PREDICTIVE MODEL FOR COVID-19 USING DEEP LEARNING

Hardev Goyal[1,2], Devahsish Attri[1,3], Gagan Aggarwal[1,4], and Aruna Bhatt[1,5]

[1] Department of Computer Science Delhi Technological University
[2] hardevgoyal@gmail.com
[3] devashishatri@gmail.com
[4] gaganag50@gmail.com
[5] aruna.bhat@dtu.ac.in

Abstract. The Coronavirus has now taken more than 2.4 million lives and infected more than 100.2 million people. The spread of Coronavirus has had an adverse effect on the global health and economy. The Coronavirus pandemic puts healthcare systems worldwide under immense pressure. With advancement in Machine Learning and, in particular Artificial Intelligence, early detection of Covid-19 can assist in rapid recovery and help to relieve strain from healthcare systems. Early results indicate that there are abnormalities in patients' chest X-rays infected with Coronavirus. In this Review paper an extensive and exhaustive guide to identify COVID virus using Chest X-Ray samples in an effective and cheap method has been presented. It highlights different CNN Architectures and Gradient Class Activation as the main approach to analyze and detect the infection. A significant amount of training and validation data/images are required for training neural networks such as Convolutional Neural Networks (CNNs) for accurate predictions on test data. Generative Adversarial Networks (GANs) especially ACGANs and Cycle GANs were used to create new images for the training dataset, which helped generalizing the classification model. The paper exhibits the application of different Convolutional Neural Networks architectures for transfer learning, including Inception and ResNet50. The paper then presents the combination of GAN and deep learning models for precise identification of COVID-19 infection. With so much research going on to detect COVID-19, this paper will help all researchers and doctors in the future.

Keywords: Coronavirus · Chest X-Ray · Deep Learning · Convolutional Neural Network · Transfer Learning · Generative Adversarial Network · CYCLEGANs

1 Introduction

COVID-19 epidemic has become a severe health crisis. "As of now the number of people infected with Coronavirus is approximately 90,988,272, with 1,947,499 deaths all over the world and 87,040,773 cases in which people have been cured successfully." [1] The main symptoms for Coronavirus are, respiratory difficulty, headache, pain, fever and cough, loss of smell and taste. The patient may also suffer from Pneumonia in critical cases. The infection can also lead to a serious intense respiratory condition, septic stun, multiple-organ failure, and ultimately, loss of life. Studies have shown that men (58Virus detection methods take lesser and lesser time as new methods are being established in various countries worldwide. The test results allow the specialists to isolate and medicate infected patients in a convenient and agreed manner [2][3].

The two ways to detect the infection is by the existence of the virus or vaccine that might have been produced in reaction to infection. Viral existence tests are used to diagnose individual cases and allow traceability and control of public health authorities' outbreaks. Instead, antibody tests indicate whether anyone once had the disease. For diagnosing current infections, they are less useful because vaccines may not develop for weeks after contamination [4]. It is used to determine the widespread presence of infections, which predicts the fatality rate of infection. X-ray is an imaging procedure used to research cracks, relocations of bones or even for chest infections. X-rays have been in dynamic use for a long time [5]. It provides a fast and effective way of examining the lungs and can help detect COVID-19 infections. The CT-Scans or MRIs give a detailed analysis of lungs but they are very expensive and out of reach of common people whereas X-Rays are cheap, and with help of neural networks and fast computing they can be used in remote locations where the healthcare infrastructure is not adequate and can also be helpful for people who cannot afford the current tests [6]. Easy transportation of portable X-ray devices is an added advantage to detecting Coronavirus using X-rays. This paper will be a guide for doctors and researchers to identify chest X-ray samples between Coronavirus patients or healthy people.

This paper is divided in 4 broad sections:

- Section II talks about various technologies used in the research.
- Section III describes existing methodology, different data sets used by researchers and the different feature extraction techniques and classification steps
- Section IV discusses the results of various models and presents the final conclusion.
- Section V we conclude the report and discuss briefly how this work can be used in future.

2 Background

2.1 Deep Learning

Deep Learning has been very prevalent over the past few years, and it finds applications in a wide range of domain such as Speech, Computer Vision and NLP. Most state-of-the-art in these areas, from even companies like Facebook, Google etc. use deep learning as the underlying solutions [7][8]. Deep learning uses approaches assisted by representative learning through neural networks for artificial learning. It learns from representative examples instead of using task-specific algorithms. For example, you need to prepare a database containing many different cat photos if you want to create a model that recognizes cats by species. [2] Today, in image/video classification, DL is rapidly becoming a vital technology. Hidden deep layers on deep learning maps given data to required labels to explore hidden patterns in complex data. Their use in medical X-ray recognition can help in Coronavirus detection.

