Training ResNet to classify pulmonary nodules

Tianwei (Owen) Li

Abstract:
Pulmonary nodules are small spots of tissue growth in the lung. There are benign(noncancerous) and malignant(cancerous) nodules. It is important to distinguish these two types of nodules, in order to provide patients with appropriate medical treatment. While this task can be performed by trained professionals, using Machine Learning can help classify pulmonary nodules more efficiently and accurately.

Keywords: pulmonary nodule, classification, ResNet, machine learning, Convolution Neural Network

1. Background Information

1.1. Definition of Pulmonary Nodules

Pulmonary nodules are small areas of growth among lung tissues. According to the definition provided by Cleveland Clinic, pulmonary nodules are "small round or oval-shaped growth in the lung"[5]. Dimension wise, pulmonary nodules should be smaller than 3 centimeters in diameter; otherwise they will be considered as a different category, "Pulmonary Mass"[5]. They can be either benign (noncancerous) or malignant (cancerous)[5]. The image below shows the CT image of a malignant pulmonary tumor, resembled as a white circular dot located on the right side of image.[10]

1.2. Frequency and Malignancy

Nodules commonly present in lungs, as Cleveland Clinic claims that lung nodules can be found within as high as 50% of CT Scans[5]; however most of them are benign. According to a series of CT scan report from Radiology Assistant, the most common type of nodule is the type with diameters of less than 4mm, among which none are malignant[11]. On the other hand, bigger nodules are considerably more malignant, such that three quarters of nodules with diameters of more than 20mm are malignant[11]. The following figure shows the Radiology Assistant's statistics of different types of nodule[11].

Nodule Size

Size	Total	Malignancy
< 4 mm	2038	0%
4 - 7 mm	1034	1%
8 - 20 mm	268	15%
> 20 mm	16	75%

1.3. Causes and Symptoms

Benign and malignant pulmonary nodules have different causes. In most cases, benign pulmonary nodules are "healed-over wounds" of lung tissues after some form of damage caused by fungal infection or tuberculosis, although there may be other less common causes[14]. Malignant pulmonary nodules, however, are early stages of lung cancer tumors, which can be caused by excessive smoking or cancer somewhere else in the body[14].

1.4. Potential influences

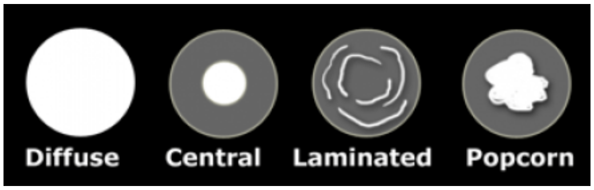
According to Health Encyclopedia at University of Rochester's Health Center, benign and malignant pulmonary nodules have distinct influences to human bodies[14]. Benign nodules rarely grow in size, and for the most part remain unchanged, which means that they will not do significant harms to human body[14]; Malignant pulmonary nodules, on the other hand, is capable of causing cancers by rapidly increasing their sizes. Malignant nodules may double their sized in as few as per 25 days[14].

Once lung cancer starts to develop within human body, the result would be severe. The 5-year survival rate of lung cancer is reported to be 18.6% by American Lung Association[2]. This number is significantly lower than the survival rate of other types of cancers, such as breast cancer (89.6%) or prostate cancer (98.2%)[2]. In addition to low survival rate, lung cancer is also extremely common. It is estimated that lung cancer caused the death of approximately 154,050 Americans in year 2018, which is about a quarter of the total cancer death counts that year[2].

Treatment of pulmonary nodules can be problematic. In some cases, a "thoracotomy" surgery is required, which means to remove an infected portion of the patient's lung[14]. Considering that only malignant pulmonary nodules are dangerous, it is not necessary nor reasonable to simply treat all types of pulmonary nodules by removal of tissues. Thus, it is beneficial to find a method that can consistently and effectively distinguish benign nodules from malignant nodules.

1.5. Traditional Methods of Diagnosis

Radiologists can diagnose the natures of pulmonary nodules after obtaining the lung images of patients using Computer Tomography (CT). However, it is known as a fact that pulmonary nodules may at times experience a process called "Calcification"[8]. Although calcification is not highly likely to happen, it may negatively interfere with the accuracy of diagnosis[8]. Radiology Assistant provides a pattern, showing that calcification pattern of nodules may be indicators of benign tumors, but these patterns also have exceptions[11].



