Pneumonia Detection and Classification using CNN and VGG16

Hrushikesh S. Rajankar Dr. S. L. Bangare Pavan S. Patil Karan V. Nakum Gopal S. Paraskar *Sinhgad Academy of Sinhgad Academy of Sinhgad Academy of Sinhgad Academy Sinhgad Academy*

*Engineering, Kondhwa Engineering, Kondhwa Engineering, Kondhwa of Engg, Kondhwa of Engg*

*bk. Pune. bk. Pune. bk. Pune. bk. Pune bk. pune*

hrushiraj08 sunil.bangare pavanadhao685 k.v.1111nakum gpara35

@gmail.com @gmail.com @gmail.com @gmail.com @gmail.com

***Abstract*—Pneumonia, an infectious disease caused by a bacterium in the lungs' alveoli, is frequently the result of pollution. A lung infection causes pus to build up in the affected tissue. Professionals conduct bodily examinations and diagnose their patients using a chest X-ray, ultrasound, or lung biopsy to determine if they have certain conditions. Misdiagnosis, incorrect treatment, and failure to recognize the disease will result in a patient's inability to lead a normal life. Deep learning's advancement helps specialists make better decisions when diagnosing patients with certain diseases. The research provides a flexible and efficient deep learning technique that uses the CNN model to predict and detect a patient who is unaffected.Using a chest X-ray photograph, the study applies a flexible and effective deep learning technique of using the CNN model in predicting and detecting a patient unaffected and affected by the illness. To demonstrate the overall performance of the CNN model being trained, the researchers used an amassed dataset of 20,000 photographs and a 224x224 photograph decision with 32 batch lengths. At some point throughout the total performance training, the trained version produced a 95 percent accuracy charge. The research study may detect and predict COVID-19, bacterial, and viral pneumonia illnesses based solely on chest X-ray photographs, according to the results of the testing.**

***Keywords—Pneumonia Detection, Adaptive Deep Learning, Deep Convolutional Neural Network Architecture***

1. Introduction

Pathogenic bacteria, physical and chemical causes, immunologic damage, and numerous drugs all cause inflammation of the lung parenchyma.

There are numerous approaches for classifying pneumonia: (1) Pneumonia is classified as infectious or non-infectious based primarily on pathogeneses, with infectious pneumonia being the more prevalent. Infectious pneumonia includes, among other things, immune-associated pneumonia, aspiration pneumonia caused by physical and chemical causes, and radiation pneumonia. Non-infectious pneumonia includes immune-associated pneumonia, aspiration pneumonia caused by physical and chemical causes, and radiation pneumonia. (2) There are three forms of pneumonia: community-acquired pneumonia (CAP), hospital-acquired pneumonia (HAP), and ventilator-related pneumonia (VAP), with CAP accounting for the majority of cases. Because of the great variety of illnesses,

HAP makes it simpler for bacteria to develop antibiotic resistance, making treatment more difficult. Pneumonia kills around 800,000 children under the age of five every year. More than 2200 individuals are killed. Pneumonia affects more than 1400 children out of 100,000. According to the Global Burden of Disease Study, pulmonary diseases such as pneumonia were the second leading cause of mortality in 2013. Pneumococcal illness has inflamed 35 percent of patients in European hospitals and 27.3 percent of patients globally. According to a recent data from John Hopkins Bloomberg College of Public Health, India had the highest rate of pneumonia mortality among children under the age of five in 2015, with almost 2.97 lac pneumonia combined diarrhoea deaths among children under the age of five. Furthermore, pneumonia's fatality rate is inversely linked to its age, and pneumonia's superiority increases significantly with age, particularly in people over 65. The high number of infant mortality due to pneumonia has prompted scientists around the world to suggest more effective and immediate ways to combat the disease. Greater and further measures are advanced as the era progresses, with radiology-based treatments being the most common and effective. Diagnostic radiological techniques for pulmonary disease include chest X-ray imaging, computed tomography (CT), and magnetic resonance imaging (MRI), with chest X-ray scanning being the most efficient and cost-effective because it is far more accessible and portable in hospitals, and it exposes sick people to lower doses of radioactivity. However, even for multiple skilled and experienced medical physicians, analysing pneumonia using X-ray snapshots remains a difficult task because X-ray images contain comparable area statistics for unique ailments, such as lung cancer. As a result, diagnosing pneumonia with conventional procedures is time-consuming and energy-intensive, and it is impossible to diagnose whether or not a patient has pneumonia in a uniform manner. As a consequence, in this study, we propose using a Convolutional Neural Network to autonomously diagnosis pneumonia using X-ray pictures, with an accuracy of 96.07 percent and an AUC of 0.9911. The remaining sections of this work are organised as follows. The literature's perspectives on medical picture processing processes are discussed in the second section. In recent times, section three has described a rapid way of Convolutional Neural Networks (CNN) architecture. The final segment included an outline of Machine Learning and Deep Learning's history. The material employed in this investigation, our proposed techniques, and the training method are all depicted in Section 4. The trials and their outcomes are presented in section 5. This observer's belief is described in section 6.

