This is the Pre-Published Version.

The following publication A. Masood, B. Sheng, P. Li, P. Yang and J. Kim, "Automatic Computer Aided System for Lung Cancer in Chest CTs Using MD-RFCN Combined with Tri-Level Region Proposal Network," 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 2019, pp. 1377-1380 is available at https://doi.org/10.1109/INDIN41052.2019.8972133.

Automatic Computer Aided System for Lung Cancer in Chest CTs Using MD-RFCN Combined with Tri-Level Region Proposal Network

Anum Masood
Department of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
anummasood@sjtu.edu.cn

Bin Sheng
Department of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
shengbin@sjtu.edu.cn

Ping Li
Department of Computing
The Hong Kong Polytechnic University
Kowloon, Hong Kong
lipingfire@ieee.org

Po Yang
Department of Computer Science
Liverpool John Moores University
Liverpool, U.K.
poyangcn@gmail.com

Jinman Kim
School of Information Technologies
The University of Sydney
Sydney, Australia
jinman.kim@sydney.edu.au

***Abstract*—Pulmonary cancer is one of the major causes of deaths caused by cancer around the globe. Early stage lung cancer detection can prove to be essential for the patients, for which the computed tomography (CT) images are analyzed by the radiologists to determine the presence of nodules and diagnose the disease. Conventional techniques used by the radiologists for nodule detection in CT images is time-consuming and inefficient; to assist in the diagnosis process and further enhance its efficiency and accuracy, decision support systems have been developed in the past few years. In our paper, we proposed a Multi-Dimension Region-based Fully Convolutional Network based decision support system for detection and classification of lung nodule. The Multi-Dimension RFCN serves as an image classifier backbone for our feature extraction step in addition to the proposed Tri-Level Region Proposal Network (3L-RPN) along with the position-sensitive score maps (PSSM) being explored. A novel median intensity projection method is used to leverage the multi-dimensional information from CT images and introduced an additional deconvolutional layer to adopt the proposed Tri-Level Region Proposal Network in our architecture to automatically identify the potential Region of Interest. We trained and evaluated our proposed decision support system using LIDC-IDRI dataset. The evaluation results demonstrated the high level performance of our proposed model in comparison to the state-of-the-art nodule detection and classification methods by attaining classification accuracy of 97.61% and sensitivity of 97.4%.**

***Index Terms*—Lung cancer, nodule classification, computer aided systems, convolutional neural network.**

This work was supported in part by the National Natural Science Foundation of China under Grant 61872241 and Grant 61572316, in part by the National Key Research and Development Program of China under Grant 2017YFE0104000 and Grant 2016YFC1300302, in part by the Macau Science and Technology Development Fund under Grant 0027/2018/A1, and in part by the Science and Technology Commission of Shanghai Municipality under Grant 18410750700, Grant 17411952600, and Grant 16DZ0501100.

I. Introduction

Pulmonary cancer is known to be one of the major causes of death around the globe. About 1.6 million deaths are caused by lung cancer within a year [1]. Main cause of lung cancer is the abnormal lung cell's uncontrollable growth. Computed tomography (CT) scan is the most effective method for early stage pulmonary nodule detection. CT scan has high resolution 3D chest images which are rich in spatial information. Early stage detection of lung cancer can increase the survival rate of the patient but it is time-consuming and complex task for radiologists to mark the nodule position without missing any overlapping nodules. Owing to the advancements in the industrial applications for lung cancer detection and diagnosis, the lung cancer death rate is comparatively reduced. Various commercially available computer-aided (CAD) systems have potential to improve the accuracy of the detection and the diagnosis process. CAD systems for pulmonary cancer are further distributed into two categories; Computer Aided Detection (CADe) and Computer Aided Diagnosis (CADx). The CADe aims to distinguish between the nodule candidates and the non-nodule (anatomical structure such as blood vessels, tissues) whereas the CADx characterize these detected nodules and classifies these into benign or malignant nodules based on their malignancy. The purpose of these CAD systems is to effectively improve the accuracy of cancer diagnosis done by the radiologists while reducing the overall CT images interpretation time.

