



Chest X-Ray Classification of Lung Diseases Using Deep Learning

Yew Fai Cheah

Faculty of Engineering and Science, Curtin University, Malaysia

*Correspondence: 700015881@student.curtin.edu.my

SUBMITTED: 6 October 2021; REVISED: 18 October 2021; ACCEPTED: 19 October 2021

ABSTRACT: Chest X-ray images can be used to detect lung diseases such as COVID-19, viral pneumonia, and tuberculosis (TB). These diseases have similar patterns and diagnoses, making it difficult for clinicians and radiologists to differentiate between them. This paper uses convolutional neural networks (CNNs) to diagnose lung disease using chest X-ray images obtained from online sources. The classification task is separated into three and four classes, with COVID-19, normal, TB, and viral pneumonia, while the three-class problem excludes the normal lung. During testing, AlexNet and ResNet-18 gave promising results, scoring more than 95% accuracy.

KEYWORDS: CNN; COVID-19; tuberculosis; viral pneumonia

1. Introduction

Corona virus disease 2019 (COVID-19) was identified in December 2019 in Wuhan, China. It has been identified as a highly infectious respiratory disease, which was already prevalent throughout the world [1]. The number of cases caused by the virus has increased dramatically. As a result, the World Health Organization (WHO) declared a pandemic on March 11, 2020. COVID-19, unfortunately, has similar characteristics to viral pneumonia and tuberculosis (TB), which makes its prevention and countermeasures more difficult. In fact, TB patients will be diagnosed similarly to COVID-19 and viral pneumonia patients as they have similar symptoms [2]. All of the above-mentioned lung diseases are characterized by a severe cough and a high fever. As a result, manually observing them is challenging.

The most frequent testing method for COVID-19 detection is a nasal and throat swab. However, due to the increased demand, the quantity of testing kits was in great demand, and supply was limited. As a result, alternative detection technologies such as chest X-ray imaging and CT-scanning have been considered [3].

The task of classifying chest X-ray images into one of the above-mentioned lung diseases has become more feasible as machine learning techniques advance. In particular, deep learning has been used to perform this task. Deep learning is a subset of machine learning that mimics the architecture of the human brain by generating patterns that can be used in decision making. A convolutional neural network (CNN) is a type of deep learning algorithm that specializes in using input images for classification purposes.

DarkCovidNet was proposed as a DarkNet-19 modification to improve its effectiveness in detecting and classifying COVID-19 disease [4]. The model was able to identify binary and multi-class problems with 98% and 87% accuracy, respectively. Furthermore, the authors [5] compared eight different models in terms of their performance in classifying normal lungs and lungs infected by pneumonia and COVID-19 based on X-ray images collected. The images used were also augmented so that the results could be compared to images that were not augmented. According to the results, Densenet201 outperformed the other networks with image augmentation, while CheXNet outperformed other networks without image augmentation.

The authors of [6] evaluated 15 CNN networks to classify three types of lung X-ray images: normal, pneumonia, and COVID-19. The findings revealed that the VGG19 network had the best testing accuracy of 89.3% while requiring the least amount of training time per epoch. In [7], the authors used combinations of both public and local data sets to train and test the MobileNets and VGG-16 models. Accuracies between 89% and 96% were obtained for the combinations. ResNet50V2 and MobileNet were used in [8] to categorize COVID-19, normal, and pneumonia chest X-ray images. The results showed that 97% accuracy could be achieved. In [9], VGG-19 was used to classify normal lungs and pneumonia lungs infected by COVID-19 and bacteria using chest X-ray images. Results showed the accuracy of more than 95%.

The objective of this paper is to classify four classes of chest X-ray images, i.e., normal, COVID-19, viral pneumonia, and TB using five CNN networks: AlexNet, GoogLeNet, ResNet-18, ShuffleNet, and SqueezeNet. Furthermore, we also classify three classes of X-ray images without the normal class. By doing this, we can observe the effectiveness of the pretrained CNN networks in classifying similar-type lung diseases. For the three-class classification testing, AlexNet gives the best accuracy of 97.89%, while for the four-class testing, ResNet-18 gives the best accuracy of 97.14%.

