COVID-19 Radiography Using ConvNets

Aryaman Gokarn
Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
gokarnaryaman09@gmail.com

Khushi Patni
Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
khushipatni18@gmail.com

Yuvraj Purohit
Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
yuvrajpurohit52@gmail.com

Dr. Reena Sonkusare
Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
reena_kumbhare@spit.ac.in

***Abstract*—The COVID-19 pandemic continues to have a negative impact on the fitness and well being of the worldwide population. A vital step in tackling the COVID-19 is a successful screening of patients, with one of the key screening approaches being radiological imaging using chest radiography. This study aims to automatically identify patients with COVID-19 pneumonia using digital x-ray images of the chest while increasing the accuracy of the diagnosis using Convolution Neural networks (CNN). The data-set consists of 5380 X-ray images consisting of 1345 X-ray images each of COVID patients, Lung Opacity, Normal patients and Viral Pneumonia. In this study, CNN based model have been proposed for the detection of coronavirus pneumonia infected patients using chest X-ray radiography and gives a classification accuracy of 93.77% (training accuracy of 99.81% and validation accuracy of 95.45%).**

***Index Terms*—COVID-19, CNN, Radiography, Deep Learning, VGG-16, VGG-19**

I. Introduction

The clinical indications for imaging, specifically chest radiography and chest CT, have advanced since the initial discovery of disease. Computed Tomography (CT) is predominant in Covid-19 radiography. However, a comprehensive description of chest radiography image in relation to the disease time course is sparse. For the foreseeable future, COVID-19 is going to be a crucial distinctive diagnosis for anyone presenting to the hospital with a flu-like illness on a full blood count or a change in their typical sense of taste or smell. Most people infected with Covid-19 do not develop pneumonia; however, chest radiography of critically ill patients with respiratory symptoms when they arrive at the hospital may be helpful in diagnosing those with Covid-19 pneumonia.

Machine learning techniques hold great promise for detecting and prognosing coronavirus disease (COVID-19) from standard-of-care chest radiographs (CXR) and chest computed tomography (CT) images. Many research studies describing new machine learning-based models for these tasks were published in 2021, but it is unclear which are of potential clinical utility. As fast assessment and reporting from an onsite or distant radiologist is not always possible, we provide guidance to non-radiologists on how to look for abnormalities on chest radiographs that may be suggestive of Covid-19, Pneumonia, and Lung opacity.

In this paper, we compare performances of Convolution Neural Network (CNN), VGG-16, and VGG-19 algorithms based on the accuracy, loss, and confusion matrix. We introduce the evaluated machine learning algorithms in Section III. The dataset for the Covid-19 radiography is described in Section IV while the experimental procedure and result analysis are given in Section V and Section VI respectively. We conclude with Section VII.

II. Related Theory

A Convolution Neural Network consists mainly of an input layer, hidden layers and an output layer. In any feed-forward neural network, the middle layers are referred as hidden layers because their inputs and outputs are masked by the activation function and final convolution.

To diagnose Covid-19 at its early stages of development, X. Ouyang, J. Huo, L. Xia developed a dual-sampling attention network to automatically identify COVID-19 from the community acquired pneumonia (CAP) in chest computed tomography (CT). In particular, they proposed a novel online attention module with a 3D convolutional neural network to focus on the infected regions in lungs when making decisions of diagnosis. Their method is evaluated upon the data for COVID-19 from 8 hospitals. In the training-validation stage, they collected 2186 CT scans from 1588 patients for a 5-fold cross-validation. In the testing stage, they employed another independent large-scale testing dataset including 2796 CT scans from 2057 patients. Results show that their algorithm can classify the COVID-19 images with the area under the receiver operating characteristic curve AUC value of 0.944, accuracy of 87.5%, sensitivity of 86.9%, specificity of 90.1 %, and F1-score of 82.0%. With this performance, the proposed algorithm could potentially aid radiologists with COVID-19 diagnosis