The recent advancements in the IoT enabled healthcare technologies [2], [3] and deep learning methods prompted researchers to develop CAD systems based on these techniques for the real-world clinical usage. These methods have significantly improved the efficiency and quality of the healthcare sector especially the screening process for the early stage lung cancer detection. As compare to the other cancer types, advancements in pulmonary cancer are slow and unsteady regarding the survival rate but deep learning methods have promising results. The lung cancer detection methods based on deep learning techniques have reduced the death rate by the factor of 22% in the last five years. In this paper, we have proposed a novel CAD system for lung cancer detection as well as classification. Our work makes two key contributions:

- We proposed a novel multi dimensional deep convolutional neural network based model for early stage lung cancer detection of lung cancer having the capability of using multi-dimensional spatial and contextual information providing broad range of discriminating feature map for nodule candidates detection.
- A Tri-Level Region Proposal Network (3L-RPN) is proposed for selection of potential Region of Interest (RoI) along with the position-sensitive score maps (PSSM) for enhanced and efficient nodule stage classification. A novel median intensity projection method is used to leverage the multi-dimensional information from CT images and introduced an additional deconvolutional layer in our architecture to automatically identify the potential RoIs.

II. Related Work

A CADe system developed for lung cancer generally has three steps; dataset pre-processing, nodule candidate detection and false positive reduction. The dataset pre-processing is used for the standardization of the CT image by spatially resampling the CT images to homogeneous resolution whereas segmentation of lung is done to optimize and limit the search space. Traditionally lung segmentation is done using 2D geometrical level set active contour, morphological features, 2D parametric deformable model, voxel clustering and multi gray-level thresholding. The nodule candidate detection phase detect candidate nodules with maximum sensitivity which leads to imbalance among the False Positives (FPs) and True Positives (TPs). The last step is the FP reduction phase which reduces the number of FPs nodules among the selected nodule candidates and produces the final detected nodules. Supervised learning approaches are developed to reduce the FPs but these methods have high computational cost [4]. Convolutional Neural Network (CNN) [5], Fully Convolutional Networks (FCN) [6], Multi-Crop Convolutional Neural Network (MC-CNN) [7], and DFCNet [8] are CAD systems based on deep learning which are capable of detecting nodules and classifying the detected nodules as benign or malignant tumors.

III. Method

A. Dataset Pre-Processing

In case of Computer Aided Diagnosis (CAD) system, the screening stage is the most crucial phase. We combined three neighboring CT scans for each axis direction, respectively. The potential nodule candidates search space was further analysed by selecting training sets of CT images. To mitigate the overfitting problem of the training dataset due to limited labeled dataset, we used data augmentation techniques. Each Region of Interest (RoI) undergoes affine transformation generating huge amount of correlated newly acquired training data samples. We augmented malignant samples by scaling, rotating, cropping, and flipping in the labeled training dataset.

B. Architecture

1) Multiview Combination: The CT images are gray-scale images which make the image classification problem different from the traditional image classification where input channels are identical as the color channels of the image. CT images are originally three dimensional where z-axis to discriminate various positions of lung nodule, while the input image sets are 2D patches. We explore Median Intensity Projection (MIP) proposed by Hessian [9] to combine information from various dimensions of CT scans. Image projected by Median Intensity Projection is presented by θ for three dimensions with input image I, θ as:

$$\begin{aligned}\theta_{(b,c)} &= \underset{a}{med}\, I_{(a,b,c)}\\ \theta_{(a,c)} &= \underset{b}{med}\, I_{(a,b,c)}\\ \theta_{(a,b)} &= \underset{c}{med}\, I_{(a,b,c)}\end{aligned} \tag{1}$$

where med shows the median operator.Multiple views provided different information, while patches in combination with multiple dimensions gave the space distribution of tumor tissues. To construct input image with three channels, we combined three Median Intensity Projection images: $\theta = [\theta_{(b,c)}, \theta_{(a,c)}, \theta_{(a,b)}]$.

