

# 2021 年臺灣國際科學展覽會 優勝作品專輯

作品編號 190022  
參展科別 電腦科學與資訊工程  
作品名稱 Deep learning on Covid-19 prediction and X-ray severity grading system

就讀學校 臺北市立建國高級中學

指導教師 張瑞峰、王鼎中

作者姓名 李昆樺

關鍵詞 COVID-19、深度學習、分級系統

## 作者簡介



我是建國中學科學班二年級的李昆樺，很榮幸能參與台灣國際科展與眾多強者切磋。無論是否得獎，我想藉此感謝台大醫學影像實驗室的張瑞峰教授、李衍緯及翁子傑學長。感謝他們帶給我研究上開闊的視野。感謝專題老師王鼎中領導與提供建議。也感謝所有在實驗路上充滿體諒的家人與老師。

# 摘要

利用深度學習解決醫學問題一直是受矚目的研究主題。鑒於近期新冠肺炎疫情上升，有關新冠肺炎檢測的研究便成了熱門研究主題。目前，最有效的檢測方法是聚合酶連鎖反應 (PCR)，然而，PCR 耗時甚久且有人為誤差。因此，以 X 光影像圖透過深度學習來診斷並分級是一個有效率且安全的做法。在研究中，我們利用深度學習進行疾病診斷，在五元分類上有相當高的準確率(84.91%)、在 COVID-19 單獨辨識時得到了極高的準確率(99.35%)、產生出疾病熱區及設計了新的分級系統( X-ray Severity Grading System , XSGS)，並將其用於嚴重程度分類，在不同分級下具有可辨別的差異。

## Abstract

Using deep learning to solve medical problems has always been a research topic that has attracted much attention. In view of the recent rise in the COVID-19 pandemic, research on COVID-19 testing has become a hot research topic. Currently, the most effective detection method is Polymerase Chain Reaction (PCR). However, PCR takes a long time and has human-effected errors. Therefore, it is an efficient and safe method to diagnose and classify X-ray images through deep learning. In the research, we used deep learning for disease diagnosis(five-classes) to obtain a acceptable high accuracy rate (84.91%), have a quite high accuracy in classification COVID-19 and Normal(99.35%), generated disease heat map and designed a new grading system ( X-ray Severity Grading System , XSGS), and used it to grading the severity, which has the distinguishable difference on different degree.

# 1. Introduction

COVID-19, caused by a novel coronavirus, SARS-CoV-2, has inflicted over one hundred millions of people worldwide with a mortality of  $\sim 4\%$  [1]. As opposed to the MERS-CoV or SARS-CoV, this disease is less deadly but higher infectious [2]. The median incubation period was estimated to be 5.1 days and the previous research has shown that the patient was also infectious during the incubation period [3][4]. After pre-symptomatic, patients would have multiple clinical symptoms like fever (88.5%), cough (68.6%), myalgia or fatigue (35.8%), expectoration (28.2%), and dyspnea (21.9%) [5].

Because of the airborne novel coronavirus and the above reasons, there is the urgency to detect earlier so that can clarify the source of infection and have the treatments instantly. There are two common methods to detect COVID. One is RTPCR the other is X-ray. However, the most accurate method (RT-PCR), which takes several days and can be influenced by many factors like inaccurate conditions, sample contaminations [6]. X-ray is cheaper, faster, and safer, so X-ray is an efficient method to solve this problem.

Unfortunately, the need of the COVID-19 pandemic has already exceeded our ability to provide sufficient numbers of adequate personal protective equipment [7]. Priority for limited resources should aim at saving the most lives [8], so a high-speed method to distinguish inflected and uninflected and separate mild cases from severe cases is important.

To solve the above dilemma, and helping people get through the difficulties, we proposed a deep learning-based model which provides the prediction with Chest X-ray image and classifies them into 5 degrees by the severity. For resolving the aforementioned issues, we used ResNet152V2 deep learning architecture to be the

base model. In order to optimize the model. Several augmentations to increase the variety of the images. After the inflected or uninflected model has been built, the probability of inflected is utilized to be the foundation of classification.

## 2. Related Works

Deep learning has been used on different areas such like traffic, Go, and natural language processing. Peppas, M. V., et al. use closed-circuit television (CCTV)'s image to analyze whether the road user to go [9]. While Silver, David, et al. using deep learning model to play the Go[10]. CNN's different variations were implemented in natural language processing by Lopez, et al[11].

Digital image processing on medical image's pioneer, AAMs, which proposed by T. F. Cootes, J. Edwards, and C. J. Taylor[12], is an important method for locating deformable objects. Then, R. Beichel et al. propose an extension of AAMs, which can be used on a larger scale of the x-ray image[13]. Sam Mavandadi et al. had the different approach on medical image analysis and diagnosis by producing a game for human to improve the model[14].

The task of distinguishing the disease and classifying has been addressed by diverse approaches. Rajkomar et al. augmented 1850 chest x-ray images into 150,000 training samples [15]. Shan et al. quantificate the lung inflection through chest CT image[16].

Wang et al. aimed to develop a deep learning-based model that can detect COVID-19 with high sensitivity, providing fast and reliable scanning [17] The task of distinguishing COVID-19 has been researched by Khalid et al [18]. Yang, Ran, et al. have assessed the severity of COVID-19 by chest CT image [19].

Nevertheless, Khalid et al still has room for improvement on their testing and

Shan et al. and Yang et al. use CT-scanning. CT scanning is expensive so can't be applied to some least developing country, and CT scanning will bring too much radiation on the body which isn't suitable to a severe patient.

## **3. Proposed Method**

### **3-1 Dataset and Environment**

#### **Dataset:**

This dataset was acquired from COVID-19 Radiography Database. We got 384 COVID-19 positive images, 1341 normal images, 1360 Bacterial images, 337 Tuberculosis images, and 1345 viral images. This dataset was collected image from the Italian Society of Medical and Interventional Radiology (SIRM) and Githubs [20]

#### **Develop Environment:**

CPU : IntelCoreTMi5-8250U 1.80GHz ◦

RAM : 8.00GB

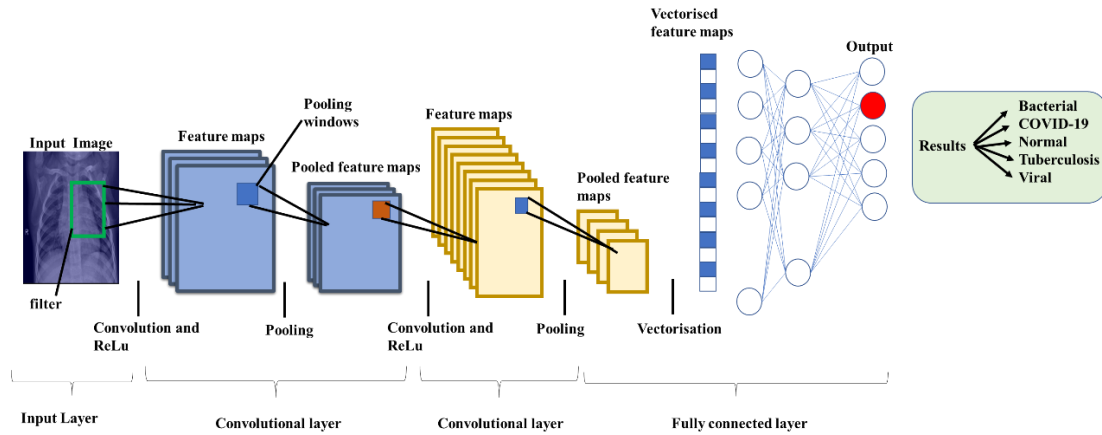
Operating System : Windows10 Home x64

I use1070ti (8GB) to be the integrated development environment.

