 0.02	87.97 ± 0.01
ST-CNN	LIDC-IDRI	77.28 ± 0.02	90.02 ± 0.03	92.47 ± 0.02
MT-CNN	LIDC-IDRI	87.74 ± 0.03	88.87 ± 0.03	$\mathbf{95.59} \pm 0.03$
Dou et al., 2017 [44] (3D CNNs + multi-level contextual)	LUNA-16	$\mathbf{92.86} \pm 0.01$	94.44 ± 0.03	93.65 ± 0.02
ST-CNN	LUNA-16	82.60 ± 0.02	92.30 ± 0.01	95.70 ± 0.03
MT-CNN	LUNA-16	84.00 ± 0.05	$\mathbf{96.80} \pm 0.04$	$\mathbf{97.30} \pm 0.03$

AUC and fewer parameters, which proves the effectiveness of our method.

To evaluate the effectiveness of auxiliary task, we also compare MT-CNN with single-task convolutional neural network (ST-CNN). ST-CNN has the same network architecture as MT-CNN except that it does not have an ImgNet branch. Compared with ST-CNN, MT-CNN has an obvious advantage in the two datasets. Although the specificity of MT-CNN is slightly lower than ST-CNN in LIDC-IDRI, the AUC of MT-CNN is higher than ST-CNN in the two datasets. These results indicate that the auxiliary task can improve the performance of the lung nodule classification system.

Compared with other methods, our method has the highest AUC (97.3%) and specificity (96.8%) in LUNA-16, and highest UAC (95.59%) in LIDC-IDRI. These results demonstrate the robustness of our approach.

B. IMAGE RECONSTRUCTION

We randomly select a malignant nodule image and a benign nodule image from each view in LIDC-IDRI dataset and visualize the input image and reconstructed image in Fig. 4. The Fig. 4 (a) and Fig. 4 (b) are the visualizations of malignant nodule images and benign nodule images, respectively. In each subfigure, the top row depicts original CT images, and the middle row depicts reconstructed images. The bottom row is the probability of the nodule belonging to the malignant class.

In Fig. 4, all reconstructed images are very similar to the original images. We can clearly see the margin and texture features of the nodules from the reconstructed image. This shows that the auxiliary branch can well complete the task of image reconstruction and most microscopic information has been preserved. The image reconstruction auxiliary task enables ShareNet to preserve pixel-level nodule information, which is beneficial to classification results. Therefore, we can retain as much information as possible to help identify malignant from benign nodules.

In general, our proposed MT-CNN method includes two branches (i.e. ClsNet branch and ImgNet branch), and the ImgNet branch is used as an auxiliary task. In this way, we can prevent the architecture from mechanically copying the input to the output, and force it to learn more valid image feature representations for nodule classification. If the encoded data can be easily restored to the original data through a decoder, we can retain more image detail features. As shown in Fig. 4, the reconstructed image is of high quality. This shows that the image reconstruction branch has a positive effect. It is successful to improve the performance of lung nodule classification system.

C. REGULARIZATION IMPACTS

Regularization is an effective way to avoid overfitting by actively discarding some information. Especially for computer visual task, model is easy to overfit due to the large amount of image data and parameters. For the branch of image reconstruction, the purpose is to improve the nodule classification performance by learning more microscopic information of nodules. And for multi-task learning, it also has the effect of regularization [23]. If regularization is still applied to the ImgNet branch, we will not be able to reconstruct the image with high quality because too many image features are lost. In Fig. 5, we conduct experiments to verify the impact of regularization on image reconstruction. And we compare the performance of our proposed models with and without l2 regularization, respectively.

