3D Inception Convolutional Neural Networks For Automatic Lung Nodule Detection

Chen Zhao
College of Computer science and technology
Xi'an University of Posts and Telecommunications
Xian, China
awp4211@gmail.com

Jungang Han
College of Computer science and technology
Xi'an University of Posts and Telecommunications
Xian, China
hjg@xupt.edu.com

Yang Jia
College of Computer science and technology
Xi'an University of Posts and Telecommunications
Xian, China
jiayang@xupt.edu.com

Abstract—**Lung cancer is the most common cause of cancer death worldwide. Early detection of lung nodules in thoracic computed tomography (CT) scans is currently the one of the most effective ways to predict and treat lung cancer. In practice, one CT image contains about 200 to 700 slices and this may cost radiologists a lot of time. In this paper, we propose two types of deep neural networks which are called MODEL-Ⅰ and MODEL-Ⅱ for automatically detecting lung nodule to help the radiologists with reading CT images. For the CT data pretreatment, we use filters to select the suspicious region for locating nodules in 2D images. we use downsampling and upsampling methods to make dataset balanced. Inspired by Google 2D Inception module, we propose the inception block for 3D convolutional neural networks to accommodate the 3D nature of CT scans, which solve the gradient vanish problems and enhance the F1 score, experiment show that The MODEL-Ⅱ acquires the best F1 score of 0.979.**

Keywords- Lung Nodule Detection; 3D Inception CNN; Data Preprocessing; CAD System

I. INTRODUCTION

Lung cancer is one of the main cause of mortality worldwide, with approximately 1.59 million deaths per year [1]. Early detection of lung nodules plays an important part in diagnosing and clinical managing of lung cancer. When the lung nodules are still in a treatable stage and radiologist can detect them, it will give more chance to survive the patient. In reality, it is difficult for radiologists with less experience to identify the nodules in a CT scan image, because there are large number of slices in one CT scan and the nodules have divers intensity [3] and many different morphological features such as solid, non-solid.

In last two decades, radiologists and computer vision researchers have developed some traditional approaches which generally require manually designed features such as morphological features, voxel clustering and pixel thresholding [22].

In the last couple of years, inspired by the remarkable progress in nature image classification, image semantic segmentation such as those outcome in the famous competition named ImageNet [4] and Pascal VOC [5] by using the convolutional neural network, various Computer Aided Diagnosis (CAD) system for lung nodule detecting have been proposed [2,3,8,9,10,11,12], but there are only a few CAD systems use 3D convolutional neural network. A comparison on 2D and 3D convolutional networks is conducted and shown that 3D convolutional network is superior than 2D convolutional network for 3D CT data [23].

In this paper, we provide a novel framework for automatic detection of nodules in CT images. We use selective search [6] and R-CNN [7] styled method to give a rough estimation of region in 2D slice images, which the noudle may locate in and then we train our 3D CNN neural networks to identify whether the candidate nodule is a malignant nodule or not. Our models get high accuracy, precision and F1 score on both training dataset and test dataset. The dataset is a relatively large dataset contains 800 CT scans with precise annotation from the experts.

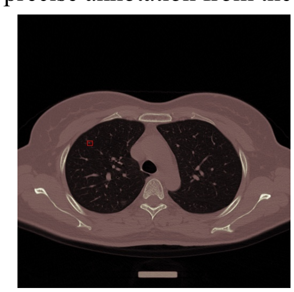


Figure 1. A close-up of a malignant nodule from lung dataset.

II. RELATED WORK

A. CAD system

In the past two decades, a large number of CAD systems have been proposed which can be categorized by traditional machine learning method and newly deep learning method. Bergtholdt used a cascaded support vector machine (SVM) to build a pulmonary nodule detection system [8]; van Ginneken developed a CAD system using 2D CNN to extract features and SVM to classify nodules [9]; Ji-kui Liu synthesizes the random forest and ensemble method to assist radiologists detect pulmonary nodules and acquire 93.2, 92.4, 94.8, 97.6% of accuracy, senstivity, specificity, area under the curve(AUC), respectively [10]; Rotem Golan designed a 3D CNN styled CAD system by sliding window scanning to find the candidate nodule's region proposal [11].