2) Tri-Level Region Proposal Network : Previously for the selection of Region of Interest (RoI), original CT scans are distributed into small sample windows: 8×8, and 32×32. For our Tri-Level Region Proposal Network, we selected the smaller sample window, such that the spatial window contains the center of the malignant nodule as per labeled by the radiologists. This technique had a disadvantage that in case the specific nodule is larger than pre-defined sample window or the targeted nodule was located such that it overlaps two sample windows, then the targeted nodule was be discarded. This resulted in the inaccurate selection of training dataset or testing dataset, and finally affected the classification performance. To solve this problem, we proposed a novel Tri-Level Region Proposal Network (3L-RPN). In the 3L-RPN, we used three level of Region Proposal Network; which generate rectangular object proposals set using the input image irrespective of its size, and further retrieve objectness value for each proposal set [10]. Malignant nodule occupies small region of any CT image therefore the traditional techniques for object recognition are not applicable. We introduced a deconvolutional layer after last feature layer, 4 kernel size and 4 stride size, to maintain the original CT image size by upsampling (see Fig. 1 for our proposed system). In Tri-Level Region Proposal Network we choose anchor sizes of 4×4, 8×8, 12×12, 20×20, 26×26, 32×32, respectively (see Fig. 2) to contain nodules of different malignant level.

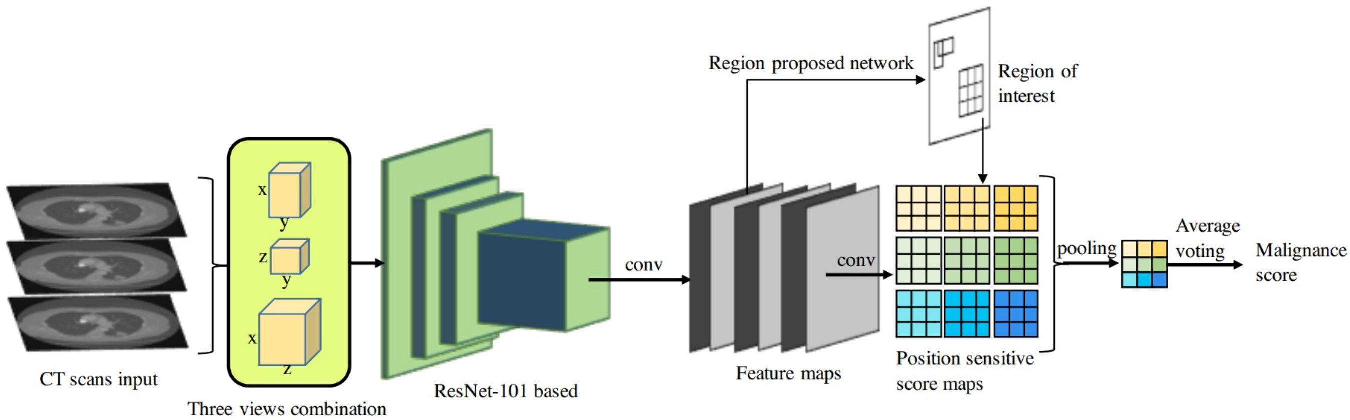

Fig. 1: Overview of the Multi-Dimension Region-based Fully Convolutional Network architecture.

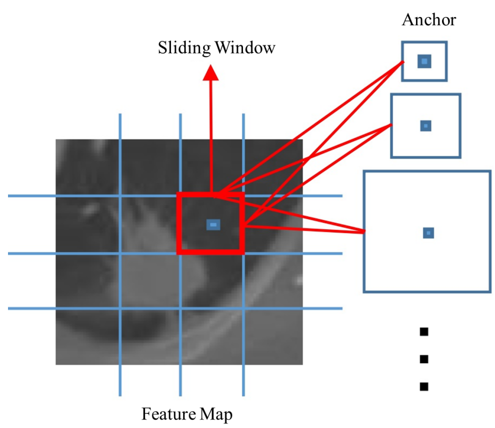

Fig. 2: Anchor in Tri-Level Region Proposal Network.