As shown in Fig. 5, MT-CNN model without l2 regularization has the best AUC. However, MT-CNN model with l2 regularization shows the lowest AUC in all considered models. From the result of MT-CNN model, the ImgNet branch without l2 regularization can greatly improve the performance of nodule classification. Therefore, if we select

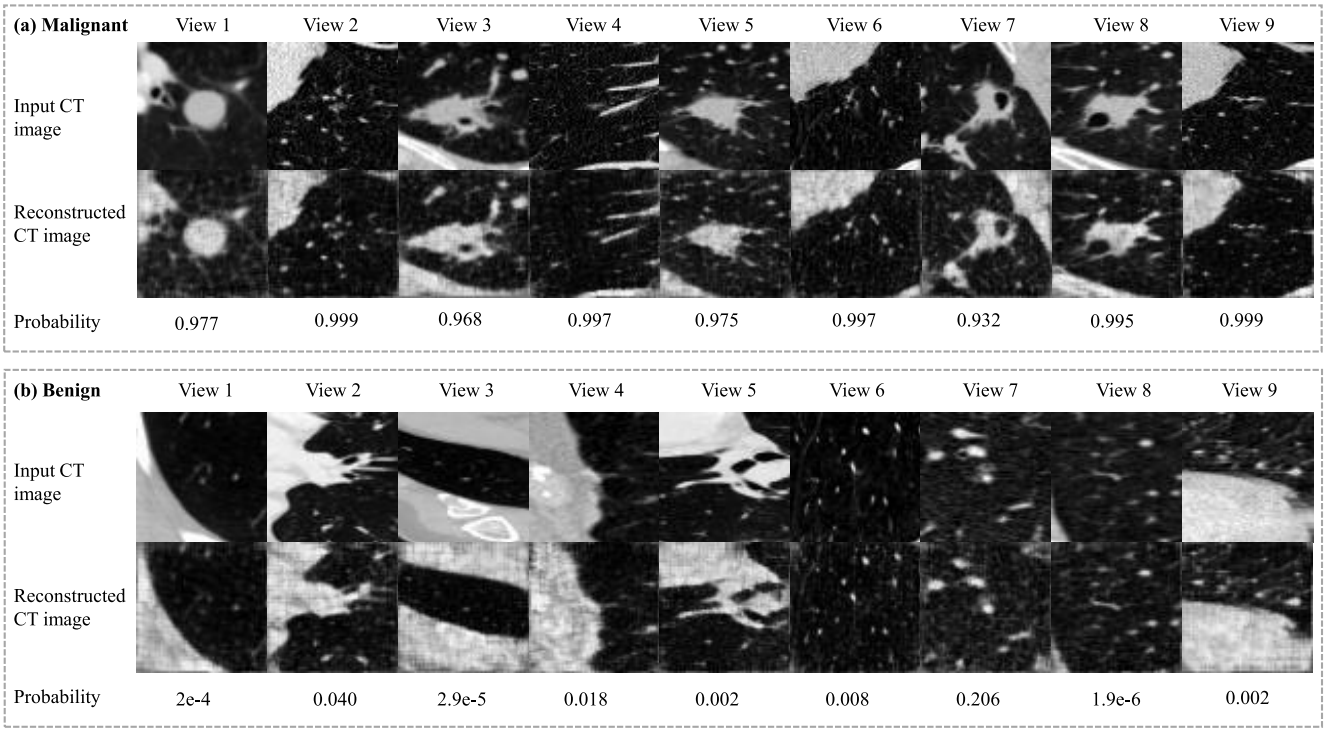

FIGURE 4. Images reconstructed by the proposed methods. (a) malignant nodule images. (b) benign nodule images. In (a) and (b), we list one original input image and the corresponding reconstructed image for each view. The top row is the original input images, the middle row is the reconstructed images, and the bottom row is the probability of predicting the nodule is a malignant nodule.

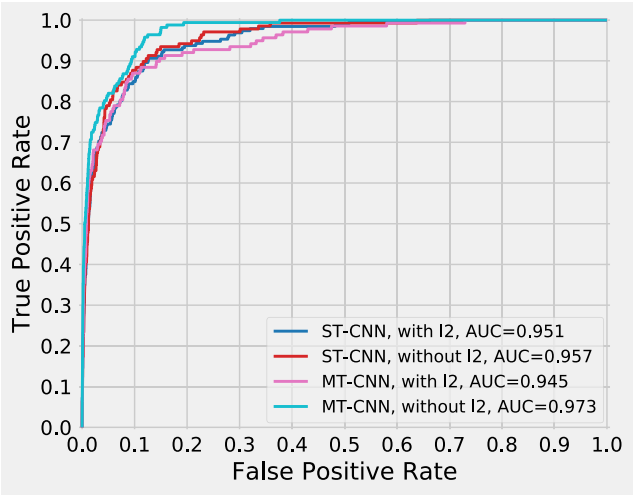

FIGURE 5. ROC of the proposed model with or without regularization in LUNA-16 dataset.

a proper parameter for auxiliary task, the performance of the model can be greatly improved. On the contrary, if the auxiliary task with improper parameters, auxiliary task may have the opposite effect and reduce the overall performance of MT-CNN model.

