ISSN (e): 2250 – 3005 || Volume, 12 || Issue, 4|| Jul. - Aug. – 2022 ||
International Journal of Computational Engineering Research (IJCER)

Impact of Covid-19 on Human Lungs and Forecasting the Fatality

*Dr. Pramod [1,] Mr. Madesha M[2,] Mr. Kedar Santosh Prabhu[3] Sabin T T [4]
[1]*Associate Professor, ISE Dept, PESITM, Shivamogga,*
[2] *Assistant Professor, ISE Dept, PESITM, Shivamogga.*
[3] *Final Year Student, ISE Dept, PESITM, Shivamogga,*
[4] *Assistant Professor, IS&E, SJCIT, Chikkaballapur*

ABSTRACT
Data science is an interdisciplinary field that applies information from data to a wide range of applications by leveraging scientific techniques, techniques, algorithms, and systems to extract knowledge and insights from a diverse data types, encompassing noisy, structured, and unstructured data. With the advent of deep learning technologies in data science, it has become possible to get into the minute details of scanned images. With the help of image segmentation, it has become possible to search for defects present in the scanned images. Here, we make use of VGG16 classifier algorithm to classify Computerized Tomography (CT) images into COVID and Non-COVID classes. We make use of training and validation data sets in this study. Using this dataset, we train our model to classify whether the patient is infected with COVID-19 or not. The model is very helpful for the medical community as it saves time required for manual diagnosing and less prone to errors which becomes crucial in saving lives. With timely knowledge about the fatality trend due to COVID-19, governments can develop a strategy to reduce the impacts and prepare well for what lies ahead. In this paper, a forecasting system is developed to predict the deaths due to adverse effect of Covid-19 on lungs using predicting models namely linear regression and Fbprophet. The outcomes of forecasting are visualized using a graph for easier comprehension. Predictions can be made for any amount of duration by setting appropriate period values.
Keywords:*VGG16, Computerized Tomography, forecasting, linear regression, Fbprophet.*

Date of Submission: 25-09-2022 Date of acceptance: 12-10-2022

I. INTRODUCTION

Deep learning enables multilevel abstraction of data using multiple layers that perform complex computations. This has helped to improve the existing architectures in computer vision, medical discoveries, speech recognition rapidly. Which has led to many breakthroughs in technology. The internal parameters that need to be computed in each layer are changed according to the structure in data sets in previous layer by using the back propagation algorithm. Deep convolution network have played a pioneering role in the processing of images, video, speech, and audio, while recurrent networks have shed light on sequential type of data like text and speech. Machine learning involves neural networks, that are at the root of deep learning algorithms. Each neural network comprises of an input layer, one or more hidden layers and an output layer which provides the computational result. Nodes are interconnected with weights and threshold being assigned to it. Whenever the output of any individual node is above threshold value, it gets activated and sends data to the next layer in the network. Convolutional, non-linear, pooling, and fully-connected layers are among the layers that make up a CNN. Pooling and non-linearity layers lack parameters, but convolutional and fully connected layers do so. Each layer has a pre-determined task and upon completion passes on the results to next layer. The CNN has an outstanding performance in machine learning problems.VGG stands for Visual Geometry Group. It comprises of multiple layers and is a standard deep Convolutional Neural Network(CNN)architecture having multiple layers. VGG exists in two forms. They have 16 or 19 convolutional layers and are termed as VGG-16 or VGG-19. The layer receives a dimensionally accurate image as input (224,224,3). The images are segregated after processing,

and the subsequent layers are then overlaid. The result at the end would be a particular class among the two.