### **3-2 Baseline CNN architecture**

Before explaining the deep learning model we use, we first introduce each unit and functions in the CNN architecture.

Fig.3-1 shows the inner structure of the model. The whole CNN is composed of Five layers, which are Input layer, convolutional layers, pooling layers, full-connection layers, and output layer.



**Fig.3-1 CNN architecture**

Input layer in the model is the X-ray image. After got the parameter of the image, we used a set of kernels to determinate a tensor of feature maps. The size of the convolution kernel is much smaller than the input image, and it acts on the input image in parallel or overlap. All the elements on the feature map were calculate through the same kernel, and it represents that all the data shared the same weight and bias.

After conv layer, all features will pass ReLU. ReLU layer (Rectified Linear Units layer) can enhance the non-linear characteristics of the decision function and the entire neural network without changing the convolutional layer itself.

Pooling layer is a nonlinear form of down sampling. There are many different forms of nonlinear pooling functions among which max pooling is the most common. The pooling layer will continuously reduce the space size of the data, so the number of parameters and the amount of calculation will also decrease, which also controls overfitting to a certain extent. Different from convolutional layer, pooling layer can be used to reduce sensitivity of the edge.

Fully connected layer does the high-level reasoning in the neural network, and each layers' neurons are connected to all activations in the previous layer. Finally output the predict results or probability of classification.

### **3-3 Deep Learning architecture**

#### **VGG-16 and VGG-19**

VGG is proposed by Visual Geometry Group in University of Oxford, which is a convolution neural net architecture. The major characteristic of this kind architectures are that they concentrated on simple  $3 \times 3$  size kernels in convolutional layers and  $2 \times 2$  size in max pooling layers and it has 2 FC (Fully Connected layers) trailed by a softmax for output. Different from VGG-16, VGG-19 has one more layer in each of the three convolutional blocks [21].

#### **ResNet50 and ResNet152V2**

ResNet (Residual Network), referring to add residual learning in traditional Convolutional Neural Network, which can solve the vanishing gradient and decreased accuracy problems in training. Resnet50 consists of five steps each with a convolution and identity block and each convolution block has 3 convolution layers and each identity block also has 3 convolution layers. Compared to ResNet50, ResNet152V2 has more building blocks in conv\_3 and conv\_4[22][23].

#### **InceptionV3 and InceptionResNetV2**

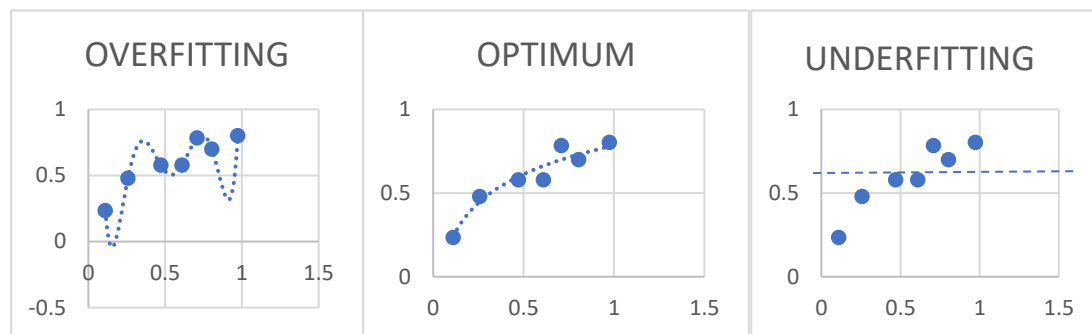
It is the new version of Inception developed by google. The concept of Inception is by designing a sparse network structure, but can generate dense data, it can not only increase the performance of the neural network, but also ensure the efficiency of the use of computing resources. It has block of parallel convolutional layers with 3 different sizes of filters (1x1, 3x3, 5x5). Additionally,  $3 \times 3$  max pooling is also performed. Being inspired by ResNet, they develop the new neural network called InceptionResNetV2, which is deeper and more intuitive[24][25].



### 3-4 Hyperparameter adjustment

#### Batch size, Epoch, and iteration

Because data is big, We can't upload all the data to the computer at once. In order to overcome this problem, we need to divide the data into smaller sizes and hand it to our computer one by one, and update the weight of the neural network at the end of each step so that it can fit the given data. At this moment, we need batch sizes, epoch, and iteration. Batch sizes is the number of training examples in a single batch. In deep learning, SGD training is generally used, that is, batch size samples are taken from the training set for each training. An iteration is equivalent to training once with batchsize samples. One epoch is equivalent to training once using all samples in the training set. It is not enough to transmit a complete data set once in a neural network, and we need to transmit the complete data set multiple times in the same neural network. If there are too many epochs will lead to overfitting (**Fig.3-2a**). On the contrary, few epochs will lead to underfitting (**Fig.3-2c**).



(a) Overfitting example    (b) Optimum example    (c) Underfitting example

**Fig.3-2 Different situation in deep learning**

#### Learning rate

The learning rate directly affects how fast our model can converge to a local minimum. Generally speaking, the greater the learning rate, the faster the neural network learns. If the learning rate is too small, the network is too slow and may likely to fall into a local optimum (**Fig.3-3**); but if it is too large and exceeds the

extreme value, the loss will stop falling and oscillate repeatedly at a certain position or even divergence (Fig.3-4).

