IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Detecting Covid-19 using Chest X-Rays: A Neural Network Based Approach

T.Prathima, Satya Shodhaka R Prabhanjan, Syed Danish Hussaini
Dept. of Information Technology, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad, Telangana, India

Abstract
The novel corona virus popularly known as COVID-19 is a pandemic which shook the entire world. The virus can be detected using X-Rays, Chest CT Scans, RTPCR, etc. In this work we used the X-Ray images of chest taken from healthy people, pneumonia effected patients and chest X-Rays of Corona effected patients. We trained the top CNN architectures and observed the results. We have also used Data Augmentation to increase the dataset. AlexNet gave us better accuracy in reasonable time.

Keywords: COVID-19, X-Rays, Data Augmentation, AlexNet, VGG, DenseNet, Squeezenet, ResNet.

1. Introduction

The novel coronavirus 2019 (COVID-2019), which first appeared in Wuhan city of China in December 2019, spread rapidly around the world and became a pandemic. It has caused a devastating effect on both daily lives, public health, and the global economy. It is critical to detect the positive cases as early as possible so as to prevent the further spread of this epidemic and to quickly treat affected patients. Coronavirus disease (COVID-19) is an extremely contagious disease and it has been declared as a pandemic by the World Health Organization (WHO) on 11th March 2020 considering the extent of its spread throughout the world [1]. The pandemic declaration also stressed the deep concerns of the alarming rate of spread and severity of COVID-19. It is the first recorded pandemic caused by any coronavirus. It is defined as a global health crisis of its time, which has spread all over the world. Governments of different countries have imposed border restrictions, flight restrictions, social distancing, and increasing awareness of hygiene. While most of the people infected with the COVID-19 experienced mild to moderate respiratory illness, some developed a deadly pneumonia [2]. There are assumptions that elderly people with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease, renal or hepatic diseases and cancer are more likely to develop serious illness. Until now, no specific vaccine or treatment for COVID-19 has been invented. However, there are many ongoing clinical trials evaluating potential treatments. More than 7.5 million infected cases were found in more than 200 countries until 11th June 2020, among which around 421 thousand deaths, 3.8 million recovery, 3.2 million mild cases and 54 thousand critical cases were reported [3,4].

The need for auxiliary diagnostic tools has increased as there are no accurate automated toolkits available. Recent findings obtained using radiology imaging techniques suggest that such images contain salient information about the COVID-19 virus. Application of advanced artificial intelligence (AI) techniques coupled with radiological imaging can be helpful for the accurate detection of this disease and can also be assistive to overcome the problem of a lack of specialized physicians in remote villages. In this project a new model for automatic COVID-19 detection using raw chest X-ray images is presented. The proposed model is developed to provide accurate diagnostics for multi-class classification (COVID vs. Normal vs. Pneumonia. RT-PCR is used as a reference method for the detection of COVID-19 patients, however, the technique is manual, complicated, laborious and time-consuming with a positivity rate of only 63% [5]. Moreover, there is a significant shortage of its supply, which leads to delay in the disease prevention efforts [6]. Many countries are facing difficulties with incorrect number of COVID-19 positive cases because of not only due to the lack of test kits but also due to the delay in the test results [7]. These delays can lead to infected patients interacting with the healthy patients and infecting them in the process, in order to combat this a cost effective and reliable alternative to RT-PCR is needed. With help of Deep learning and AI we can develop a screening tool which can automate the process of detecting Covid-19 effectively.

The purpose of this work is to evaluate the effectiveness of state-of-the-art pre-trained convolutional neural networks proposed by the scientific community, regarding their expertise in the automatic diagnosis of Covid-19 from thoracic X-rays. This is due to the fact that the state-of-the-art CNNs are sophisticated model requiring large-scale datasets to perform accurate feature extraction and classification. With transfer learning, the retention of the knowledge extracted from one task is the key to perform an alternative task.