2. Related Works

**[1]**There are various challenges in developing a long-term version of COVID-19 detection from X-ray images. To begin with, there are just a limited number of photos available. Because the epidemic has only been spreading for a few months, not enough datasets have been collected and released for scientists to use. Second, there may be a compelling need for scientists to create a smart device that can quickly identify and analyse viral infections utilising breast X-ray images, which can convey the seriousness of the problem to the human body. Following an assessment of existing technology and the difficult conditions we face, it was found that method transfer is viable and reasonable for this research issue. This method involves using images of pneumonia to train deep neural networks on illnesses in one or both lungs, which can be caused by bacteria, infections, or fungus. The condition causes irritation in the lungs' alveoli, which are tiny air sacs. Breathing becomes difficult as the respiratory system fills with fluids or pus. To identify the type of virus, the process uses transfer learning on well-studied deep learning methods, validates the designs on a wide range of data sets, and then transfers the models and insights learned during training and validation to a new data set in a similar region, in this case, a new infectious fast-growing illnessThe coronavirus, commonly known as COVID-19, is a virus of the coronavirus family. The proposed method tackles a difficult problem in deep learning: how to develop a credible model from a small data set that hasn't been extensively analysed and has undiscovered properties.

[2] Deep CNNs do actually perform better when a large dataset is used rather than a smaller one. While there is a large number of infected COVID-19 individuals worldwide, the number of publicly available chest X-ray photographs online is small and distributed. As a result, the authors of this study defined a rather large dataset of COVID-19 infected chest X-ray photos, despite the fact that normal and pneumonia photographs are readily available and used in this study.

**[3]**The observations, which use a chest X-ray image, leverage flexible and sophisticated deep learning approaches, including the use of six CNN models to forecast and identify whether the patient is unaffected or affected by the ailment. GoogLeNet, LeNet, VGG-16, AlexNet, StridedNet, and ResNet-50 models with a dataset of 28,000 images and a 224x224 decision with 32 and sixty-four batch sizes are used to monitor the accuracy of each version being learnt. The study additionally employs Adam as an optimizer, giving all of the models a 500-epoch learning rate and an adjusted 1e-four learning rate. Both GoogLeNet and LeNet models received a 98 percent accuracy rate during development, VGGNet-16 received a 97 percent accuracy rate, AlexNet and StridedNet models received a 96 percent average accuracy rate and the ResNet-50 model received an 80% accuracy rate. For total performance training, GoogleNet and LeNet fashions have the highest average accuracy. The six models identified were possible to perceive and predict a pneumonia illness, which included a normal chest X-ray.

**[4]**The authors of this paper classify the three types of X-rays. Image writers employed the ensemble technique at some point in the prediction process, and each image is sent through the type layer, where they verify whether the image is COVID-19, pneumonia, or ordinary.

**[5]**In the first phase of statistical preprocessing, the X-rays dataset is divided into training and test subsets. At this level, character X-ray images are also normalised. Data augmentation of training photos was finished for a strong version development. The age of a version that divides x-rays into two groups is the second degree (consolidation and non-consolidation). The accuracy of this model was demonstrated by the use of an okay-fold cross-validation approach (where okay is the number of folds). The final stage is the implementation of the explainable AI technique that we decided to improve our Machine's comprehensibility. The heatmap has arrived.To measure the quality of our Machine, we utilise two unique methods: 1) creating the heatmap using the simplest version, and 2) generating the heatmap using an ensemble of styles with the same structure but educated with unique record folds. The second method allows us to calculate an uncertainty level for each pixel (given by the same old deviation), allowing us to analyse the heatmap's robustness.