B. *Convolutional Neural Networks*

Convolutional neural network (CNN) is a type of feed forward neural network. CNN is widely used in image feature extraction while the raw image needs very little pretreatment work [12], and CNN rules over the image classification competition such as ImageNet, Pascal VOC 2007-2012 and MS COCO. CNN is an effective approach on image segmentation, a success case is the U-Net [13] for biomedical image segmentation. When training a CNN network, back propagation algorithm is often used and sometime gradient vanish often happens. Deep Residual network [14] plays an important role in handling the gradient vanish problem. In recent years, 3D CNN has successfully applied in vedio classification [15]. CNN is a good feature extractor compared to traditional image feature extraction method such as SIFT, SURF and Laplacian Pyramid. A key question in our challenge is whether the success deep CNN networks in nature images tasks can also apply to medical image process tasks.

C. *Object Detection*

Recently deep Convolutional nets have significantly improved in object detection [7, 16]. Compared to image classification, it is difficult to find object in images. In many object detection tasks, the first step is to offer region proposals. Example includes selective search [6]. Selective search uses [17] to create initial regions, calculates the similarities between the regions and group the most similar regions. Finally selective search can generate some coarse grained object regions. R-CNN [7] performs a convnet to forward pass the proposal regions and use a SVM classifier to classify whether the region contains an object.

III. 3D CNN LUNG NODULE DETECTION ALGORITHM

In this section, we will describe our steps of data preprocessing, model setting and training trick. At first, we introduce the dataset.

A. *Dataset Introduction*

The dataset has 800 3D CT images and have been manually divided into training set with 600 images and test set with other 200 images. In training set there are 975 malignant nodules and each nodule is annotated with the position (3D position represented by coordX, coordY and coordZ) in full images and diameter (mm). The nodules sizes' distribution is shown in Table.1.

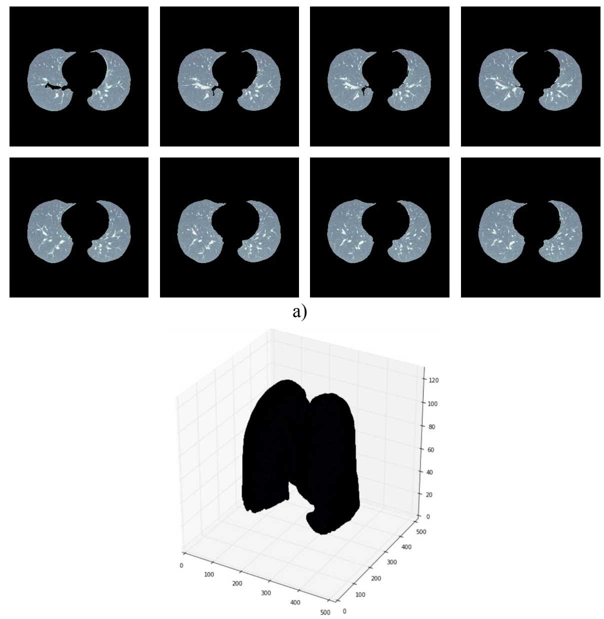
TABLE I. THE DISTRIBUTION OF MALIGNANT NODULES SIZE

Distribution	***5-10mm***	***10-30mm***
Distribution	50%	50%

The CT images have the size of SLICES*512*512, where SLICES range from about 100 to 700 and implies the number of the slices in one CT image.

B. *CT Data Pretreatment*

In order to make computer easy to find the volumes surrounding suspicious nodule, we first resize each voxels in CT images into $1mm^3$ volume. The HU value of voxel in mhd files represents ranges from -2048 to 400, where -2048 represents the area with no CT scan. Air has the value of -1000 and bone has the value of 400. We take a window to map float value (0, 1) to the HU values (-1000, +400). Meanwhile, we only focus on the area with HU value greater than -320 to segment lung parenchyma so that we can get the mask of lung parenchyma binary images. The segmentation result is shown in Fig.2.



a)

b)

Figure 2. Lung segmentation results. a) 2D slices for lung parenchyma; b) 3D lung parenchyma binary mask.

