

# COVID-19 Detection based on Lung CT Images using Convolutional Neural Network Architecture

Vangala Hymavathi<sup>1</sup> and P. Sreekanth<sup>2</sup>

<sup>1</sup>PG Scholar, CVR College of Engineering/ECE Department, Hyderabad, India  
Email: vangala.hymavathi94@gmail.com

<sup>2</sup>Asst. Professor, CVR College of Engineering /ECE Department, Hyderabad, India  
Email: sreekanth.isoft@gmail.com

**Abstract:** A CNN (Convolutional Neural Network) is a class of deep neural networks that are most used for analyzing visual imagery. This is the most widely used deep learning algorithm for image and video recognition, image classification, image segmentation, medical image analysis and natural language processing, and many more. But, training the Neural Network from scratch requires more training time and requires a lot of data, and the performance of the model is affected. ‘Transfer learning’ is the best approach where existing architectures that have been trained on the large datasets are then fine-tuned for our customized dataset. This reduces the time to train and often results in better performance.

The main objective of this paper is to detect COVID-19 from lung CT images using the VGG16 CNN (Convolutional Neural Network) model from the ImageNet.

**Index Terms:** CNN, Transfer learning, VGG16, and ImageNet.

## I. INTRODUCTION

The real-time Reverse Transcription Polymerase Chain reaction (RT-PCR) is the test used for detecting the covid-19 virus which is a time-consuming process and doesn't give accurate results with tiny quantities of RNA taken from a person's throat and nose. Computed Tomography (CT) imaging gives more accurate results. But it involves radiology experts to analyze the image as it takes a considerable amount of time. So, automated analysis of CT images is desirable. CNN's (Convolution Neural Networks) are used to detect COVID-19 infected patients positive or negative using lung CT images.

CNN (Convolutional Neural Network) is a type of feed-forward Artificial Neural Network in which the connectivity pattern between the neurons resembles the visual cortex of the living beings, most probably human beings.

Any CNN consists of the following:

1. The input layer is an image.
2. The output layer is binary or multi-class labels.
3. Hidden layers consist of convolution layers, ReLu (Rectified Linear Unit) layers, pooling layers, and a fully connected Neural Network.

ReLu function: It is a linear activation function that has thresholding at zero. The merging of gradient descent can be enhanced by applying ReLu.

$$f(x) = \max(0, x)$$

There are various architectures available in ImageNet out of which VGG16 Architecture has been chosen because of its unique structure. It has 16 convolution layers that have

tunable parameters and other layers cannot be tuned. It uses small filters also called kernels which are of 3\*3 sizes.

Example of CNN Architecture:

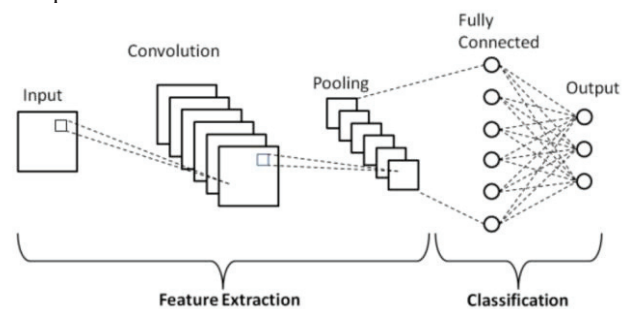


Figure 1. Basic architecture of Convolutional Neural Networks-image from Researchgate.net [6].

There are several traditional architectures available in the ImageNet which are trained on about 15 million labeled high-resolution images. The weights of these available networks can be used to train our model on our own dataset or for our problem. Which is the added advantage for improving the accuracy of the model instead of training the model from scratch and time to train the model and computing power is reduced. In this paper VGG16, architecture is used, which was the runner up in the ImageNet Large Scale Visual Recognition Competition (ILVRSC) in 2014 with an error rate of 7.3.

The VGG-16 is one of the most popular pre-trained models for image classification. Introduced in the famous ILSVRC 2014 Conference, it was and remains the model to beat even today. Developed at the Visual Graphics Group at the University of Oxford, VGG-16 beat the then standard of AlexNet and was quickly adopted by researchers and the industry for their image Classification Tasks. This model is simple and intuitive in nature hence, its function is commendable. Trained on the ImageNet corpus, and thus cemented its place in the list of top pre-trained models for image classification.

