**Detection And Classification Of COVID-19 Using CNN**

Mr. Prashanth Ch1, Sai Dharani V2, Sai Harika T 3

*1,2,3Gayatri Vidya Parishad College of Engineering (Autonomous), Visakhapatnam*

*prashanth.deva.777@gmail.com1, saidharani.dharani@gmail.com2,saiharika.tankala@gmail.com3*

***ABSTRACT***

*The novel coronavirus(COVID-19) with a starting point at china, has spread over the other countries and is approximately 305,275 cases in world wide according to the statistics of European Center for Disease prevention control . The COVID-19 cases was increases day by day the hospitals have only limited number of COVID-19 test kits therefore it is necessary to implement an automatic detection system as quick diagnosis and given better prescription to patients This study will helps us to detecting and classification of COVID-19 patients to prevent the spread of the diseases to other people The image segmentation and neural network are used for automatic detection and classification of COVID-19 It takes the image as input and apply techniques to that image like preprocessing segmentation and then the input is given to the convolution neural network in order to classify the diseases*

***Keywords:*** Image processing , K-means clustering, HOG, CNN

**1. INTRODUCTION**

Today, all around the globe the severe public health problem arises is called COVID-19 that appeared in Wuhan, China in December 2019, it is also called as Severe Acute Respiratory Syndrome CoronaVirus (SARS-CoV-2). CoV is a combination of a number of diseases such as MERS-CoV & SARS-CoV. A new species that was discovered in 2019 and not identified in humans is called COVID-19. Corona viruses which are zoonotic in nature due to infection from animals to humans . COVID-19 virus is presumed to be contaminated from bats to humans . COVID-19 causes milder symptoms in about 82 percent of cases, the others are severe or critical but respiratory transmission of the disease from person to person caused rapid spread of the epidemic.Coronaviruses (CoV) are a large family of viruses that can cause respiratory infections in humans. These viruses can be found in mammals and aves. The symptoms include runny nose, sore throat, cough, fever and in more serious cases, pneumonia or breathing difficulties. Since December 8, 2019 several cases of pneumonia have been reported in Wuhan, Hubei Province, China and it is known to be caused by the novel Coronavirus (2019). Coronavirus can cause multiple system infections such as severe acute respiratory syndrome (SARS) and Middle-East Respiratory Syndrome (MERS). For healthy people’s respiratory rates of cases with COVID-19 pneumonia have been shown to be faster. Also in different distinct developed countries the health system, due to the increasing demand for intensive care units, is at the point of collapse. Severe care units are filled with patients who get worse with COVID-19 pneumonia. All over the world the confirmed cases are increasing these days.

By this research paper focus on covid-19 if person is infected or not by using deep learning and image processing to detection and classification occurred image processing is a method to perform some operation on an image in order to get an enhanced image or to extract some useful information from it .It is a type of signal processing in which input is an image and output may be image features associated with that image deep provides the healthcare industry with the ability to analyze data at exceptional speeds without compromising on accuracy .The benefits of deep learning in healthcare are fast,efficient,accurate and by using this techniques to detection the covid-19 to predicted the probability

of virus in image using CNN algorithm The method to denoising the desises first step to remove the noise in images and equalize the histogram then segmentation would done by using k-means clustering and feature extraction by image in hog technique and then trained our model.In this paper presenting the image collection, image preprocessing, image segmentation, feature extraction, and classification of covid-19 disease

**2. RELATED WORK**

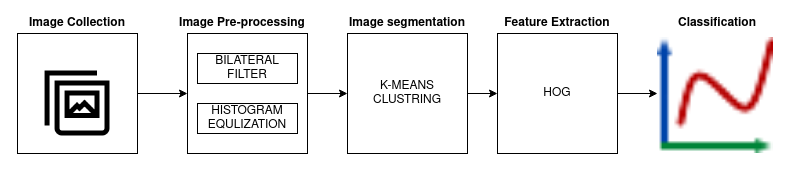
So far novel coronavirus has been detected in at least 188 countries and territories, internationally. The disease it causes has been named as Coronavirus disease (2019) or “COVID-19”. Early on, patients in Wuhan, China had some or the other link to a sea-food which suggests that animal to person spread could have happened. But, later, it was reported that the patients did not have exposure to animal markets, thus, indicating person to person spread. Similarly, person to person spread has been reported in the US and other locations outside China. The clinical picture concerning COVID-19 is not yet fully understood. However, a person to person spread will continue to occur. Widespread transmission of COVID-19 internationally would result in large numbers of people needing medical care at the same time. Hence, public healthcare systems may become overloaded, with increased prices of hospitalizations and death.