**[6]**In this study, the author suggests a contemporary technique based on an ensemble of RetinaNet and masks R-CNN. The size and ratio of pneumonia regions (Figs. 3 and 4) are known, with a mean top of 304 pixels (29.6% of the photograph peak) and a median width of 219 pixels (21.three percent of picture width). Pneumonia shows on a chest x-ray in a very small area, making it difficult to identify with contemporary item detectors.We used FPN as the backbone of both techniques to solve this problem since it delivers multi-scale feature maps with higher first-rate statistics than the standard feature pyramid. The FPN structure, as shown in Fig. 1, combines low-resolution, i.e. linguistic in-depth, with improved capacities. Through a top-down channel and lateral linkages, semantically weak functions. In comparison to the primary stacked convolutional layers, we employed residual networks as a base spine version because they reduced the effect of the degrading hassle and allowed us to develop deeper designs.

3. Background

Over the last few years, researchers have become increasingly interested in machine learning (ML) techniques. This method can take full advantage of the enormous potential to develop computer calculators in image processing with pre-determined algorithm phases. On the other hand, traditional machine learning methods for partitioning projects need the use of manual design algorithms or the manual setup of output layers to separate images. In response to the aforementioned situation, LeCun et al. **[7]**offered a CNN approach, which can automatically extract features with the use of constantly stacking features and exit that the included photos may not be in any class. The shallow networks are quite deep and focus on the low-level aspects of the image. As the number of network layers grows, the CNN model reveals more complex capabilities. CNN uses a back-propagation technique to update and record learned parameters after integrating and evaluating these priority characteristics to learn the differences between different images. The premise behind CNN is to employ a certain convolution kernel to filter a previous picture or map component before building the next layer map element, as well as merge functions such as merging functions to reduce feature map scale and mitigation to count. The non-line activation function is then applied to the newly generated component.

a mapping to improve the model's simulation capabilities The most common integration tasks include mid and high integration. The plural of integration denotes that the element sent to the integration layer is split into many regions, with each sub-region having a different size in terms of horizontal and vertical steps. The sole distinction between high and medium integration is a lower region where the aggregation rate yields the average of each sub-region. ReLU (Rectified Linear Units) and Sigmoid are two common activation events. Image elements are automatically extracted using segmentation and a continual accumulation of convolutional processes, integration functions, indirect opening functions, and other completely integrated layers. Then, by evaluating these derived characteristics, it is possible to extract pneumonia from the photos processed by the model. The model's general capacity is increased by fully utilising pixel-level image information. The most prominent neural framework has been proposed in past few decades for in-depth learning development., such as AlexNet **[8]** and VGGNet **[9]**. However, when the number of layers in the network increase, Instead of learning the numerous productive features, the neural network will be modified to particular parts of the training image, which makes the model similar to the capacity declines and creates congestion. The remaining communication framework was proposed to overcome the problem of network depth. Since then, neural networks have advanced, garnered a lot of attention and research, and have formed the foundation for a lot of occupations. We also looked at the efficiency of residual connections in our reduced CNN architecture with only a few layers in this study.

4. Materials and Methods

4.1 Data

The proposed database, which will be used to assess the model's performance, contains 5863 X-ray images obtained through Kaggle.

In 2017, Dr. Paul Mooney held a Kaggle competition to classify viral and bacterial pneumonia. It stands out from the other datasets since it includes 5,863 paediatrics photographs. We're discussing the most recent version of this dataset**[6].**

The database is further divided into three folders (train, test, and val) with subfolders for each image category (Pneumonia / General). Figure 1 shows a few instances of common and pneumonia photos that have been scaled to a static size. Due to the low amount of exposure in patients, chest X-ray images always show symptoms of limited brightness, and chest X-ray images always have black, white, and grey pants. The lungs are on both sides of the thoracic cavity, and the lung area is plainly visible on an X-ray since it is virtually black. The heart, which is situated between the lungs, appears practically as white as X-rays can go through it entirely. Because bones are comprised of protein and are exceedingly dense, X-rays cannot pass through them, leaving the bones virtually white. Furthermore, the bones have distinct edges.