1.1 Motivation

Accurate and rapid diagnosis of COVID-19 suspected cases plays a crucial role in timely quarantine and medical treatment. Developing a deep learning-based model for automatic COVID-19 diagnosis on chest CT is helpful to counter the outbreak of SARS-CoV-2. COVID- 19 is an acute resolved disease, but it can also be deadly, with a 2% case fatality rate. Severe disease onset might result in death due to massive alveolar damage and progressive respiratory failure. The early and automatic diagnosis of Covid-19 may be beneficial for countries for timely referral of the patient to quarantine, rapid intubation of serious cases in specialized hospitals, and monitoring of the spread of the disease. Although the diagnosis has become a relatively fast process, the financial issues arising from the cost of diagnostic tests concern both states and patients, especially in countries with private health systems, or restricted access health systems due to prohibitive prices.
In March 2020, there has been an increase in publicly available X-rays from healthy cases, but also from patients suffering from Covid-19 [9]. This enables us to study the medical images and identify possible patterns that may lead to the automatic diagnosis of the disease. The development of deep learning applications over the last 5 years seems to have come at the right time. Deep Learning is a combination of machine learning methods mainly focused on the automatic feature extraction and classification from images, while its applications are broadly met is object detection tasks, or in medical image classification tasks. Machine learning and deep learning have become established disciplines in applying artificial intelligence to mine, analyze, and recognize patterns from data. Reclaiming the advances of those fields to the benefit of clinical decision making and computer-aided systems is increasingly becoming nontrivial, as new data emerge.

1.2 Objective

The main objective of the work is to develop a Convolutional Neural Network (CNN) model which can distinguish between Normal, Viral Pneumonia and Covid-19 from Chest X-Ray with high accuracy and is cost effective. Section 2 presents the existing work, section 3 its implementation is detailed, Section 4 presents testing and analysis of the results, Section 6 conclusion and future scope is discussed.

2. Literature Survey

In [10] authors presented a new model for automatic COVID-19 detection using raw chest X-ray images is presented. The proposed model is developed to provide accurate diagnostics for binary classification (COVID vs. No-Findings. The proposed model has an end-to-end architecture without using any feature extraction methods, and it requires raw chest X-ray images to return the diagnosis. This model is trained with 125 chest X-ray images, which are not in a regular form and were obtained hastily. Diagnostic tests performed after 5–13 days are found to be positive in recovered patients. This crucial finding shows them that recovered patients may continue to spread the virus. Therefore, more accurate methods for the diagnosis are needed. One of the most important disadvantages of chest radiography analyses is an inability to detect the early stages of COVID-19, as they do not have sufficient sensitivity in GGO detection. However, well-trained deep learning models can focus on points that are not noticeable to the human eye, and may serve to reverse this perception. Instead of initiating a deep model development from scratch, a more rational approach is to construct a model using already proven models. Therefore, while designing the deep model used in this study, the Darknet-19 model is chosen as the starting point. Darknet-19 is the classifier model that forms the basis of a real-time object detection system named YOLO (You only look once). This system has the state-of-the-art architecture designed for object detection. The DarkNet classifier is used on the basis of this successful architecture. The DarkCovidNet architecture inspired by the DarkNet architecture that has proven itself in deep learning, instead of building a model from scratch. They have used fewer layers and filters as compared to the original DarkNet architectures. They gradually increased the number of filters such as to 8, 16, 32. To better understand this new model, it is helpful to understand the basics of the Darknet-19, which consists of 19 convolutional layers and five pooling layers, using Maxpool. These layers are typical CNN layers with different filter numbers, sizes, and stride values. In this study, X-ray images obtained from two different sources were used for the diagnosis of COVID-19. A COVID-19 X-ray image database was developed by Cohen JP using images from various open access sources. This database is constantly updated with images shared by researchers from different regions. At present, there are 127 X-ray images diagnosed with COVID-19 in the database. There are 43 female and 82 male cases in the database that were found to be positive. In this dataset, a complete metadata is not provided for all patients. The age information of 26 COVID-19 positive subjects is given, and the average age of these subjects is approximately 55 years. Also, the ChestX-ray8 database provided by Wang et al. was used for normal and pneumonia images. In order to avoid the unbalanced data problem, they used 500 no-findings and 500 pneumonia class frontal chest X-ray images randomly from this database. In this study, they have proposed a deep learning-based model to detect and classify COVID- 19 cases from X-ray images. They reported an accuracy of 98.08%. the limitations are model takes too much training time to reach optimal accuracy and it is not efficient with noisy database.

