ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse292952020.pdf
https://doi.org/10.30534/ijatcse/2020/292952020

Detection of Pneumonia Using Deep Transfer Learning

K Reddy Madhavi[1], G Madhavi[2], B Rupa Devi[3], Padmavathi Kora[4]
[1]Sree Vidyanikethan Engineering College, AP, India, kreddymadhavi@gmail.com
[2]JNTUK, AP, India, madhavi.researchinfo@gmail.com
[3]Annamacharya Institute of Technology, AP, India, rupadevi.aitt@annamacharyagroup.org
[4]GRIET, Telangana, India, padma386@gmail.com

ABSTRACT
Lung diseases are most prevalent cause of death in humans of all ages, and this disorder may occur for different reasons. The newest WHO (World Health Organization) report states that just in the United States over 1 million citizens seek care due to pneumonia and there are nearly ten million cases of tuberculosis worldwide. Perhaps some part of it is lethal due to lack of medical staff or human mistake. The manual analysis of x-ray images is a long process requiring radiological expertise and a large volume of time. Deep learning can play a crucial role in exceeding decision making, detecting marks of disease as well as conducting the initial examination and suggesting urgent cases.

Key words: Lung image, X-ray, Transfer learning, VGG-16

1. INTRODUCTION

This work focuses on lung diseases (pneumonia) classification, we first want to discuss selected screening approaches briefly. All techniques use radiography, a medical imaging method that used to be obsolete. However, machine learning and digital advances revived this method and its significance in diagnosis of lungs diseases [1]-[4]. Computer vision supported with deep neural networks finds its usefulness in any area of life starting from facial emotion recognition to disease detection. Thanks to the recent technological advances, computer aided image analysis algorithms compete with professionals in terms of accuracy yet remain unchallenged in speed and volume of reviewed cases. Unlike doctors, computers make quick rational decisions unaffected by emotions and tiredness.

Especially, they allow detecting multiple forms of cardiothoracic lessons on X-ray scans. The growing popularity of machine learning models in medical diagnosis correlated with their accuracy is considered as a great success as it leads to better disorder recognition. Recent encouraging results in deep learning applied in the field of lung diagnosis led to the usage of a GPU-based platform which is able to process a large volume of images in high-resolution within seconds and thus exceed the work of radiologists. Previous approaches required both a large volume of data and strong computing power computers [5]. Instead of following the popular trend of creating new algorithms to solve a problem, we decided to leverage existing tools and show their high accuracy in medical tasks outperforming the suggested solutions [6]. The motivation of this project was to create a pipeline allowing us to detect pulmonary diseases using tiny datasets ($< 10^3$ images per class) and limited computational resources. Also, we want to show the importance of segmentation as a tool focusing our algorithms on information that is crucial in diagnosis. Should models search through pulmonary changes within lungs images omitting redundant data (bones, internal organs) [7], the decision is considered reasonable. To prove that statement, we generate class activation maps marking regions of interest (regions that vastly contributed to the final classification).

As this research focuses on lung diseases (pneumonia and tuberculosis) classification, we first want to discuss selected screening approaches briefly. All techniques use radiography, a medical imaging method that used to be obsolete. However, machine learning and digital advances [8] and its significance in diagnosis of lungs diseases [9], [10], [11]. Especially, they allow detecting multiple forms of cardiothoracic lessons on x-ray scans. The growing popularity of machine learning models in medical diagnosis correlated with their accuracy is considered as a great success as it leads to better disorder recognition. Recent encouraging results in deep learning applied in the field of lung diagnosis led to the usage of a GPU-based platform which is able to process a large volume of images in high-resolution within seconds and thus exceed
the work of radiologists.

2. MATERIALS AND METHODS

The enormous need for a vast number of Pneumonia picture datasets is discussed in the proposed work using data augmentation techniques. To detect class labels (Pneumonia and Normal), the pre-trained model extracts characteristics from trained augmented images and integrates multi-scale discriminant characteristics. Figure 1 displays the hierarchical representation of the suggested technique for the three-step diagnosis of pneumonia using lung X-ray slices.

2.1. X-ray dataset

This work combines two relatively small datasets ($< 10^3$ images per class) datasets for the classification (pneumonia and normal detection). We selected 306 examples per "diseased" class (306 images containing marks of normal and 306 images with pneumonia) and 306 of healthy patients contributing to a set of 918 samples coming from different patients.