C. *Proposal of Candidate Nodules Region*

We tried to use 3D U-NET [13] to segment nodules. But the largest nodule has only a 30*30*30 mm^3 volume compared to the entire lung 3D image with a 400*400*400 mm^3 volume. And the U-NET architecture can hardly search for a needle in a haystack. So we turned to use the selective search algorithm to obtain suspicious region. Selective search algorithm produce diverse regions in 2D slices and a number of regions produced are not the candidate nodules. We employ filter to remove the most fake nodules according to the two strategies: a) Only the areas larger than 5*5 pixels and smaller than 30*30 pixels in 2D slices are preserved ; b) selective search will return the rectangles with the object inside and the ratio of the object area to rectangle area. The ratio below 0.1 and above 0.8 are removed. An example of selective search result is shown in Fig.3.

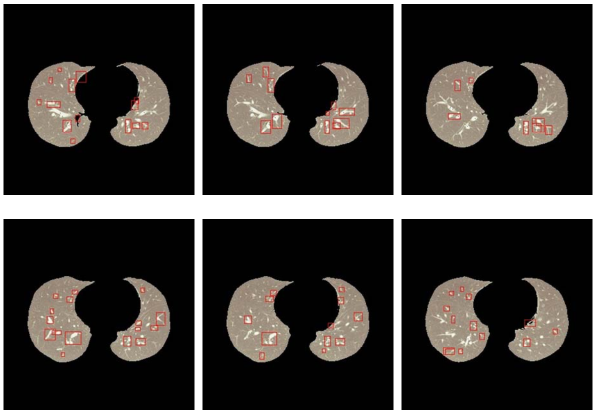


Figure 3. Selective search result for suspicious nodules.

In order to create training dataset, we first make positive nodule masks from the annotated nodules. Each cube surrouding the annotated nodule with size of 48*48*48 mm^3 is cropped from lung CT image (the cube is larger than the suspicious nodules) such that the center of nodule is coincide the center of the cube.

As we known, it is difficult to train a dataset with unbalanced type of data. If we don't balance the positive and negative samples, the neural networks will learn to be biased to negative samples. The original dataset only has 975 annotated nodules, To augment the positive data set we rotate the nodules in z axis from 10 degree to 350 degree with step of 10 degree.

The proposed regions in Fig.3 are 2D images. To create negative 3D sample for training 3D-CNN, we crop 3D volume with size of 48*48*48. Let (x, y, z) being the center of the volume, where (x, y) is the 2D center of related 2D regions and z is the index of the slice containing the 2D region. Note that we have to remove the volumes coincide with positive samples and randomly down sample the created negative samples with the probability of 0.2.

At last, the training data set includes 35100 positive volumes with label 1 and 42748 negative volumes with label 0. In the same manner, we create 21714 volumes as test dataset. These volumes will be sent to neural networks described below for training and testing.

D. 3D CNN Network Structure

We design two types of 3D CNN networks named MODEL-Ⅰ and MODEL-Ⅱ. MODEL-Ⅰ contains 6 3D CNN layers, 2 Maxpool layers and 3 Fully Connected layers FC1, FC2 and FC3. Inspired by the Inception structure [18], we design MODEL-Ⅱ with residual convolutional block and spatial reduction block, both of the blocks use inception structure. The details of network architecture are shown in Fig.4. for MODEL-Ⅰ and Fig.5. for MODEL-Ⅱ. To prevent overfitting, both networks use Dropout layer [19]. At the end of each convolutional layers we use ReLU [20] as activation function.

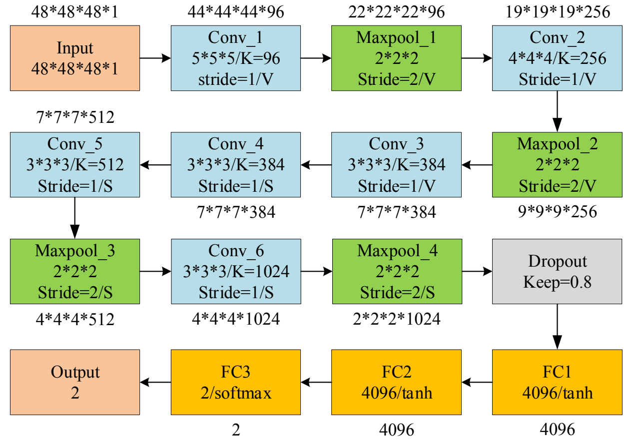


Figure 4. Architecture of MODEL-Ⅰ.