1.6. Computer Aided Diagnosis

As mentioned previously, traditional ways of medical diagnosis have certain limitations, even with trained professionals. A study published on American Journal of Roentgenology, however, shows that computer may be the ultimate solution to these current issues[1]. The study compared the performance of human radiologists with and without the aid from computer programs in terms of differentiating benign pulmonary nodules from malignant pulmonary nodules. The result is significant, showing an average increase of accuracy from 0.785 to 0.853 ”by a statistically significant level” ($p = 0.016$)[1]. This suggests that computer programs may be exceptionally helpful in diagnosis of pulmonary nodule malignancy.

2. Conducting Experiments

Since Computer Tomography (CT) is widely used to scan for pulmonary nodules, I decided to develop and train a neural network to process the scanned CT images of nodules. The network should take in a feature vector derived from CT images of pulmonary nodules, and output the result of ”benign” or ”malignant”. Overall, I choose to use Convolution Neural Network (CNN) with ResNet structure, because it is suitable for image processing, and can have great network depth without losing too much training performance.

2.1. Data set

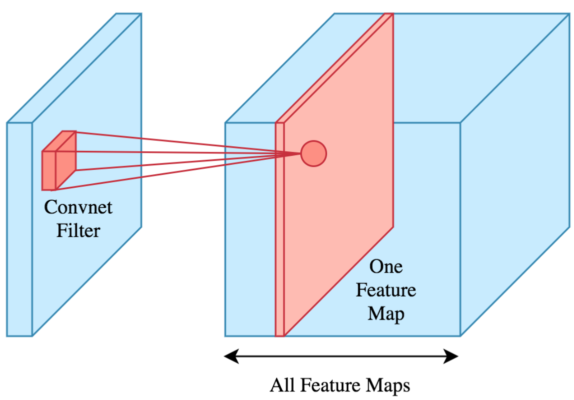
The data-set I used is published from the LIDC-IDRI collection from Cancer Imaging Archive. It is packaged in an hdf5 file with images and labels inside.[4]

2.2. Programming Setup

For the convenience of experiment, I used Tensor-Flow CPU version to conduct my experiment. It is installed under a virtual environment within ”PyCharm” IDE, to prevent possible interference. Other python packages, including Keras and Numpy, are also used in this project.

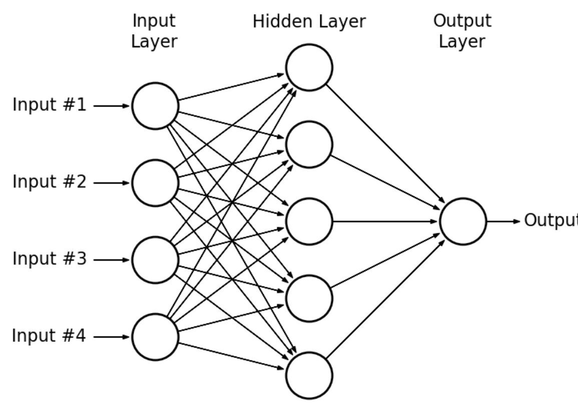
2.3. Convolutional Neural Network

Convolutional Neural Network typically has several ”convolutional layers” followed by a few fully connected layers[12]. Convolutional layers are layers such that a number of ”kernels” or ”filters” each convolve with the previous layer and output results in forms of 2-dimension "feature maps"[12]. These feature maps then stack up together to form the next "layer"[12]. This convolutional process is particularly suitable for image processing, because pixels that are near to each other in an image are likely to be related, and those "kernels" serve well to capture such local relations, or features[12]. Below is a diagram that shows the fundamental principle of Convolutional Network, with the kernel being called "Convnet Filter"[7].



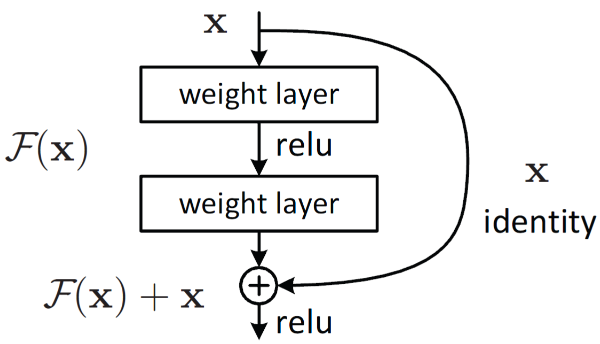
2.4. Fully Connected Networks

Although convolutional neural networks (CNN) are suitable to extract features from various images, they cannot effectively classify the images themselves. Thus a solution is to add a few fully-connected layers of neurons to the output end of CNN, in order to classify the visual images based on extracted features. In these fully-connected layers, every neuron from the previous layer is connected to every neuron of the next layer, making up a very fundamental form of neural networks. The image below is a simple illustration of the fully-connected layers[7]. Inputs in this model should be connected to the outputs of CNN, and output of this fully connected neural network provides a number for the classification results of images.