The rest of this paper is structured as follows. In Section 2, the fundamentals of CNN are revisited. In Section 3, simulation results and analysis are performed. Finally, Section 4 concludes this paper.

2. Material and Method

2.1. Convolutional Neural Network (CNN)

A CNN architecture's structure consists of many layers that must be assembled into blocks. The convolution layer, fully connected layer, and pooling layer are the three basic layers in CNN. In implementation, the CNN architecture employs these layers in repetition. CNN requires data input in order to learn the patterns and characteristics that will allow it to discriminate between various classes during testing. Images are the most commonly utilized input data for CNN construction. Forward propagation refers to the stage at which input is converted into output by the layers of the CNN. Meanwhile, backpropagation modifies the network's weights during the training stage based on the rate of error received from each previous epoch. The backpropagation sequence begins with the input arriving at the output layer. Each neuron from the input, hidden, and output layers will be calculated at the output layers based on the randomly determined weight inside the layer.

As the name implies, the convolution layer becomes the main layer in CNN. Two parameters that define CNN are the kernel and its size. The kernel size is usually 3x3 or 5x5. The kernel will scan the image from left to right, top to bottom. Subsequently, an element-wise

product of the image and kernel will be performed and summed up to result in a convolution value. The output of convolution layers is then passed to a nonlinear activation function, such as Rectified Linear Unit (ReLU).

The pooling layer is always put after the nonlinear layer. This layer can be considered as the "summary" layer which summarizes the features from the previous layers. There are two common types of pooling layer: average pooling and maximum pooling. Maximum pooling computes the maximum value of each feature map, whereas average pooling computes the average value of each feature map. There is also a global type of pooling, such as global average pooling and global maximum pooling, which implements average and maximum operations on the entire feature map [10].

Finally, the fully connected layer converts each of the preceding layers' output into a one-dimensional array of integers. These numbers are then linked to one other and to additional completely connected layers, producing the dense layer, in which each input is coupled to each output. In this paper, five CNN networks are used: AlexNet, GoogLeNet, ResNet-18, ShuffleNet, and SqueezeNet. The summary of their architectures is shown in Table 1.

Table 1. CNN architectures

	AlexNet	GoogLeNet	Resnet-18	ShuffleNet	SqueezeNet
Input size	224x224x3	224x224x3	224x224x3	224x224x3	227x227x3
Activation function	ReLU	ReLU	ReLU	ReLU	ReLU
Convolution kernel	3x3, 5x5, 11x11	1x1, 3x3, 5x5, 7x7,	1x1, 3x3, 7x7,	1x1, 3x3	1x1, 3x3
No of learnable layers	8	22	18	50	18

2.2. Research Methodology

The number of images used for training and validation in each class is 200. The chest X-ray images were collected from various sources. The COVID-19, viral pneumonia, and normal lung images are taken from [11], which contains chest X-ray images from various sources, such as Qatar, Bangladesh, Pakistan, and Malaysia. Meanwhile, the TB chest X-ray images are taken from the database of the National Library of Medicine, which contains images from Montgomery County, USA and Shenzhen No. 3 People's Hospital, Guangdong, China. Fig. 1 shows a sample of COVID-19, viral pneumonia, and TB chest X-ray images. Furthermore, the images are separated into a training set and a validation set with a ratio of 70:30. In this paper, we use an independent data set for the testing stage. The testing data set is an independent set of 100 images categorized into 3-class and 4-class sets, except for COVID-19 (85 images) due to the limitation of the images. The data set details are summarized in Table 2.

An image preprocessing process is performed before training the images into the networks. The X-ray images are grayscale and need to be converted to three color channels to be suitable for the input. A computer with an AMD Ryzen 5 5400H @3GHz processor, 8GB RAM, and a 4-GB NVIDIA GeForce GTX 1650 Ti Graphic Processing Unit (GPU) is required for simulations. The training and validation settings are shown in Table 3.

Table 2. Summary of data set details.