In [11] the authors worked on feature extraction and fine-tuning are used to train on multiple variants of convolutional neural networks, namely InceptionResNetV2, Xception, DenseNet201, and VGG19. With a small amount of data, a higher detection accuracy is attained on the chest X-rays pneumonia detection task. Convolutional neural network simulates feature differentiation through convolution, then reduces the order of magnitude of network parameters through convolutional weight sharing and pooling, and finally completes classification and other tasks through traditional neural networks. Multiple variants of convolutional neural networks are used here, namely InceptionResNetV2, Xception, DenseNet201, and VGG19. *InceptionResNetV2:* An important feature of Inception V3 is the decomposition of convolution. It solves the convolution integral of n × n into the convolution of 1 × n and n × 1, which reduces the amount of parameters and increases the depth and calculation speed of the network. The important feature of ResNet is to allow shortcuts in the model and use the residual connection to make the output of a previous layer as the input of a later layer. In this way, the output of the previous layer is added to the activation of the latter layer, which largely solves the problem of network degradation and vanishing gradient. *VGG19:* It contains 16 convolutional layers, 3 full connected layers, and 5 pooling layers. The size of the convolutional kernel used in the convolutional layer is 3x3. By setting stride to 1 and padding to the same, each convolutional layer can keep the same width and height as the previous layer. The max-pooling layer with a stride and kernel size of 2 is used in the pooling layer. The whole network is divided into five blocks, each block is composed of several convolutional layers and a pooling layer, and in the same block, each convolutional layer has the same number of channels. *Xception:* is another improvement of Inception V3 proposed by Google. Generally, convolution on a set of feature maps requires a three-dimensional convolution kernel, that is, the convolution kernel needs to learn the spatial correlation and the correlation between channels meanwhile. It uses separable convolutions (also called extreme Inception modules) to replace the convolution operations so that the correlation between the channels and the spatial correlation is separated. *DenseNet:* The basic idea of DenseNet is the same as ResNet, but it proposed a more radical dense connection mechanism than ResNet. It establishes a dense connection between all the front layers and the back layer. Each layer will accept all the layers in front of it as its additional input. Another major feature of DenseNet is feature reuse through the connection of features on channels. These features allow DenseNet to achieve better performance than ResNet in the case of fewer parameters and computational costs.

First, the convolution is instantiated. Then feature extraction is performed without data augmentation. The extracted feature shape is flattened and input to a densely connected classifier for training. After that, data augmentation is used for feature extraction. A densely connected classification is added to the convolutional base. The frozen convolutional base is used to train the model end-to-end. Finally, the last three layers of the convolutional base are fine- tuned and the model is evaluated on the test set. The input tensor of InceptionResNetV2, Xception and VGG19 is set to (150,150,3), and the input tensor of DenseNet201 is set to (221,221,3). All experiments were carried out according to the above steps. During the experiment, the RMSprop optimizer is used, the loss function selected is binary cross-entropy, the learning rate is set to 2e-5, the number of steps per epoch is set to 100, and the number of verification steps is set to 50. In the convolutional base, the bottom layer encodes more general reusable features, while the top layer encodes more specialized features. Fine-tuning these more specialized features is more useful.

Besides, the more training parameters, the greater the risk of over fitting. For these two reasons, the last three convolutional layers are selected for fine-tuning in the experiment. In this paper, a variety of variants of convolutional neural networks are used to deal with chest X-ray pneumonia detection tasks, including InceptionResNetV2, Xception, DenseNet201, and VGG19. InceptionResNetV2 achieved the highest accuracy of 94.20%. The limitations are accuracy of feature extraction without data augmentation is slightly lower, Fine-tuning achieves higher accuracy but with high training time.

In [12] authors presented a model for COVID prediction from chest X-rays using CheXNet is presented in this paper. CheXNet is a deep Convolutional Neural Network consisting of 121 layers which is shown in Fig. 2. This network produces a heatmap that localizes the areas in which disease symptoms are highly indicative in the image along with the prediction probability. This was developed to predict the pneumonia from chest X-rays. This model used chestX- ray14 dataset containing X-rays of 14 different pathologies. Four practicing radiologists classified the images in the test set, on which the performance of the model is compared to that of radiologists. CheXNet was primarily developed for pneumonia prediction. They build a CheXNet model by using pre-trained model of DenseNet121. It has five convolutional layers and average pooling is used. The weights of the pre- trained model taking the weights file and loaded those weights in our model. They achieved more accuracy compared to the models which used transfer learning cause this CheXNet model was trained on chest x-ray images itself unlike other models. The model is built with five conv layers and is trained with pre-trained weights. They also visualized feature maps in CheXNet model to observe how the given input chest x-ray is considered by each layer of the proposed CheXNet model. This gives insights about the internal representation of how the model considers input and its evaluation. This feature maps or activations maps provide the preferential regions of the image that are used by convolutional neural networks to classify covid chest x-ray with that of normal one. The differences in activation maps of both depict that covid x-ray contains major portion of area affected when compared to that of normal one. Finally, by these the model can clearly classify the chest x-rays into respective Covid and Non-Covid Classes. They performed their experiment to predict COVID19 from chest x-ray images. By training our developed DenseNet121 model on 1824 images and achieved results showing 99.9% accuracy. But the Limitations are Small Dataset and cannot classify between multiple classes.