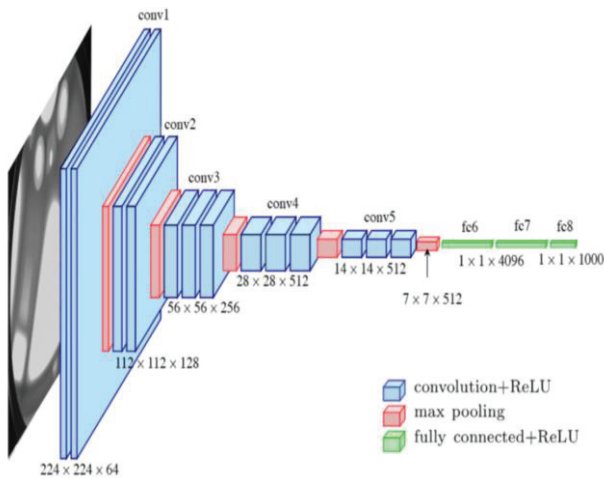


Figure 2. VGG16 Architecture-image from Researchgate.net [5].

Our work is organized by downloading the VGG16 model with the weights of the ImageNet and freezing those weights to be made useful for our problem and then designing the model by tuning the architecture of the base model by freezing the layers that are needed, which means the layers of the model are made as non-trainable. Then add the features, that are needed for the base model to get the desired output for the problem. After the tuning of the architecture, the model is trained with the lung CT images of COVID-19 positive and negative patients on the tuned architecture and tested the model by giving the random image to detect whether it is a COVID-19 positive or negative image.



Figure 3a. COVID-19 positive lung CT image.



Figure 3b. COVID-19 negative lung CT image.

The data is collected from the GitHub platform which consists of 348 COVID-19 infected lung CT images and 397 negative lung CT images collected from 216 patients confirmed by senior radiology experts. The data used for the proposed model consists of 250 COVID-19 infected lung CT images taken out of 348 COVID-19 infected lung CT images and 250 COVID-19 negative lung CT images taken out of 397 negative lung CT images till the 348 range for training the model (<https://github.com/UCSD-AI4H/COVID-CT>).

For validating the model 50 images of both COVID-19 infected lung CT images and negative lung CT images are collected. For testing the model 48 lung CT images of both COVID-19 infected and negative are used.

The preprocessing of the images is done using the Keras preprocessing library.

## II. RELATED WORK

### A. Convolutional Neural Networks

Convolutional Neural Networks are widely being used for object recognition and used in other domains like object tracking and now it is being used for medical diagnosis using available architectures of CNN like AlexNet, GoogleNet, VGGNet, Inception v4, ResNet.

(Khan & Yong, 2017) proposed convolution Neural Network after experimenting on LeNet, AlexNet, and GoogleNet architectures for classifying the medical anatomy images which outclassed the three architectures with an accuracy of 81% and validation accuracy of 76% [1].

(HORRY, et al., 2020) contributed to provide a theoretical transfer learning framework to support COVID-19 detection with the use of image classification using deep learning models for multiple imaging modes including X-Ray, Ultrasound, and CT scan. The acquisition of a sufficiently large, publicly available corpus of medical image sample data for fully training deep learning models is challenging for novel medical conditions such as COVID-19. since the collection and labeling of images requires significant time and resources to compile. They evaluated different traditional CNN architectures using Transfer learning on multimodal images of COVID-19 infected and normal.

The selected model performed better in classifying the X-ray with 86% precision and ultrasound images with a precision of 100% when compared with the CT images with a precision of 84%. But CT scan imaging is the best method for the accurate diagnosis as it gives clear visualization of bones, soft tissues, and blood vessels [2].

(Yang, et al., 2020) contributed to building an open-source dataset of COVID-19 images to encourage the development of AI models which predict whether a person is affected with COVID-19 by analyzing his/her CTs. Using this dataset, they developed an approach based on multi-task learning and contrastive self-supervised learning that achieved an accuracy of 0.89, on a test set of original CT images donated by hospitals by using two models DenseNet-169 and ResNet-50 [3].