Source of X-ray images	
Normal	[11]
COVID-19	[11]
Viral pneumonia	[11]
TB	National Library of Medicine
Number of images for training and validation	
Normal	200
COVID-19	200
Viral pneumonia	200
TB	200
Number of images for testing	
Normal	100
COVID-19	85
Viral pneumonia	100
TB	100

**Figure 1.** X-ray image for COVID-19 (left), viral pneumonia (middle), and TB (right).**Table 3.** Training setting.

Setting	Value
Initial learning rate	0.0001
Validation frequency	35
Max epoch	35
Mini batch size	35

For testing task, the metrics used in this research are accuracy, precision, recall, and F1-score which are defined by

$$\text{Accuracy} = \frac{\text{Number of correct prediction}}{\text{Total number of predictions}} \quad (1)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Note that “True Positive” refers to correct prediction of the positive class and “True Negative” refers to correct prediction of the negative class. On the contrary, “False Positive” is the incorrect prediction of the positive class, and “False Negative” is the incorrect prediction of the negative class. This can be easily visualized using the confusion matrix of 2-class problems as shown in Fig. 2. For multiclass problems, we can find overall precision, recall, and F1-score metrics by averaging the values for each class.

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

Figure 2. Two-class confusion matrix.

3. Simulation Results and Analysis

For three-class classification, the training accuracies for the five networks are shown in Table 4. It can be seen that ResNet-18 has the highest accuracy. In addition, the model does not overfit, as seen from the validation loss.

Table 4. Training results of three-class problem.

Network	Validation Accuracy	Validation Loss
AlexNet	97.22%	0.0806
GoogLeNet	93.89%	0.1427
ResNet-18	98.33%	0.0897
ShuffleNet	93.89%	0.1728
SqueezeNet	97.78%	0.0646

The testing results for this problem are depicted in Table 5. From testing results, it can be seen that AlexNet has the best accuracy, although it is relatively comparable with ResNet-18.

Table 5. Testing results of three-class problem.

Network	Accuracy	Precision	Recall	F1-Score
AlexNet	97.89%	97.67%	98.00%	98.00%
GoogLeNet	93.33%	93.33%	93.67%	93.33%
ResNet-18	97.54%	97.67%	97.33%	97.67%
ShuffleNet	95.09%	95.00%	95.00%	95.00%
SqueezeNet	96.84%	96.67%	95.80%	97.00%

For the four-class classification problem, the training and testing results are shown in Table 6 and 7, respectively. As seen, ResNet-18 gives the best result for both training and testing. Moreover, ShuffleNet has the same results as SqueezeNet in terms of accuracy, and they have the lowest accuracy.

Table 6. Training results of four-class problem.

Network	Validation Accuracy	Validation Loss
AlexNet	95.83%	0.0916
GoogLeNet	96.25%	0.0965
ResNet-18	97.08%	0.1308
ShuffleNet	92.92%	0.1909
SqueezeNet	93.75%	0.1770

Based on the two tabulated findings (Tables 5 and 7), it is possible to conclude that AlexNet and ResNet-18 are two stronger model networks in this task. On the other hand, ShuffleNet is less accurate and has a lower F1-score. Table 8 shows the comparison between

AlexNet and ResNet-18 in three-class problems. A precision of 93% indicates that the model can accurately predict the COVID-19 images 93% of the time. A recall of 100% indicates that the model correctly identifies all COVID-19 images. Lastly, the F1-score of 96% indicates a good measure, which shows the balance between precision and recall.

Table 7. Testing results of four-class problem.

Network	Accuracy	Precision	Recall	F1-Score
AlexNet	96.62%	96.50%	96.75%	96.75%
GoogLeNet	95.58%	95.50%	95.50%	95.50%
ResNet-18	97.14%	97.25%	97.00%	97.00%
ShuffleNet	94.03%	94.00%	94.00%	94.00%
SqueezeNet	94.03%	94.00%	94.00%	93.75%

Table 8. Comparison between AlexNet and ResNet18 in three-class problem.