(a)



(b)

Fig. 1. Examples from the dataset. (a) normal cases, (b) pneumonia cases

4.2 Data Preprocessing

Table 1 lists the tactics employed throughout this article. Rescale is a value that we will multiply the data by before any other processing in our investigation. Our original photos had RGB coefficients ranging from 0 to 255, but values like this would be too high for our models to handle (given a typical learning rate), so we scale them down by a factor of 1./255. shear range is used to apply shearing transformations at random. When there are no assumptions of horizontal asymmetry, zoom range is used to randomly zoom inside photographs, and horizontal flip is used to randomly flip half of the images horizontally (e.g. real-world pictures)

Data pre-processing techniques used in this study

|  |  |
| --- | --- |
| Rescale | 1./255 |
| Zoom Range | 0.2 |
| Shear Range | 0.2 |
| Horizontal\_Flip | True |

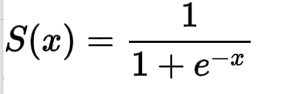
Table 1

4.3 Proposed Network

In this study, we designed a VGG-based CNN model to extract the features of chest X-ray images and use those features to detect if a patient suffers from pneumonia. In Our CNN Architecture, We

started with a lower filter setting of 32 and worked our way up layer by layer. A layer of Conv2D was used to build the model, followed by a layer of MaxPooling. An odd number, such as 3x3, is desirable for kernel size.

The activation functions Tanh, ReLU, and others can be employed, but ReLU is the most popular. input shape accepts the width and height of an image, with the last dimension serving as a colour channel. After that, we flattened the input and added ANN layers.



f(x)=max(0,x)

*S(x) = Sigmoid*

*f(x) = ReLU*

For the last layers (ANN Layers), I used a softmax activation function and defined units as the total number of classes. For binary classification, I used a sigmoid and set the unit to 1.

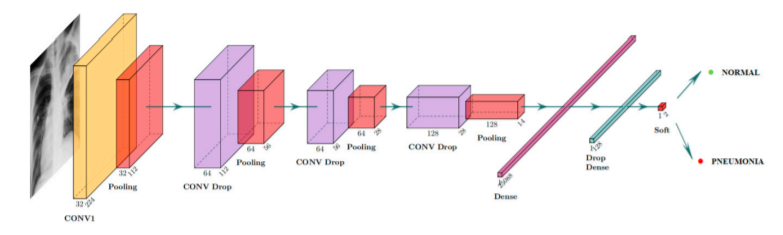


Fig 2. Details of proposed DL model.

5. Implementation

Convolutional neural networks are a type of neural network that shares all of the characteristics of other neural networks. CNN, on the other hand, was created specifically to process input images. The architecture of their organisation is therefore more specific: it is made up of two primary parts.