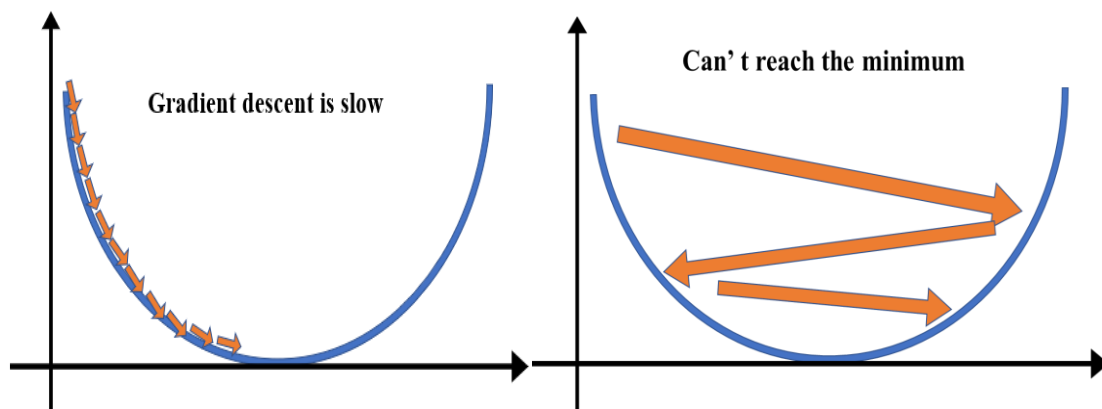


Fig.3-3, Fig.3-4 Gradient decent with too small & too large learning rate

### SGD (Stochastic Gradient Descent)

In machine learning/deep learning, the loss function of the objective function usually takes the average of the loss functions of each sample, then suppose the objective function is:  $f(x) = \frac{1}{n} \sum_{i=1}^n f(x_i)$ . Among them  $F(x_i)$  is the objective function of the  $x_i$  sample, then the gradient of the objective function at  $x$  is:

$$\nabla f(x) = \frac{1}{n} \nabla \sum_{i=1}^n f(x_i).$$

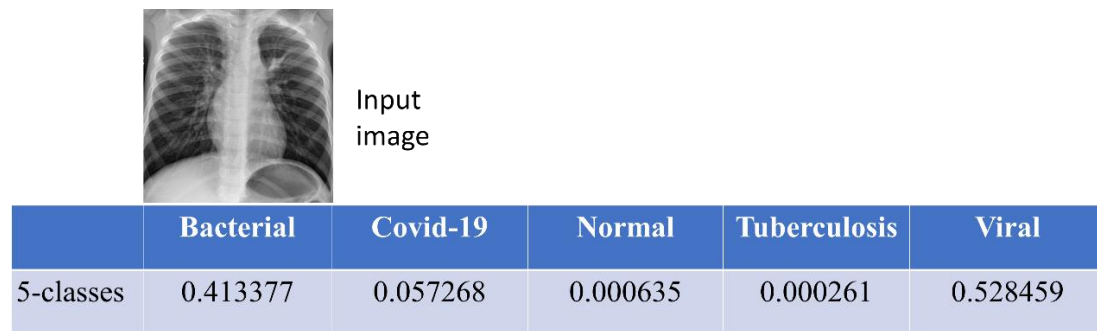
If we choose original gradient descent our calculation will be  $O(n)$ . In contrast, the idea of stochastic gradient descent is to randomly sample a sample  $f(x_i)$  to update the parameters and the calculation will become  $O(1)$ .

### Pre-trained Model

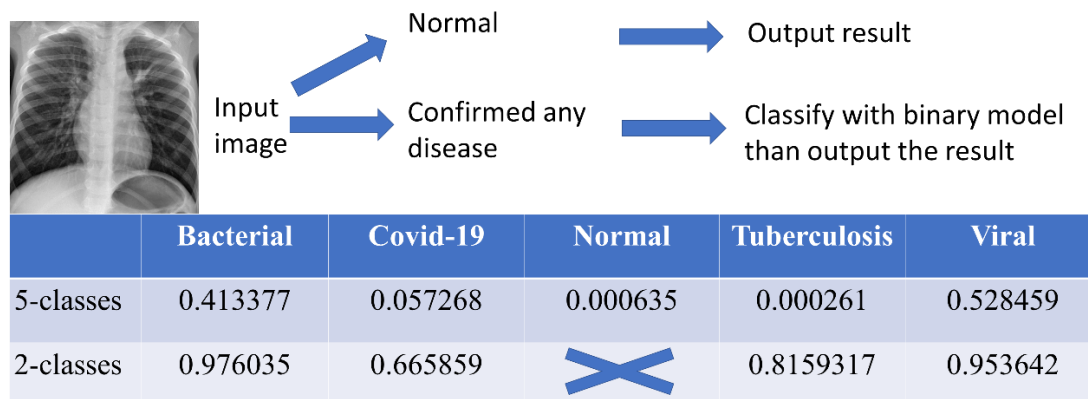
Pre-trained Model is the model created by predecessors to solve similar problems. So that we don't need to train a new model from scratch and start with a model trained in similar problems. At the same time, the biggest advantage of using pre-trained model is that it greatly reduces the training time. It only needs to train for the dense layer, and it takes very little time.

### 3-5 Double validation

Sometimes, the model had a poor performance on sorting two of five-class classification (**Fig.3-5**). They were two of them had higher probabilities than others. However, both of them had probabilities lower than 75%. In order to increase accuracy on the total model. I trained binary classification for bacterial-normal, COVID-19-normal, Tuberculosis-normal, and Viral-normal so that we can double checked the prediction. We used models trained on same pre-trained model to show the accuracy and precision increase from Double validation. The final model of Double validation will choose the highest accuracy model of six models.



**Fig. 3-5 poor performance example**



**Fig. 3-6 the process of double validation**

### 3-6 X-ray Severity Grading System (XSGS)

I developed the X-ray Severity Grading System (XSGS) to assess the burden of COVID-19 during the initial scan at admission. This score uses lung opacity as a substitute for the expansion of the disease in the lungs. According to the testing data, we can find that there are different probabilities in different cases, and the severity and probability have positive correlation. We will output the probability value of positive and negative on the last layer. And we will classify the severity according to the probability value. Therefore, I divided into five degrees according to their probability. Like Fig. 3-7, we sort the positive image data according to probability and grading them in five degrees from mild to severe by percentile. From 0 to 20 is the first degree and 20 to 40 is the second degree etc.