Class	Precision AlexNet	Recall	F1-score
COVID-19	0.93	1	0.96
TB	1	0.95	0.98
Viral pneumonia	1	0.99	0.91
	ResNet-18		
COVID-19	0.98	0.95	0.97
TB	0.98	0.97	0.98
Viral pneumonia	0.97	1	0.98

4. Conclusion

The goal of this paper is to use deep learning to diagnose more lung diseases that have similarities with COVID-19 based on chest X-ray images. This leads to more complicated tasks. To do these tasks, five pre-trained networks were trained, verified, and tested using data sets supplied online from hospitals all around the world. The acquired data sets were divided into three sets: training, validation, and testing. The data sets obtained for each class utilized in the network were independent from one another. The training network was separated into 3-class and 4-class training. The three-class training consisted of COVID-19, TB and viral pneumonia. The four-class training consisted of normal, COVID-19, TB, and viral pneumonia. In testing, AlexNet and ResNet-18 were the two best networks in terms of accuracy.

References

- [1] Mohapatra, R.K.; Pintilie, L.; Kandi, V.; Sarangi, A.K.; Das, D.; Sahu, R.; Perekhoda, L. (2020). The recent challenges of highly contagious COVID-19, causing respiratory infections: Symptoms, diagnosis, transmission, possible vaccines, animal models, and immunotherapy. *Chemical Biology & Drug Design*, 96, 1187 – 1208. <https://doi.org/10.1111/cbdd.13761>.
- [2] Togun, T.; Kampmann, B.; Stoker, N.G.; Lipman, M. (2020). Anticipating the impact of the COVID-19 pandemic on TB patients and TB control programmes. *Annals of Clinical Microbiology and Antimicrobials*, 19, 1-6. <https://doi.org/10.1186/s12941-020-00363-1>.
- [3] Benmalek, E.; Elmhamdi, J.; Jilbab, A. (2021). Comparing CT scan and chest X-ray imaging for COVID-19 diagnosis. *Biomedical Engineering Advances*, 1, 100003. <https://doi.org/10.1016/j.bea.2021.100003>.
- [4] Ozturk, T.; Talo, M.; Yildirim, E.A.; Baloglu, U.B.; Yildirim, O.; Acharya, U.R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in Biology and Medicine*, 121, 103792. <https://doi.org/10.1016/j.compbiomed.2020.103792>.

- [5] Chowdhury, M.E.H., Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.; Mahbub, Z.; Islam, K.; Khan, M.S.; Iqbal, A.; Al-Emadi, N.; Reaz, M.B.I.; Islam, M. (2020). Can AI help in screening viral and COVID-19 pneumonia?. in *IEEE Access*, 8, 132665-132676. <http://dx.doi.org/10.1109/ACCESS.2020.3010287>.
- [6] Rahaman, M.M.; Li, C.; Yao, Y.; Kulwa, F.; Rahman, M.A.; Wang, Q.; Qi, S.; Kong, F.; Zhu, X.; Zhao, X. (2020). Identification of COVID-19 samples from chest X-ray images using deep learning: A comparison of transfer learning approaches. *Journal of X-Ray Science and Technology*, 28, 821-839. <https://doi.org/10.3233/xst-200715>.
- [7] Khasawneh, N.; Fraiwan, M.; Fraiwan, L.; Khassawneh, B.; Ibnian, A. (2021). Detection of COVID-19 from chest X-ray images using deep convolutional neural networks. *Sensors*, 21, 5940. <https://doi.org/10.3390/s21175940>.
- [8] Shorfuzzaman, M.; Masud, M. (2020). On the detection of COVID-19 from chest X-ray images using CNN-based transfer learning. *Computers, Materials & Continua*, 64, 1359-1381. <https://www.techscience.com/cmcc/v64n3/39434>.
- [9] Gao, T. (2020). Chest X-ray image analysis and classification for COVID-19 pneumonia detection using deep CNN, medRxiv, 2020. Link: <https://www.medrxiv.org/content/early/2020/10/12/2020.08.20.20178913.full.pdf>
- [10] Brownlee, J. (2021). Accessed on 2 October 2021, A gentle introduction to pooling layers for convolutional neural networks. *Machine Learning Mastery*, <https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/>.
- [11] COVID-19 Radiography Database. Accessed on 17 August 2020. <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>.



© 2021 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).