(Simonyan & Zisserman, 2014) contributed to the evaluation of the networks increasing depth by proposing an architecture with very small (3x3) convolution filters, which shows that significant improvement on the prior-art configurations can be achieved by pushing the depth to 16-19

weight layers. These findings were the basis of their ImageNet Challenge 2014 submission, where the team secured the first and the second places in the localization and classification tracks respectively [4].

### III. PROPOSED MODEL

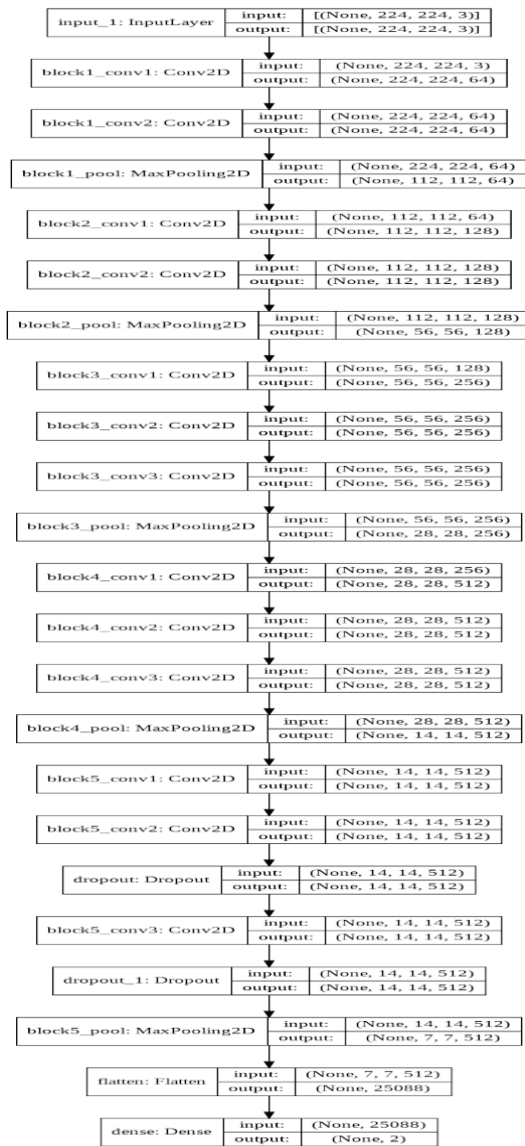


Figure 4. Proposed VGG16 CNN architecture.

The proposed model is the tuned version of the VGG16 model by modifying the top of the model by removing the fully connected layers and applying the flatten method right after the max-pooling layer of block 5 of the VGG16 model and applying the Softmax activation function for the classification of the COVID-19 CT images.

Softmax activation function is given as:

$$\sigma(Z)_i = \frac{e^{Z_i}}{\sum_{j=1}^k e^{Z_j}}$$

Where  $Z$  = input vector

$e^{Z_i}$  = standard exponential function for input vector

$K$  = number of classes in multi-class classifier

$e^{Z_j}$  = standard exponential function for input vector

The existing VGG16 architecture is modified in the top layers to have the dropout of 0.5 in block5\_conv2 and block5\_conv3 layers and a second dropout layer between the block5\_conv3 layer and block5\_pool layers which prevents the overfitting of the model.

The dataset used for training the proposed model consists of COVID-19 infected lung CT images collected from 216 patients confirmed by a senior radiologist who has been diagnosing and treating COVID-19 patients since the outbreak of this pandemic which is collected from the GitHub platform (<https://github.com/UCSD-AI4H/COVID-CT>).

### IV. IMPLEMENTATION

The code is implemented using Google colab environment which gives free access to GPUs which is required for doing the high computation tasks. Google's TensorFlow is the up-to-date and fast-growing and open source software library that can do high computational jobs and uses data flow graphs where edges denote tensors. The library is publicly available. Keras is the second fast-growing framework and open-source library written in Python and capable of running on top of TensorFlow.