For image classification model, the performance of models with regularization will be better than models without regularization due to overfitting. In the proposed ST-CNN model, we have used dropout in the first fully connected layer. Dropout plays the same role as l2 regularization, they can reduce the complexity of the model and solve the overfitting problem. If we continue to apply regularization to a model, the number of model parameters will continue to decrease, and the learning ability of model will be discounted. But in our proposed ST-CNN model, we only apply dropout to the first fully connected layer, so the impact on the model will not be too great even if we use regularization in other layers. Therefore, regularization does not have too much impact on ST-CNN model, and the performance of model does not drop by a large scale.

Although we did not apply regularization to our proposed model, severe overfitting does not occur. This is because not only we apply dropout to the first fully connected layer in ClsNet branch, but multi-task learning has a regularization effect. Multi-task learning acts as a regularizer by introducing an inductive bias. In this way, it reduces the risk of overfitting as well as the ability to fit random noise of the model.

D. MODELS WITH DIFFERENT CONVOLUTIONAL LAYERS

In MT-CNN, we apply several convolutional layers to implement ShareNet. Our purpose is to verify the ability of multi-task learning and the effect of fusing multiple views. The basic network in our proposed approach is depicted

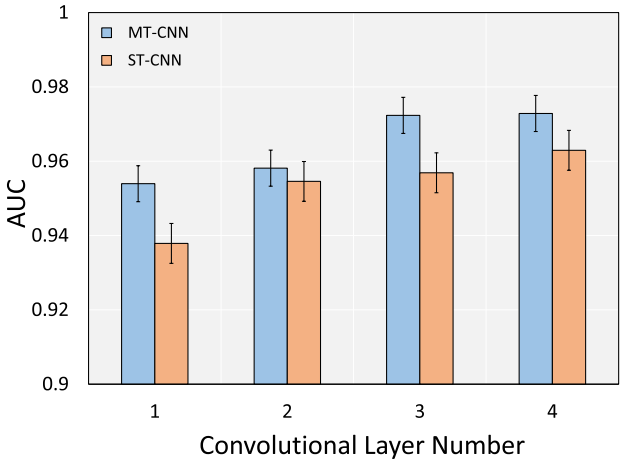

FIGURE 6. AUC of the proposed MT-CNN model and ST-CNN model with different number of convolutional layers in LUNA-16 dataset.

in Fig. 2. We use three convolutional layers and three max-pooling layers in ShareNet. Fig. 6 shows the effectiveness of exploiting different number of convolutional layers and deconvolutional layers. We vary the size of the convolutional layer from 1 to 4. When the number of convolutional layer equal to five, the size of the obtained feature map is $64 \times 1 \times 1$, it is equivalent to the first fully connected layer in the nodule classification branch. The value of α is manually defined. We design five parameters (0.2, 0.4, 0.6, 0.8, 1.0) with 0.2 as the step size. When $\alpha = 0.2$ and $\alpha = 0.4$, we get the same result. Therefore, we make α equal to 0.3 and rerun experiment, and we get the best result. Therefore, we finally set $\alpha = 0.3$.

As shown in Fig. 6, for the MT-CNN model, the better the performance, the more the number of the convolutional layer in ShareNet. And the performance of the model with more than two convolutional layers is much better than the performance of model with no more than two convolutional layers. It is effective to improve the performance by adding the number of convolutional layers for MT-CNN model.