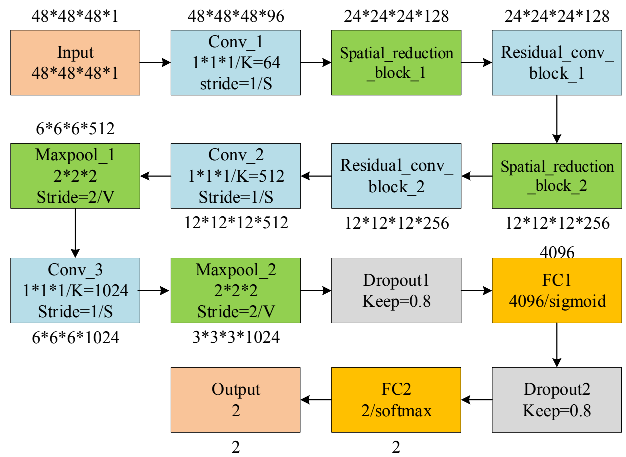
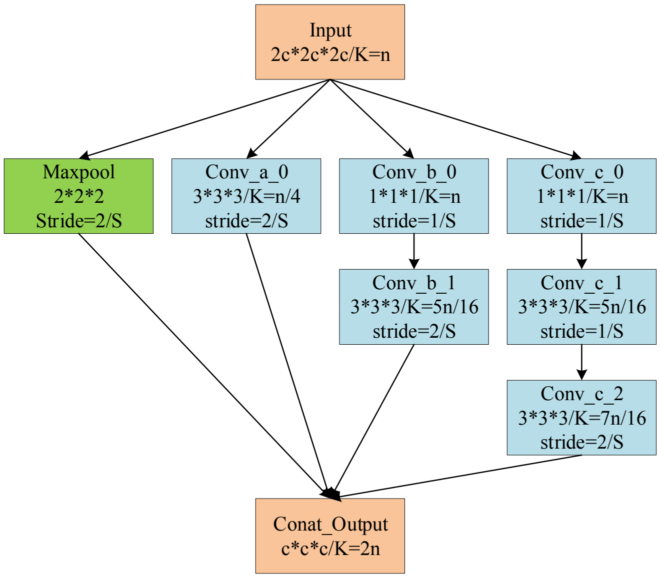
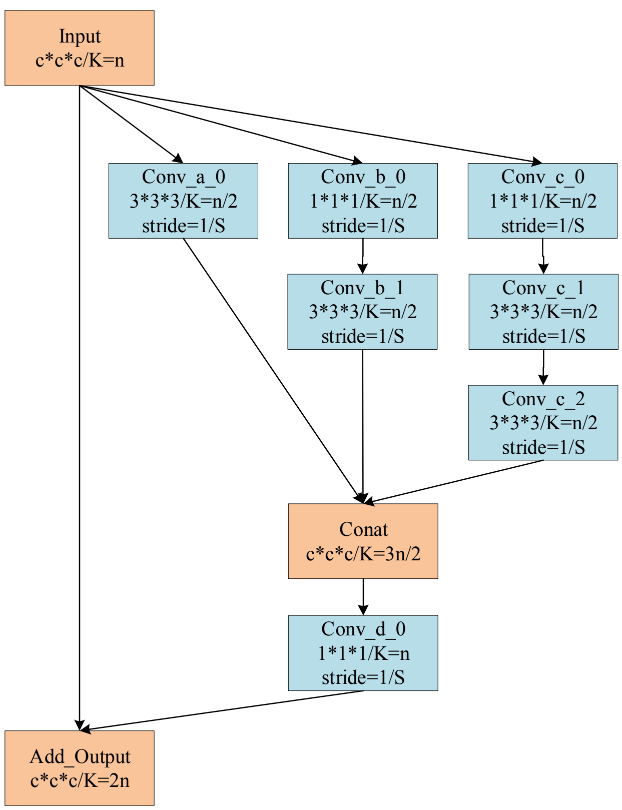


Figure 5. Architecture of MODEL-Ⅱ.

