

# Classification of Malignant and Benign Lung Nodule and Prediction of Image Label Class using Multi-Deep Model

Muhammad Bilal Zia<sup>1</sup>, Zhao Juan Juan<sup>\*2</sup>  
Ning Xiao<sup>4</sup>, Jiawen Wang<sup>5</sup>, Ammad Khan<sup>6</sup>  
School of Information and Computer  
Taiyuan University of Technology, Taiyuan, China

Xujuan Zhou<sup>3</sup>  
School of Management and Enterprise  
University of Southern Queensland  
Toowoomba, Australia

**Abstract**—Lung cancer has been listed as one of the world's leading causes of death. Early diagnosis of lung nodules has great significance for the prevention of lung cancer. Despite major improvements in modern diagnosis and treatment, the five-year survival rate is only 18%. Before diagnosis, the classification of lung nodules is one important step, in particular, because automatic classification may help doctors with a valuable opinion. Although deep learning has shown improvement in the image classifications over traditional approaches, which focus on handcraft features, due to a large number of intra-class variational images and the inter-class similar images due to various imaging modalities, it remains challenging to classify lung nodule. In this paper, a multi-deep model (MD model) is proposed for lung nodule classification as well as to predict the image label class. This model is based on three phases that include multi-scale dilated convolutional blocks (MsDc), dual deep convolutional neural networks (DCNN A/B), and multi-task learning component (MTLc). Initially, the multi-scale features are derived through the MsDc process by using different dilated rates to enlarge the respective area. This technique is applied to a pair of images. Such images are accepted by dual DCNNs, and both models can learn mutually from each other in order to enhance the model accuracy. To further improve the performance of the proposed model, the output from both DCNNs split into two portions. The multi-task learning part is used to evaluate whether the input image pair is in the same group or not and also helps to classify them between benign and malignant. Furthermore, it can provide positive guidance if there is an error. Both the intra-class and inter-class (variation and similarity) of a dataset itself increase the efficiency of single DCNN. The effectiveness of mentioned technique is tested empirically by using the popular Lung Image Consortium Database (LIDC) dataset. The results show that the strategy is highly efficient in the form of sensitivity of 90.67%, specificity 90.80%, and accuracy of 90.73%.

**Keywords**—Lung nodule classification; dilated blocks; dual DCNNs; multi-task learning; multi-deep model

## I. INTRODUCTION

Lung cancer is the world's most prevalent and deadly type of cancer. The failure to diagnose the early stages of lung cancer is one cause of higher mortality induced by lung cancer

because symptoms usually appear in the final stages [1]. In 2018, an estimated 154,000 deaths were recorded in the USA from lung cancer, which is 1/4 of all cancer death. In developed countries, there is a 16% chance for lung cancer patients of a five-year survival rate [2]. The detection of lung cancer in the initial phase is difficult due to dime-sized lesion growth. The small lesions can just detectable by computerized tomography (CT) scan, and it takes the amount of effort of radiologists to detect and label them as benign or malignant. Computer-aided diagnosis (CAD) systems help the radiologist to decrease the burden as a second option. Lung Nodule classification is also a challenging task in computer-aided diagnosis.

Nowadays, deep learning, which is rebranded from neural networks, is considered as one of the best solutions to solve many strenuous problems of computer vision and pattern recognition like medical diagnostics, natural language processing, etc.

Even though deep learning techniques perform better in the state-of-art in different medical images tasks, but still, there are many challenges to solve. For instance, facing issues with small medical dataset and biological variations. In [16,17], the pre-trained DCNN architectures have been used due to robust learning capability from large scale datasets like ImageNet to solve the generally small amount of data visual recognition problems.

The main issue of classifying lung nodules is inter-class ambiguity and intra-class variations [18], which pose complex challenges in different modalities in the differentiation of benign lesions from malignant ones. A clear example that shows the complication is shown in Fig. 1. In which there is a massive difference between (a) benign lesions, and (b) malignant lesions. Although both benign and malignant lesions are similar, respectively, in both color and shape. Even though several neural networks are productive sufficient to memorize all the training samples [19] due to useful ability of deep learning models, the uncertainty formed by intra-class variation and inter-class similarity can disturb the neural networks and make them fall into misperception.