2.5. ResNet

Residual Neural Network, or ResNet in short, is implemented in this experiment, because it is capable of having a large amount of layers without the problem of "Vanishing Gradients" that can slow down training process[6]. ResNet have such characteristic because after every block of network, it "feeds" its value to the next unit via simple addition[6]. That way, a deeper neural network may be built to classify pulmonary nodules. The diagram below illustrates the fundamental principle of ResNet —- adding its values to the next unit in the sequence[6].

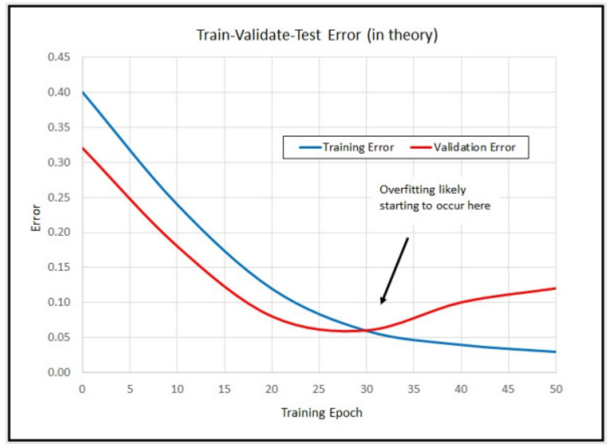


3. Adjustments

A Convolutional Neural Network is built, but effective training still require a number of adjustments on parameters, as well as the initial selection of threshold between benign and malignant nodule.

3.1. Over-fitting and Early Stopping

Over-fitting can cause problems in the training process by learning random features from the training data set as useful features[3]. As a result, after a certain point in the training process, the model might have learnt too much randomness to be effective in classification of images outside of training set[3]. The following image illustrates this issue: the red line represents accuracy from test data, which starts to increase after prolonged training[9].



To overcome this issue, I separate the training data into three different sets: training set, validation set, and test set. Training set and test set are self-explanatory; while the validation set is a separate set of data used to supervise the accuracy of model. Specifically, in this case I set an Early Stopping Patience of 20 in my program, so that once the accuracy of model decreases after 20 continuous batches, the training of this model is stopped, and training result is saved.

3.2. Learning Rate and Epochs

These factors are determined based on training time required. On the only device available, when using a learning rate of 0.01 with

SGD optimizer, the program takes 50 to 60 epochs to finish training, either reaching accuracy goal or an early stoppage occurring to prevent the model from over-fit. Considering this amount of training is practical and the accuracy is reasonable, I choose 0.01 and 60 as my learning rate and number of epoch. Training time marginally exceeds one hour.

3.3. Threshold of classification

The neural network outputs a decimal number. The input is classified as malignant if the output decimal number is greater than a preset threshold, and is classified as benign otherwise. Through a few trials, I find 0.5 to be a reasonably accurate threshold; however, considering the different risks of misdiagnosis, this threshold may need to be adjusted based on the testing result of this classification model.

4. Results

The testing show that after 60 epochs of training, the model has approximately 93 percent accuracy while classifying pulmonary nodule images from the testing data set. The following table is a confusion matrix generated after evaluating the classification model, which counts the number of correct benign and malignant nodule classification, as well as misdiagnosis in either way. Note that in the table, Ben and Mal respectively equal to Benign and Malignant.

Thresh=0.5	actual Ben	actual Mal
diagnsd Ben	814	53
diagnsd Mal	28	444

5. Optimization

One important point to notice is according to the confusion matrix, more malignant nodules are diagnosed as benign than benign nodules are diagnosed malignant. Based upon information in previous sections, misdiagnosis of malignant nodule as benign nodule may cause lung cancer to develop within patients, while misdiagnosis of benign nodule as malignant nodule may result in unnecessary lung surgeries like thoracotomy. The 5-year death rate of lung cancer, however, is 81.4%, considerably higher than 6.4%, the eventual death rate of lung surgeries, according to MDLinx[2][13].

Misdiagnosis of malignant pulmonary nodule as benign pulmonary nodule is much more dangerous than other way around. Thus it is important to change the diagnosing process, in order to achieve the crucial goal of minimizing the chance of malignant-as-benign misdiagnosis; on the other hand, minimizing benign-as-malignant diagnosis should have far lower priority, because they are not as dangerous.