For the ST-CNN model, it has better performance when the model has two or three convolutional layers. If we increase the number of convolutional layers, the problem of overfitting may occur and the model performance will decrease. On the contrary, for the MT-CNN model, even if it has a large number of parameters, the model does not have the problem of overfitting due to the regularization effect of multi-task learning.

Compared MT-CNN model with ST-CNN model, the performance of the former is no worse than the performance of the latter when they have the same number of convolutional layers. The performance difference also increases as the number of convolutional layers increases.

For the MT-CNN model, the AUC score with four convolutional layers is higher than three convolutional layers. Because the difference is too small, there is no difference between the two models from statistical significance. But training a model with four convolutional layers takes much more time and memory than training a model with three convolutional layers. In order to balance various factors, no matter for multi-task learning or single-task learning, three convolutional layers are one of the best choices for our proposed model.

E. IMAGE RECONSTRUCTION LOSS WEIGHT

In our proposed approach, we use the mean binary cross-entropy loss function for ImgNet branch. In the proposed MT-CNN model, the nodule classification task is the main task, and the effect of ImgNet branch is to improve the performance of nodule classification. Therefore, in order to train a predictive model with superior classification performance, it is necessary to ensure that the two losses have appropriate weights in the combined loss. In order to constraint the effect of ImgNet branch, we set a threshold α for ImgNet loss in Equ. (1). And the α is the ratio of the nodule classification loss weight to the ImgNet loss weight. But for ST-CNN model, it has the classification task, we do not need to consider the impact of auxiliary task.

To determine the threshold for image reconstruction loss, we perform experiments among the interval [0, 1]. If the threshold equals to 0, the MT-CNN model cannot obtain any information from ImgNet branch, so the model is equivalent to ST-CNN model. If the threshold equals to 1, this means that there is no difference between the primary task and the auxiliary task for MT-CNN, and the importance between these two tasks is equal.

As shown in Fig. 7, when the threshold is in a certain range, the performance of MT-CNN will be improved as the threshold increases. When the threshold is greater than a certain value, the performance of MT-CNN will decrease. It can be seen from these results that the performance of the ImgNet branch can greatly affect the final performance of the MT-CNN model.

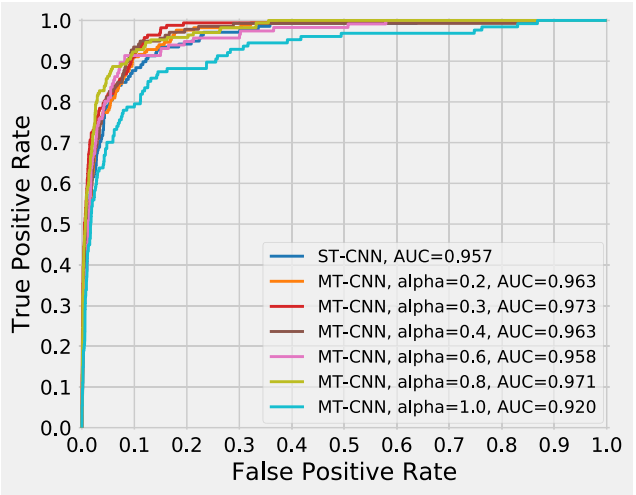

FIGURE 7. ROC curves of the proposed model with different image reconstruction loss weight in LUNA-16 dataset. If the weight equals to 0, the MT-CNN model is equivalent to ST-CNN model.

Compared MT-CNN model with ST-CNN model, although the MT-CNN model has more parameters and it can learn more characteristics due to the ImgNet branch, not all models perform better than ST-CNN model. Like threshold equals to 1, the performance of MT-CNN model much worse than ST-CNN model and it has the worst performance in all models. When the threshold equals to 1, the image reconstruction loss is greater than nodule classification loss, although we use mean binary cross-entropy loss function. And dropout only was applied to the first fully connect layer of nodule classification branch, there is no regularization in other parts. As a result, the model learns more features about image reconstruction and has an overfitting problem when training this model. But if we select a proper threshold for image reconstruction, the trained model learns more features used to identify malignant and benign nodules, and do not occur overfitting. Therefore, the performance of the MT-CNN model with proper image reconstruction loss threshold is far better than ST-CNN model.