978-1-6654-5659-3/22/$31.00 ©2022 IEEE

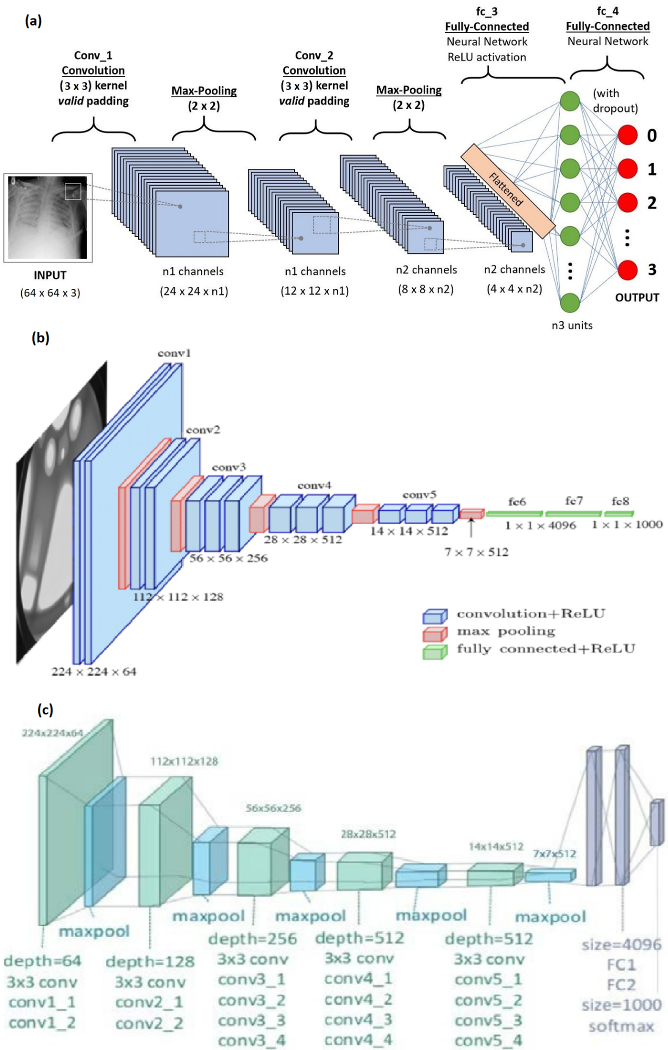

Fig. 1. Initial and final layers of Convolution Neural Network architecture employed in this study. (a) CNN (b) VGG-16 (c) VGG-19

from CAP, especially in the initial stage of the COVID-19 outbreak.[1]

Author H. Mike, C. Subrata and et.al. have selected the VGG19 CNN model for the image modalities to show how the models can be used for the highly scant and challenging COVID-19 datasets. Their approach is specifically aimed to reduce undesirable noise from the images so that deep learning models can focus on detecting diseases with particular features from them. VGG19 model used in this study has been extensively tuned with appropriate parameters, performs in considerable stages of COVID-19 detection against pneumonia or normal for all three lung image modes with the accuracy of up to 86% for X-Ray, 100% for Ultrasound and 84% for CT scans.[2]

Due to the emerging nature of the COVID-19 pandemic, a systematic collection of CXR dataset for deep neural network training is difficult. To address this problem, Y. Oh, S. Park, J. C Ye proposed a patch-based convolutional neural network approach with a relatively small number of trainable parameters for COVID-19 diagnosis. [3]

III. Proposed System Design

A. Convolution Neural Network

Deep learning has piqued the interest of researchers in recent years. Convolutional Neural Network (CNN) is a type of artificial neural network, which is a sub-field of deep learning that has outperformed traditional machine learning methods in every domain.

The hidden layers in a convolutional neural network, as shown in Fig. 1 (a) include convolutional layers. The convolution operation generates a feature map as the convolution kernel slides along the input matrix for the layer, which then contributes to the input of the next layer. This is followed by a pooling layer, which reduces data dimensions by combining neuron cluster outputs from one layer into a single neuron in the next layer. Global pooling affects all neurons in the feature map. The Max pooling technique is used, which takes the maximum value of each local cluster of neurons in the feature map. This layer is then followed by a fully connected layer that connects every neuron.