In this paper the author focuses and explains the scenario of first occurrence of Covid-19 in Wuhan with confirmed cases around above 4.7 million, for which radiography images used to speed up diagnostic process through decision support mechanism by built on Machine learning algorithms. This work explains a CNN architecture for a small number of parameters to distinguish COVID-19 from normal X-rays by performing a quantitative analysis to evaluate 12 off- the-shelf CNNs for the purpose of COVID-19 X-ray image analysis. Chest X- ray images used in this work are taken from publicly available sources in which COVID-19 X-ray image datasets and a large dataset of other COVID-19, PNEUMONIA bacterial infections,PNEUMONIA viral and then normal X-rays utilized. Final conclusion shows that CNN decisions not be taken into consideration, in the face of their high classification accuracy, until input image visually inspected by that in turn leads to its prediction . The Authors in this paper focus and give the study to evaluate the performance of CNN architectures proposed over the recent years for medical image classification. The author performed experimentation by considering datasets which collected from public medical repositories as X-ray images. The final result shows the probability of diseases by using CNN algorithm

**3. METHODOLOGY**

In this paper, the system presents basic steps. The first step starts with taking a gathering of CT images and X-rays from the available database. The second step applies bilateral filter for image pre-processing to get best level of quality and clearness.The third step is image enhancement which uses the histogram equalization The fourth step is image segmentation which uses k-means clustering method and the fifth step contains the calculation of feature extraction using hog .The final step is the classification of covid-19 disease Fig:1 illustrates the system architecture modules ,feature extraction and classification of this system.

**FIGURE1:SYSTEM ARCHITECTURE**

**A . IMAGE COLLECTION**

The foremost image processing in collection of images are lungCT images and X-ray images are collected from kaggle and internet resources.The medical data in usually in DICOM format which is the standard for storage and transfer of medical images into computed tomography(CT) images have better clarity ,low noise and distortion for lung diagnosis.in this paper we consider four diseases pneumonia viral ,pneumonia bacteria, covid-19 ,abnormal images, CT scan and X-ray images of lung images are given as input for this system.Dimensions of images are different but we are change the dimensions into 512x512 pixels in size Figures 2,3,4,5 shows the original CT image and X-ray images with nodule.The input of X-rays images and CT image contain noises such as white noise, salt and pepper noises etc.Therefore image preprocessing stage is needed to eliminated the noises.