\*Corresponding Author

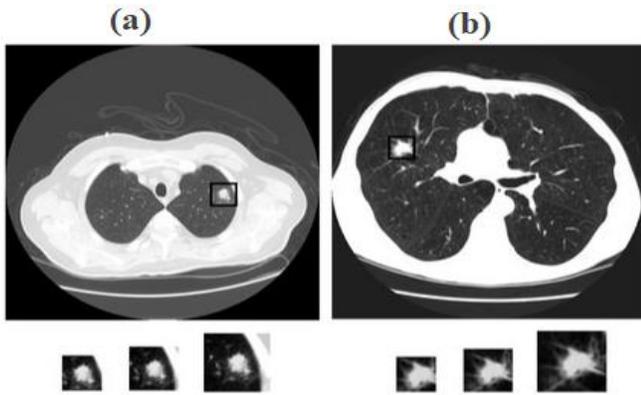


Fig 1. Example of (a) Benign nodule and (b), Malignant nodule.

To address the present challenges, in this paper, a multi-deep model is used for classification of a lung nodule in LIDC datasets, which consists of multi-scale dilated convolutional layer blocks, dual DCNNs, and multi-task learning component. The multi-scale method is applied on LIDC datasets, which helps to view and figure out the dataset at different scales. It worked as a speckle noise filter by eliminating the high-frequency constituent at every scale where no edges were spotted. Moreover, the dual pre-trained DCNNs extract the parameters from the ImageNet dataset [20], and these parameters were fine-tuned with our dataset. Both dual DCNNs simultaneously learn the representation of images from a pair of images, along with two similar images in various categories and two different images in the same category due to strong learning ability. The multi-task learning technique is applied on input pair of images where it will classify the multi-scale dilated Convolutional layers (MsDc) paired images which belong to the similar class or not and also helps to classify between malignant and benign. In addition, the multi deep model is easily trained “end-to-end” in classification supervision. During the test stage, each sample is given with the precision probabilities of a dual neural network together as the probability for a joint decision. The MD model is tested on LIDC datasets. Hence it is proved that the proposed model is state-of-the-art on lung nodule classification problem

## II. RELATED WORK

Nowadays deep learning is helping in many medical fields to improve image classification accuracy. From the last few years, a number of solutions are being published to solve the image classification problems [3, 4, 5, 6, and 7]. These solutions contain handcrafted feature extraction and classifier learning process. For just a classification task, it is very difficult to design such a handcrafted feature. So, deep learning models help for superior classification and to minimize the need for manual feature design. It also helps to enhance the performance of medical image detection, classification [8, 9], and segmentation [10, 11]. For instance, Jung H et al. [12] adopt a 3D-DCNN with dense connection and shortcut connection technique for lung nodule classification using LUNA 16 datasets to capture the 3D features. Xu Y et al. [13] use a deep convolutional neural

network algorithm in their paper to minimize the manual annotation and produce superior feature portrayal for classification and segmentation of colorectal cancer using Cellular pathology. In machine learning, multi-task learning techniques are used in many applications like in drug detection [21], speech recognition [22], spatio-temporal event forecasting [28] and in natural language processing [23], etc. Li X et al. [14] proposed a multi-task learning framework that captures all the lung nodule assortment by taking out all the distinctive features using Convolutional Neural Network (CNN) from alternately stacked layers. For improvement in the final result, they train the CNN and form multi-tasking learning which shares information among nine different nodule features at the same time. Wu Z et al. [15] developed a multi-scale convolutional neural network (CNN) for removing lesion surface from CT scans which rely on 3D context fusion named M3DCF. An improved Multi-scale algorithm is used by Qingyuan [24] for image enhancement by using a canny operator to divide the image into edge region and non-edge region zone. Wang Y et al. [25] introduced a succinct and powerful multi-scale dilated convolutional method which used the dilated filters to integrate situational multi-scale data without reducing the receptive field efficiently. The logic behind this approach was to focus on the phenomenon that the dilated convolutional can efficiently extend the correct receptive area while retaining the useful contextual information. In the meantime, they also use residual methods to improve the learning process. Another dilated convolutional approach is used by Wei Y [26] in which they used several multiple dilated convolutional blocks with different dilated levels to create dense position maps of objects for weakly or semi-supervised manner for semantic segmentation networks. It has been also found that dilated convolutional [27], which by increasing the respective field size of kernels, offers a promising solution. In which they used different dilated levels and generated the localization maps at these dilated rates to enhance the discriminative ability.

## III. METHOD

### A. Multi-Deep Model

The proposed multi-deep (MD) model contains three main modules, i.e. Input block, the dual-DCNN unit, as shown in Fig. 2 and Multi-task learning component. The MD architecture takes a pair of images that are arbitrarily selected from the training data. The Dual DCNN contain DCNN-A and DCNN-B, which is the main two sequence learning module. The DCNN-A is pre-trained with the VGG16 network, and DCNN-B is pre-trained with ResNet50, and also both are fine-tuned to monitor the correct input sequence labels. The input block contains a pair of images where multi-scale dilated convolutional layers (MsDc) strategy is applied to both pairs. Then these pairs of images individually enter into DCNN-A and DCNN-B. A multi-task learning component is not only used to predict whether these pairs belong to the same class or not, but it also classifies the image pairs between benign and malignant. Furthermore, this component also helps to provide positive guidance if there is a synergistic mistake from both DCNNs.