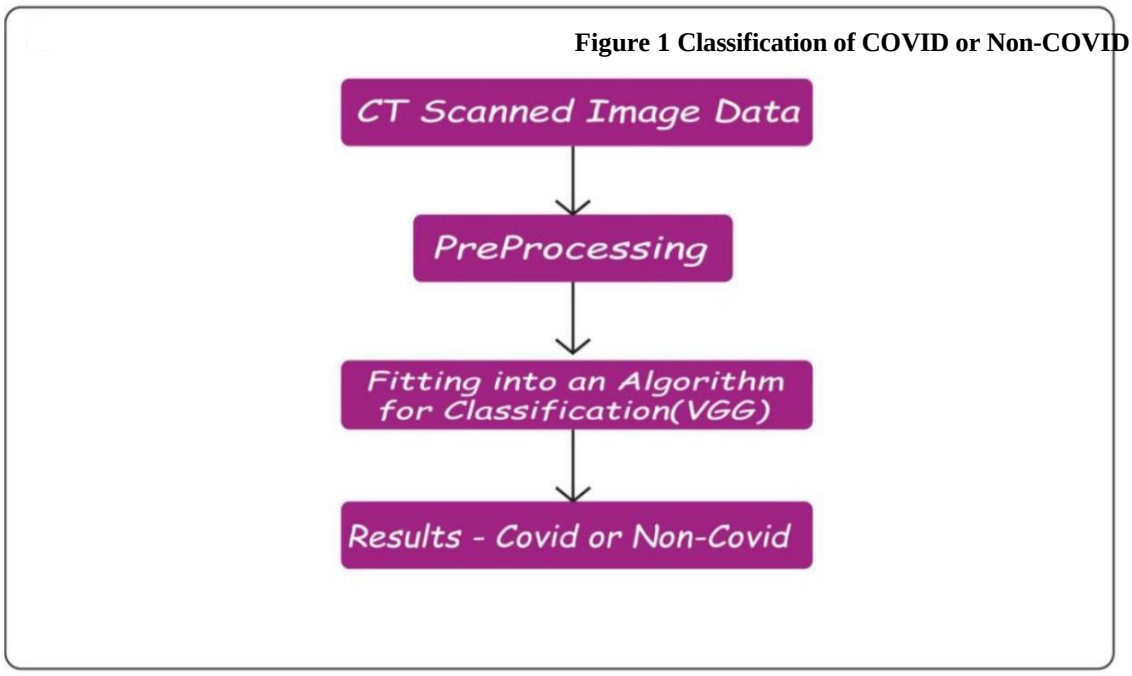
Linear regression is used for forecasting as it is a simple model to work with. Here, the independent variables predict the dependent variables. Regression analysis estimates dependent '*y*' variable value with the help of independent variable values '*x*'. Since this algorithm is simple, we have made use of another algorithm which is superior than a linear regression model which is Fbprophet. Fbprophet is an open-source, univariate forecasting framework which is used for forecasting time series data that takes trend, seasonality and holiday effects which can make the task of forecasting more comprehensive and easier to carry out. It works best with time series that have strong seasonal effects and several seasons of historical data. Fbprophet can handle outliers in data and it is also quite robust to shifts in the trend and missing data. Upon passing the dataset and performing normalization, we fit it into the algorithm. The results are obtained as a graph with better and comprehensive visualizations. This helps us to gain an easier understanding of results obtained without creating confusion or misinterpretation.

II. LITERATURE SURVEY

Author	References	**Contribution**	**Findings or Drawbacks**
Asif, S., Wenhui, Y., Jin, H., & Jinhai, S	[1]	The prediction system of COVID-19 infected patients used chest X-ray images. The normal, pneumonia and COVID-19 classes were obtained.	Since a basic model was developed, the accuracy was low with anomalies.
Yasin, R., Gouda,W	[2]	The study was done using Chest X-Rays (CXRs) to predict the type of disease in lungs. The use of AI techniques to detect the COVID-19 patients using X-Ray images in an automated manner.	CXRs are a simple way of analyzing chest X- Ray images and diagnosis requires strong parameters for analysis.
Kwee, T.C., & Kwee, R. M.	[3]	The system runs proposed image processing algorithm that analyses the view of the lung to discard those CT images that is not properly visible in them. Chest CT is performed to report the findings suggestive of COVID-19 pneumonia detected in patients.	No methods to classify the types of abnormalities in the lungs were discussed.
Guefrechi, S., Jabra, M. B., Ammar, A., Koubaa,A., & Hamam, H.	[4]	Aa deep learning algorithm that analyses chest X-ray pictures to extract characteristics and find COVID-19. ResNet50, InceptionV3, and VGG16, three powerful networks, have been optimized on an upgraded dataset that was created by compiling COVID-19 and typical chest X-ray pictures from many public databases.	The X-Rays do not show a detailed structure of lungs and even with complex classification algorithms, we might be misclassifying the classes.