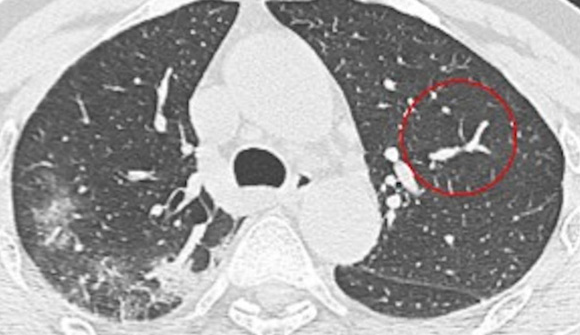
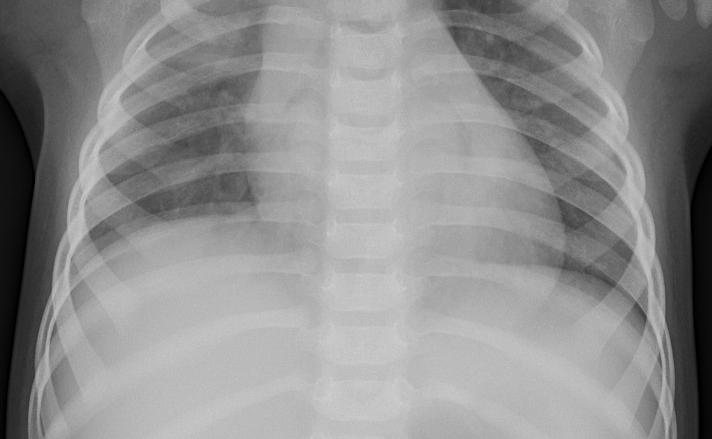
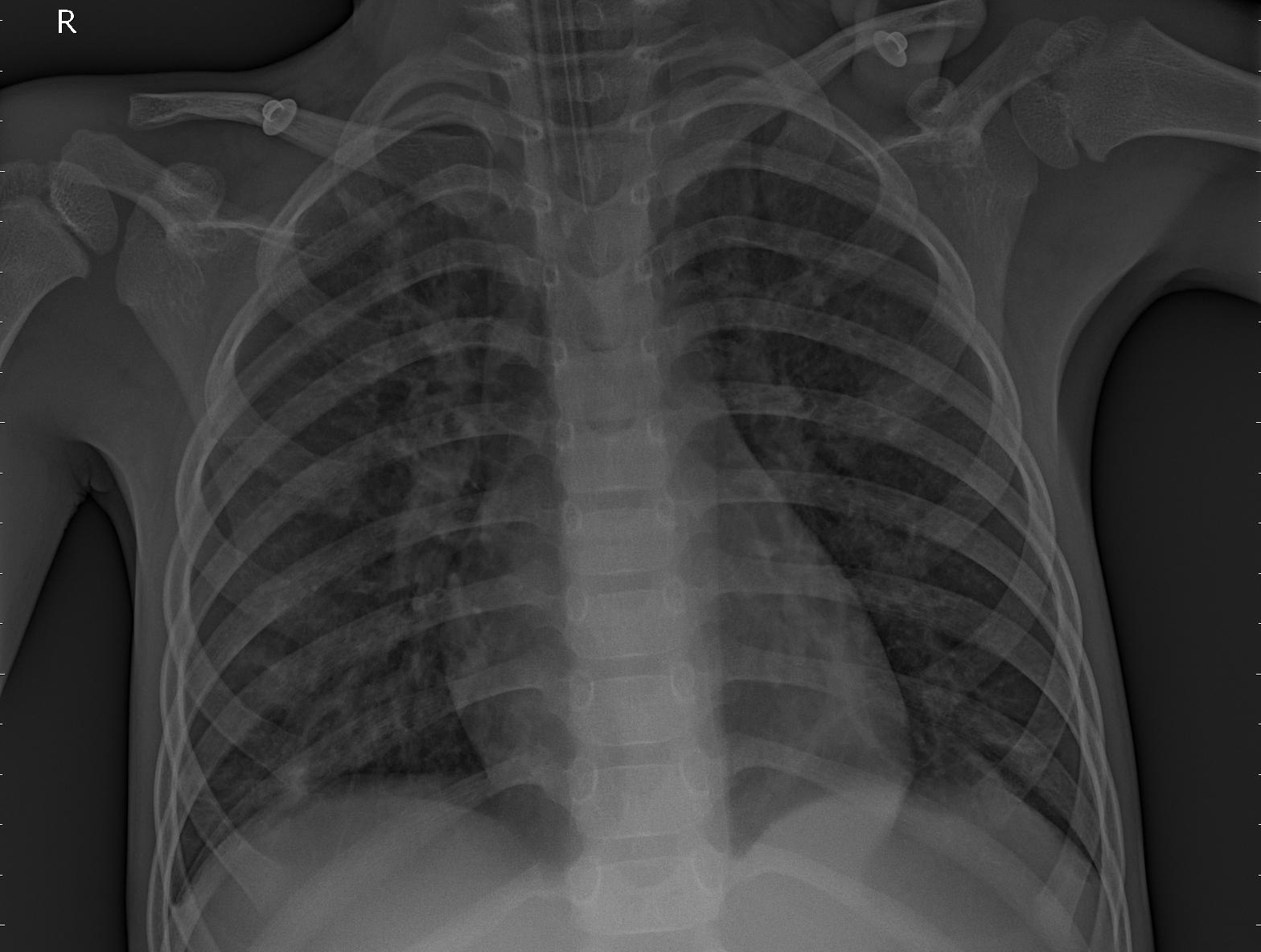
   

Fig2:Abnormal Fig3: covid Fig4: bacteria Fig5: virus

**B. BILATERAL FILTER**

Bilateral filter is one of the filter methods of image pre-processing. Image preprocessing is a way to improve the quality of image, so that the filtered image is better than the original one. The bilateral filter is a non-linear image technique. Bilateral filtering is a simple, intuitive and easy to implement smoothing image, i.e. reducing the amount of intensity variation between one pixel and the next, than other filters. The bilateral filter is normally used to reduce salt-and-pepper noise in an image. It often does a better job than the other filters of preserving useful detail in the image.The bilateral filter in Gaussian smoothing we take a weighted average of pixel values in the surrounding. The weights are inversely proportional to the distance from the center of the surrounding pixels. These spatial weights, the bilateral filter added the tonal weight(tonal weight are digital weights measuring the speed, acceleration ,force, power and range) such that pixel values that are close to the pixel value in the center are weighted more than pixel values that are more different. This tonal weighting makes that the bilateral filter is capable of preserving edges while smoothing in the more flat regions

g(x)=(f∗Gs)(x)=∫Rf(y)Gs(x−y)dy

The weight for f(y) equals Gs(x−y) and is only dependent on the spatial distance ∥x−y∥. The bilateral filter adds a weighting term that depends on the tonal distance f(y)−f(x) Observing that because the weights explicitly depend on the image values, we need an explicit normalization such that the ‘sum’ of all weights equals one. After filtering the images by using bilateral remove the salt and noise as shown images below Figure 6,7,8,9