Google colab is the short term used for the Collaboratory which is provided by Google for writing the python code where one can use the free GPUs and TPUs for free for debugging and executing their codes by simply log in to their google accounts and clicking on the authentication link to get access to the files in their google drive. To save the work the user needs to save the files to their google drive because the colab environment is refreshed.

### V. RESULTS

In this paper, the model is trained for 50 epochs and a batch size of 32 with the 500 training samples consisting of COVID-19 positive and Non-COVID images each of 250 samples belonging to their respective category. Initially, the model is trained by downloading the weights from the ImageNet and downloading the architecture without the top and tuning the top of the VGG16 architecture which gained an accuracy of 99% and validation accuracy of 85%.

The below figures showed the accuracy and loss plots of training and validation.

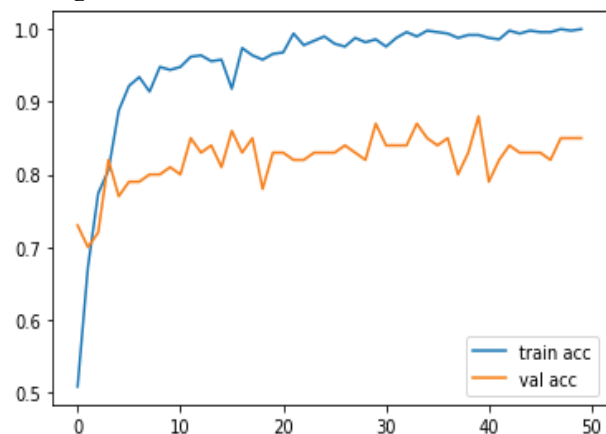


Figure 5. Training and Validation accuracy without dropout.

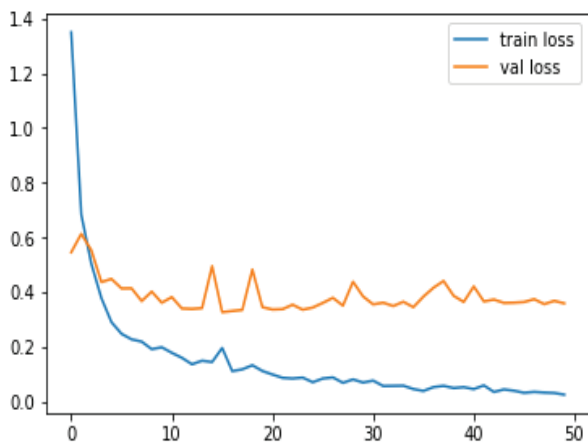


Figure 6. Training and Validation loss without the dropout.

From the above figures, it is observed that the model is overfitting which means the model performs well with the training set but doesn't seem to perform well with the testing image which is unseen by the model is provided.

To overcome the overfitting of the model dropout is added to the proposed architecture and the model is again trained for about 50 epochs and batch size is 32 with 500 images consisting of COVID and Non-COVID lung CT images and achieved the training accuracy of 91% and validation accuracy of about 81%.

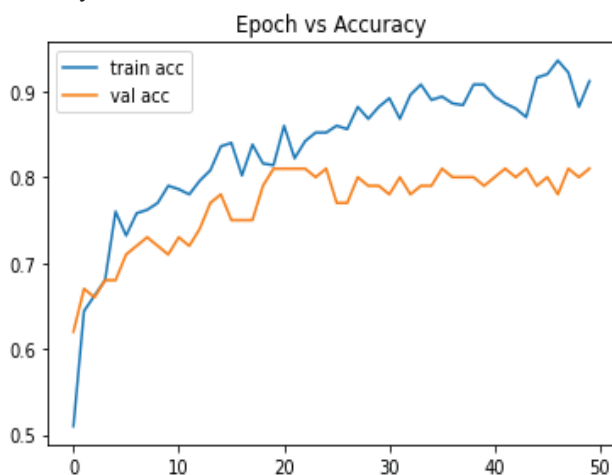


Figure 7. Training and Validation accuracy with dropout.