V. DISCUSSION

In this paper, a novel approach for lung nodule classification using multi-task learning convolutional neural network is proposed. We use an image reconstruction task as the auxiliary task to improve classification performance of lung nodules. We conduct our experiments on PyTorch (the version is 1.3) and Tesla V100 GPU. Python 3.7 is used to implement this method. Compared with published methods that use 3-D CNN or multi-view for lung nodule classification, our proposed approach achieves the best performance on LIDC-IDRI dataset (see Table 1). These results indicate that a proper auxiliary task can improve the performance of the main task. Image reconstruction task allows the model to remember more microscopic information, such as shape, size and texture. Therefore, the image reconstruction task is suitable to solve the problem of identifying malignant and benign nodules. And these results suggest that it is important for MT-CNN methods to design an appropriate auxiliary task.

We applied a multi-task learning method to identify malignant and benign nodules. The advantages of the proposed approach mainly have two aspects: a suitable auxiliary task and multiple views extracted from candidate nodules. Compared with ST-CNN model, MT-CNN model can share features obtained from each task among multiple tasks. For each task, they learn different features from the same input images, due to their output and network structure are different. For example, the nodule classification model may learn more margin features of nodules. But for image reconstruction, model may learn more texture features of images. Therefore, we can learn more detailed features from an input image by integrating multiple related tasks.

In this paper, we extract nine views from different planes of a candidate nodule cube. And we construct a 2-D MT-CNN model for each view. We learn 3-D spatial information of nodules by fusing the nine 2-D models. Compared with 2-D images, although 3-D images include more spatial information, we need more computation and time to train a 3-D model. We solve the problem by fusing multiple 2-D models. Since we reduce the information that is not related to nodules, the fusion model produces representations with higher discriminative ability while reducing the amount of calculation and time. In our paper, the main contribution is image reconstruction branch to improve nodule classification performance. For different views, the prediction ability of the model is different, and we are able to get better prediction results by fusing different models. However, our aim is to compare the classification ability between single-task CNN model and multi-task CNN model, and the comparison of ensemble approach has been studied [18], [45]. Therefore we only use a weighted integration approach by assigning a big weight to the model with a strong classification ability, and we use AUC to represent the weight of each view.

CNN has been increasingly applied to medical image processing and many methods have shown excellent effects [46]. Numerous researches have demonstrated that the performance of 3-D CNN models is far better than the performance of 2-D CNN models. However, training a 3-D model is more difficult than train a 2-D model especially when high quality training samples are absent. It is difficult to train a 3-D model to handle the problem of lung nodule classification. In LUNA-16, we have 1,186 positive samples, but the number of negative samples is more than 500 times that of positive samples. Although we have used data augmentation to balance the number of positive and negative samples, it is difficult to obtain a high-quality balanced dataset.

In multi-task learning network, the sharing of low-level semantic information helps reduce the amount of computation. At the same time, the shared presentation layer enables several related tasks to better combine related information. Multi-task learning models must learn information that can represent all tasks, and therefore can help reduce overfitting problems of single task. Although we have used data augmentation methods to increase positive samples, positive samples are still far fewer than negative samples. For multi-task learning, it provides implicit data augmentation since different tasks have different noise patterns. A multi-task learning model that learns two tasks simultaneously can learn more general representations. Therefore, it is an effective way to reduce the impact of data imbalance by multi-task learning [23], [47].

In this paper, we focus on reducing false positives rather than developing a whole automatic lung nodule detection system. We assumed that candidate nodules have been detected. This means that we can combine our proposed approach with any candidate nodule detector by simply setting the output size of the detector to our input size. Therefore, the final results of our proposed method will depend on the performance of a candidate nodule detector. The proposed approach will achieve a better result if the provided candidate detector with higher performance. In contrast, the proposed approach will obtain a worse result if the provided candidate detector has low sensitivity.

To summarize, our proposed approach has achieved a good performance for lung nodule classification by using multi-task learning and combining multiple views. For each view, we train one MT-CNN model and finally obtain the final result by integrating the prediction probability from all views. There is no information communication between different views before fusion. According to our proposed approach, one interesting direction that might also improve performance is to share parameters and features among all the 2-D MT-CNN models.