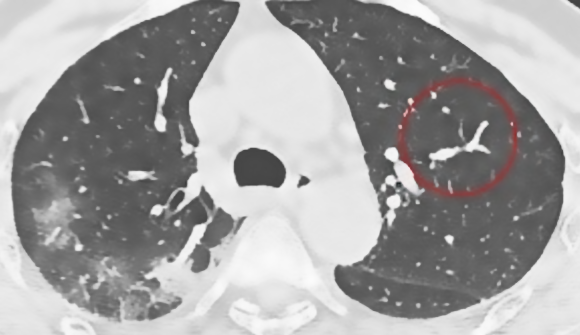
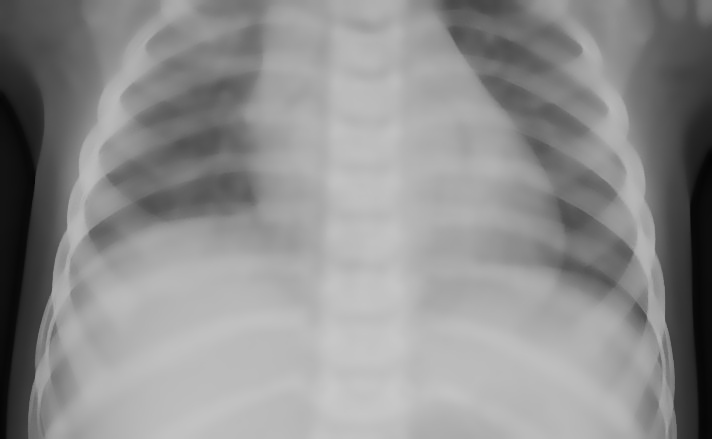
   

Fig6:Abnormal Fig7: covid Fig8: bacteria Fig9: virus

**HISTOGRAM EQUALIZATION**

Histogram Equalization is a picture handling method that changes the differentiation of a picture by utilizing its histogram. To improve the picture's differentiation, it spreads out the most continuous pixel force esteems or loosens up the power scope of the picture. By achieving this, histogram evening out permits the picture's regions with lower difference to pick up a higher differentiation we can observe in the Figure 10,11 below. This histogram normally increases the contrast of the image

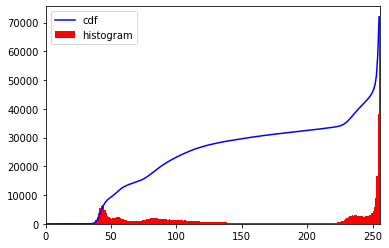
 

Fig10:Original image Fig11:Hist curve

We can see the histogram graph lies between the brighter regions so we require the full spectrum we need to transform the function which maps the input pixels in brighter regions to output pixels in the full region. Then we find the minimum values of histogram (excluding 0) and apply the histogram equalization in CT images and X-rays images Here, I has used concepts masked the images array and numpy we can see the Figure 12,13 below

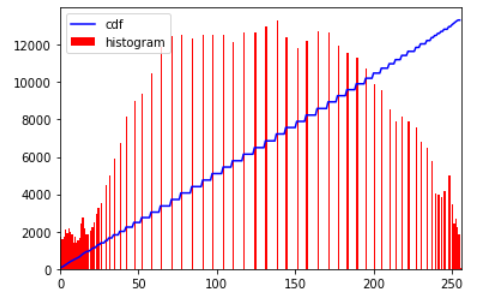
 

Fig12:Hist image Fig13:Hist curve image

Histogram Equalization can be utilized when you have pictures that are cleaned out on the grounds that they don't have adequate difference. In such photos, the light and dull territories mix together making a compliment picture that needs features and shadows show the Figure14,15,16,17