One way to counter this risk is to alter the threshold of diagnosis. The fully-connected neural networks uses the extracted graphical features to classify pulmonary nodules, and outputs a decimal number ranging from 0 to 1. If the output exceeds our preset threshold of 0.5, a nodule is classified as malignant. Thus by decreasing the threshold, it will be more likely for the model to classify nodules as malignant than previous, thus effectively avoid malignant-as-benign type diagnosis. The result is significant, that with a threshold of 0.3, the confusion matrix appears as following:

Thresh=0.5	actual Ben	actual Mal
diagnsd Ben	826	22
diagnsd Mal	53	438

Acknowledgements

This research about pulmonary nodule is supported by Mr.Junshen Xu from Massachusetts Institute of Technology and also supported by Embark China research projects.

References

[1] Feng Li, Masahito Aoyama, Junji Shiraishi, Hiroyuki Abe, Qiang Li, Kenji Suzuki, Roger Engelmann, Shusuke Sone, Heber MacMahon and Kunio Doi. *Radiologists' Performance for Differentiating Benign from Malignant Lung Nodules on High-Resolution CT Using Computer-Estimated Likelihood of Malignancy.* American Journal of Roentgenology. 2004;183: 1209-1215. 10.2214/ajr.183.5.1831209

[2] American Lung Association. *Lung Cancer Fact Sheet.* December 14, 2018.

[3] AWS Documentation. *Model Fit: Underfitting vs. Overfitting.*

[4] Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., Kazerooni, E.A., MacMahon, H., van Beek, E.J., Yankelevitz, D., Biancardi, A.M., Bland, P.H., Brown, M.S., Engelmann, R.M., Laderach, G.E., Max, D., Pais, R.C., Qing, D.P., Roberts, R.Y., Smith, A.R., Starkey, A., Batra, P., Caligiuri, P., Farooqi, A., Gladish, G.W., Jude, C.M., Munden, R.F., Petkovska, I., Quint, L.E., Schwartz, L.H., Sundaram, B., Dodd, L.E., Fenimore, C., Gur, D., Petrick, N., Freymann, J., Kirby,

J., Hughes, B., Casteele, A.V., Gupte, S., Sallam, M., Heath, M.D., Kuhn, M.H., Dharaiya, E., Burns, R., Fryd, D.S., Salganicoff, M., Anand, V., Shreter, U., Vastagh, S., Croft, B.Y., Clarke, L.P., 2015. *Data from LIDC-IDRI.* doi:10.7937/K9/TCIA.2015.LO9QL9SX.

[5] Cleveland Clinic. *Pulmonary Nodule.* 9500 Euclid Avenue, Cleveland, Ohio, 2019.

[6] Akshay Mishra, Hong Cheng. *Advanced CNN.* http://slazebni.cs.illinois.edu/spring17/.

[7] Luciano Strika, MercadoLibre. *What are Convolutional Neural Networks?* KD Nuggets.

[8] Ali Nawaz Khan, Hamdan H. Al-Jahdali,1 Carolyn M. Allen, Klaus L. Irion,2 Sarah Al Ghanem,1 and Shyam Sunder Koteyar. *The calcified lung nodule: What does it mean?* 2010 Apr-Jun; 5(2): 6779.

[9] James D. McCaffrey. *Neural Network Train-Validate-Test*. Jun 28, 2018.

[10] Alex A. Balekian, MD, MSHS, and Michael K. Gould, MD, MS. *The Lung Nodule That Refused to Grow.* December 2012.

[11] Ann Leung and Robin Smithuis. *Solitary pulmonary nodule: benign versus malignant differentiation with CT and PET-CT* Department of Radiology, Stanford University Medical Center, Stanford, California and the Department of Radiology, Rijnland Hospital, Leiderdorp, the Netherlands. Publicationdate May 20, 2007.

[12] UFLDL Tutorial. *Convolutional Neural Network*. Stanford Supervised Learning and Optimization.

[13] Liz Meszaros, MDLinx. *Most deaths after resection for lung cancer happen post-discharge*. November 21, 2018. (Note that the 6.4% mortality rate is calculated from "Among patients undergoing resection for lung cancer, 50% of deaths occur within 90 days after hospital discharge" and "Dr. Quero-Valenzuela and colleagues found that the overall 30-day mortality rate was 1.6% compared with 3.2% for overall 90-day mortality".)

[14] Health Encyclopedia. *Pulmonary Nodules*. 2019 University of Rochester Medical Center Rochester, NY.