3) Multi-Dimension Region-based Fully Convolutional Network (MD-RFCN) : With Tri-Level Region Proposal Network (3L-RPN) proposing RoI, we apply Region-based Fully Convolutional Network [11] to predict the presence of lung nodule or its malignancy level if a nodule is detected. The architecture of MD-RFCN has ResNet-101 [12] as backbone, containing 3×3 filters with the convolutional network. In original ResNet-101, in case the filters and layers are comparable then the size of the feature map is the same but for maintaining the time complexity of every layer, the number of the filters are doubled when the activation map size is reduced to half. In our Multi-Dimension Region-based Fully Convolutional Network, we excluded the fully convolutional layer as well as the mean pooling layer for more efficient feature maps layer while improving computation cost. Following the annotations from the radiologists, we used a $k \times k(5+1)$-channel convolutional layer (4×4) as the output layer to produce position sensitive score maps (PSSM) to leverage position sensitive score in each grid. PSSM phase is almost cost-free as there is no supervised learning layer after the RoI layer.

C. Loss Function

For the optimization of training process to provide region proposals, we defined our loss function by using focal loss $\mathbf{FL}$ instead of cross-entropy loss function which is defined as:

$$\mathbf{FL} = -\sum_{m=1}^{M}\sum_{c=1}^{C}\sum_{j=1}^{N}\mathbf{d}_x(1-q)^{\gamma}\log q \tag{2}$$

where, M, C and N are the mini-batch size, number of classes and number of pixels, respectively. d is a vector and $\mathbf{d}_x = 1$ where x denotes ground truth values. The term $(1-q)^{\gamma}$ controls the contrast of loss value. As the γ becomes larger then the contrast between loss value of benign and malignant classes become larger as well. But if $\gamma = 0$ then the focal loss equals to cross-entropy loss.

D. Implementation Details

The network was initialized using LIDC-IDRI pre-trained basic network and weight of layers of our proposed model. In the initial step, we freeze all basic model layers with only training layers of Multi-Dimension Region-based Fully Convolutional Network combined with Tri-Level Region Proposal Network (3L-RPN). In the next step, the whole proposed model is trained in two phases by decreasing learning rate. Input images were normalized and then MD-RFCN settings were employed. The experiments were conducted on Ubuntu 16.04.3 LTS with four processors, Intel® Xeon® CPU E5-2686 and 64GB memory space. The model was trained on Tesla K-80 along with 12 GB memory. For MD-RFCN, we used Intel Extended Caffe and common libraries like sklearn 0.18.2, SimpleITK 1.1.0, numpy 1.13.1, and pandas 0.19.2. Training process used standard back-propagation using stochastic gradient descent (SGD) with momentum of 0.12, weight decay of 0.00045, and 0.02 learning rate with 15 increasing factor after 400 iterations.

IV. Results and Discussions

Approximately 1.59 million deaths are caused by lung cancer. CAD systems have been developed to help radiologists

TABLE I: CAD System's Comparison to Detect Lung Nodule based on Leave-One-Out Validation Method.

Classifiers	Accuracy (%)	Sensitivity (%)	FPs/Scan
CNN [5]	80.8	79.6	3.63
MTANNs [13]	88.6	86.53	2.62
FCN [6]	91.2	91.14	3.16
MC-CNN [7]	87.14	77.3	2.97
MD-RFCN	92.1	94.4	2.21
MD- RFCN (Using 3L-RPN)	97.61	97.4	2.39

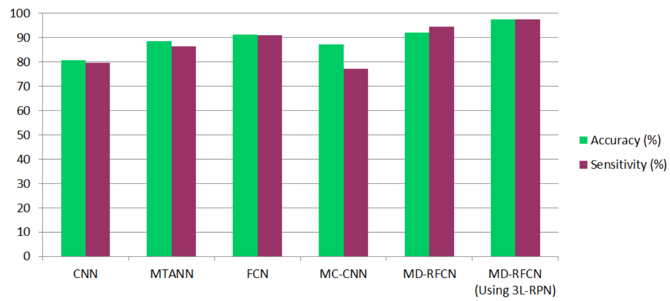

Fig. 3: Performance Evaluation of proposed MD-RFCN (Using 3L-RPN) with state-of-the-art CAD systems.