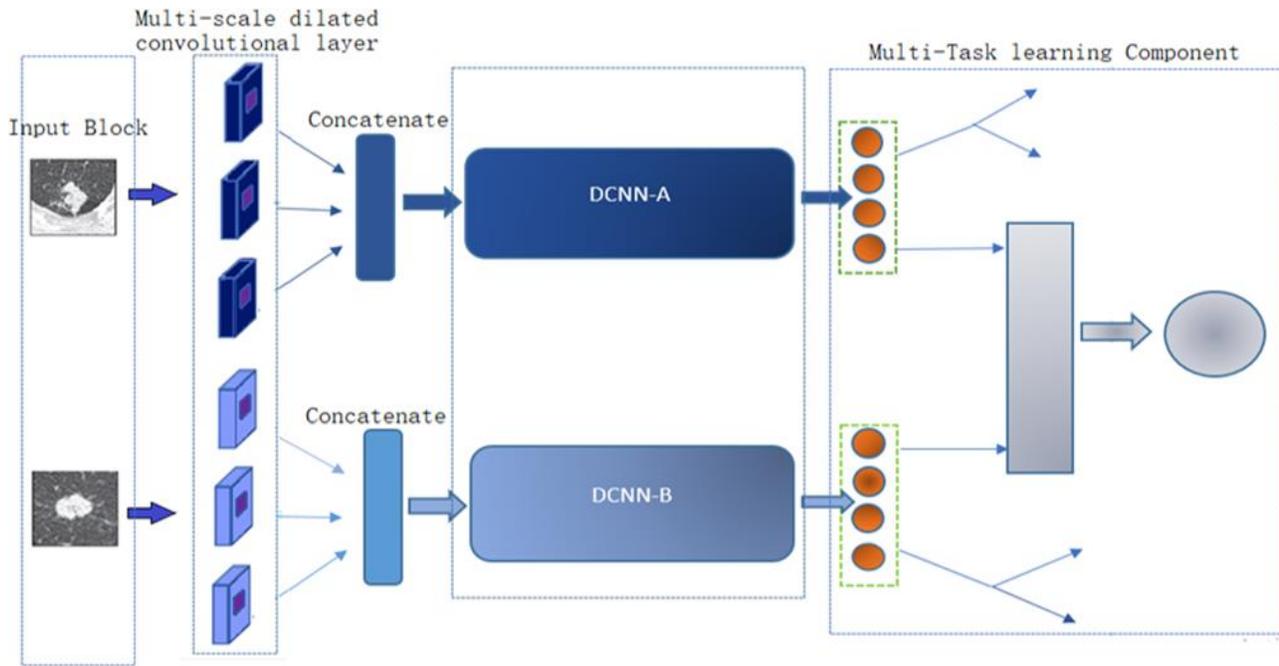


Fig 2. Architecture of Proposed MD Model.

### B. Input Block and Multi-scale Dilated Convolutional Layer

Not quite the same as the traditional DCNNs, the suggested model randomly selects the MsDc pair of images as input from training data. Each image has its class label, and then it transfers to each DCNN. Each image is redimensioned to  $224 \times 224 \times 1$  by using a bicubic interpolation algorithm to unify the image size. Inside the input block, there are pair of grayscale images as input. The MsDc strategy is applied to the paired images before getting into DCNN (A/B). After employing this technique, the 3 channels image with different scale dilated convolution operation. Then 3 one channel images got from different receptive fields to one three-channel image is concatenated as the input of DCNN (A/B) as presented in Fig. 3. This approach uses different scale dilated filters to integrate the multi-scale contextual information systematically by extending the receptive area of the convolutional layers. The logic behind this approach is focused on the phenomenon that the dilated convolution could effectively extend the correct receptive area while maintaining useful contextual information. The MsDc composed of different dilation rates, which leads to various receptive filed within the input images. So, the dilation rate 3, 5, and 7 is used in MsDc. In Fig. 4, it is clearly shown how dilation allows the transfer of information. For the classification process to recognize this as a “Lung cancer” image, the circle region in the green cycle is most discriminating. To learn the corresponding feature representation at the area shown by the red cycle, a  $3 \times 3$  convolutional kernel is implemented.