III. CLASSIFICATION OF CT SCAN IMAGES INTO COVID OR NON-COVID

Figure 1 Classification of COVID or Non-COVID



For the classification model primarily, we collect the CT scanned image dataset having classes as Covid and

Non Covid. The dataset contains 746 CT scanned images of both classes. The pre- processing is done for the image dataset where the images are resized. For normalization the RGB values are encoded as 8bit integers. The image dataset is given as input to the VGG16 algorithm which is a type of CNN architecture. VGG16 refers to 16 convolutional layers, that are provided as input to the network. Each image is of dimensions (224,224) [3]. The hidden layer performs the complex computations. At the end of this stage, we get classification result as Covid or Non Covid. Also, we get the accuracy as percentage. The classification of CT scan images into COVID or Non COVID classes begins with the collection of datasets which in this case is CT scan images of the lungs. We have collected 746 such samples. Then image resizing is performed. We need to make sure all images have similar data distribution among them. So, we perform normalization which ensures equal data distribution for each input parameter (pixel, in this case). This makes convergence faster while training the network. Images with RGB values are encoded as 8bit integers which makes it easier for the algorithm. Here, we make use of VGG 16 classification algorithm. VGG16 is a type of Convolutional Neural Network(CNN) which is very efficient on computer vision and image classification. In this architecture of VGG 16, very small convolution filters (3 x 3)were used which increased the depth of model. This magnified its performance when compared to previous computer vision architectures. We pass the input to this algorithm. Various layers process this input and provide the classification of a particular input image into either of the two classes. Along with this, the accuracy with which the model can classify the image is displayed.

Snapshot of results obtained:

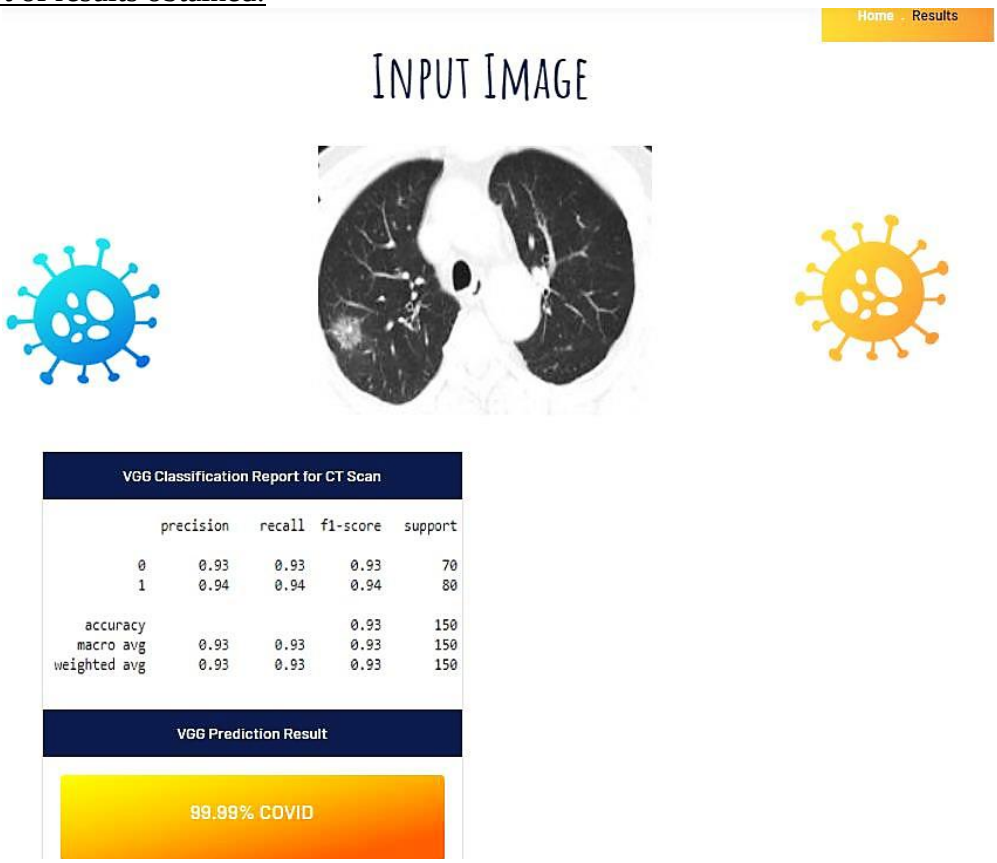


Figure 2 ClassificationresultforCOVID-19

IV. FORECASTING OF FATALITY

The dataset for the forecasting consists of attributes namely date, Country/region, confirmed, recovered, deaths. For the prediction we consider the attributes dates and deaths. The date parameters are formatted to 'YYYY-MM-DD'. The Fbprophet algorithm considers the attributes as parameters 'ds' and 'y'. For prediction, a data frame needs to be created with the future dates. These are termed to as periods by FBprophet. The settings for the periods specify how many days should be produced. According with methods employed, 365 days function as the forecasting period. Results are visualized using scatter plots. The Fbprophet is based on decomposable time-series. The equation is as follows:

$$y(t)= g(t) + s(t) + h(t) + \varepsilon t$$

The three main components are:

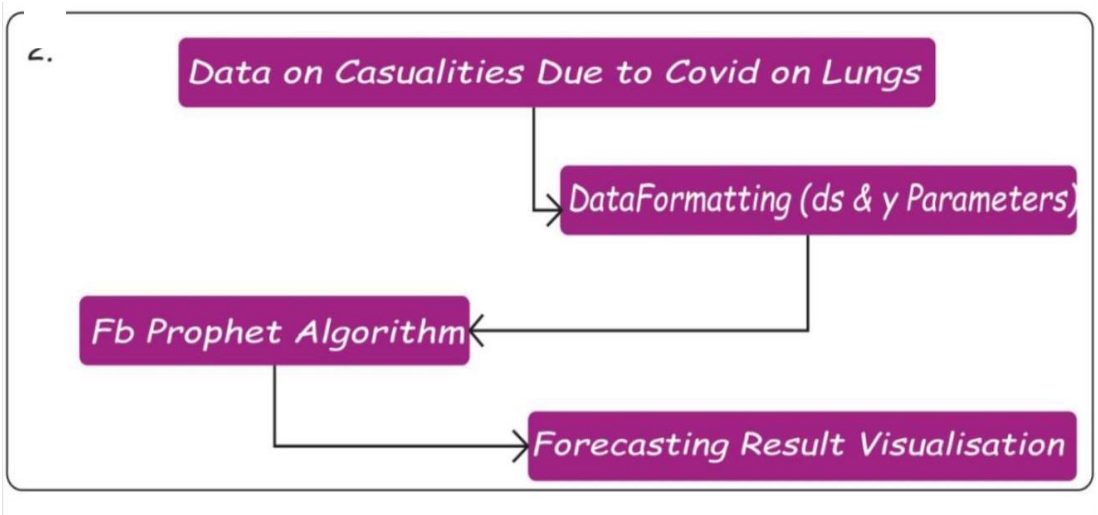
g(t): represents the trend function, which might be piecewise linear or logical to accommodate the changes in the time series' value which aren't periodic..

s(t): depicts the seasonality that occurs on a weekly and/or annual basis..

h(t): represents the holiday effects.