In [13] the authors proposed multimodal deep learning using two modalities Chest CT- Scan and X-Ray Images based on the concatenation of extracted features from two different transfer learning models with each network using the same parameters. The original size and number of the channels of the data images from both datasets are varied. The input size images are resized to 150 x 150 pixels and the number of channels is set to 3 (RGB) for the feed of the input layer. The main idea of transfer learning is to use previously gained knowledge and applying it to a target task that is still related. Some of the Transfer learning models that are used in this work are DenseNet, Mobile Net, Xception, Inception, ResNet and VGG . Transfer learning models that are trained using ImageNet dataset can be used to make quicker and more accurate models for classifying images. A concatenated neural network is constructed by concatenating the extracted feature of the neural network from two different transfer learning models and then connecting the concatenated layer to output layers. The activation function on the fully connected layers is ReLu and the output layers are softmax with two classes. They have used ResNet50, DenseNet121, and Xception model for CT-Scan dataset (left input layer) and VGG16, MobileNet, and InceptionV3 model for the X-Ray dataset (right input layer). Our model architecture which is the concatenation of two different transfer learning models . It is shown that the concatenated network performs better for classifying COVID-19 Pneumonia but the individual network of VGG16 gives a slightly better result compared to the concatenate network of Xception and InceptionV3. It is reported that the concatenation of ResNet50- VGG16 and Densenet121-MobileNet gives the same confusion matrices. Sensitivity and Specificity gives a balanced score between each other and that is because we use balanced data. In this research, the combined computational time on two individual networks is nearly identical to its concatenated network. The computer specifications for running the algorithm are as follows: AMD Ryzen 5 3600X 6-Core Processor, 16384 MB of RAM, and NVIDIA GeForce GTX 1660 SUPER. The concatenation of DenseNet121- MobileNet gives an Accuracy 99.87%, Sensitivity 99.74%, and Specificity 100%. Then, the computational time for this network is quicker than the concatenation of ResNet50-VGG16 which had the same result. The limitations are higher number of parameters in transfer learning models does not guarantee higher accuracy, Cannot classify between multiple classes.

In [14] authors proposed FT-CNN for chest x-ray image classification. Fine-Tune is a process based on the concept of learning transfer. Here our proposed model is beginning to train CNN to learn the characteristics of a broader domain with a segmentation function aimed at reducing error on that domain. After that, it will replace the partition function and use the network and reduce the error in another domain. Under this setting, transfer network features and parameters from a broad domain to a specific one. Classification function uses SoftMax classification in this network learning. All preparation pictures are made to a size of 256X256 pixels. Also, to maintain a strategic distance from exaggeration in the model information expansions are utilized which incorporate arbitrary editing and irregular flat filtering. The precision of the confirmation and the comparing seasons of every one of the three sorts are set down. Fundamental testing and convincing proof for COVID-19 have assumed a significant part in forestalling the spread of this new infection around the world. Time, cost, and precision with a couple of components considered inside any piece of overabundance COVID-19 assurance. To address these issues, FT-CNN-based model is proposed in this paper to recognize COVID-19 cases in X-bar chest patients. There are an aggregate of 1560 X-bar chest pictures that are partitioned into two classes: 'Coronavirus Case' and 'Typical Case', which are utilized to set the model. To be sure, a 20% chest X-beam picture is utilized to address the model. This proposed works with a precision and exactness of 70.32% and 90.70% each. Limitations of this model are it requires extra occasion adjustment and clinical testing to work, model can be improved by proceeding with the accessibility of enormous informational collections.

3. Proposed System

In the proposed system, a deep learning model is proposed for the automatic diagnosis of COVID-19. The proposed model has an end-to-end architecture without using any feature extraction methods, and it requires raw chest X-ray images to return the diagnosis.