VI. CONCLUSION

In this work, we present a multi-task learning convolutional neural network model for malignant and benign nodule classification on chest CT. In this method, an image reconstruction auxiliary task is exploited to improve the performance of nodule classification (main task) by learning more chest CT image detailed characteristics. The experiments results manifest the effectiveness of the proposed model. In the future, we plan to improve the proposed method in two aspects: (1) we will continue to improve the performance of lung classification system by exploring the relationships among different views. (2) We will try to evaluate the proposed approach to clinical data, and apply it to the clinic.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA, Cancer J. Clinicians*, vol. 69, no. 1, pp. 7–34, Jan./Feb. 2019.
[2] M. Tan, R. Deklerck, B. Jansen, M. Bister, and J. Cornelis, "A novel computer-aided lung nodule detection system for CT images," *Med. Phys.*, vol. 38, no. 10, pp. 5630–5645, Oct. 2011.
[3] M. Firmino, A. H. Morais, R. M. Mendoça, M. R. Dantas, H. R. Hekis, and R. Valentim, "Computer-aided detection system for lung cancer in computed tomography scans: Review and future prospects," *Biomed. Eng. OnLine*, vol. 13, no. 1, p. 41, 2014.
[4] J. Zhang, Y. Xia, H. Cui, and Y. Zhang, "Pulmonary nodule detection in medical images: A survey," *Biomed. Signal Process. Control*, vol. 43, pp. 138–147, May 2018.
[5] P. Monkam, S. Qi, H. Ma, W. Gao, Y. Yao, and W. Qian, "Detection and classification of pulmonary nodules using convolutional neural networks: A survey," *IEEE Access*, vol. 7, pp. 78075–78091, 2019.
[6] A. Teramoto, H. Fujita, O. Yamamuro, and T. Tamaki, "Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique," *Med. Phys.*, vol. 43, no. 6, pp. 2821–2827, May 2016.
[7] N. Tajbakhsh and K. Suzuki, "Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs," *Pattern Recognit.*, vol. 63, pp. 476–486, Mar. 2017.
[8] Y. Xie, J. Zhang, Y. Xia, M. Fulham, and Y. Zhang, "Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT," *Inf. Fusion*, vol. 42, pp. 102–110, Jul. 2018.
[9] S. Chen, J. Qin, X. Ji, B. Lei, T. Wang, D. Ni, and J.-Z. Cheng, "Automatic scoring of multiple semantic attributes with multi-task feature leverage: A study on pulmonary nodules in CT images," *IEEE Trans. Med. Imag.*, vol. 36, no. 3, pp. 802–814, Mar. 2017.
[10] L. Srensen, S. B. Shaker, and M. de Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 559–569, Feb. 2010.
[11] F. Zhang, Y. Song, W. Cai, Y. Zhou, M. Fulham, S. Eberl, S. Shan, and D. Feng, "A ranking-based lung nodule image classification method using unlabeled image knowledge," in *Proc. ISBI*, Apr. 2014, pp. 1356–1359.
[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. NIPS*, 2012, pp. 1097–1105.
[13] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep features learning for medical image analysis with convolutional autoencoder neural network," *IEEE Trans. Big Data*, early access, Jun. 20, 2017, doi: 10.1109/TBDATA.2017.2717439.
[14] F. Ciompi, B. de Hoop, S. J. van Riel, K. Chung, E. T. Scholten, M. Oudkerk, P. A. de Jong, M. Prokop, and B. V. Ginneken, "Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box," *Med. Image Anal.*, vol. 26, no. 1, pp. 195–202, Dec. 2015.
[15] X. Zhao, L. Liu, S. Qi, Y. Teng, J. Li, and W. Qian, "Agile convolutional neural network for pulmonary nodule classification using CT images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 4, pp. 585–595, Apr. 2018.
[16] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.
[17] X. Huang, J. Shan, and V. Vaidya, "Lung nodule detection in CT using 3D convolutional neural networks," in *Proc. ISBI*, Apr. 2017, pp. 379–383.
[18] A. A. A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. van Riel, M. M. W. Wille, M. Naqibullah, C. I. Sanchez, and B. van Ginneken, "Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1160–1169, May 2016.
[19] Q. Dou, H. Chen, Y. Jin, H. Lin, J. Qin, and P.-A. Heng, "Automated pulmonary nodule detection via 3D convnets with online sample filtering and hybrid-loss residual learning," in *Proc. MICCAI*, 2017, pp. 630–638.
[20] W. Zhu, C. Liu, W. Fan, and X. Xie, "DeepLung: Deep 3D dual path nets for automated pulmonary nodule detection and classification," in *Proc. WACV*, Mar. 2018, pp. 673–681.
[21] D. Ardila, A. P. Kiraly, S. Bharadwaj, B. Choi, J. J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D. P. Naidich, and S. Shetty, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Med.*, vol. 25, no. 6, pp. 954–961, Jun. 2019.
[22] Y. Xie, Y. Xia, J. Zhang, Y. Song, D. Feng, M. Fulham, and W. Cai, "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT," *IEEE Trans. Med. Imag.*, vol. 38, no. 4, pp. 991–1004, Apr. 2019.
[23] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017, *arXiv:1706.05098*. [Online]. Available: http://arxiv.org/abs/1706.05098
[24] S. G. A. Iii, G. Mclennan, L. Bidaut, M. F. Mcnittgray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, and E. A. Hoffman, "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, 2011.
[25] A. A. A. Setio *et al.*, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," *Med. Image Anal.*, vol. 42, pp. 1–13, Dec. 2017.
[26] D. Manos, J. M. Seely, J. Taylor, J. Borgaonkar, H. C. Roberts, and J. R. Mayo, "The lung reporting and data system (LU-RADS): A proposal for computed tomography screening," *Can. Assoc. Radiologists J.*, vol. 65, no. 2, pp. 121–134, May 2014.
[27] D. P. Naidich, A. A. Bankier, H. Macmahon, C. M. Schaefer-Prokop, M. Pistolesi, J. M. Goo, P. Macchiarini, J. D. Crapo, C. J. Herold, J. H. Austin, and W. D. Travis, "Recommendations for the management of subsolid pulmonary nodules detected at CT: A statement from the fleischner society," *Radiology*, vol. 266, no. 1, pp. 304–317, Jan. 2013.
[28] K. Murphy, B. van Ginneken, A. M. R. Schilham, B. J. de Hoop, H. A. Gietema, and M. Prokop, "A large scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification," *MedIA Image Anal.*, vol. 13, no. 5, pp. 757–770, 2009.
[29] C. Jacobs, E. M. van Rikxoort, T. Twellmann, E. T. Scholten, P. A. de Jong, J. M. Kuhnigk, M. Oudkerk, H. J. de Koning, M. Prokop, C. Schaefer-Prokop, and B. van Ginneken, "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images," *MedIA Image Anal.*, vol. 18, no. 2, pp. 374–384, 2014.
[30] A. A. A. Setio, C. Jacobs, J. Gelderblom, and B. van Ginneken, "Automatic detection of large pulmonary solid nodules in thoracic CT images," *Med. Physics.*, vol. 42, no. 10, pp. 5642–5653, 2015.