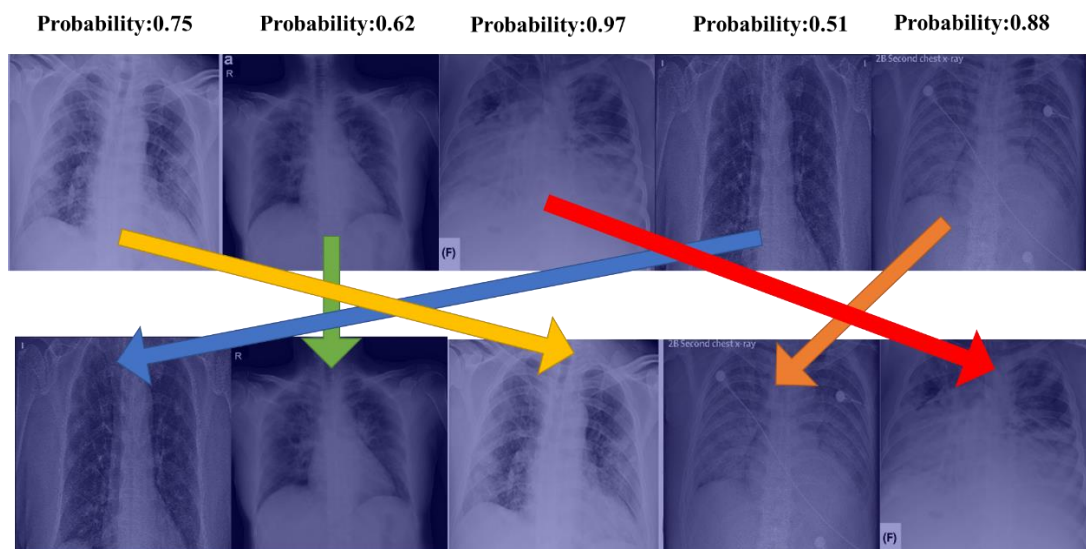


Fig. 3-7 XSGS sort example

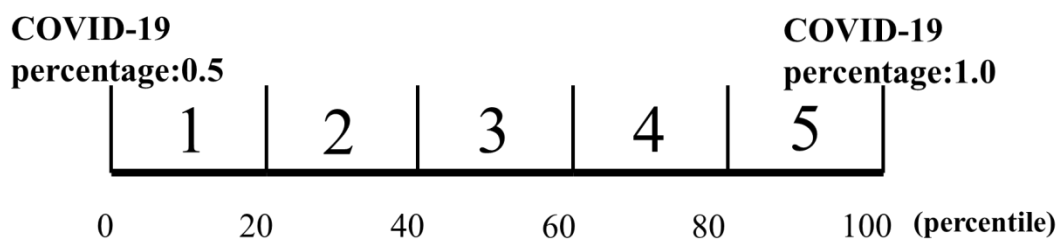


Fig. 3-8 XSGS classification

### 3-7 Grad-CAM

Although the network architecture has become more complete and the accuracy rate is even higher than that of humans, we cannot be sure whether its judgment mechanism is reliable or meets the truly diseased location. In order to address this problem, we need Grad-CAM. But if the image is a normal image, it won't focus on the lung so fierce.

Grad-CAM's method is Backpropagation. First we Add the pixel values to get the average.

$$F^k = \frac{1}{X} \sum_i \sum_j A_{ij}^k$$

$F^k$  is feature map obtain the result of global average pooling

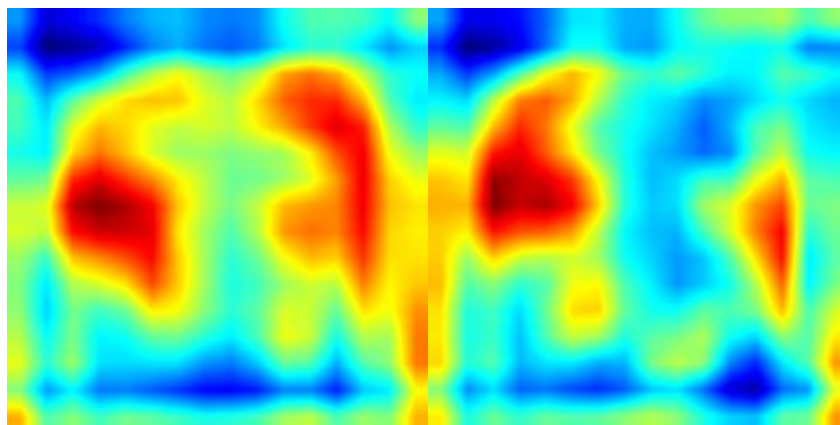
$X$  is the number of pixels in feature map

$A_{ij}^k$  is the pixel value of feature map

Then, we can got CAM-result in below formula.

$$Y^c = \sum_k w_k^c \cdot F^k$$

Finally, we'll generate a map with heat map (Fig.3-9).



**Fig. 3-9 Grad-CAM heatmap**

# 4. Experiment result

## 4-1 Training result of five classes

### VGG-16

Fig.4-1, Fig.4-2 shows the accuracy and loss curve of VGG-16. It converges to 76.12%. For the train data's loss starts from 3.1028 and converges to 0.6451. The test accuracy is 74.80%

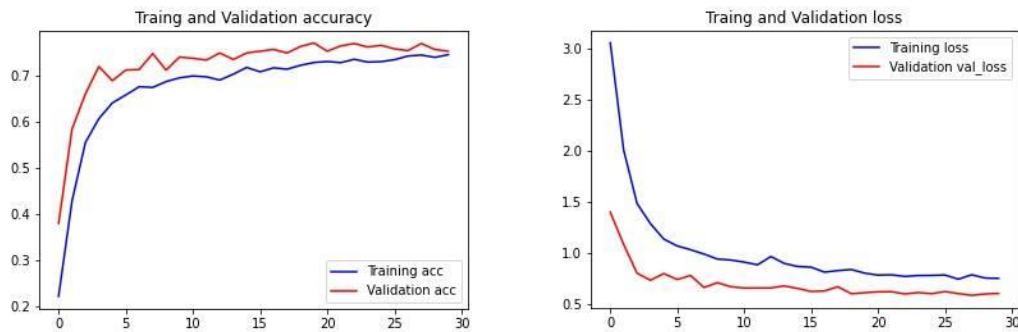


Fig.4-1 & Fig.4-2 VGG-16's accuracy and loss curve

### VGG-19

Fig.4-3, Fig.4-4 shows the accuracy and loss curve of VGG-19. It converges to 72.34%. For the train data's loss start from 2.9567 and converges to 0.3717. The test accuracy is 69.08%