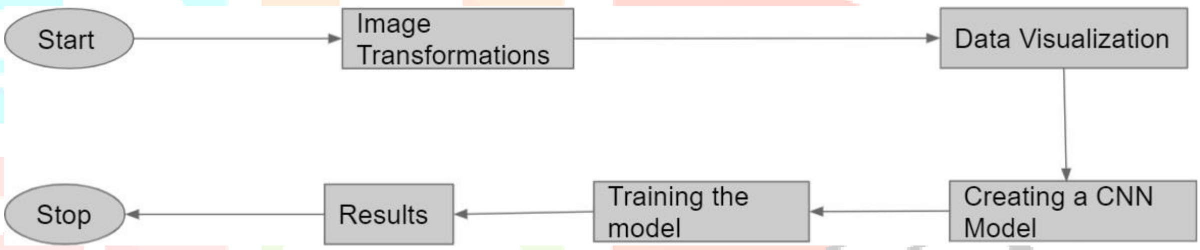

Fig 1: System flow diagram

The above image Fig.1 describes the flow of the system starting from giving chest x ray image as an input to predicting the class which that image belongs to.

- First, we load the dataset that contains 5724 images with 3 classes (COVID 19 and normal X-rays) for multiclass classification.
- As the image models takes input images of size 224 x 224, resize the images in our dataset to 224 x 224.
- Take the image models network with pre-trained weights of ImageNet dataset without including the fully connected (FC)layer as head.
- Then freeze the CONV weights of image classification model such that only head network will be trained.
- Now, pass a new X-ray image to detect whether the patient is having COVID-19 or Viral Pneumonia or not.

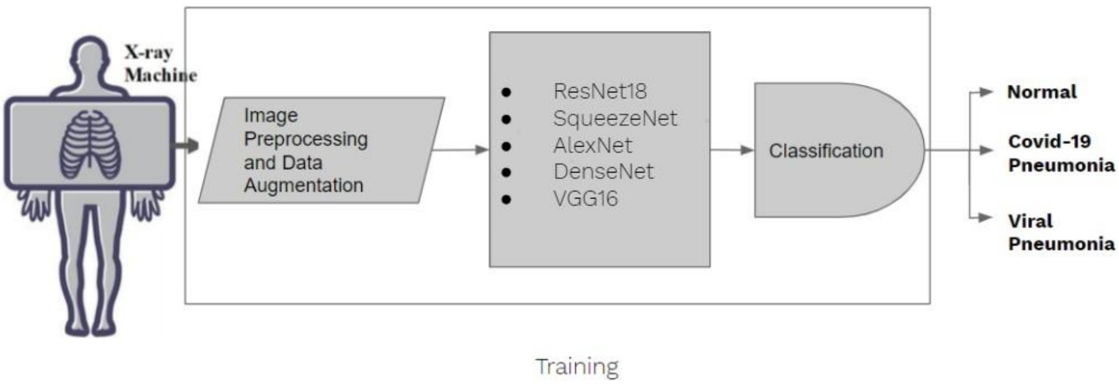

Fig. 2: System Architecture

Transfer learning is a strategy wherein the knowledge mined by a CNN from given data is transferred to solve a different but related task, involving new data, which usually are of a smaller population to train a CNN from scratch. In deep learning, this process involves the initial training of a CNN for a specific task (e.g., classification), utilizing large-scale datasets. The availability of data for the initial training is the most vital factor for successful training Physical and Engineering Sciences in Medicine (2020) since CNN can learn to extract significant characteristics (features) of the image. Depending on the capability of the CNN to identify and extract the most outstanding image features, it is judged whether this model is suitable for transfer learning. The batch normalization operation is used to standardize the inputs, and this operation has other benefits, such as reducing training time and increasing stability of the model. The Maxpool method is used in all the pooling operations. Maxpool downsizes an input by taking the maximum of a region determined by its filter. If three different classes of images are used in the input, the same model performs the classification task to determine the labels of the input chest X-ray images as COVID-19, Pneumonia, or Normal.

3.1 Image Transformations

Chest X-ray is a fairly structured examination, taken the same way every time, and the body is outwardly symmetric across the midline (bilateral symmetry). The lungs are reasonably symmetric, and the outer soft tissues and bones of the chest very symmetric.