[31] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, and J. Tian, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," *Pattern Recognit.*, vol. 61, pp. 663–673, Jan. 2017.
[32] A. Ng, "Sparse autoencoder," in *Proc. CS294A Lect. Notes*, 2011, p. 72.
[33] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. AISTATS*, 2011, pp. 315–323.
[34] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015, pp. 1–41.
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. ICCV*, Dec. 2015, pp. 1026–1034.
[36] X. Yan, J. Pang, H. Qi, Y. Zhu, C. Bai, X. Geng, M. Liu, D. Terzopoulos, and X. Ding, "Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2d and 3d strategies," in *Proc. ACCV*, 2016, pp. 91–101.
[37] D. Kumar, A. Wong, and D. A. Clausi, "Lung nodule classification using deep features in CT images," in *Proc. CRV*, Jun. 2015, pp. 133–138.
[38] F. Han, H. Wang, G. Zhang, H. Han, B. Song, L. Li, W. Moore, H. Lu, H. Zhao, and Z. Liang, "Texture feature analysis for computer-aided diagnosis on pulmonary nodules," *J. Digit. Imag.*, vol. 28, no. 1, pp. 99–115, Feb. 2015.
[39] T. N. Shewaye and A. A. Mekonnen, "Benign-malignant lung nodule classification with geometric and appearance histogram features," 2016, *arXiv:1605.08350*. [Online]. Available: https://arxiv.org/abs/1605.08350
[40] X. Yan, J. Pang, H. Qi, Y. Zhu, C. Bai, X. Geng, M. Liu, D. Terzopoulos, and X. Ding, "Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2D and 3D strategies," in *Proc. ACCV*, 2016, pp. 91–101.
[41] R. Dey, Z. Lu, and Y. Hong, "Diagnostic classification of lung nodules using 3D neural networks," in *Proc. ISBI*, Apr. 2018, pp. 774–778.
[42] Y. Xie, J. Zhang, and Y. Xia, "Semi-supervised adversarial model for benign–malignant lung nodule classification on chest CT," *Med. Image Anal.*, vol. 57, no. 13, pp. 237–248, Oct. 2019.
[43] S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," *Expert Syst. Appl.*, vol. 128, pp. 84–95, Aug. 2019.
[44] Q. Dou, H. Chen, L. Yu, J. Qin, and P.-A. Heng, "Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1558–1567, Jul. 2017.
[45] B. Zhang, S. Qi, P. Monkam, C. Li, F. Yang, Y.-D. Yao, and W. Qian, "Ensemble learners of multiple deep CNNs for pulmonary nodules classification using CT images," *IEEE Access*, vol. 7, pp. 110358–110371, 2019.
[46] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1153–1159, May 2016.
[47] A. Argyriou, T. Evgeniou, and M. Pontil, "Multi-task feature learning," in *Proc. NIPS*, 2007, pp. 1–8.