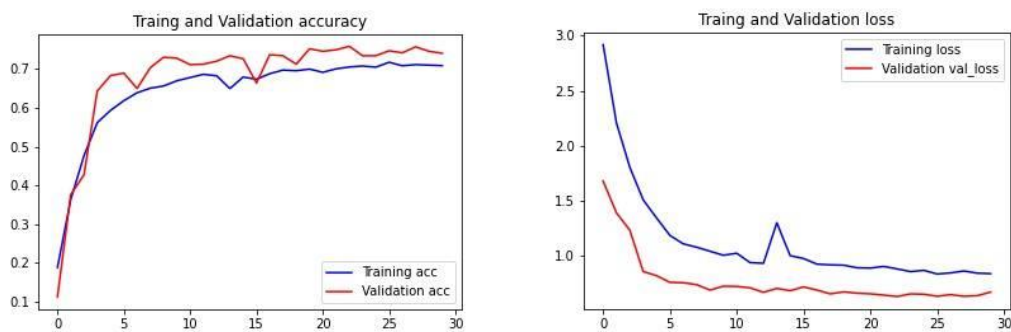
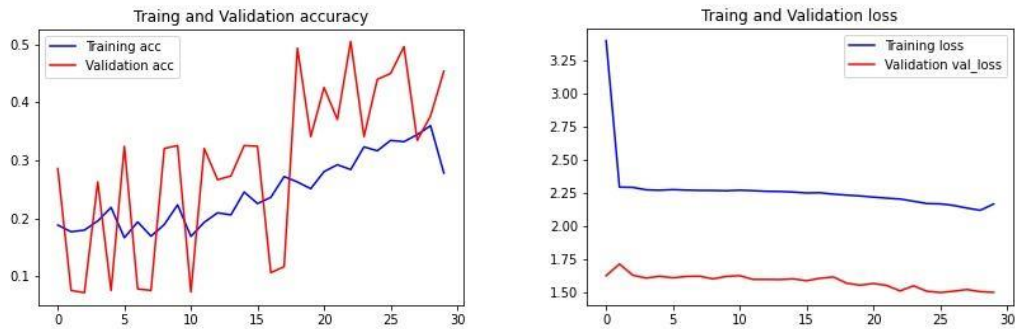


Fig.4-3 & Fig.4-4 VGG-19's accuracy and loss curve

## ResNet50

**Fig.4-5, Fig.4-6** shows the accuracy and loss curve of ResNet50. It converges to 40.05%. The train data's loss start from 3.4792 and converges to 2.2259.

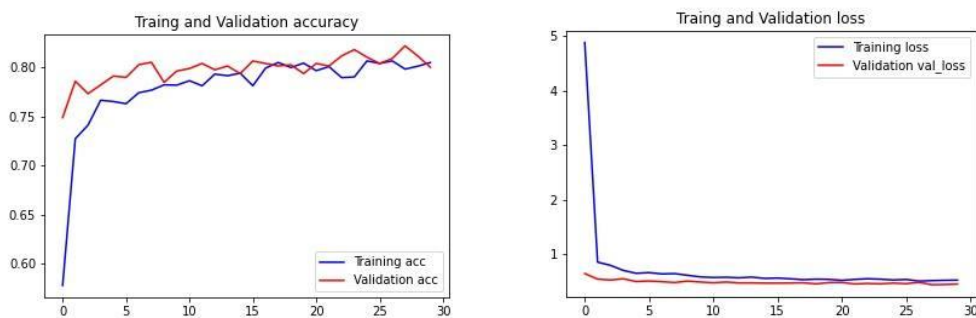
The test accuracy is 37.50%



**Fig.4-5&4-6 ResNet50's accuracy and loss curve**

## ResNet152V2

**Fig.4-7, Fig.4-8** shows the accuracy and loss curve of ResNet152V2. train data is obvious quickly increasing and converges to 82.35%. Train data's loss starts from 4.9385 and converges to 0.4391. The test accuracy is 79.35%

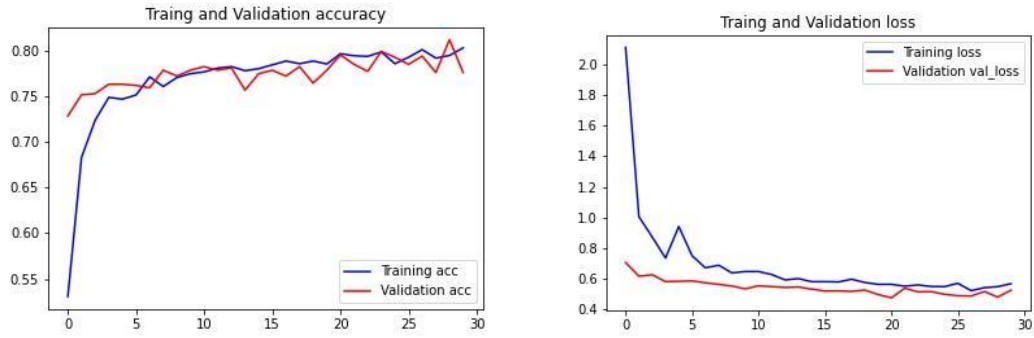


**Fig.4-7&4-8 ResNet152V2's accuracy and loss curve**

## InceptionV3

**Fig.4-9, Fig.4-10** shows the accuracy and loss curve of InceptionV3. It converges to 79.86%. The train data's loss start from 2.2468 and converges to 0.5732.

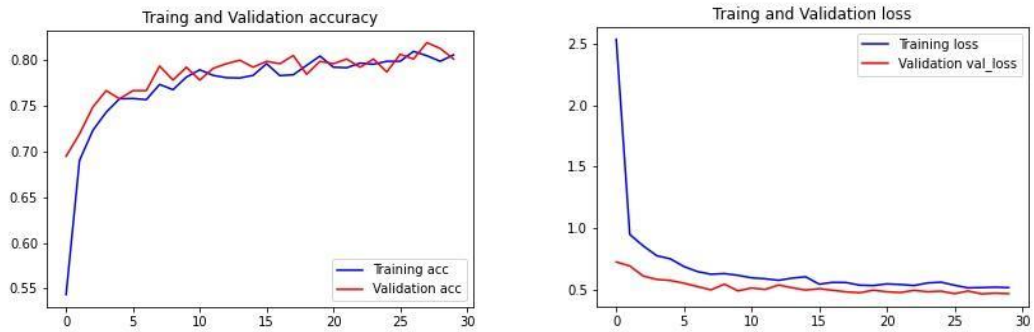
The test accuracy is 77.49%



**Fig.4-9&4-10 InceptionV3’s accuracy and loss curve**

### InceptionResNetV2

**Fig.4-11, Fig.4-12** shows the accuracy and loss curve of Inception ResNetV2. It converges to 79.73% in epoch 30. The train data’s loss starts from 2.5124 and converges to 0.6739. The test accuracy is 75.15%



**Fig.4-11&4-12 InceptionResNetV2’s accuracy and loss curve**

### 4-2 Training and Testing result discuss of five classes

**Table 4-1 Testing result**

	Accuracy	Bacterial precision	COVID-19 precision	Tuberculosis precision	Viral precision
VGG-16	0.7480	0.6467	0.8527	0.9174	0.6174
VGG-19	0.6908	0.5855	0.8605	0.9083	0.6072
ResNet50	0.3750	0.2357	0.8630	0.2569	0.0029
ResNet152V2	0.7935	0.7100	0.9612	0.9755	0.7174
InceptionV3	0.7749	0.6878	0.8811	0.9205	0.6906
InceptionResNetV2	0.7515	0.6558	0.9121	0.9297	0.6746



According to the **Table 4-1**, all of the deep learning architecture got almost 69% accuracy except ResNet50. Among them, VGG-16 and VGG-19's model has the lower accuracy (<75%) than others (75%~79%). All of the architecture have a higher validation accuracy (>60%) since epoch 5 except ResNet50. ResNet50 has a very poor result during the training process, and it may represent that architecture misses the central issue or the feature vanished during the training process so that it doesn't have good performance. VGG-16 and VGG-19 got little bit lower precision in COVID-19 and it will lead to false positive not detected. False positive will lead to local transmission or even community spread. ResNet152V2 is the best architecture in this research, which accuracy is 0.7935. Its precision is great so that it can reduce unnecessary waste of medical resources and also prevent most of false positive. We will choose ResNet152V2's model to be the five classes classification model.