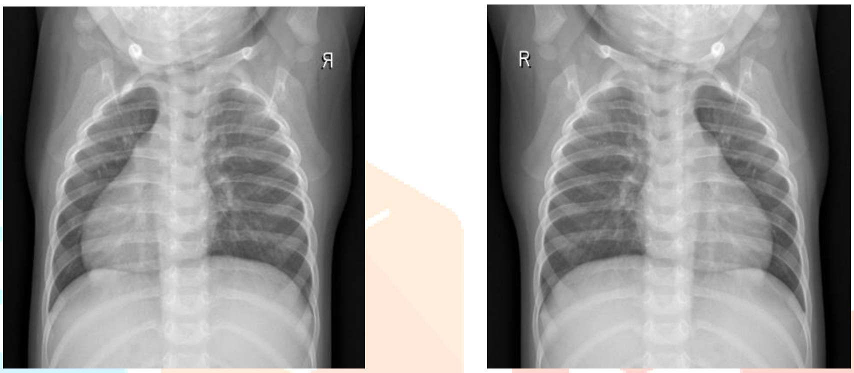

Fig. 3 Horizontal Flip on a Chest X-Ray Image

From the above figure (Figure 3), we can induce that applying horizontal image transformation technique which helps the network to "forget" some of the overtrained detail it has learned making it more robust. The only area of relative asymmetry is the heart/mediastinum, which is not the direct focus of this exam.

System Implementation

A typical CNN structure has a convolution layer that extracts features from the input with the filters it applies, a pooling layer to reduce the size for computational performance, and a fully connected layer, which is a neural network. By combining one or more such layers, a CNN model is created, and its internal parameters are adjusted to accomplish a particular task, such as classification or object recognition. We have researched and used various models for image classification which are mentioned below: -

- Resnet-18
- DenseNet
- SqueezeNet
- AlexNet
- VGG16

ResNet-18 is a convolutional neural network that is 18 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by- 224. Mostly in order to solve a complex problem, we stack some additional layers in the Deep Neural Networks which results in improved accuracy and performance. The intuition behind adding more layers is that these layers progressively learn more complex features. For example, in case of recognising images, the first layer may learn to detect edges, the second layer may learn to identify textures and similarly the third layer can learn to detect objects and so on. But it has been found that there is a maximum threshold for depth with the traditional Convolutional neural network model. Here is a plot that describes error% on training and testing data for a 20 layer Network and 56 layers Network.

3.2 Residual Block:

This problem of training very deep networks has been alleviated with the introduction of ResNet or residual networks and these Resnets are made up from Residual Blocks.

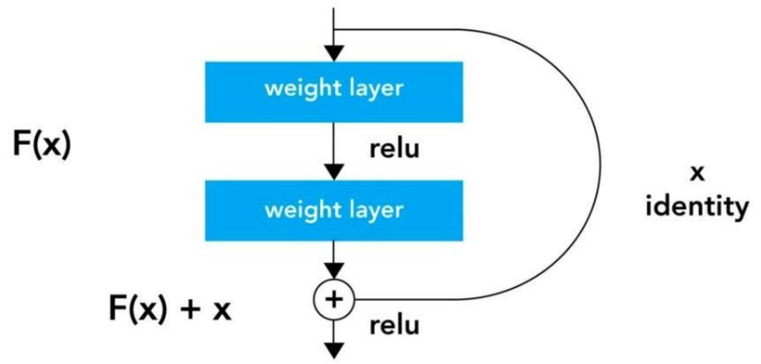

Fig. 4. Residual learning: a building block

The very first thing we notice to be different (Figure 4) is that there is a direct connection which skips some layers(may vary in different models) in between. This connection is called 'skip connection' and is the core of residual blocks. Due to this skip connection, the output of the layer is not the same now. Without using this skip connection, the input 'x' gets multiplied by the weights of the layer followed by adding a bias term.

DenseNet

A DenseNet is a type of convolutional neural network that utilizes dense connections between layers, through Dense Blocks, where we connect all layers (with matching feature- map sizes) directly with each other To preserve the feed-forward nature, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent layers. In DenseNet, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent layers. Concatenation is used. Each layer is receiving a "collective knowledge" from all preceding layers. Since each layer receives feature maps from all preceding layers, network can be thinner and compact, i.e. number of channels can be fewer. The growth rate k is the additional number of channels for each layer. Since each layer in DenseNet receive all preceding layers as input, more diversified features and tends to have richer patterns.

SqueezeNet: The building brick of SqueezeNet is called fire module, which contains two layers: a squeeze layer and an expand layer. A SqueezeNet stacks a bunch of fire modules and a few pooling layers. The squeeze layer and expand layer keep the same feature map size, while the former reduce the depth to a smaller number, the later increase it. The squeezing (bottleneck layer) and expansion behavior is common in neural architectures. Another common pattern is increasing depth while reducing feature map size to get high level abstract.

Architectural Design Strategies:

Strategy 1. Replace 3×3 filters with 1×1 filters

- Given a budget of a certain number of convolution filters, we can choose to make the majority of these filters 1×1, since a 1×1 filter has 9× fewer parameters than a 3×3 filter.