### C. The Dual Deep Convolutional Neural Network

Within this suggested model, the dual DCNN is an essential component with two full training units, DCNN-A and DCNN-B. Although an arbitrary structure of any DCNN such as GoogleNet or AlexNet can be implemented as a DCNN part in the recommended method. Both DCNNs is

trained using  $X = \{x(1), x(2), \dots, x(M)\}$  image sequence and the corresponding  $Y = \{y(1), y(2), \dots, y(M)\}$  label class. For DCNN-A, a pre-trained VGG-16 model is used, which has thirteen convolutional layers and three fully-connected (FC) layers. Then a pre-trained ResNet 50 is adopted to initialize the DCNN-B because of the high representation capability of the popular residual network. It is concocted of 50 learnable layers, and it also trained for classification tasks on ImageNet datasets. In order to adopt the above models to our dataset, all fully connected (FC) layer is supplant to FC of 1024 neurons for DCNN(A/B) and then fine-tuning the ResNet-50 and VGG16 parameters by utilization our own training data. During optimization, the parameters of DCNN (A/B) are not shared mutually, which is denoted by  $\theta_A$ , and  $\theta_B$ . The uniform  $U(-0.05, 0.05)$  distribution of the weights of new FC layers is initialize. Both DCNNs is defined as the cross-entropy loss function as:

$$l(\theta) = \frac{1}{Q} \sum_{i=1}^Q [\log(\sum_{j=1}^k e^{\theta_j^t x^{(i)}}) - \theta_{y^{(i)}}^t x^{(i)}] \quad (1)$$

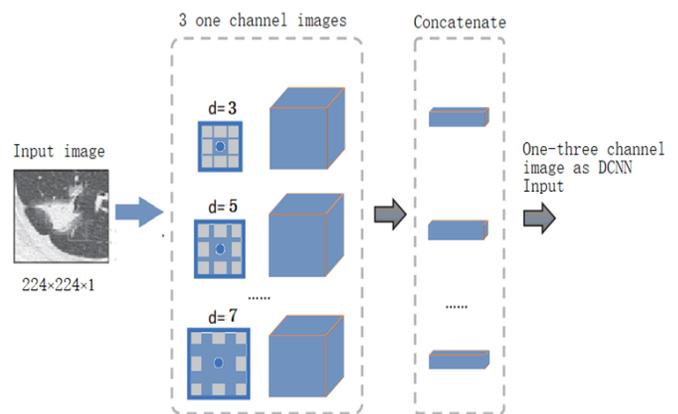


Fig 3. Structure of Multi-Scale Dilated Convolutional Layer Block.

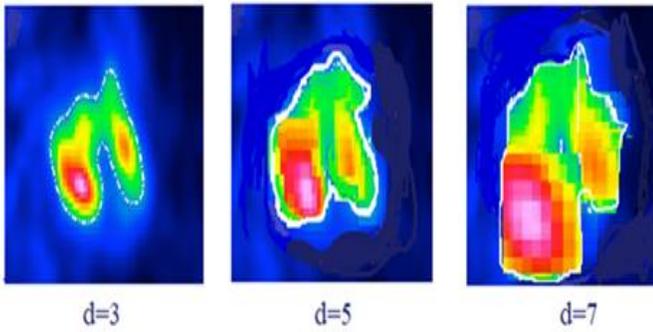


Fig 4. Reason for this Approach: the information can be moved through changing Dilated Levels from an Originally biased Area to other Regions.

Where  $Q$  is the number of training data. The Adam optimizer (mini-batch Adam) is used to optimize the  $\theta$ . Input from a pair of images is accepted by both DCNN, which seek to control the training process with the true labels of the respective input sequence in each learning unit. While both DCNN has the ability to determine the input image label class, so, a multi-task learning component is produced that breaks the learning independency of the dual DCNNs by integrating activations from the last two FC layers in both DCNNs.

#### D. Multi-task Learning Component

To further track the training of each pair of images of each DCNN part, a multi-task learning component (MTLc) is designed, which consists of fully connected layers and vector concatenation layer, as shown in Fig. 5. Let the DCNNs components (A/B) have an image pair  $(x_c, x_d)$ , respectively. The second to last fc layer performance in the DCNN is described as the deep image features learned from that DCNN that can be accomplished through forwarding computation, as shown formally:

$$f_c = F(x_c, \theta(i)) \quad (2)$$

$$f_d = F(x_d, \theta(i)) \quad (3)$$

In the training data, the image pairs are arbitrary selected and denoted the attribute of a pair  $(x_c, x_d)$  as:

$$B(x_c, x_d) = \begin{cases} 1 & \text{if } y_c = y_d \\ 0 & \text{if } y_c \neq y_d \end{cases} \quad (4)$$