B. VGG-16

VGG16 is a CNN architecture proposed by K. Simonyan and A. Zisserman [12], which consists of 21 layers (13 convolutional layers, 5 max-pooling layers, and 3 fully connected layers). Therefore, the number of layers having tunable parameters is 16 out of which thirteen are convolutional layers and three are fully connected layers. The first block contains 64 filters, which are then doubled in subsequent blocks until the total reaches 512. Two fully connected hidden layers and one output layer round out this model.

C. VGG-19

VGG-19 is an advanced CNN architecture with pre-trained layers that is 19 layers deep. It has been trained on millions of diverse images with complex classification tasks.

As we can see in Fig. 1 (c), it accurately depicts the VGG-19 architecture. This architecture is basically composed of 3 types of layers i.e. Convolution layer to extract the feature from the image by employing different number and types of filters, Max-pooling layer to decrease the image size and to extract the feature from the feature map obtained from these filters present in the Convolution layer , Flatten layer to turn the batches of feature maps into 1D tensor and finally 3 Fully-Connected where first two have a dense unit of size 4096, and final classification layer has 1000-way ILSVRC classification. The final layer is the soft-max layer. The configuration of the fully connected layers is the same in all networks. The final classification layer has a softmax activation for predicting the probabilities of each class. [10]

TABLE I
COMPARING AND ANALYSING

Author	Dataset	No. of Images	Method	Accuracy
Ouyang1001[1]	Internal	2796	Attention Mechanism	87.5%
J. Horry[2]	COVID-CT dataset	746	VGG19	84%
Y.Oh[3]	JSRT	502	ResNet-18	88.9%
Zhang[4]	CC-CCI	61775	3D Resnet-18	92.49%
Proposed Method	COVID-19 Radiography	5380	VGG-19	93.77%

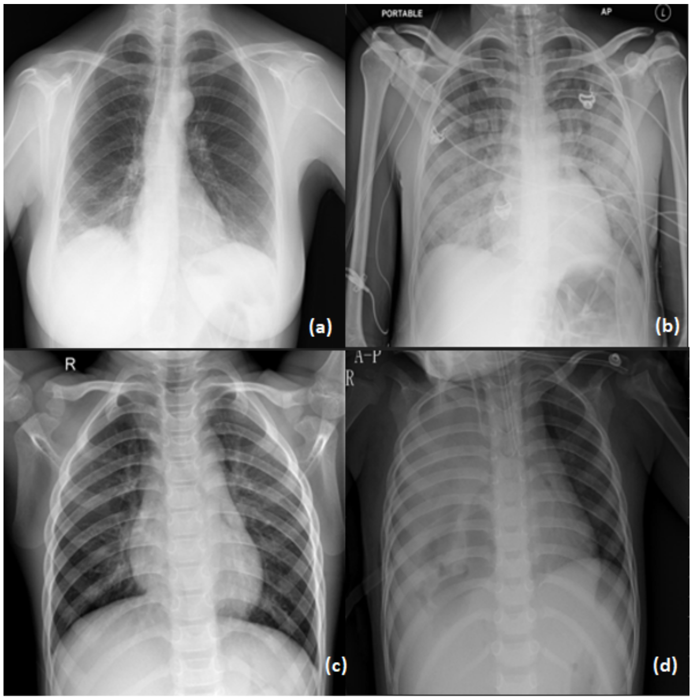

Fig. 2. Example images from the Dataset for four categories such as (a) Covid-19, (b) Lung Opacity, (c) Normal and (d) Viral Pneumonia

IV. Description Of Dataset

The Kaggle dataset [7] was obtained by a team of researchers from Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh along with their collaborators from Pakistan and Malaysia in collaboration with medical doctors. The dataset consists of chest X-ray images for COVID-19 positive cases along with Normal and Viral Pneumonia images.

The dataset contains 5380 radiography images divided into four categories: Covid-19, Lung Opacity, Normal, and Pneumonia as shown in Fig. 2. Each category contains 1345 images. For the evaluation, we split the images in a 6:2:2 ratio for the train/test/validation set per category.