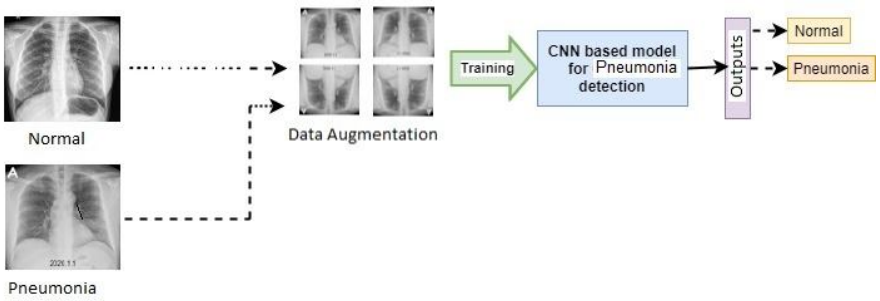

Figure 1: Flow diagram for COVID-19 detection using X-ray

3. DEEP LEARNING

3.1. Deep Learning Model

After publishing AlexNet [12] in 2012, convolutional neural networks (CNN) renewed the interest of the research community. CNN and subsequent deep models such as VGG [13] proved their usefulness, especially in computer vision related tasks. The contribution of published work demonstrates that those models are suited much better at capturing different features than traditional algorithms which heavily rely on different feature engineering methods In a classical formulation of a convolutional neural network used in classification, CNN consists of multiple convolution layers followed by pooling operators. The convolution layers are made of kernels, small tensors compared to windows which process input and output information. Those operators can successfully capture the spatial and temporal dependencies in an image and (e.g., gradient change) thus learn different local features like straight lines (horizontal or vertical) and curves while upper layers (hidden) can perform detection of more sophisticated information like rectangles or circles based on the received input, therefore, understand it better. As processed data flows higher to deeper layers, a network learns more "abstractive" combinations.

3.2 Transfer Learning in Lung Diseases Classification

Transfer Learning is a special approach to tasks associated with computer vision using, where data resources are limited, deep neural networks [14]-[16]. Therefore, we integrate pre-trained templates that are experienced at solving related problems

to establish a starting point for a new mission. Due to the lack of sample volume, this method is crucial in medical image processing. VGG model is a deep convolutional neural network proposed by researchers A. Zisserman and K. Simonyan from the University of Oxford [13] containing over 138 million parameters. This model was able to achieve 7.4% error rate on the ImageNet dataset. It improved the Alex-Net [12] network by changing the kernel size and instead of 11x11 and 5x5 filters in the first two layers, it implemented multiple smaller ones 3x3 filters one after another.

VGG-16: The VGG-16 model had 16 weight layers and the design is similar to LeNet and AlexNet. As we go deeper into the network, the size of feature maps (Convolution layers in succession) also increases. The convolution filter size is 3x3 for all the layers. The standard VGG-16 consists of two convolution layers in succession followed by a max-pooling layer, which reduces the feature map into 112x112 as shown in Figure 2. Each of these convolution filters has 64 feature maps. And followed by another set convolutions and max-pooling(two convolutions layers). Then two layers with 3 convolutions in succession followed by max-pooling and three FC layers. The VGG-16 model is the best performing model thus its architecture is explained in detail [12]. Figure 3 visualizes the architecture of the VGG model with 16 layers.

Type-of-Layer	Output-shape	Parameters
InputLayer	[224, 224, 3]	0
Conv-2D	[224, 224, 64]	1792
Conv-2D	[224, 224, 64]	36928
MaxPooling2D	[112, 112, 64]	0
Conv-2D	[112, 112, 128]	73856
Conv-2D	[112, 112, 128]	14586
MaxPooling2D	[56,56,128]	0
Conv-2D	[56,56,256]	295168
Conv-2D	[56,56,256]	590080
Conv-2D	[56,56,256]	590080
MaxPooling2D	[28,28,256]	0
Conv-2D	[28,28,512]	1180160
Conv-2D	[28,28,512]	2359808
Conv-2D	[28,28,512]	2359808
MaxPooling2D	[14,14,512]	0
Conv-2D	[14,14,512]	2359808
Conv-2D	[14,14,512]	2359808
Conv-2D	[14,14,512]	2359808
MaxPooling2D	[7,7,512]	0
Avg-Pooling2d(3)	[1,1,512]	0
Flatten	[None,512]	0
Dense	[None,64]	32832
Linear	[None,3]	130
Total params	14,747,650	
Trainable params	32,962	
Non-trainable params	14,714,688	

Figure 2: The layers and parameters of the proposed model (VGG-16)

4. RESULTS

As in the previous chapter, training of model repeatedly ten times for 150 epochs took about one day. This short time is a result of setting all parameters of pre-trained models as non-trainable. Therefore, the gradients flow only through the concatenated layers. The images were resized, which also beneficially influences the execution duration. The biggest problem related to training our models was the maximum platform usage time, which is up to twelve hours.