εt: represents any unusual change which is not accommodated by the model.

Figure3:Forecastingof Fatality



For the forecasting of fatality, we collect death data of each day in India since the beginning of pandemic. The dataset has "Date" in the format of "DD-MM-YYYY" and Death values as features. A simplistic predictive model called linear regression is used to predict the value of a variable based on the value of another variable. The dependent variable is the one you want to be able to predict. The independent variable is the variable you are using to predict the value of the other variable. We fit the data into linear regression model using the function Linear Regression(). After fitting, we plot using matplotlib for easier comprehension.

The figure below shows forecasting using Linear Regression:

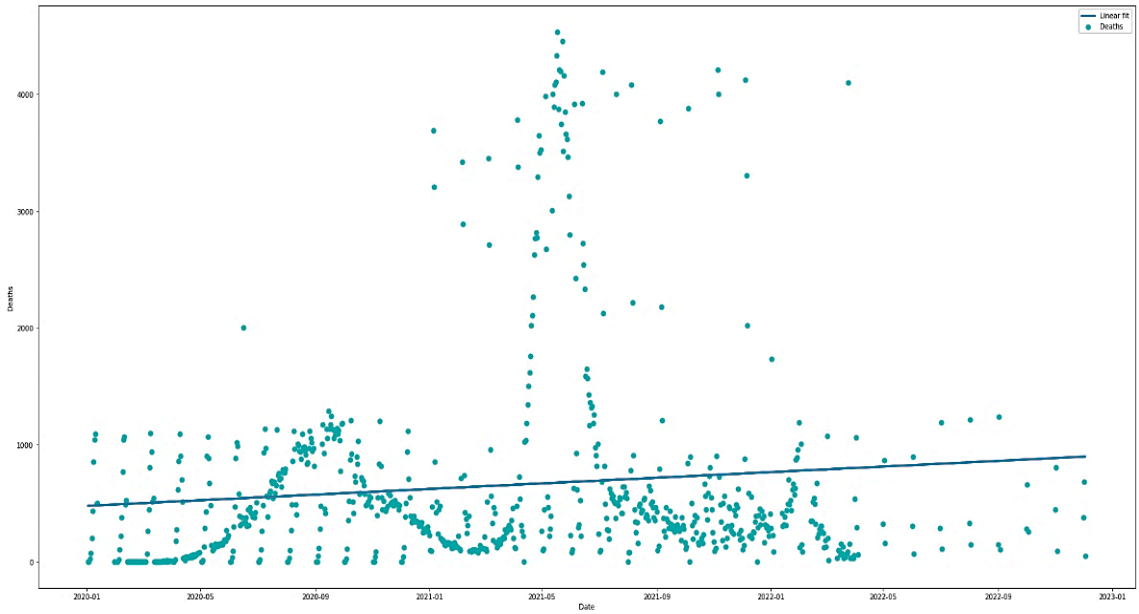


Figure 4Forecasting using Linear Regression

But since this algorithm is not much accurate and it assumes a linear relationship between the input variables and the single output variable, we make use of a better algorithm known as Fbprophet. Also, Fbprophet is more interesting, feature rich and visually pleasing than linear regression. The dataset is imported and converted into "YYYY-MM-DD" format as Fbprophet uses this as standard format. The algorithm requires feature names to be renamed as "ds" and "y". Here, "ds" represents Date and "y" represents Death values. Daily seasonality is true as we have daily death data. We set the period for forecasting as 200 days. The algorithm forecasts for the next 200 days. Upon using the predict(), the complete table of various useful parameters like yhat_upper, yhat_lower, trend are displayed. Also, the predict() fits and performs the forecasting for future dates. fbprophet.plot is used to plot the results obtained with the help of a graph. This graph is easy to understand as the values are displayed on top of each point in graph. The period values can be adjusted and zoomed in or out. This model has improved accuracy when compared to other predicting algorithms.