in detection and diagnosis of lung cancer. To develop a reliable CAD system, CAD system should have high performance in terms of accuracy, diagnosis speed, and low error rate. In state-of-art, the prevalent criteria for performance evaluation of CAD system are sensitivity and false positive per scan (FP/scan). For the testing of our proposed CAD system on LIDC dataset we used Leave-one-out Validation method in order to validate our CAD classifier's performance. We randomly divided the CT scans from LIDC-IDRI dataset into training sets and testing sets, training sets were used to train our CAD system and testing sets were used to validate its performance. The effectiveness of MD-RFCN is verified by comparing with state-of-the-art methods CNN [5], MTANNs [13], FCN [6], and MC-CNN [7], the results are depicted in Table I.

From Table I, we can deduce that MD-RFCN's performance is highest attaining an accuracy of 94.4% and 92.1% sensitivity with a lowest FP per scan rate of 2.21 among the existing state-of-art CAD system. The results indicate the superiority of MD-RFCN as CAD system for nodule detection. The comparison between "MD-RFCN" and "MD- RFCN (Using 3L-RPN)" showed that the proposed CAD system with Tri-Level Region Proposal Network setting performs better for RPN in nodule candidate selection both in terms of accuracy and sensitivity, which are recorded to be 97.4% and 97.61%, respectively. The experimental results presented in Fig. 3 demonstrate the superiority in generalization of our proposed MD-RFCN combined with 3L-RPN in comparison to existing CAD systems.

V. Conclusion

Main purpose of a CAD system designed for lung cancer is to assist the radiologist in detection of nodule and serve as a second opinion. In this research paper, we have proposed a novel CAD system based on the Multi-Dimension Region-based Fully Convolutional Network where we applied novel median intensity projection to obtain useful spatial information from the nodule dataset combining multi dimensional views. We proposed a Tri-Level Region Proposal Network (3L-RPN) in our architecture to improve the performance of RPN. Our CAD system is capable of indicating the presence of nodule, its location and outlines the possible structure of the nodule. LIDC-IDRI dataset was used to train and evaluate the proposed CAD system. Experimental results showed the high performance of our proposed system attaining classification accuracy of 97.61% and the sensitivity of 97.4%.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7–30, 2018.
[2] J. Qi, P. Yang, G. Min, O. Amft, F. Dong, and L. Xu, "Advanced internet of things for personalised healthcare systems: A survey," *Pervasive and Mobile Computing*, vol. 41, pp. 132–149, 2017.
[3] P. Yang, D. Stankevicius, V. Marozas, Z. Deng, E. Liu, A. Lukosevicius, F. Dong, L. Xu, and G. Min, "Lifelogging data validation model for internet of things enabled personalized healthcare," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 1, pp. 50–64, 2016.
[4] 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," *Medical Image Analysis*, vol. 42, pp. 1–13, 2017.
[5] M. Tan, R. Deklerck, B. Jansen, M. Bister, and J. Cornelis, "A novel computer-aided lung nodule detection system for CT images," *Medical Physics*, vol. 38, no. 10, pp. 5630–5645, 2011.
[6] B. van Ginneken, A. A. A. Setio, C. Jacobs, and F. Ciompi, "Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans," in *IEEE International Symposium on Biomedical Imaging*, pp. 286–289, 2015.
[7] 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 Recognition*, vol. 61, pp. 663–673, 2017.
[8] A. Masood, B. Sheng, P. Li, X. Hou, X. Wei, J. Qin, and D. Feng, "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images," *Journal of Biomedical Informatics*, vol. 79, pp. 117–128, 2018.
[9] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset, "Model based multiscale detection of 3D vessels," in *Workshop on Biomedical Image Analysis*, pp. 202–210, 1998.
[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *NIPS*, pp. 91–99, 2015.
[11] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object detection via region-based fully convolutional networks," in *NIPS*, pp. 379–387, 2016.
[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE CVPR*, pp. 770–778, 2016.
[13] D. Kumar, A. Wong, and D. A. Clausi, "Lung nodule classification using deep features in CT images," in *Conference on Computer and Robot Vision*, pp. 133–138, 2015.