V. Experimentation

The workflow of this study begins with collection of primary dataset containing four image classes: first class attributing to chest X-rays of COVID-19 patients, second belonging to chest X-rays of Viral Pneumonia, next class belonging to chest X-rays of Lung Opacity and the final class belonging to the healthy patients. We performed various operations on the dataset like rescale, zoom, changing the size of the image and other image transformation steps. The resulting dataset was used to train the model in the next phase.

The Proposed CNN Architecture consists of 6 layers in which 2 are convolutional (Conv2D) layers, 2 max pooling layers, 3 activation layers, 1 flatten layer, and 1 fully connected layers; CNN model input image shape is (64, 64, 3), i.e., 64-by-64 RGB image. In all Conv2D layers, a 3×3 size kernel has been used but the filter size after every Con2D layer increases. At the 1st layer of Con2D, 128 filters have been used to learn from input and the 2nd layer of Con2D uses 64 filters. After each Con2D layer, the max pooling layer with 2×2 pooling size has been used.

In the proposed VGG-16 Architecture, the input to the conv1 layer is of a fixed size of 224 x 224 x 3, i.e., RGB image. The input image is processed by a stack of convolutional (conv.) layers, where the filters were used with a very small receptive field: 3×3. For one of the configurations, it utilizes 1×1 convolution filters, which can be act as a linear transformation of the input channels. The convolution stride is set to one pixel, and the spatial padding of the convolution layer input is set to keep the spatial resolution after convolution, i.e. 1-pixel padding for 33 conv. layers. 5 max-pooling layers follow some of the convolution layers and perform spatial pooling. The Max-pooling is done with stride 2 over a 22 pixel window. Following a stack of convolutional layers, three Fully-Connected (FC) layers are added: the first two have 4096 channels each, while the third performs 1000-way ILSVRC classification and thus has 1000 channels (one for each class), as shown in Fig. 1 (b) and the final layer is the soft-max layer. At last the configuration of the fully connected layers is the same in all networks, while all hidden layers are equipped with the rectification (ReLU) non-linearity.

The proposed VGG-19 Architecture takes a RGB image of size (224 * 224) as an input to this network, implying that the matrix was of shape (224,224,3). The network employs kernels of (3 * 3) size with a stride size of 1 pixel to cover the entire image concept, and spatial padding is employed to preserve the image's spatial resolution. Max pooling is done with stride size of 2 over a 2 * 2 pixel window. It is followed by Rectified linear unit (ReLu) to bring non-linearity into the model to enhance classification and computational speed. The model is concluded by three fully linked layers, first two of which are 4096 in size, followed by a layer with 1000 channels for 1000-way ILSVRC classification, and the final layer being a softmax function.

TABLE II
THE BEST PERFORMANCE OF CNN, VGG-16 AND VGG-19 MODELS: KAGGLE DATASET

Model	Accuracy	Val Accuracy	Test Accuracy
CNN	96.93 %	88.57 %	84.75 %
VGG-16	94.14 %	91.45 %	89.77 %
VGG-19	99.81 %	95.45 %	93.77 %