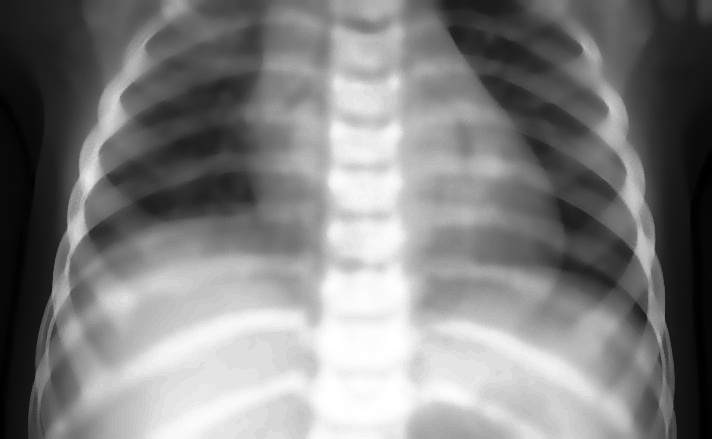
   

Fig14:Abnormal Fig15: covid Fig16: bacteria fig17: virus

**C. K-MEANS CLUSTERING**

Clustering is the task of partitioning the dataset into groups, called clusters. The goal is to split up the data in such a way that points within a single cluster are very similar and points in different clusters are different. Similarly to classification algorithms, clustering algorithms assign (or predict) a number to each data point, indicating which cluster a particular point belongs to k-means clustering is one of the simplest and most commonly used clustering algorithms. It tries to find cluster centers that are representative of certain regions of the data. The algorithm alternates between two steps: assigning each data point to the closest cluster center, and then setting each cluster center as the mean of the data points that are assigned to it. The algorithm is finished when the assignment of instances to clusters no longer changes the Euclidean distance used on all points for high dimensions to perform the clustering. In very high dimensional data, we could do a trick, which is an "engulfing sphere" i.e. starting at every centroid (k = 4 spheres in total), start growing a spherical space (circle in 2D, sphere in 3D, etc), outward radially. Until the spheres intersect, everything in the sphere belongs to the same cluster. All the other points can be assigned to clusters by calculating the perpendicular distance between the centroids. The Euclidean distance in n dimensions to compute the WCSS (p and q are two observation points, and their coordinates p 1 , ..., p n , q 1 , ..., q n are the features of these observation points): After compilation of k-mean clustering the images will show below Figures 18,19,20,21

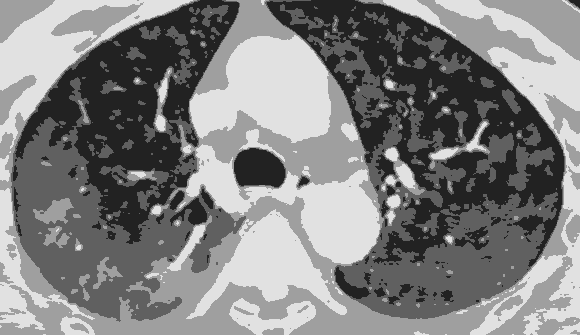
   

Fig18:Abnormal Fig19: covid Fig20: bacteria Fig21: virus

**D. FEATURE EXTRACTION**

Histogram of oriented gradients (or) HOG is a feature descriptor i.e., used to extract features from an image and mainly focuses on the structure or shape of an object .It is widely used in computer vision tasks for object detection . It decomposes an image into small squared cells and computes an histogram of oriented gradients in each cell. The normalizes using a block wise pattern and returns a descriptor for each cell. It is mainly able to provide the edge direction as well.HOG is done by extracting the gradient and orientation (magnitude and direction) of the edges. The “localized portions” is called orientation. It means the image is splitted into smaller regions and for each region gradients and orientation are calculated .Finally, histograms for each of these regions separately are created using the gradients and orientations of the pixel values. Hence the Histograms of oriented gradients.

Step 1: Preprocess the Data (64 x 128)

### Step 2: Calculating Gradients (direction x and y)

Step 3: Calculate the Magnitude and Orientation

Total Gradient Magnitude = √[(Gx)2+(Gy)2]

the value of the angle would be:

Φ = atan(Gy / Gx)

After complication of HOG feature extraction the images will show below Figures 21,22,23,24