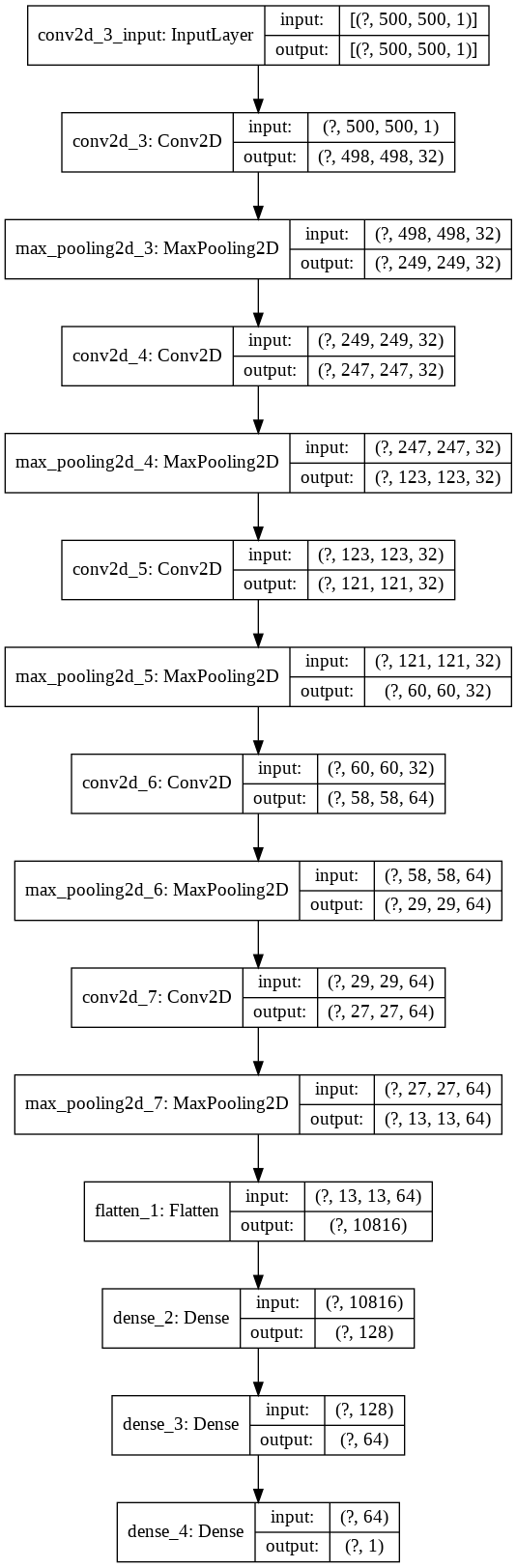
***Conv Layers***

Because it works as a feature extractor, the first block establishes the uniqueness of this sort of neural network. It accomplishes this by adding convolution filtering processes to template matching. The first layer uses many convolution kernels to filter the image and produce "feature maps," which are then normalised (using an activation function) and/or shrunk.

***Pool Layers***

The second block is not unique to CNNs; in fact, it appears at the end of all categorization neural networks. To return a new vector to the output, the input vector values are transformed (using various linear combinations and activation functions). The chance that the image corresponds to class I is represented by element I of this last vector, which has as many elements as there are classes. As a result, each element has a value between 0 and 1, and the total value is 1. The last layer of this block (and hence of the network) calculates these probabilities using a Sigmoid function (binary classification) or a RELU function (multi-class classification) as an activation function.

**Visualization of CNN model Sequence**



**Fitting Model**

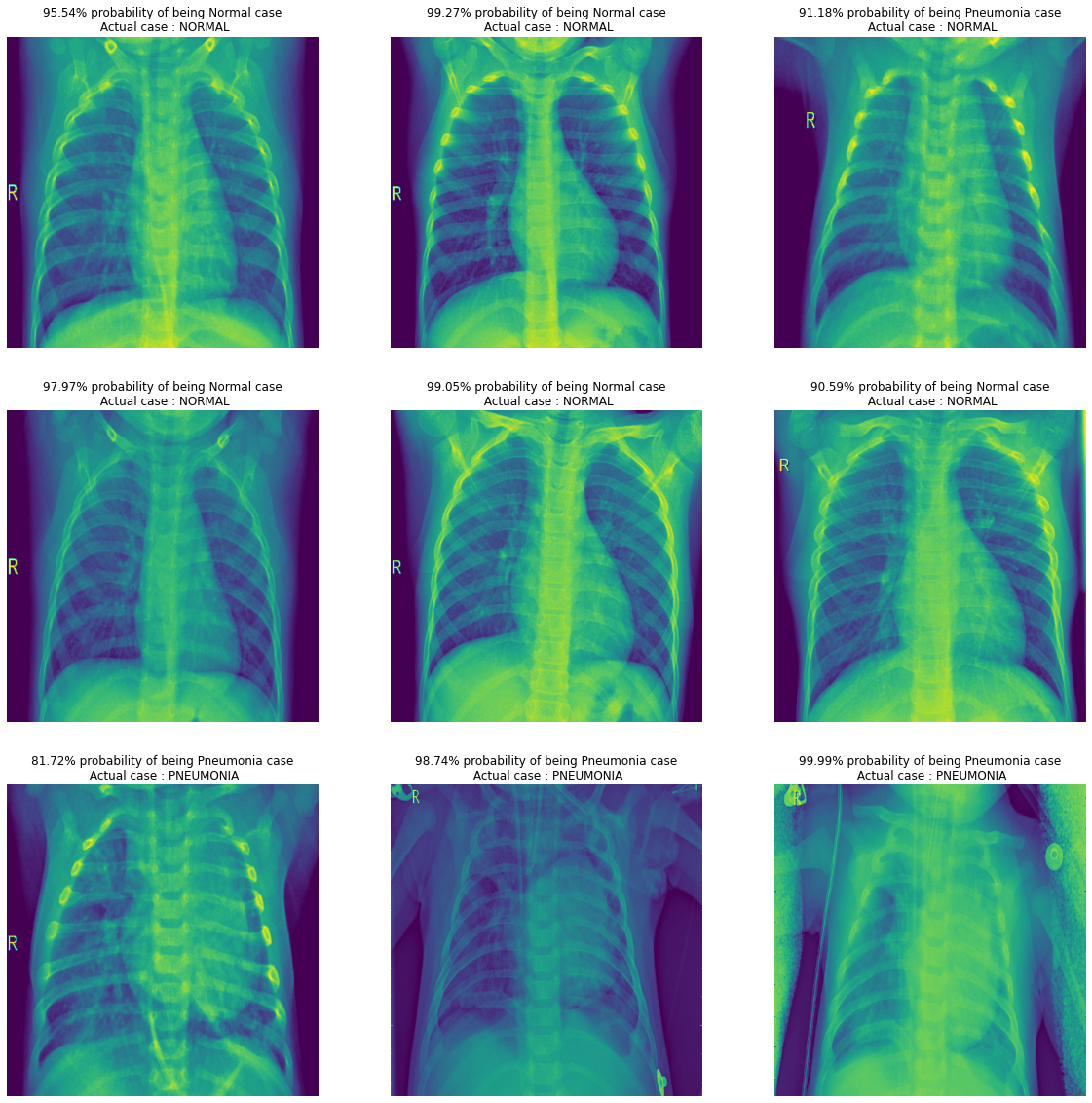
EarlyStopping is a function that allows you to halt the epochs early based on a measure. It aids in the avoidance of overfitting the model. We're urging you to stop based on the val loss statistic since it needs to be as low as possible.