#### 4-3 Testing result discuss of two classes classification

**Table 4-2 Testing result (Bacterial, Normal)**

	Accuracy	Recall	Specificity	Precision	F1 score
VGG-16	0.9630	0.9575	0.9686	0.9682	0.9628
VGG-19	0.9603	0.9538	0.9667	0.9663	0.9600
ResNet50	0.9418	0.9482	0.9353	0.9361	0.9422
ResNet152V2	0.9806	0.9834	0.9778	0.9780	0.9806
InceptionV3	0.9612	0.9501	0.9723	0.9716	0.9607
InceptionResNetV2	0.9741	0.9723	0.9760	0.9759	0.9741

**Table 4-3 Testing result (COVID-19, Normal)**

	Accuracy	Recall	Specificity	Precision	F1-score
VGG-16	0.9839	0.9592	0.9885	0.9400	0.9736
VGG-19	0.9806	0.9400	0.9885	0.9400	0.9636
ResNet50	0.9645	0.8980	0.9770	0.8800	0.9358
ResNet152V2	0.9935	0.9800	0.9962	0.9800	0.9880
InceptionV3	0.9613	0.8364	0.9882	0.9388	0.9060
InceptionResNetV2	0.9871	0.9423	0.9961	0.9800	0.9685

**Table 4-4 Testing result (Tuberculosis, Normal)**

	Accuracy	Recall	Specificity	Precision	F1 score
VGG-16	0.9807	0.9821	0.9792	0.9792	0.9807
VGG-19	0.9762	0.9732	0.9792	0.9790	0.9761
ResNet50	0.9583	0.9554	0.9613	0.9610	0.9582
ResNet152V2	0.9911	0.9881	0.9941	0.9940	0.9910
InceptionV3	0.9747	0.9643	0.9851	0.9848	0.9744
InceptionResNetV2	0.9807	0.9792	0.9821	0.9821	0.9806

**Table 4-5 Testing result (Viral, Normal)**

	Accuracy	Recall	Specificity	Precision	F1 score
VGG-16	0.9647	0.9647	0.9647	0.9647	0.9647
VGG-19	0.9600	0.9572	0.9628	0.9626	0.9600
ResNet50	0.9480	0.9498	0.9461	0.9463	0.9481
ResNet152V2	0.9768	0.9796	0.9740	0.9741	0.9768
InceptionV3	0.9582	0.9535	0.9628	0.9625	0.9580
InceptionResNetV2	0.9805	0.9833	0.9777	0.9778	0.9805

According to the **Table 4-2**, **Table 4-3**, **Table 4-4**, **Table 4-5** all of the deep learning architecture got at least 95% accuracy in four different kinds of classification except ResNet50. ResNet152V2 is the best architecture in Bacterial-Normal, COVID-19-Normal, and Tuberculosis-Normal classification. InceptionResNetV2 is the best architecture in Viral-Normal classification. ResNet152V2 has better accuracy in most categories, so we used ResNet152V2' models to be the Double validation models.

#### 4-4 Double validation result

After the modify of Double validation, the revised data **Table 4-6** showed that the final result after Double validation. All of the models have grown at least 3% accuracy. ResNet50 had poor performance in five-class. Through the validation it has increased a lot of course. The best model ResNet152V2 has increased 5.6% and ends up to quite high accuracy 84.91%. **Table 4-7** showed the increments of all accuracy and precision.

**Table 4-6 Final result after Double validation**

	Accuracy	Bacterial precision	COVID-19 precision	Tuberculosis precision	Viral precision
VGG-16	0.7777	0.6829	0.9328	0.9602	0.6522
VGG-19	0.7390	0.6259	0.9096	0.9602	0.6232
ResNet50	0.7031	0.6259	0.8630	0.7737	0.6333
ResNet152V2	0.8491	0.7615	0.9897	0.9755	0.7609
InceptionV3	0.8265	0.7761	0.9638	0.9602	0.7123
InceptionResNetV2	0.8095	0.7184	0.9690	0.9847	0.7123

**Table 4-7 Accuracy and precision increased after Double validation**

	Accuracy	Bacterial precision	COVID-19 precision	Tuberculosis precision	Viral precision
VGG-16	0.030	0.036	0.080	0.043	0.035
VGG-19	0.048	0.040	0.049	0.052	0.016
ResNet50	0.328	0.390	0.000	0.517	0.630
ResNet152V2	0.056	0.051	0.028	0.000	0.043
InceptionV3	0.052	0.088	0.083	0.040	0.022
InceptionResNetV2	0.058	0.063	0.057	0.055	0.038


#### 4-5 XSGS result

According to the test data, we have 50 samples can be used on XSGS classification standard including one false negative. By using percentages to split, we separate 11 datas at 1 degree, 10 datas at 2,3,4 degree, and 9 datas on 5 degree. First

degree contains 1 false positive (Table 4-8).