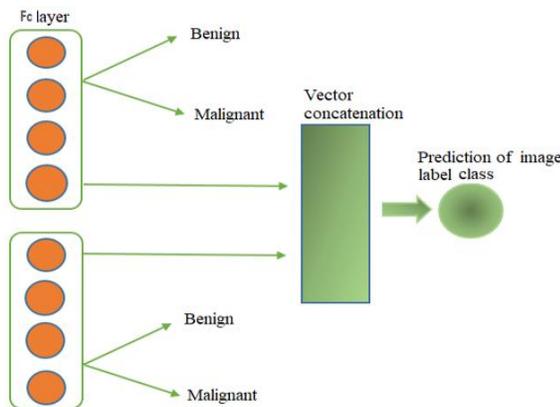


Fig 5. Diagram of Multi-Task Learning Component.

Where  $x_c$  and  $x_d$  are lung nodule images,  $y_c$  and  $y_d$ , respectively represent the true labels of  $x_c$  and  $x_d$  (like benign or malignant). The number of positive pairs, which is  $S=1$  in the LIDC datasets is approximately 45% - 55% to prevent the unbalance data issue. After the output got from both DCNN A/B, two 1024 dimensional FC vectors, we copy and break them into 2 pieces for the multi-task learning part. In order to improve classification across learning tasks, the aim of multi-task is to accomplish mutual training by leveraging dependencies in the functionality in order to improve the performance of one task using the other. Moreover, softmax function is used as the non-linear activation function for the final prediction layer and use ReLU function as the non-linear activation for other fully connected layers. For each multi-task component, the hidden layers have 1000 neurons to solve the more complex problem. Furthermore, by using MTLc technique the two tasks is intended for prediction. The first task distinguishes among malignant and benign tumor while the other task will concatenate the vector from another DCNN network to predict the image label class. The analysis of the organized signal is useful to expedient the following binary entropy loss of the MTLc.

#### IV. EXPERIMENT

This section includes details of the implementation and analysis of the proposed model as well as a comparison of the experimental results. Section “A and B” illustrates how the hyperparameters are initiated and it also describes the dataset used in experiments. Section “C” presents the experiment on the LIDC dataset. The impact of  $\lambda$  is mainly tested on different values and compares the classification accuracy of the proposed model, and it also describes the comparison of MD model with several recently effective methods.

##### A. Dataset and Hyperparameter Setting

The LIDC dataset consists of 1018 CT scan patients of pulmonary cancer, and these CT images are divided into five categories. Detailed information for each nodule (diameter, coordinate, malignancy, texture, etc.) is indicated by 4 professional radiologists. The diameter of the nodules is between 3 mm and 30 mm. The interpolation spline technique is used as a method to separate with  $1 \text{ mm} \times 1 \text{ mm}$  due to the difference in resolution. In this study, the presumption is examined in the malignancy of nodules. Each nodule has an annotated malignancy by radiologists ranking from 1 to 5. As the final decision, the voting strategy is considered. If the nodule value was over 3 annotated by more than two radiologists, the nodule is considered as malignant. In contrast, the nodule is deemed benign. There are approximately 195 malignant nodules and 158 benign nodules. The nodules with the same votes are rejected. The central transect for each nodule voxel is extracted to minimize computational complexity. To alleviate the deep learning model overfitting, the Data Augmentation (DA) method is applied to enlarge the data by adding variants to a dataset. In particular, by zooming 0.2, the image is randomly flip and magnify. The selection of the translation step is from  $[-6, 6]$  voxels, and the angle of rotation was selected randomly from  $[90^\circ, 180^\circ, \text{ and } 270^\circ]$ . Finally, there are 1956 malignant nodules, 1862 benign nodules.

### B. Hyperparameter Setting

For hyperparameter setting the number of iteration steps is set to 12000, and the learning rate is modified using an exponential decay process, and the initialize learning rate  $r = 0.0002$ , with the learning rate attenuation of 0.95. The mini-batch Adam was acknowledged as an optimizer with a batch size 64. To stop the training procedure when the model is overfitting, 20% of training data were randomly selected to form a validation set, which was used to monitor the performance of prospective model.

As the training model, Keras is used to distinguish the benign and malignant lung nodules and Tensorflow as Keras backup. Python versions 2.7–3.6 are available in Keras. And it offers the following benefits: strong modularity, reduced simplicity, and scalability.