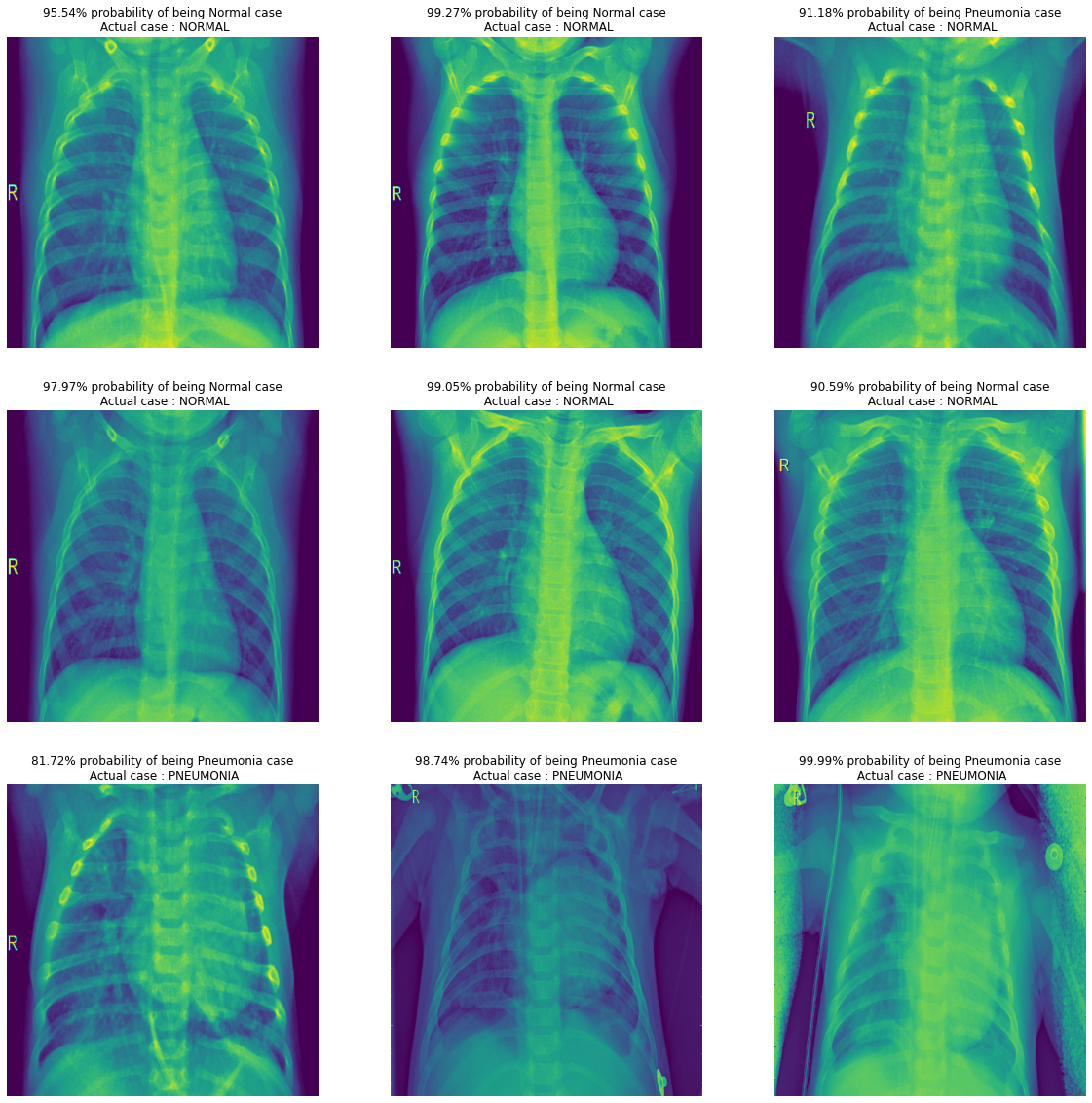
Patience states that once a minimum val loss is attained, if the val loss grows in any of the next three iterations, the training will end at that epoch.