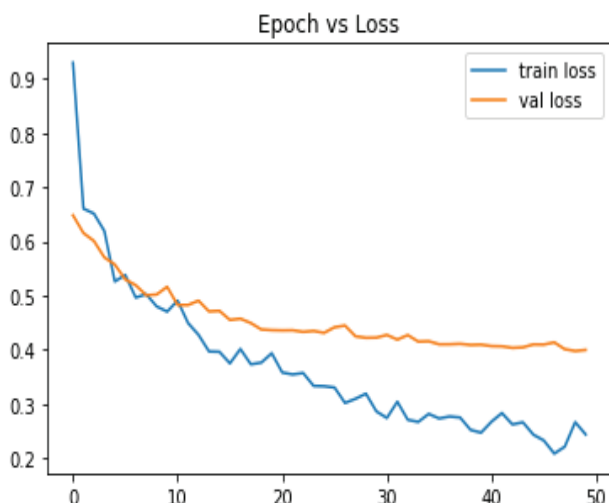


Figure 8. Training and Validation loss with dropout.

The VGG16 architecture acquired 99% accuracy without dropout, ResNet-50 acquired 83% and Inception-V3 with 96% accuracy on the dataset of COVID-19 infected Lung CT images. In this paper, an overview of the only top 3 pre-trained models for image classification is provided. AlexNet is the predecessor of these models and has 62 Million parameters that use different kernel sizes and takes huge amounts of training time as well as computation time, VGG16 is a much more efficient model.

Model	No. of Parameters	Accuracy Achieved
VGG-16	138 Million	99%
ResNet-50	25 Million	83%
Inception V3	24 Million	96%

Note: However, this is a continuously growing domain and there is always a new model to look forward to and push the boundaries further.

## VI. CONCLUSIONS

This paper proposes the tuned version of the existing VGG16 architecture which can be helpful in screening the COVID-19 patients through the lung CT images. The model can be further modified to increase the accuracy, which can be later useful for the automated diagnosis of the COVID-19 when deployed in any hospital database. It can be used for diagnosing other diseases as well if provided with the appropriate dataset consisting of images.

The existing architectures were most easily and consistently degraded in performance through image blurring, which is similar in nature to real-life scenarios of attempting face recognition from low-resolution imagery. Other covariates found to have a considerable effect on the verification performance where noise, image brightness, and JPEG compression impacted the performance of the models only marginally. No specific architecture was found to be significantly more robust than others to all covariates [7]. Even though there are modern architectures VGG16 architecture seemed to perform better with the low-resolution images although it requires much memory and is time hungry. Because of the unique structure of VGG16, it is widely used for object recognition. As the images collected are donated by the hospitals for research and development different CT machines have different resolutions. So, VGG16 architecture is preferred for the proposed model for diagnosing COVID-19 affected patients using lung CT images.

The VGG16 model is tuned in fully connected layers. A dropout layer is added with a dropout rate of 0.5 in block5\_conv2 and another dropout layer between the block5\_conv3 layer and block5 pool layers to prevent the Co-adaptation of neurons that means multiple neurons extracting the similar features are dropped randomly and the output layer is modified to have 2 classes which are COVID and NON-COVID.

### REFERENCES

- [1]. S. Khan and S.-P. Yong, "A deep learning architecture for classifying medical images of anatomy object," in 2017
- [2]. M. J. HERRY, S. CHAKRABORTY, M. PAUL, A. ULHAQ, B. PRADHAN, M. SAHA and N. SHUKLA, "COVID-19 Detection through Transfer Learning using Multimodal Imaging Data," 2020.
- [3]. X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang and P. Xie, "COVID-CT-Dataset: A CT Scan Dataset about COVID-19," 2020.
- [4]. k. simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 2014.
- [5]. M. Ferguson, "Automatic localization of casting defects with convolutional neural networks," IEEE International Conference on Big Data (Big Data) At Boston, MA, USA, December 2017.
- [6]. V. H. Phung and E. J. Rhee, "A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets," no. Machine Learning Techniques Applied to Geospatial Big Data, 2019.
- [7]. Grm, K., Štruc, V., Artiges, A., Caron, M., & Ekenel, H. K. (2018). Strengths and weaknesses of deep learning models for face recognition against image degradations. (pp. 81-89). IET Biometrics.