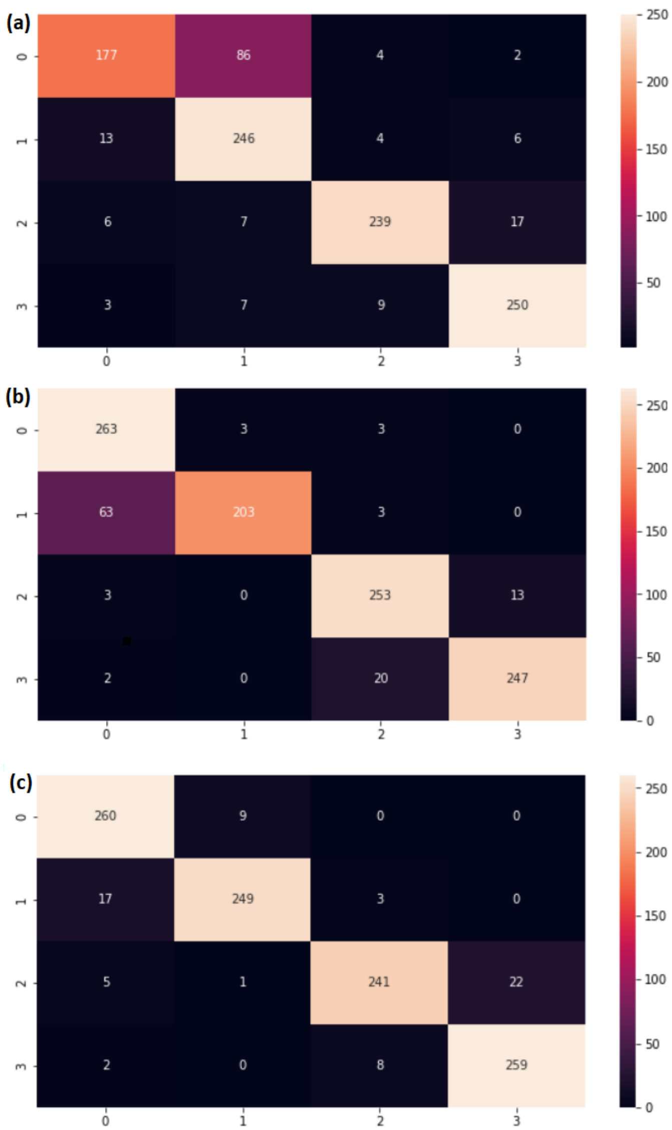

Fig. 3. Confusion Matrix (a) CNN, (b) VGG-16, (c) VGG-19

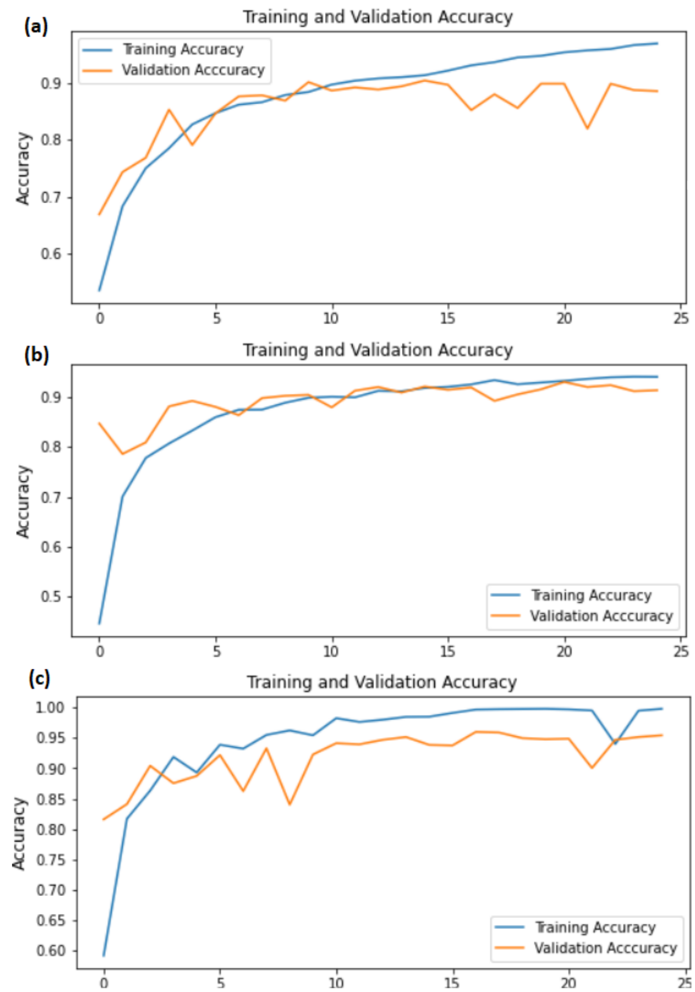

Fig. 4. Training and Validation Accuracy (a) CNN, (b) VGG-16, (c) VGG-19

VI. RESULT ANALYSIS

Our dataset consists of 5380 X-ray images consisting of 1345 X-ray images each of COVID patients, Lung Opacity, Normal patients and Viral Pneumonia. The comparison between various model approach and our proposed method is listed in Table 1.