### C. Experimental Result and Analysis

To evaluate the MD model performance on the LIDC dataset, first, the efficiency of the MD model is assessed by taking different values from the hyperparameter  $\lambda$ . The results in Fig. 6 indicate the best performance of suggested model with  $\lambda = 15$ . Then the classification accuracy of the proposed MD model is compared with the same experimental setting against the VGG-16, ResNet-50 model, and the proposed MD model is also evaluated without multi-scale dilated convolutional layer. The classification accuracies of the above models in each DCNN (A/B) are displayed in Fig. 7, and it is clearly shown that the recommended model attains the highest accuracy of 90.73% on the validation set. This reveals that each component of designed model, which is also VGG-16, ResNet-50 achieves an improvement in the precision of over more than 3% relative to the standard VGG-16 and ResNet-50 norm since integrating the multi-task learning component into a dual-DCNN architecture.

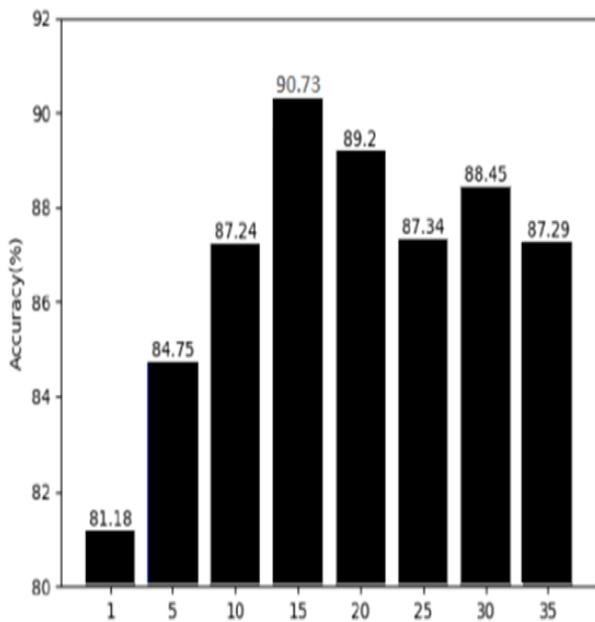


Fig 6. MD Model Performance with different  $\lambda$  Values.

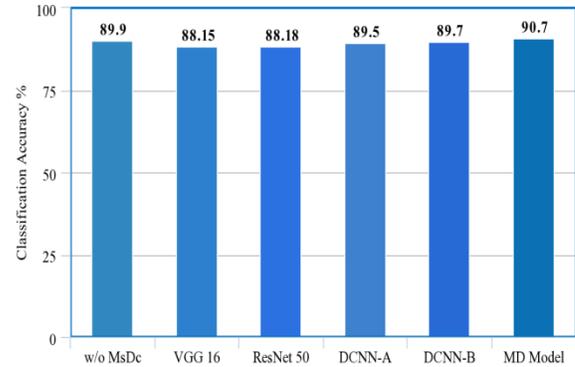


Fig 7. Classification Accuracy of Proposed MD Model w/o MsDc, with VGG 16, ResNet 50, Dual DCNNs on the Validation set.

To assess the designed model's performance against the related methods, it is also compared to certain other effective methods with recent good results shown in Table I. Kumar et al. [29, 30] introduced a new ensemble system for the classification of medical images. The ensemble method has used many advanced CNNs as tailored extractors for image features that have captured the different information in medical images of different types. Zhang J et al. [31] have suggested a dual deep-convolutional neural network that is equipped with a synergic signal network to learn the representation of the image jointly and as well as a synergic signal system helps to verify the pair of the image belongs to the same category or not. Yuhai et al. [32] used an ensemble of 5 pretrained VGG and ResNet- 50 models, and 5 completely trained DCNN models, the ImageCLEF-2013 data set is augmented, and the baseline ResNet-50 model is improved by far. A hierarchical learning mechanism called Multi-scale Convolutional Neural Network (MCNN) [33] which is used to explore nodule complexity by eliminating discriminatory features from alternating stacked layers. The architecture uses multi-sized nodule patches to learn a collection of class features at the same time through concatenating response neuron activations of each input level at the last layer. For the classification of lung nodule [34], Haralick, Gabor, and LBP (local binary patterns) features are extracted and also implemented SVM classifier. They achieved 89.5% sensitivity and 86.02% specificity, respectively. Dhara et al. [35] based on the edge, shape, and textures features to ensure a benign and malignant classification.

However, the manual feature set is insufficiently suited as to whether the difference between the various pulmonary nodules types can be described accurately. Devinder Kumar et al. [36] suggested a CAD method utilizing deep functions separated from an auto-encoder for classifying pulmonary nodules as malignant or benign. In the analysis of 9 semantic features in CT images for lung nodules, Chen et al. [37] utilized three multi-task learning (MTL) systems for the use of various computational features extracted from deep learning systems of the convolutional neural network (CNN) and stacked denoising auto-encoder.