Strategy 2. Decrease the number of input channels to 3×3 filters

- Consider a convolution layer that is comprised entirely of 3×3 filters. The total quantity of parameters in this layer is:
- (number of input channels) × (number of filters) × (3×3)
- We can decrease the number of input channels to 3×3 filters using squeeze layers, mentioned in the next section.

Strategy 3. Downsample late in the network so that convolution layers have large activation maps

- The intuition is that large activation maps (due to delayed downsampling) can lead to higher classification accuracy.

AlexNet: AlexNet. The architecture consists of eight layers: five convolutional layers and three fully-connected layers. But this isn't what makes AlexNet special; these are some of the features used that are new approaches to convolutional neural networks:

- ReLU Nonlinearity. AlexNet uses Rectified Linear Units (ReLU) instead of the tanh function, which was standard at the time. ReLU's advantage is in training time; a CNN using ReLU was able to reach a 25% error on the CIFAR-10 dataset six times faster than a CNN using tanh.

- Multiple GPUs. Back in the day, GPUs were still rolling around with 3 gigabytes of memory (nowadays those kinds of memory would be rookie numbers). This was especially bad because the training set had 1.2 million images. AlexNet allows for multi-GPU training by putting half of the model's neurons on one GPU and the other half on another GPU. Not only does this mean that a bigger model can be trained, but it also cuts down on the training time.
- Overlapping Pooling. CNNs traditionally "pool" outputs of neighboring groups of neurons with no overlapping. However, when the authors introduced overlap, they saw a reduction in error by about 0.5% and found that models with overlapping pooling generally find it harder to overfit. The Overfitting Problem. AlexNet had 60 million parameters, a major issue in terms of overfitting. Two methods were employed to reduce overfitting:
- Data Augmentation. The authors used label-preserving transformation to make their data more varied. Specifically, they generated image translations and horizontal reflections, which increased the training set by a factor of 2048. They also performed Principle Component Analysis (PCA) on the RGB pixel values to change the intensities of RGB channels, which reduced the top-1 error rate by more than 1%.
- Dropout. This technique consists of "turning off" neurons with a predetermined probability (e.g. 50%). This means that every iteration uses a different sample of the model's parameters, which forces each neuron to have more robust features that can be used with other random neurons. However, dropout also increases the training time needed for the model's convergence.

VGG16: VGG16 is a convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large- Scale Image Recognition". The model achieves 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes. It was one of the famous model submitted to ILSVRC-2014. It makes the improvement over AlexNet by replacing large kernel-sized filters (11 and 5 in the first and second convolutional layer, respectively) with multiple 3×3 kernel-sized filters one after another. VGG16 was trained for weeks and was using NVIDIA Titan Black GPU's. VGG16 has a large number of hyper-parameters and focuses on having convolution layers of 3x3 filter with a stride 1 and always used same padding and Maxpool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC (fully connected layers) followed by a SoftMax for output. The 16 in VGG16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million (approx.) parameters.

The dataset consists of 2,905 images with 1341 images containing chest x rays which are normal, 1345 images containing chest x rays which are of viral Pneumonia and 219 images containing chest x rays which are of covid pneumonia.

Image data augmentation is a technique that is used to artificially expand the size of a training dataset by creating modified or different versions of images in the dataset. Training deep learning neural network models on more data can result in more skillful models, and the augmentation techniques can create variations of the images that can increase the ability of the fit models to generalize what they have learned to new images. The dataset is virtually increased by applying data augmentation techniques, the technique used in our project is Horizontal Flip which flips the image horizontally. The optimizer controls the learning rate. In this project the 'Adam' optimizer is used as the optimizer. Adam is a good optimizer that is used for many cases. The Adam optimizer adjusts the learning rate throughout training. Adam is another method that compute adaptive learning rates for each parameter. Adam is a replacement optimization algorithm for a stochastic gradient descent for training deep learning models. Adam combines the best properties of the AdaGrad and RMSP algorithms to provide an optimization algorithm that can handle sparse gradients on noisy problems. This project uses 'binary_crossentropy' for the loss function. This is the most common choice for classification. A lower score indicates that the model is performing better. To make things even easier to interpret, the 'accuracy' metric is used to see the accuracy score on the validation set when the model is trained. As a binary classifier is used, binary cross entropy is used as the loss function. Binary cross entropy is a loss function that is used in binary classification tasks. These are tasks that answer a question with only two choices (yes or no, A or B, 0 or 1, left or right).