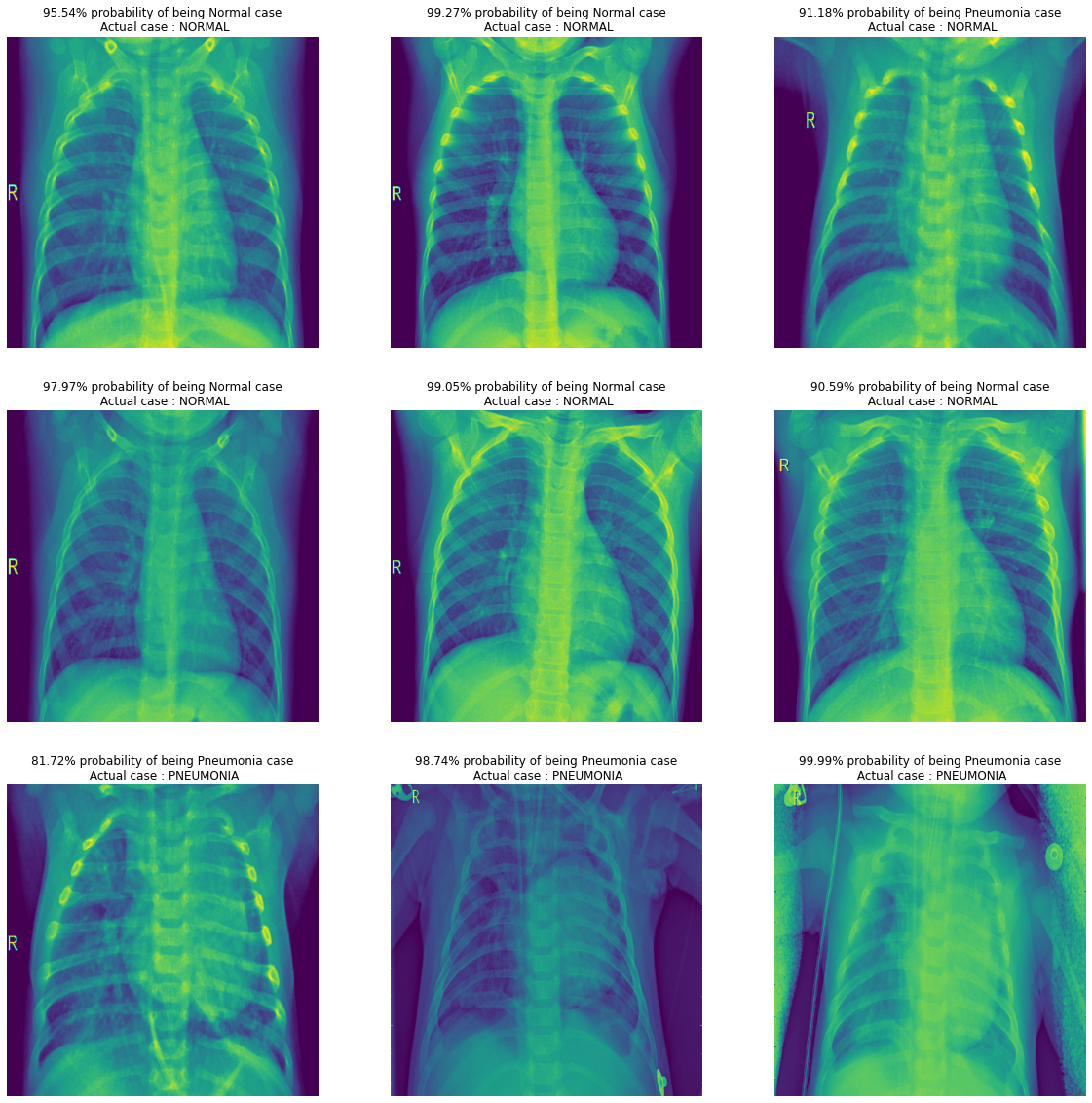
EarlyStopping in our training session came to a halt in the 9th epoch, with val loss = 39.8% and val accuracy = 68.7%. when the patience level is 3