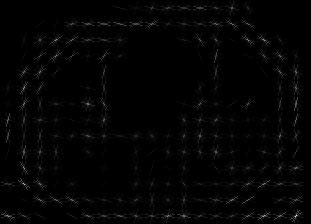
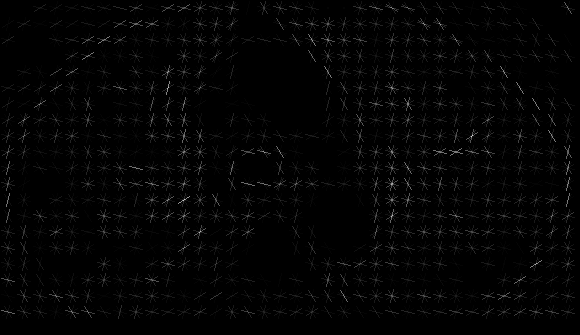
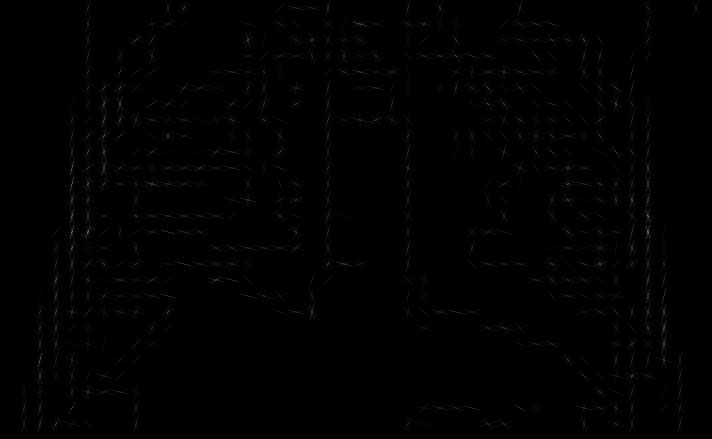
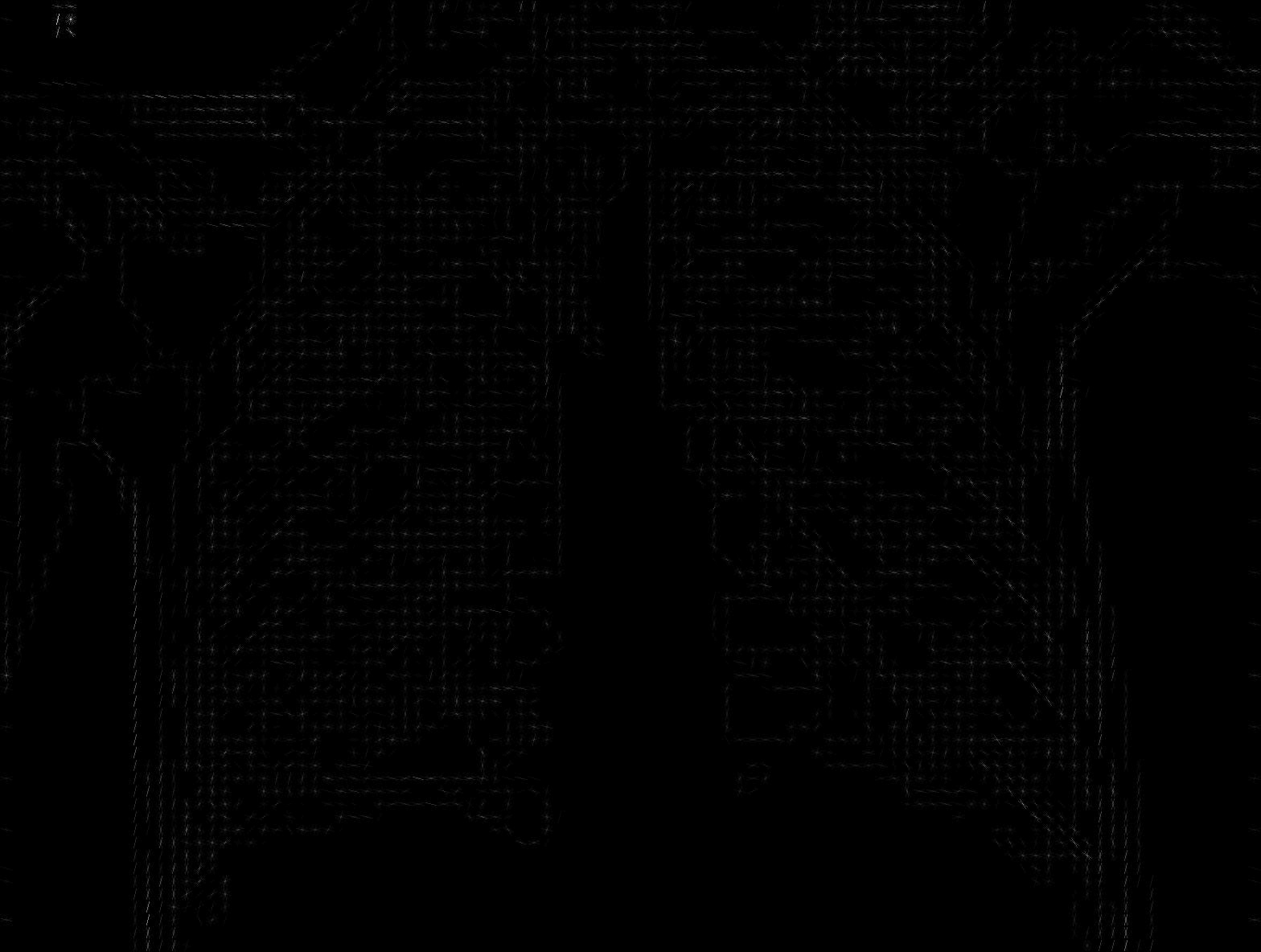
   