The figure below shows forecasting using Fbprophet:

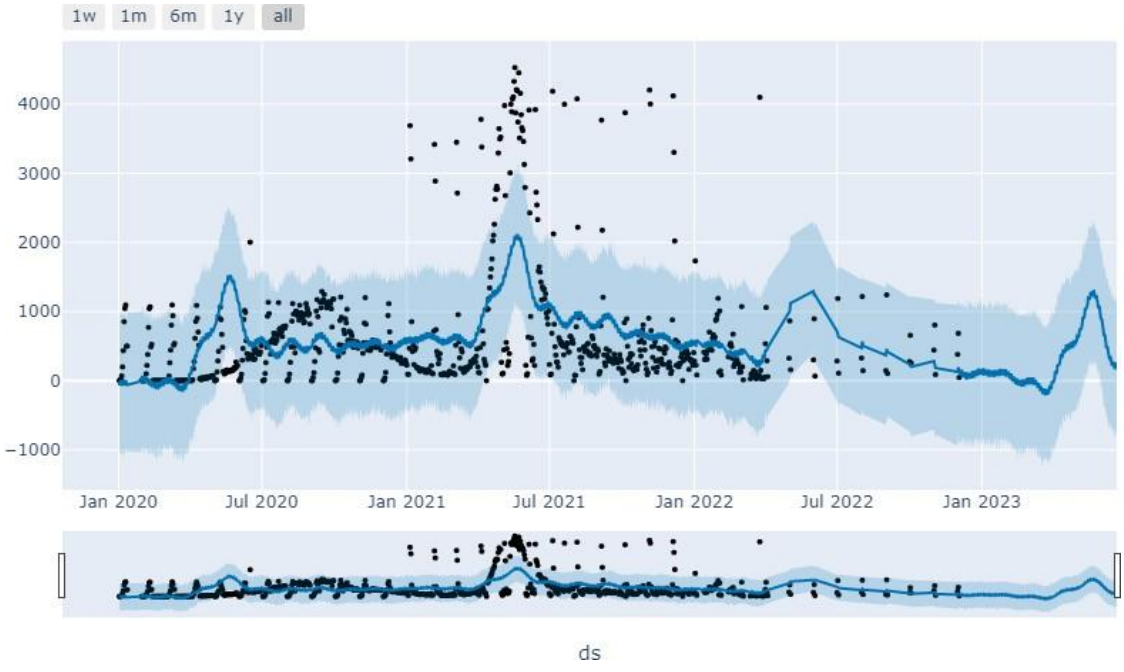


Figure 5 Visualization of Fbprophet

V. CONCLUSION

COVID-19 is still posing a threat to the world's healthcare system and has already killed millions of people. While diagnosing a patient, doctors have limited time and large number of patients can overwhelm them. So, computer-aided analysis could rescue lives through early screening and timely appropriate care can be provided. A deep learning based model is developed that uses CT scan images of lungs of patients and automatically classifies whether the patient has COVID or not. The classification model developed is used to detect COVID-19 using a VGG model. With this model, we believe it is quite useful as it saves time and resources thus paving way for healthcare workers to make a better and informed clinical decision. This in turn reduces burden and provides an accurate diagnosis. In this work, a forecasting model for COVID-19 has been proposed to predict the confirmed and death cases of COVID19. The obtained data is visualized as a graph to show the trend in fatality. The future scope would be to increase the accuracy of the model so that the predictions become much more accurate. As forecasting algorithms require large amount of data for generalization of model and to increase accuracy score of the model, data we acquired is quite less. But still the model has showed realistic predictions of the trajectory of pandemic. With time, more data is generated and this will help to increase the accuracy of model. The model can be used to forecast the peak values of the pandemic which can assist the healthcare sector to come up with a strategy to flatten the curve to controlling the spread of virus.