In Fig.4, 5 and 6, the 'K' in Conv block indicates the number of convolutional kernels; the 'S' and 'V' in Conv and Maxpool block indicate the padding method of 'SAME' and 'VALID' respectively. Fig.6. shows the detail of the spatial reduction block and residual conv block in Fig.5.



a)



b)

Figure 6. a) the architecture of spatial reduction block, b) the architecture of residual convolutional block.

IV. Experiments And Conclusions

In this section, we describe the hyper parameters settings, training evaluation, training details and training results. Finally we give the conclusion of the 3D CNN modeling for lung nodule detection.

A. Hyper parameter setting

In order to systemically explore the effect of proposed model on lung cancer dataset and get higher evaluation scores, we use a large scale grid search of different hyper parameters. We list the ranges of search in Table II.

TABLE II. Grid Search For Hyper Parameters

Parameter	***model***	***learning rate***	***dropout***
Value	[MODEL-Ⅰ,MODEL-Ⅱ]	[1e-4, 1e-6]	[0.5, 0.8]

To evaluate the performance of our model, we define the true positive (TP) value to be the number of lung nodules that is annotated in annotation file, FP denotes the number of false positive and FN denotes number of false negative nodules. The F1 score is the harmonic average of precision value and recall value:

$$F1 = 2TP / (2TP + FP + FN) \quad (1)$$

In the experiment, we chose the hyper parameters with best F1 score rather than accuracy as the final parameters settings.

B. Training Details And Results

We adopted Adam Optimizer to train our neural networks. The least square is used as lose function to metric the true labels and predict labels. All the code was implemented with the most popular deep learning library TensorFlow 1.0.0. Our code is available at https://github.com/awp4211/lung. The batch size was set according to the GPU memory: for MODEL-Ⅰ, the batch size is 30; for MODEL-Ⅱ, the batch size is 15. Each model runs 20 epochs with the speed of 2 epochs per hour for MODEL-Ⅰ and 1 epoch per hour for MODEL-Ⅱ on a single GTX 1080 from Nvidia.

The test results are shown in Table.3 and the training process is shown in Fig. 7.

TABLE III. Evaluation Values On Test Data

Model	***learning rate***	***dropout***	***F1 score***	***accuracy***
MODEL-Ⅰ	1e-4	0.5	**0.973**	93.3%
	1e-4	0.8	0.964	**93.6%**
	1e-6	0.5	0.969	92.0%
	1e-6	0.8	0.923	90.0%
MODEL-Ⅱ	1e-4	0.5	0.964	96.4%
	1e-4	0.8	0.971	93.6%
	1e-6	0.5	**0.979**	**98.3%**
	1e-6	0.8	0.968	96.3%

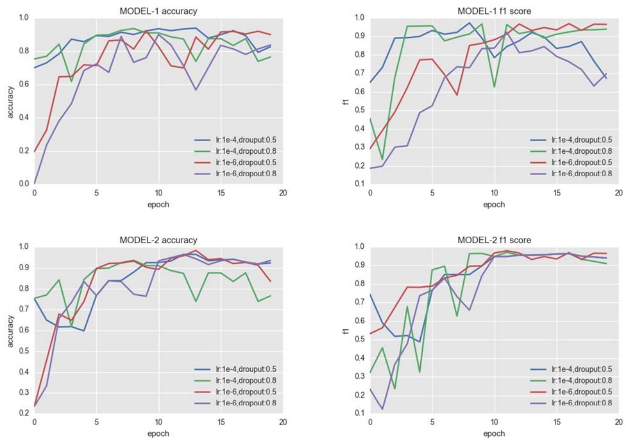


Figure 7. Architecture of MODEL-Ⅱ.