Models can be evaluated using different criteria, such as classification accuracy, sensitivity (true positive rate), and specificity. The accuracy, sensitivity, and specificity are calculated as given in Equation (1), Equation (2), and Equation (3), respectively:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$Sensitivity = TP/(TP + FN) \quad (2)$$

$$Specificity = TN/(TN + FP) \quad (3)$$

where TP and TN denote the true-positive and true-negative values, respectively; and FP and FN represent false-positive and false-negative values, respectively.

Macro-averaged F1 score is a measure of model performance for multi-class (multi-label) problems that have more than two output classes, when the data is imbalanced, and the accuracy is not reliable. It considers the harmonic mean of recall and precision scores of all classes separately, and measures the capacity of the model for the correct detection of samples.

Confusion Matrix, as shown in Fig. 3, gives a comparison between Actual and predicted values. The confusion matrix is a 4 x 4 matrix, where 4 is the number of classes (covid-19, pneumonia, lung opacity and normal).

The performance of the model was evaluated using the training curve as plotted in Fig. 4. The accuracy after the 25 epochs for CNN, VGG-16 and VGG-19 model was found to be 84.75%, 89.77% and 93.77% respectively, as shown in Table II, in which a value of 100% indicates a perfect accuracy. The loss, on the other hand, keeps on decreasing with each epoch and then it reaches critical point where it almost remains constant.

All experiments were performed on a Windows 11 64-bit operating system, Intel Core i7-11800H CPU @2.30 GHz × 12, 16 GB RAM, NVIDIA GeForce RTX3050 GPU.

VII. Conclusion

We compared performances of CNN, VGG-16 and VGG-19 architectures using the Kaggle dataset. Performance metrics such as accuracy, F-score, precision, and confusion matrix were used. We have demonstrated that with current limited and challenging COVID-19 datasets, our model could be used to develop suitable deep learning-based tools for COVID-19 detection. The model is capable of classifying Normal, COVID-19, Pneumonia and lung opacity conditions from X-Ray. Our study uncovers the challenging characteristics of the limited COVID-19 image datasets. This should be helpful for practitioners aiming to use these datasets for their research and development.

References

[1] X. Ouyang, J. Huo, L. Xia et al., "Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia," IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2595–2605, 2020.

[2] H. Mike, C. Subrata, M. Paul et al., "COVID-19 detection through transfer learning using multimodal imaging data," IEEE Access, vol. 8, pp. 149808–149824, 2020.

[3] Y. Oh, S. Park, J. C Ye, and ", "Deep learning COVID-19 features on CXR using limited training data sets," IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2688– 2700, 2020.

[4] K. Zhang, X. Liu, J. Shen et al., "Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography," Cell, vol. 181, no. 6, pp. 1423–1433, 2020.

[5] Shervin Minaee, Rahele Kafieh, Milan Sonka, Shakib Yazdani, Ghazaleh Jamalipour Soufi, Deep-COVID: Predicting COVID-19 From Chest XRay Images Using Deep Transfer Learning, Medical Image Analysis (2020).

[6] Very Deep Convolutional Networks for Large-Scale Image Recognition (ICLR 2015).

[7] COVID-19 Radiography Database. https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database. Accessed: Aug. 30, 2022

[8] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, "Can AI help in screening Viral and COVID-19 pneumonia?" IEEE Access, Vol. 8, 2020, pp. 132665 - 132676.

[9] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S. and Chowdhury, M.E., 2020. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images. arXiv preprint arXiv:2012.02238.

[10] Transfer Learning — Part — 4.0!! VGG-16 and VGG-19. https://becominghuman.ai/transfer-learning-part-4-0-vgg-16-and-vgg-19-d7f0045032de . Accessed: Aug. 30, 2022

[11] Cleverley J, Piper J, Jones M M. The role of chest radiography in confirming covid-19 pneumonia BMJ 2020; 370 :m2426 doi:10.1136/bmj.m2426

[12] VGG 16 Easiest Explanation. https://medium.com/nerd-for-tech/vgg-16-easiest-explanation-12453b599526