TABLE I. PERFORMANCE COMPARISONS WITH OTHER STATE-OF-ART METHODS FOR LUNG NODULE CLASSIFICATION TASK

Methods	Sensitivity %	Specificity %	Accuracy %
Davinder K et al. [36]	83.3	-	75.0
Kumar et al.[30]	-	-	77.5
Kumar et al.[29]	-	-	82.48
Zhang et al. [31]	-	-	86.58
Chen et al. [37]	60.3	95.4	86.8
Shen et al. [33]	-	-	86.84
Shen et al. [38]	77.0	93.0	87.1
Yuhai et al. [32]	-	-	87.37
Han et al. [34]	89.3	86.02	-
Dhara et al. [35]	89.7	86.3	-
<b>Proposed MD Model</b>	90.67	90.80	90.73

## V. CONCLUSION

In this paper, a Multi-deep model is proposed for lung nodule classification and to resolve the problem posed by intra-class variation and inter-class similarity. In the first step, the MsDc method is applied to pair of images, which could help to increase the performance of lung nodule classification afterward our technique uses dual DCNNs with a multi-task learning component to allow dual DCNNs to learn from one another. It promotes the ability of the suggested model to distinguish between interclass samples that are easily ignored and the clear diversity of intra-class samples. The experimental result on the LIDC-IDRI dataset demonstrates that the proposed model attains the state-of-the-art performance in the lung nodule classification problem.

## REFERENCES

[1] Zhao, B., Gamsu, G., Ginsberg, M. S., Jiang, L., & Schwartz, L. H. (2003). Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. *Journal of applied clinical medical physics*, 4(3), 248-260.

[2] Valente, I. R. S., Cortez, P. C., Neto, E. C., Soares, J. M., de Albuquerque, V. H. C., & Tavares, J. M. R. (2016). Automatic 3D pulmonary nodule detection in CT images: a survey. *Computer methods and programs in biomedicine*, 124, 91-107.

[3] Lazebnik, S., Schmid, C., & Ponce, J. (2006, June). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 2, pp. 2169-2178). IEEE

[4] Fei-Fei, L., & Perona, P. (2005, June). A bayesian hierarchical model for learning natural scene categories. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 2, pp. 524-531). IEEE.

[5] Wang, J., Yang, J., Yu, K., Lv, F., Huang, T., & Gong, Y. (2010, June). Locality-constrained linear coding for image classification. In 2010 IEEE computer society conference on computer vision and pattern recognition (pp. 3360-3367). IEEE.

[6] Wang, Z., Hu, Y., & Chia, L. T. (2013). Learning image-to-class distance metric for image classification. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(2), 1-22.

[7] Yang, Y., Yang, L., Wu, G., & Li, S. (2012, October). A bag-of-objects retrieval model for web image search. In Proceedings of the 20th ACM international conference on Multimedia (pp. 49-58).

[8] Sirinukunwattana, K., Raza, S. E. A., Tsang, Y. W., Snead, D. R., Cree, I. A., & Rajpoot, N. M. (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE transactions on medical imaging*, 35(5), 1196-1206.

[9] Xie, F., Fan, H., Li, Y., Jiang, Z., Meng, R., & Bovik, A. (2016). Melanoma classification on dermoscopy images using a neural network ensemble model. *IEEE transactions on medical imaging*, 36(3), 849-858.

[10] Li, R., Zeng, T., Peng, H., & Ji, S. (2017). Deep learning segmentation of optical microscopy images improves 3-D neuron reconstruction. *IEEE transactions on medical imaging*, 36(7), 1533-1541.

[11] Chen, H., Qi, X., Yu, L., Dou, Q., Qin, J., & Heng, P. A. (2017). DCAN: Deep contour-aware networks for object instance segmentation from histology images. *Medical image analysis*, 36, 135-146.

[12] Jung, H., Kim, B., Lee, I., Lee, J., & Kang, J. (2018). Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method. *BMC medical imaging*, 18(1), 48.

[13] Xu, Y., Mo, T., Feng, Q., Zhong, P., Lai, M., Eric, I., & Chang, C. (2014, May). Deep learning of feature representation with multiple instance learning for medical image analysis. In 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 1626-1630). IEEE.

[14] Li, X., Kao, Y., Shen, W., Li, X., & Xie, G. (2017, March). Lung nodule malignancy prediction using multi-task convolutional neural network. In *Medical Imaging 2017: Computer-Aided Diagnosis* (Vol. 10134, p. 1013424). International Society for Optics and Photonics.