C. Conclusions

In this paper, we have presented a complete automatic program for lung nodule detection. For the initial stage, we set a threshold on HU value to segment the lung parenchyma, and use selective search to locate the suspicious nodule regions. Then we use downsampling and upsampling to make dataset balanced. Finally we propose two architectures of 3D CNN neural networks and set some hyper parameters to grid search. In the Table 3 and Fig. 7 we can see that MODEL-Ⅱ with the parameter of learning rate at 1e-6 and dropout keep prob at 0.5 get highest F1 score. Through our study process, we can conclude that large learning rate (1e-4) will lead neural networks towards local minimum and small dropout keep probability may cause neural networks instable. The 3D Inception block can solve the gradient vanish problems and enhance the F1 score and accuray on both training data and test data.

ACKNOWLEDGMENT

The authors would like to thank the developers of the Tensorflow for providing deep learning library. The authors gratefully acknowledge the support of University of Posts and Telecommunications and College of Computer science and technology with the donation of a Pascal GPU.

REFERENCES

[1] Swensen SJ, Jett JR, Hartman TE, Midthun DE, Sloan JA, Sykes A-M, Aughenbaugh GL, Clemens MA. Lung cancer screening with ct: Mayo clinic experience. Radiology. 2003;226(3):756–61.

[2] Jacobs C, van Rikxoort E M, Scholten E T, et al. Solid, part-solid, or non-solid?: classification of pulmonary nodules in low-dose chest computed tomography by a computer-aided diagnosis system[J]. Investigative Radiology, 2015, 50(3):168-73.

[3] Ciompi F, Chung K, Riel S J V, et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning[J]. Scientific Reports, 2017, 7:46479.

[4] Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3):211-252.

[5] Hoiem D, Divvala S K, Hays J H. Pascal VOC 2008 Challenge[J]. World Literature Today, 2009.

[6] Uijlings J R R, Sande K E A V D, Gevers T, et al. Selective Search for Object Recognition[J]. International Journal of Computer Vision, 2013, 104(2):154-171.

[7] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[J]. 2013:580-587.

[8] Wiemker R, Klinder T. Pulmonary nodule detection using a cascaded SVM classifier[C]// SPIE Medical Imaging. 2016:978513.

[9] Ginneken B V, Setio A A A, Jacobs C, et al. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans[C]// IEEE, International Symposium on Biomedical Imaging. IEEE, 2015:286-289.

[10] Liu J K, Jiang H Y, Gao M D, et al. An Assisted Diagnosis System for Detection of Early Pulmonary Nodule in Computed Tomography Images[J]. Journal of Medical Systems, 2017, 41(2):30.

[11] Golan R, Jacob C, Denzinger J. Lung nodule detection in CT images using deep convolutional neural networks[C]// International Joint Conference on Neural Networks. IEEE, 2016:243-250.

[12] Xu K, Ba J, Kiros R, et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[J]. Computer Science, 2015:2048-2057.

[13] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015:234-241.

[14] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015:770-778.

[15] Diba A, Pazandeh A M, Gool L V. Efficient Two-Stream Motion and Appearance 3D CNNs for Video Classification[J]. 2016.

[16] Sermanet P, Eigen D, Zhang X, et al. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks[J]. Eprint Arxiv, 2013.

[17] Felzenszwalb P F, Huttenlocher D P. Efficient Graph-Based Image Segmentation[J]. International Journal of Computer Vision, 2004, 59(2):167-181.

[18] Szegedy C, Ioffe S, Vanhoucke V, Wentao Zhu, Chaochun Liu,Wei Fan, Xiaohui Xie ınception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[J]. 2016.

[19] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1):1929-1958.

[20] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[J]. Jmlr W & Cp, 2011, 15.

[21] Wentao Z, Chaochun L,et al DeepLung: 3D Deep Convolutional Nets for Automated Pulmonary Nodule Detection and Classification.

[22] Murphy, K.; van Ginneken, B.; Schilham, A. M.; De Hoop, B.; Gietema, H.; and Prokop, M. 2009. A large-scale evaluation of automatic pulmonary nodule detection in chest ct using local image features and k-nearest-neighbour classification. Medical image analysis 13(5):757–770. Jacobs et al. 2014.

[23] Yan, X.; Pang, J.; Qi, H.; Zhu, Y.; Bai, C.; Geng, X.; Liu, M.; Terzopoulos, D.; and Ding, X. 2016. Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2d and 3d strategies. In ACCV, 91–101. Springer.