Actual test results were 91.98 percent accurate on the model.

**Visualizing some predicted images with percentage %**







* This provides you a percentage estimate of the individual image, which you may load straight from your hard drive by specifying its path.
* After importing the image as we did previously, we must recreate all of the data pretreatment procedures in order to input the test set into the model and obtain a forecast. Importing the tensorflow.keras.preprocessing.image class is required for pre-processing.
* Import an image with dimensions of (500,500) and a grayscale colour channel.
* To predict the case, convert the image to an array, rescale it by dividing it by 255, and extend dimension by axis = 0 as shown earlier.

6. Conclusion and Future Work

This paper presents an automated diagnosis of pneumonia in X-ray scans using deep CNN. The research was conducted utilising the X-Rayscan dataset, which contains 5863 scans. Experiments yielded a variety of scores, including accuracy, recall, precision, and AUCranking, proving the efficiency of our network model. The proposed framework was successful in reaching a 91 percent categorization accuracy. To improve the model's efficiency, hyper-parameter optimizations were examined, and multiple optimization techniques, such as stochastic gradient descent, Adagrad, and Adamoptimizer, were used. The customizedVGG16 model's effectiveness in detecting pneumonia reveals that the model surpasses other optimizers when compared to Adam.In the future, this research will be broadened to include the detection and differentiation of multi-class X-ray images. The efficiency could also be improved by applying more complex feature extraction approaches based on many recently established deep learning models for biomedical picture segmentation.

7. Reference

[1] Vandecia Fernandes et al., “Bayesian convolutional neural network estimation for pediatric pneumonia detection and diagnosis”, Computer Methods and Programs in Biomedicine, Elsevier, 2021

[2] Hongen Lu et al., “Transfer Learning from Pneumonia to COVID-19”, Asia-Pacific on Computer Science and Data Engineering (CSDE), 2020 IEEE

[3] Sammy V. Militante et al., “Pneumonia and COVID-19 Detection using Convolutional Neural Networks”, 2020 the third International on Vocational Education and Electrical Engineering (ICVEE), IEEE, 2021

[4] Nanette V. Dionisio et al., “Pneumonia Detection through Adaptive Deep Learning Models of Convolutional Neural Networks”, 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC 2020), 8 August 2020

[5] Md. Jahid Hasan et al., “Deep Learning-based Detection and Segmentation of COVID-19 & Pneumonia on Chest X-ray Image”, 2021 International Information and Communication Technology for Sustainable Development (ICICT4SD), 27-28 February 2021

[6]<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

[7] LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1989, 1, 541–551.

[8] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 2012, 25, 1097–1105.

[9] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556

[10] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 618-626.

[11] L. Wang and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images," arXiv:2003.09871, 2020.

[12] https://towardsdatascience.com/medical-x-ray-%EF%B8%8F-image-classification-using-convolutional-neural-network-9a6d33b1c2a