[15] Wu, Z., Chen, J., Wang, Z., Su, J., & Cai, G. (2019, November). Multi-scale Convolutional Neural Network Based on 3D Context Fusion for Lesion Detection. In Chinese Conference on Pattern Recognition and Computer Vision (PRCV) (pp. 573-585). Springer, Cham.

[16] Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and transferring mid-level image representations using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1717-1724).

[17] Mettes, P., Koelma, D. C., & Snoek, C. G. (2016, June). The imagenet shuffle: Reorganized pre-training for video event detection. In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval (pp. 175-182).

[18] Song, Y., Cai, W., Huang, H., Zhou, Y., Feng, D. D., Wang, Y., ... & Chen, M. (2015). Large margin local estimate with applications to medical image classification. *IEEE transactions on medical imaging*, 34(6), 1362-1377.

[19] Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2016). Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*.

[20] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.

[21] Ramsundar, B., Kearnes, S., Riley, P., Webster, D., Konerding, D., & Pande, V. (2015). Massively multitask networks for drug discovery. *arXiv preprint arXiv:1502.02072*.

[22] Deng, L., Hinton, G., & Kingsbury, B. (2013, May). New types of deep neural network learning for speech recognition and related applications: An overview. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 8599-8603). IEEE.

[23] Collobert, R., & Weston, J. (2008, July). A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th international conference on Machine learning (pp. 160-167).

[24] Meng, Q., Bian, D., Guo, M., Lu, F., & Liu, D. (2012). Improved multi-scale retinex algorithm for medical image enhancement. In *Information Engineering and Applications* (pp. 930-937). Springer, London.

[25] Wang, Y., Wang, G., Chen, C., & Pan, Z. (2019). Multi-scale dilated convolution of convolutional neural network for image denoising. *Multimedia Tools and Applications*, 78(14), 19945-19960.

[26] Wei, Y., Xiao, H., Shi, H., Jie, Z., Feng, J., & Huang, T. S. (2018). Revisiting dilated convolution: A simple approach for weakly-and semi-

- supervised semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7268-7277).
- [27] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv:1412.7062.
- [28] Gao, Y., Zhao, L., Wu, L., Ye, Y., Xiong, H., & Yang, C. (2019, July). Incomplete Label Multi-Task Deep Learning for Spatio-Temporal Event Subtype Forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, pp. 3638-3646). Kumar, A., Lyndon, D., Kim, J., & Feng, D. (2016, September). Subfigure and Multi-Label Classification using a Fine-Tuned Convolutional Neural Network. In CLEF (Working Notes) (pp. 318-321).
- [29] Kumar, A., Kim, J., Lyndon, D., Fulham, M., & Feng, D. (2016). An ensemble of fine-tuned convolutional neural networks for medical image classification. IEEE journal of biomedical and health informatics, 21(1), 31-40.
- [30] Kumar, A., Lyndon, D., Kim, J., & Feng, D. (2016, September). Subfigure and Multi-Label Classification using a Fine-Tuned Convolutional Neural Network. In CLEF (Working Notes) (pp. 318-321).s
- [31] Zhang, J., Xia, Y., Wu, Q., & Xie, Y. (2017). Classification of medical images and illustrations in the biomedical literature using synergic deep learning. arXiv preprint arXiv:1706.09092.
- [32] Yu, Y., Lin, H., Meng, J., Wei, X., Guo, H., & Zhao, Z. (2017). Deep transfer learning for modality classification of medical images. Information, 8(3), 91.
- [33] Shen, W., Zhou, M., Yang, F., Yang, C., & Tian, J. (2015, June). Multi-scale convolutional neural networks for lung nodule classification. In International Conference on Information Processing in Medical Imaging (pp. 588-599). Springer, Cham.
- [34] Han, F., Wang, H., Zhang, G., Han, H., Song, B., Li, L., ... & Liang, Z. (2015). Texture feature analysis for computer-aided diagnosis on pulmonary nodules. Journal of digital imaging, 28(1), 99-115.
- [35] Dhara, A. K., Mukhopadhyay, S., Dutta, A., Garg, M., & Khandelwal, N. (2016). A combination of shape and texture features for classification of pulmonary nodules in lung CT images. Journal of digital imaging, 29(4), 466-475.
- [36] Kumar, D., Wong, A., & Clausi, D. A. (2015, June). Lung nodule classification using deep features in CT images. In 2015 12th Conference on Computer and Robot Vision (pp. 133-138). IEEE.
- [37] Chen, S., Qin, J., Ji, X., Lei, B., Wang, T., Ni, D., & Cheng, J. Z. (2016). 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(3), 802-814.
- [38] Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., & Tian, J. (2017). Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recognition, 61, 663-673.