2.2 Convolution Neural Network

Vision is a huge part of human life. There is three central part to a CNN:

- Convolution: This is used for extraction of features from a given image or to select most relevant features from the previous layer..
- Non-linearity: It allows to deal with non-linear data
- Pooling: Allows your image's spatial resolution to be sampled down so that the necessary features can be magnified.

A CNN powerfully uses adjacent pixel information to successfully down sample the image first by convolution and then use classification layer(s) to give the final

After filters are applied on a set of features the most important features will be highlighted which can then be used to make further predictions. A filter is basically a matrix used to magnify the most important features. [10] For medical image/video classification and identification, Convolutional Neural Networks have achieved remarkable success. Convolutional neural networks (ConvNets or CNNs) are key groups for image classification in neural networks. One of the most readily accessible and realistic approaches to the diagnosis of Coronavirus from X-ray is the Convolutional Neural Network (CNN). Several reviews are being carried out to highlight recent contributions to the identification of Coronavirus. In medical images, Convolutional neural networks (CNNs) have achieved state-of-the-art efficiency, given sufficient data. Training on a labelled date is crucial to fine-tune its millions of parameters. Due to many parameters, Convolutional Networks can easily overfit/underfit to a data having less samples along with same kind of samples, so the degree of generalization is highly dependent on the size of the labelled dataset. With a restricted number of samples, small datasets are the most significant challenge in medical imaging. Provided sufficient data Convolutional Neural Networks (CNNs) have achieved near perfect efficiency in classification of medical images. To fine-tune its millions of parameters, training on a labelled data is necessary. CNN's can easily over fit small datasets due to several parameters, so a large dataset is required for deep learning models. Small databases are the most critical problem in medical imaging, with a limited number of samples. After reviewing the current models, a trend is highlighted that even though deep learning and CNN are successful in the domain of computer vision, the accuracy of COVID detection using x-rays is still low because of the limited number of training dataset. A large dataset is required to train an efficient and accurate model. [12]

2.3 Generative Adversarial Networks

The aim of a GAN is to learn to produce data that is indistinguishable from the data being used to training. The reason we call the training process for generative adversarial network adversarial training is that the other player generates a losing scenario for the first person. Even though, one of the players is performing at its worst the other player is being benefitted because it is learning to perform well on the inputs provided by the other network. Generative Adversarial networks are mostly intended to solve the task of generative modelling. [12] The idea behind generative modelling is that we have a huge dataset of training examples usually of large multidimensional examples. The particular approach the generative adversarial network take to generative modelling is to have two different models playing a game against each other. One of these agents is the generative network which tries to generate data, and the other model is a discriminator network that examines data and determines whether it is real or fake. The goal of the generator is to fool the discriminator. As the players compete with each other to win they get so good at their jobs that eventually the generator reaches a level where it is capable for producing realistic images, which are indistinguishable from the images in the training dataset. In the first half of the training process, we take a random set of images from the training dataset and call this set as X. The discriminator is represented by D, which can also be considered as the first player. The discriminator is a neural network, i.e. it is a differentiable function whose parameters define the shape of the function. We apply the function D, the discriminator neural network, to the set X, and aim of D is to make D(x) as close as possible to unity i.e. 1. [4][5] In the second half of the training process, we sample some random noise Z from a prior distribution over latent variables in our generative model. Z is just a sort of randomness that allows the generator to output many different images instead of outputting only one realistic image. After we sample the input noise Z, we apply the generator function which is similar to the discriminator function. The generator is a differentiable function controlled by some set of parameters, and in other words, it's usually a deep neural network. After applying the function G to input noise Z, we obtain a value of X sampled from the model. We apply the discriminator function to the fake example that we pull from the generator. The discriminator tries to make its output D(G(Z)) be near to 0. Earlier, when we use the discriminator and real data, we wanted D(X) to be near 1 and now the discriminator wants D(G(Z)) to be near null/0 to distinguish that the input is false. At the same time, the generator aims to make D(G(Z)) close to unity/1, thus proving the discriminator wrong. [2][3] The current data set doesn't have sufficient amount of images to train an efficient and accurate model. This is because of lack of medical images. To have a data set of significant size, Auxiliary Classifier Generative Adversarial Network (ACGAN) which is an extension of GAN, it predicts the label of an image instead of getting the label as an input. It is capable of producing high quality images which are indistinguishable from real chest X-ray images.