PENGHUA ZHAI received the M.S. degree in automatic control from the Shenyang Institute of Automation, Chinese Academy of Sciences, in 2019. He is currently working with the Ningbo Institute of Life and Health Industry, University of Chinese Academy of Sciences. His research interests include machine learning, deep learning, and their applications in medical image analysis.

YALING TAO received the M.S. degree in immunology from the Institute of Zoology, Chinese Academy of Sciences (CASIZ), in 2016. She is currently working an Assistant Researcher with the HwaMei Hospital, University of Chinese Academy of Sciences. Her research interests include immunology, oncology, and medical imaging.

HAO CHEN received the M.S. degree from the University of Southampton, in 2019. He is currently working with the Ningbo Institute of Life and Health Industry, University of Chinese Academy of Sciences. His research interests include machine learning, deep learning, and their applications in medical image analysis.

TING CAI received the M.Med. degree in clinical medicine from Nankai University, China, and Flinders University, Australia, in 2012. He is currently the President of the HwaMei Hospital, University of Chinese Academy of Sciences, and an Executive Director of the Institute of Life and Health Industry, University of Chinese Academy of Sciences. He is also a Professor and the Ph.D. Supervisor with Ningbo University, China, and the Director of the Zhejiang Key Laboratory of Tumor Molecular Biology, China. He has presided 13 provincial and municipal research projects. He has published more than 30 articles and medical monographs. His research interests include cancer prevention and treatment and critical care system in ICU.

JINPENG LI received the B.E. and M.E. degrees in automatic control from the University of Science and Technology, Beijing (USTB), China, in 2012 and 2015, respectively, and the Ph.D. degree in pattern recognition and intelligent systems from the National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences (CASIA), in 2019. He is currently a Researcher with the HwaMei Hospital, University of Chinese Academy of Sciences, and the Institute of Life and Health Industry, University of Chinese Academy of Sciences. He has authored or coauthored more than ten peer-reviewed articles in prestigious international journals and conferences. His research interests include pattern recognition, machine learning, deep learning, transfer learning algorithms, and their applications in brain–computer interface and medical image analysis.