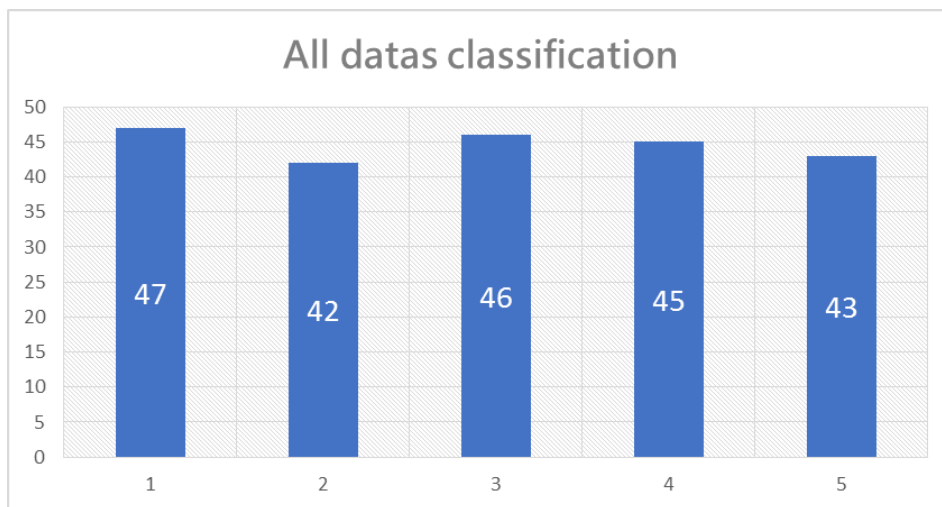
**Table 4-8 XSGS standard established**

Categories	1	2	3	4	5
Data numbers	11	10	10	10	9



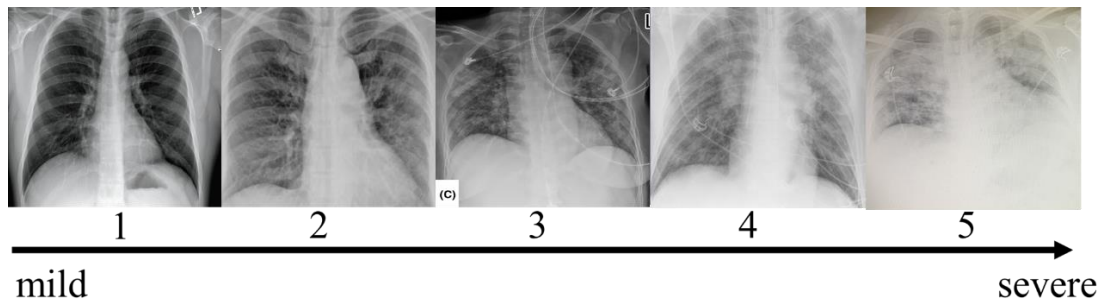
mild severe

Check this data's probability we can get that the boundary lines of XSGS are 50%, 70%, 80%, 85%, 90%. Then we put all the training datas, validation datas, and testing datas into XSGS. We got 47 at 1 degree, 42 at 2 degree, 46 at 3 degree, 45 at 4 degree, 43 at 5 degree.



**Fig. 4-13 XSGS classification result**

Then, we chose data randomly to see the picture is genuine severity or deep learning's misjudgment. Lining up the following pictures, we can find significant difference on each degree. Therefore, model provides accurate classifications.



**Fig. 4-14 XSGS result sample B**

#### 4-6 Grad-CAM result

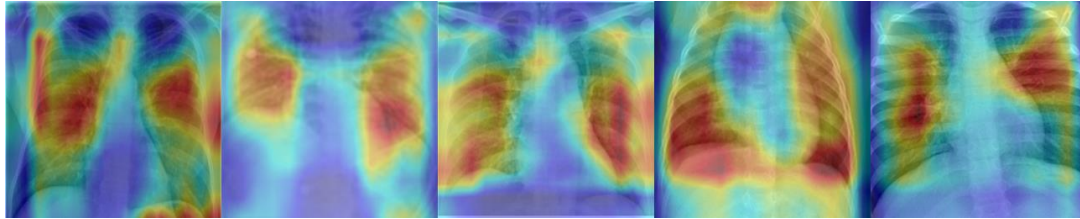
**Fig 4-15** shows us the most Bacterial image result’s types of the model watching on Chest X-ray image. It can focus on the diseased part precisely. Therefore, we can proof the model is having accurate weight on the right side instead of rib or medical equipment and other non-pathology factor. And this can also be the disease location reference provides to doctors on the therapy process. For example, the middle image shows that the patient’s right upper lobe and left lower lobe have a more serious infection.



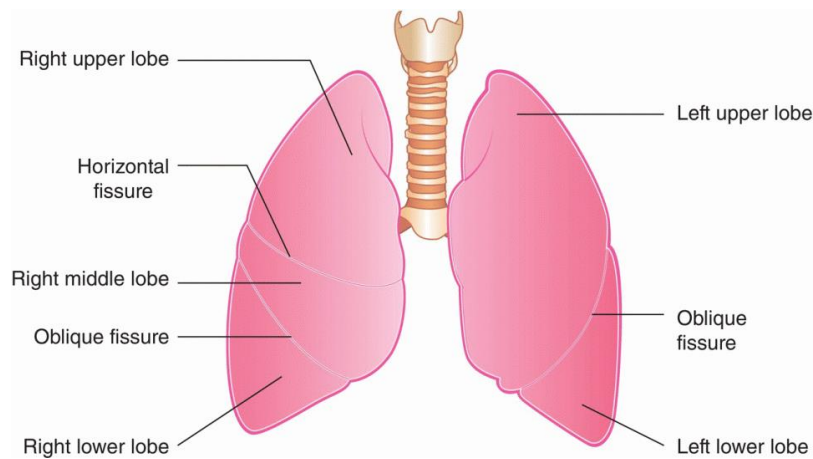
**Fig. 4-15 Bacterial image Grad-CAM result example**

**Fig 4-16** shows us the most COVID-19 image result’s types of the model watching on Chest X-ray image. It can focus on the diseased part precisely. Therefore, we can proof the model is having accurate weights on the right side instead of rib or medical equipment and other non-pathology factor. According to **Fig 4-17**, we could manual calculate five lobes of the lung’s infection rate. Finally, we statistics the COVID-19 dataset and get the result. Right lower lobe has the highest infection rate

(87%) and Right middle lobe has the least inflection rate. Compared to Bao et al's research[26], we got similar result with them on inflection rate of lobe. Therefore, we can prove that our model focusses on the lesion area exactly.



**Fig. 4-16 COVID-19 image Grad-CAM result example**

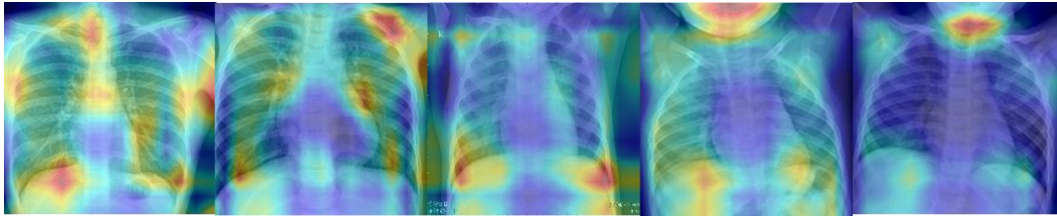


**Fig. 4-17 Lung structure [26]**

**Table 4-9 COVID-19 region inflected result**

Region	Right upper lobe	Right middle lobe	Right lower lobe	Left upper lobe	Left lower lobe
rate	62%	47%	87%	65%	70%
Bao et al[26]	65%	55%	87%	69%	81%

**Fig 4-18** shows us the normal images' different types of the model watching on Chest X-ray image. It didn't focus on the lung so much so that can assure the image didn't have the pathology factor in lung.