Fig21:Abnormal Fig22: covid Fig23: bacteria Fig24: virus

**E . CLASSIFICATION**

Depending on the present scenarios and review of various researchers, to implement an automatic detection system as a fast alternative diagnosis choice to prevent COVID-19 spreading among people which is performed through different classification techniques to detect disease using convolution neural network models. Convolutional neural network is a feed forward neural network i.e., generally used to analyze visual images by processing data.CNN is also known as ConvNet. A CNN uses a system much like a multilayer perceptron that has been designed for reduced processing requirements.In CNN every image is represented in the form of arrays in pixel values.The input image is scanned multiple times to generate the input feature map.Pooling layers connected one after another that carry out feature extraction.The rectified feature map now goes through a pooling layer.pooling is a down sampling operation that reduces the dimensions of the feature map pooling layer uses different filters to identify different parts of the image like edges,corners,body..etc., Pooling is a Subsampling pixels will not change the object ,We can subsample the pixels to make image smaller fewer parameters to characterize the image .

There are a number of deep learning models like Convolutional Neural Network (CNN), AlexNet, GoogleNet, VGG (Visual Geometry Group) generally applied for disease detection. CNN technique is supervised type, which mixes convolutional layers, ReLU Layer, fully connected layers, pooling layers and activations layers.

**CONVOLUTION LAYER**

Convolution is that the starting layer which is mostly accustomed extracts features from the input. It applies a convolution operation to input. During this layer applies different filter layers to make a feature map (kernel).

**RECTIFIED LINEAR UNIT(RELU)**

It generally wants to increase non linearity on the feature map in deep learning models as an activation function. It removes negative values setting back to returns 0. This can be represented as below where x is input to neuron, f(x) = x+ = max (0, x) (1)

**POOLING LAYER**

Pooling layer want to reduce the dimensionality, it reduces different parameters when image is large, and it control the overfitting problems

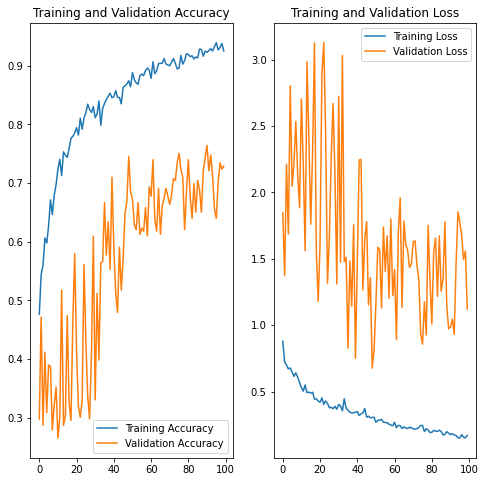
**FULLY CONNECTED LAYER**

This is the ultimate layer where actual classification happens, here each neuron within the input is connected to every neuron within the output, here add artificial neural network to combines features and attributes which predict classes

**TRAIN CNN WITH TENSORFLOW**

For numerical computation, Tensor Flow is a software library to code for creating convolutional neural networks. Mathematical operations are represented through nodes while the multidimensional data arrays are represented through graph edges (tensors) communicated between them.The convolution neural network using two library files tensorflow and keras we given the input to CNN algorithm input files belonging to 4 classes they splitting into training and testing dataset and labeling the batch size. The CNN used total parameters 3,989,156 and trained the 100 Epochs . we can see how much data fit in algorithm will plot a graph off training and validation(testing) accuracy and also plot graph of data loss plot graph of data loss in training and validation(testing) loss the graph shown below Figure :25 and finally model predict in softmax layers to given the confidence of image diseases

Figure25: Training and loss data graph

****

**RESULT ANALYSIS**

The CNN layer is processed through the sequence, the output layer is fed into a dense layer connected to a layer of softmax neurons for (soft) classification giving a probability for each class. The label for the images is again determined by the sum of the soft classification. Finally the model predicts the softmax layer to give the confidence of image diseases . we can see below Figure:26 how our image show in output