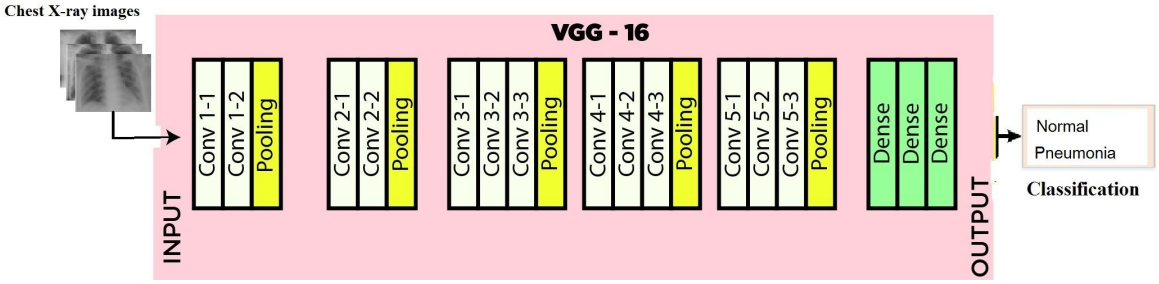

Figure 3: The architecture of the proposed model(VGG-16)

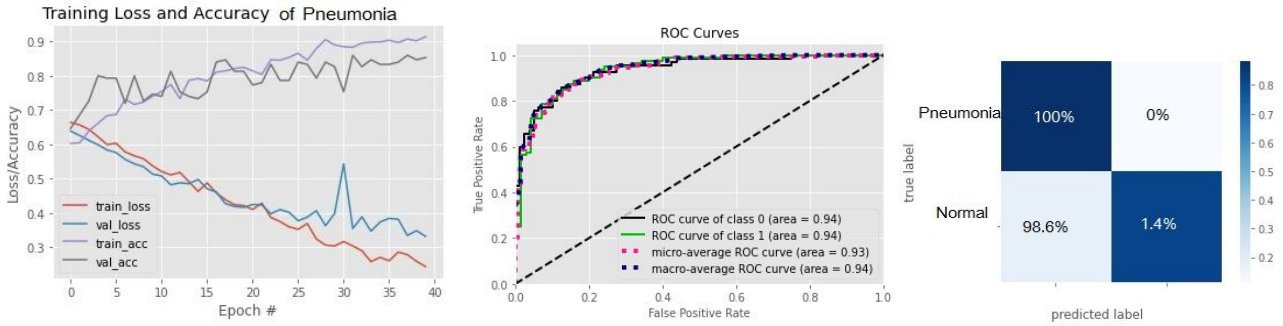

Figure 4: [a] Accuracy curve of VGG-16 using X-ray with fine tuning [b] ROC curve of VGG-16 using X-ray with fine tuning [c] Confusion matrix of VGG-16 using X-ray with fine tuning for the detection of Pneumonia

Table 1: Pneumonia detection performance metrics

Model	Sensitivity	Specificity	Precision	F1-score	Accuracy
VGG-16(Fine tuning)	100%	98.6%	99.0%	99.5%	99.4%
ReseNet-18	89%	87.2%	87.9%	88.9%	88.8%
GoogleNet	87.9%	85.8%	86.9%	87.4%	87.0%
SqueezeNet	85.8%	84.4%%	85.8%	85.7%	85.2%