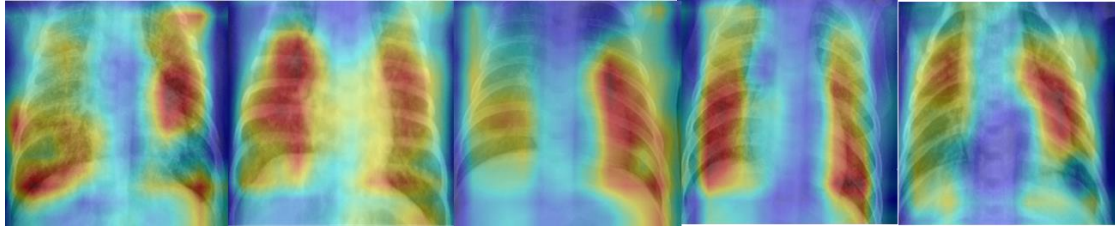
**Fig. 4-18 Normal image Grad-CAM result example**

**Fig 4-19** shows us the tuberculosis images' distinct types of the model watching on Chest X-ray images. Tuberculosis may cause the cavity in the lung and it showed like a round shadow in x-ray image. For example, the second one from the right's image has an apparently round heat spot at right upper lobe. The one on the far left image has an apparently round heat spot at left upper lobe.



**Fig. 4-19 Tuberculosis image Grad-CAM result example**

**Fig 4-20** shows us the most Viral image result's types of the model watching on Chest X-ray image. It can focus on the diseased part precisely. Therefore, we can proof the model is having accurate weight on the right side instead of rib or medical equipment and other non-pathology factor. And this can also be the disease location reference provides to doctors on the therapy process. For example, The one on the far right image shows that the patient's right upper lobe and left upper lobe have a more serious infection.



**Fig. 4-20 Viral image Grad-CAM result example**

## **5. Conclusion and Future works**

Early prediction and triage are vital to save people in the world. In this research, we proposed a method that can integrate both things, which can reduce frontline medical workers' workload and provide medical assistance to more people. First, we trained an acceptable accuracy model. Next we used double validation to increase the accuracy to 84.91%. Then, we put forward XSGS, (X-ray severity grading system). Final, we proved our model can focus on diseased part precisely by using Grad-CAM and got similar percentage on COVID-19 infected area.

In future study, we'll segment the lung in the Chest x-ray to lobes, so that we can count the infected percentage more precisely. We will apply this system on other diseases which can be detected in Chest X-ray image in order to make this system more functional. Moreover, we will turn the local model into a website and an application so that more frontline staff can test suspected cases rapidly. The final objective is to combine several diagnostic procedures, such as rapid test, CT, RT-PCR(COVID-19). For other diseases, we will include Thoracoscopy, Bronchoscopy and many more.



## References

- [1] Sariol, Alan, and Stanley Perlman. "Lessons for COVID-19 immunity from other coronavirus infections." *Immunity* (2020).
- [2] Petersen, Eskild, et al. "Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics." *The Lancet infectious diseases* (2020).
- [3] Lauer, Stephen A., et al. "The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application." *Annals of internal medicine* 172.9 (2020): 577-582.
- [4] Wölfel, Roman, et al. "Virological assessment of hospitalized patients with COVID-2019." *Nature* 581.7809 (2020): 465-469.
- [5] Li, Long-quan, et al. "COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis." *Journal of medical virology* 92.6 (2020): 577-583.
- [6] Laventhal, Naomi, et al. "The ethics of creating a resource allocation strategy during the COVID-19 pandemic." *Pediatrics* 146.1 (2020).
- [7] Emanuel, Ezekiel J., et al. "Fair allocation of scarce medical resources in the time of Covid-19." (2020): 2049-2055.
- [8] Pan, Lei, et al. "Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China: a descriptive, cross-sectional, multicenter study." *The American journal of gastroenterology* 115 (2020).
- [9] Peppas, M. V., et al. "URBAN TRAFFIC FLOW ANALYSIS BASED ON DEEP LEARNING CAR DETECTION FROM CCTV IMAGE SERIES." *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42.4 (2018).
- [10] Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." *nature* 529.7587 (2016): 484-489.

- [11] Lopez, Marc Moreno, and Jugal Kalita. "Deep Learning applied to NLP." arXiv preprint arXiv:1703.03091 (2017).
- [12] Beichel, Reinhard, et al. "Robust active appearance models and their application to medical image analysis." *IEEE transactions on medical imaging* 24.9 (2005): 1151-1169.
- [13] Beichel, Reinhard, et al. "Robust active appearance models and their application to medical image analysis." *IEEE transactions on medical imaging* 24.9 (2005): 1151-1169.
- [14] Mavandadi, Sam, et al. "Distributed medical image analysis and diagnosis through crowd-sourced games: a malaria case study." *PloS one* 7.5 (2012): e37245.
- [15] Rajkomar, A., Lingam, S., Taylor, A. G., Blum, M., & Mongan, J. (2017). High-throughput classification of radiographs using deep convolutional neural networks. *Journal of digital imaging*, 30(1), 95-101.
- [16] Shan, Fei, et al. "Lung infection quantification of covid-19 in ct images with deep learning." arXiv preprint arXiv:2003.04655 (2020).
- [17] Wang, Linda, and Alexander Wong. "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images." arXiv preprint arXiv:2003.09871 (2020).
- [18] Asnaoui, Khalid El, Youness Chawki, and Ali Idri. "Automated methods for detection and classification pneumonia based on x-ray images using deep learning." arXiv preprint arXiv:2003.14363 (2020).
- [19] Yang, Ran, et al. "Chest CT severity score: an imaging tool for assessing severe COVID-19." *Radiology: Cardiothoracic Imaging* 2.2 (2020): e200047.
- [20] Chowdhury, Muhammad EH, et al. "Can AI help in screening viral and COVID-19 pneumonia?." arXiv preprint arXiv:2003.13145 (2020).
- [21] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks

for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[22] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[23] He, Kaiming, et al. "Identity mappings in deep residual networks." European conference on computer vision. Springer, Cham, 2016.

[24] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[25] Szegedy, Christian, et al. "Inception-v4, inception-resnet and the impact of residual connections on learning." arXiv preprint arXiv:1602.07261 (2016).

[26] Bao, Cuiping, et al. "Coronavirus disease 2019 (COVID-19) CT findings: a systematic review and meta-analysis." Journal of the American college of radiology 17.6 (2020): 701-709.

[27] <https://thoracickey.com/2-embryology-anatomy-and-physiology-of-the-lung/>

## 【評語】 190022

本研究主題清楚，且可用科學方法檢驗研究成果。實驗設計尚屬完整，資料與分析。建議可多針對相關研究進行比較分析，