**This image most likely belongs to COVID with a 98..31 percent confidence**

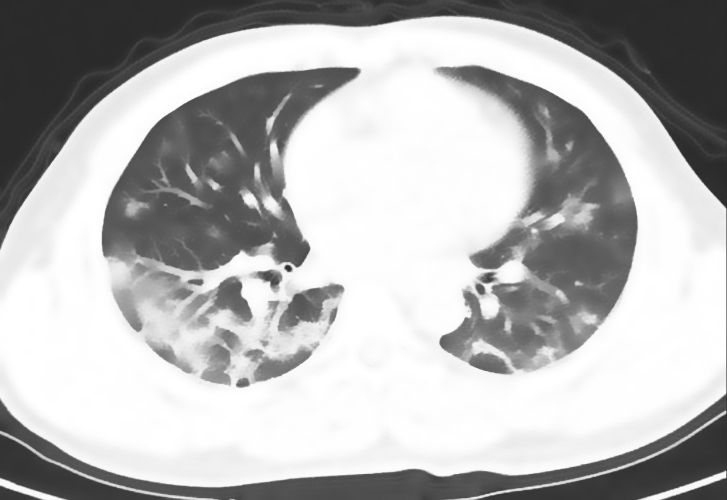


Fig :26

We are conducting an experiment randomly given the images to our model they predicting the images diseases percent based on that we compared each class in prediction power as shown below figure:27in bar graph

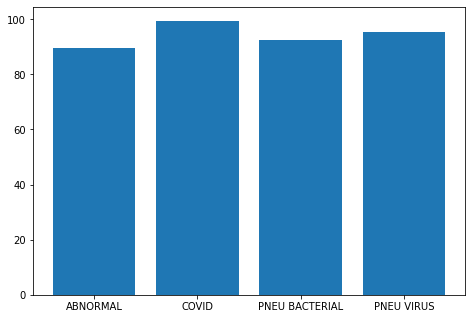


Fig:27

**CONCLUSION AND FUTURE**

The deep learning model proposed by the current situation gives importance to the necessity of early prediction of COVID-19 patients images because it is significant to forestall the spread of the diseases to people . During this study would be proposed a highest classification performance convolutional neural network model to detect COVID-19 from patients detected COVID-19 at an early stage using deep learning . Hence it concentrates on testing models which perform using large dataset and predicting the diseases of CT images and X-rays images . In future research do this experiment on the new strain virus CT images

**REFERENCES**

[1] https://www.cdc.gov/coronavirus/2019-cov/index.html

[2] https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30183-5/fulltext

[3] Anne Bonner, The Complete Beginner’s Guide to Deep Learning: Convolutional Neural Networks ImageClassification, Feb 2, 2019.

[4] Khari Johnson, How people are using AI to detect and fight the coronavirus, March 3, 2020.

[5]https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks

[6] Reza Ashen, Mansour Ebrahimi, The First Implication of Image Processing Techniques on Influenza A Virus Sub-Typing Based on HA/NA Protein Sequences, using Convolutional Deep Neural Network,October 19, 2018.Available: https://www.biorxiv.org/content/10.1101/448159v1.fullDOI: <https://doi.org/10.1101/448159>.

[7] Taban Majeed, Raspberry Rashid, Dashti Ali, Aras Asaad, “Covid-19 detection using CNN transfer learning from X-ray Images”,doi: https://doi.org/10.1101/2020.05.12.20098954, May 19, 2020.

[8] Ioannis D. Apostolopoulos, Tzani A. Mpesiana, “Covid 19: automatic detection from X ray images utilizing transfer learning with convolutional neural networks”, Physical and Engineering Sciences in Medicine (2020) 43:635–640 https://doi.org/10.1007/s13246-020- 00865-4.

[9] K. Roosa a, Y. Lee a, R. Luo a, A. Kirpich a, R. Rothenberg a, J.M. Hyman b, P. Yan c, G. Chowell, “Real-time forecasts of the COVID- 19 epidemic in China from February 5th to February 24th, 2020”, Infectious Disease Modelling 5 (2020) 256-263.

[10] Chaolin Huang et. al., “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China”, [www.thelancet.com](http://www.thelancet.com) Vol 395 February 15, 2020.

[11] Yan, L., Zhang, H. T., Xiao, Y., Wang, M., Guo, Y., Sun, C., Tang, X., Jing, L., Li, S., Zhang, M., Xiao, Y., Cao, H., Chen, Y., Ren, T., Jin, J., Wang, F., Xiao, Y., Huang, S., Tan, X., Huang, N., Jiao, B., Zhang, Y., Luo, A., Cao, Z., Xu, H., and Yuan, Y. Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan. medRxiv 2020.02.27.20028027, 2020.

[12] Coronavirus. World Health Organization: https://www.who.int/health- topics/coronavirus, 2020.