The examination of training and validation error curves allowed us to find a relatively good number of epochs after which models were over-fitting. the training data. The training process was then stopped, and the final results were measured as an average of all results obtained at that step. The following results were generated for ten independent training runs to observe a similar training pattern. Each of the ten training and validation curves were plotted on the same charts based on the tape (training or validation). So as to keep up a high state of readability, the results were isolated. The wider, dotted curve is an averaged result of all obtained at the particular epoch. The lowest loss value on training and validation data sets, whereas on it corresponds to the maximum accuracy obtained. The model stops dropping the validation error around the 70th epoch and maintains its level throughout the remaining roughly 80 epochs. A similar behavior is experienced when examining Table I. Here, the average validation accuracy slows down and only slightly increases. Eventually, the models were evaluated on the test

set and scored an average accuracy of 99.4%, which is over 6 percentage of improvement comparing to the results To visualize results, we selected a network which obtained the best accuracy score after 150 epochs. The confusion matrix in Figure 4 shows that the VVG -16 model has the biggest advantage with classifying 'pneumonia images' to the corresponding class and normal as healthy. One problem is related to labeling healthy images as healthy, although in this case, it is not a major issue since the cost of assigning healthy patients to a category of sick ones is more acceptable than otherwise. Additionally, the majority of samples containing marks of pneumonia were correctly classified. Table I shows that accuracy score for binary classification for the detection of pneumonia.

5. CONCLUSION

In this paper, we focused on exploring pneumonia disease classification problem using deep neural networks under the supervision of the small size dataset (nearly 300 examples). Moreover, we examined class activation maps to explore the reasoning of our models and investigate which regions are determinative. In Chapter 3 we first introduced VGG-16 deep learning architectures and compared to ImageNet challenge. Then, we use those networks as feature extractors to train our shallow algorithms. The results are summarized in section 4, here we only use accuracy to evaluate the performance since the test set is class equally distributed. Here we also show how our solutions outperform deeper models trained on the same data. After comparing class activation maps, we conclude that VGG-16 not only improved the accuracy score yet also the reasoning behind the classification. Pre-processed Chest X-Ray images with remaining lungs force networks to explore only those areas which lead to improvements in the inter-pretability of present model as the pronounced areas might helpful for the treatment of sick patients.

REFERENCES

M. Kraft, Approach to the patient with respiratory disease, in: Goldman's Cecil Medicine: Twenty Fourth Edition, Elsevier Inc., 2011, pp. 512–516.

S. H. Ralston, I. D. Penman, M. W. Strachan, R. Hobson, Davidson's Principles and Practice of Medicine E-Book, Elsevier Health Sciences, 2018.

3. K. I. Zheng, G. Feng, W.-Y. Liu, G. Targher, C. Byrne, M.-H. Zheng, Extra-pulmonary complications of covid-19: a multi-system disease?, Journal of Medical Virology.

4.https://www.lung.org/lung-health-diseases/lung-disease-lookup (2020)

5. S. Stirenko, Y. Kochura, O. Alienin, O. Rokovyi, Y. Gordienko, P. Gang, W. Zeng, Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation, in: 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO), IEEE, 2018, pp. 422–428.

6. S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani, et al., Automatic tuberculosis screening using chest radiographs, IEEE transactions on medical imaging 33 (2) (2013) 233–245.

7. T. Ishigaki, S. Sakuma, M. Ikeda, One-shot dual-energy subtraction chest imaging with computed radiography: clinical evaluation of film images., Radiology 168 (1) (1988) 67–72.

8. A. Karargyris, J. Siegelman, D. Tzortzis, S. Jaeger, S. Candemir, Z. Xue, K. Santosh, S. Vajda, S. Antani, L. Folio, et al., Combination of texture and shape features to detect pulmonary abnormalities in digital chest x-rays, International journal of computer assisted radiology and surgery 11 (1) (2016) 99–106.

9. C. S. Pereira, H. Fernandes, A. M. Mendonça, A. Campilho, Detection of lung nodule candidates in chest radiographs, in: Iberian Conference on Pattern Recognition and Image Analysis, Springer, 2007, pp. 170–177.

10. G. Coppini, S. Diciotti, M. Falchini, N. Villari, G. Valli, Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms, IEEE Transactions on Information Technology in Biomedicine 7 (4) (2003) 344–357.

11. R. C. Hardie, S. K. Rogers, T. Wilson, A. Rogers, Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs, Medical Image Analysis 12 (3) (2008) 240–258.

12. A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

13. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

14. Kora, Padmavathi, and K. Sri Rama Krishna. "ECG based heart arrhythmia detection using wavelet coherence and bat algorithm." Sensing and Imaging 17.1 (2016): 12.

15. Padmavathi, K., and K. Sri Ramakrishna. "Detection of atrial fibrillation using autoregressive modeling." International Journal of Electrical and Computer Engineering (IJECE) 5.1 (2015): 64-70.

16. Padmavathi, K., and K. Sri Rama Krishna. "Myocardial infarction detection using magnitude squared coherence and Support Vector Machine." 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom). IEEE, 2014.