4. Training and Results

The accuracy obtained by training the AlexNet model against the dataset gave an accuracy of about 98.89% with a training time of 5 mins and 36 secs along with the confusion matrix when run on a system with Intel i7 processor with 16 GB of RAM and 8GB GPU.

The accuracy obtained by training the DenseNet model against the dataset gave an accuracy of about 95.56% with a training time of 3 mins and 16 secs when run on a system with Intel i7 processor with 16 GB of RAM and 8GB GPU. The accuracy obtained by training ResNet -18 model against the dataset gave an accuracy of about 97.78% with a training time of 4 mins and 22 secs when run on a system with Intel i7 processor with 16 GB of RAM and 8GB GPU.

The accuracy obtained by training SqueezeNet model against the dataset when run on a system with Intel i7 processor with 16 GB of RAM and 8GB GPU resulted in overfitting and the accuracy of the model declined as the training time increased

The accuracy obtained by training VGG16 model against the dataset gave an accuracy of about 97.78% with a training time of 8 mins and 3 secs when run on a system with Intel i7 processor with 16 GB of RAM and 8GB GPU.

Comparison of the results:

The following table (Table 1) shows the comparison of accuracy and the efficiency of the five different CNN models which were trained against the data

Table 1: Result Analysis.

Name of the model	Accuracy	Training Time
AlexNet	98.89	5mins 31 secs
VGG16	97.78	8mins 3secs
SqueezeNet	92.22	-
ResNet-18	97.78	4mins 22secs
DenseNet	95.56	3mins 16secs

5. Conclusion & Future Scope

This work presents deep CNN based transfer learning approach for automatic detection of COVID- 19 pneumonia. Five different popular and previously reported efficient CNN based deep learning algorithms were trained, validated and tested for classifying normal, viral and covid pneumonia patients using chest X-ray images. It was observed that AlexNet outperforms other different deep CNN networks while image augmentation was used for training the CNN models. AlexNet being the one that outperformed other classification models it has given a classification accuracy of about 98.89%. Thus this system developed can be used as a reliable and cost effective screening tool to RT-PCR test.

Future Scope

With the continuous collection of data, we aim in the future to extend the experimental work validating the method with larger datasets. We also aim to add an explainability component to enhance the usability of the model. Finally, to increase the efficiency and allow deployment on handheld devices, model pruning, and quantization will be utilized.

References

1. (2020). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. Available: https://www.who.int/dg/speeches/detail/who-director-general-sopening- remarks-at-the-media-briefing-on-covid-19 11-march-2020
2. (2020). Coronavirus Disease 2019 (COVID-19). Available: https://www.cdc.gov/coronavirus/2019-ncov/need-extraprecautions/people-at-higher-risk.html
3. W. H. Organization, "Global COVID-19 report," March 25,2020 2020
4. J. H. U. MEDICINE. (2020). Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Available: https://coronavirus.jhu.edu/map.html
5. Diagnosing of COVID-19 Pneumonia using Chest X-Rays Available: https://pubs.rsna.org/doi/10.1148/radiol.2020202944
6. Chest X-ray Pneumonia Detection Based on Convolutional Neural Network Available:https://ieeexplore.ieee.org/document/9196403
7. Automated detection of COVID-19 cases using deep neural networks with X-ray images Available:https://www.sciencedirect.com/science/article/abs/pii/S0010482520301621?via%3Dih ub
8. COVID-19 Radiography Dataset Available:https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
9. A. J. NEWS. (2020). Bangladesh scientists create $3 kit. Can it help detect COVID-19? Available: https://www.aljazeera.com/news/2020/03/bangladesh-scientists-create3-kit-detect- covid-19-200323035631025.html
10. Tulin Ozturka, Muhammed Talob, Eylul Azra Yildirimc, Ulas Baran Baloglud, “Automated detection of COVID-19 cases using deep neural networks with X-ray images”, 2020
11. Zebin Jiang, “Chest X-ray Pneumonia Detection Based on Convolutional Neural Networks”, 2020
12. D. Haritha, M. Krishna Pranathi, “COVID Detection from Chest X-rays with Deep Learning: CheXNet”, 2020
13. Naufal Hilmizen, Alhadi Bustamam, “The Multimodal Deep Learning for Diagnosing COVID-19 Pneumonia from Chest CT- Scan and X-Ray Images”, 2020
14. Dhairya Vyas, Harsh Dave, “Classification of COVID-19 cases using Fine-Tune Convolution Neural Network (FT- CNN)”, 2021