GAN has two main components:

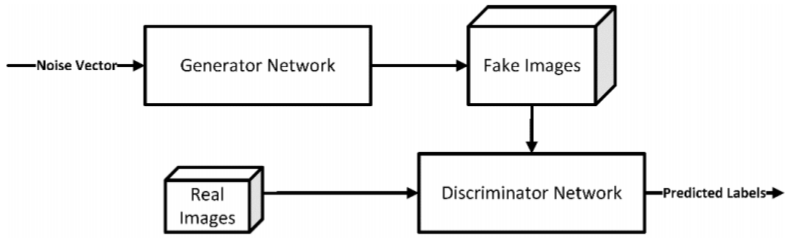

Fig. 1. Generative Adversarial Network model

- The generative part is responsible for taking N-dimensional uniform random variables (noise) as input and generating fake images. The generator captures the probability P(X), where X is the input.
- The discriminatory part is a simple classifier that evaluates the images produced and separates them from the images. The discriminator takes the conditional condition P (Y — X), where X is the input variable and the label is Y.

In GANs, the input for one network is produced by another network. The quality of the input depends on the performance of that particular network. We can think of the Generator and Discriminator as being counterfeiters and police. The police would like to allow people with real money to spend their money without being punished. Still, they would also like to capture counterfeit money, remove it from circulation and punish the counterfeiter. Simultaneously the counterfeiter would want to fool the police and successfully use their money. But if the counterfeiters are not very good, they will get caught. Overtime the police learn to be better at catching counterfeit money and counterfeiter learn to produce it. [6] So in the end, to examine this case, we can use game theory, we discover that if both the police and the counterfeiter, or in other words, both the Discriminator and Generator have infinite capabilities, then this game's Nash Equilibrium corresponds to the generator generating perfect samples from the same distribution as the training results. The counterfeiters, in other words, make counterfeit cash that is indistinguishable from real money. And at that point, the discriminator or, in other words, the police cannot distinguish between the two data sources and say that each input is half likely to be true and half likely to be false. We can formally describe the learning process using the Minimax Game.

3 EXISTING METHODLOGY

Abdul Waheed, Muskan Goyal et al [1] used CNN and ACGANs for COVID-19 detection in this analysis. They combined Covid X-Ray images produced using ACGAN and the original images so as to make the dataset generalized and large and then used with their proposed CNN architecture. These were the following contributions to this study:

- ACGAN was used for the first time to produce synthetic Covid images in form of X-Rays.
- For COVID-19 detection, [1] developed a CNN-based model.
- Used ACGANs for generation of training dataset with the different CNN models like Resnet ,Xception , Inception , VGG for better detection of Coronavirus.

Elene , Gabriel et al. [2] used CNN along with heat activation AND CyclicGANs for detecting this infection. They also used MobileNet With Support Vector Machine. Pamuk et al used a mixture of Pneumonia and Coronavirus Dataset which resulted in 98.7M.Talo et al in their paper used 3 classes for the model – Pneumonia , Covid and Normal with more than 1127 images combined which resulted in 98

3.1 Data Generation

Radiologists and Researchers participation is required in order to make a good database of medical images like for the X-Rays. It is expensive and requires repeating process again and again. COVID-19 is very new and there has not been so much testing with X-Rays, so it is difficult to collect adequate Chest X-ray (CXR) image data. By using synthetic data augmentation, we suggest alleviating the disadvantages. To create the dataset, most Researchers obtained the images from these available datasets:

- IEEE X-ray dataset for COVID Chest
- X-Ray Database COVID-19 and
- COVID-19 Initiative on the X-ray Covid Data

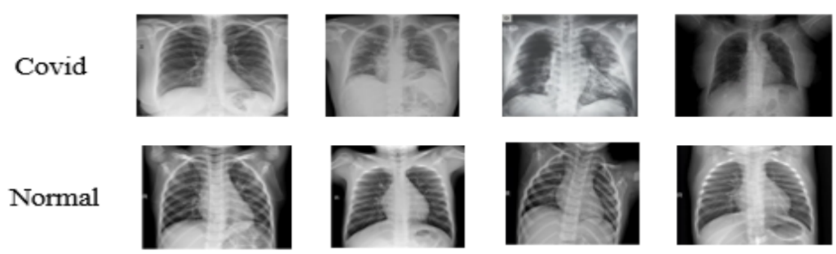

Fig. 2. Sample of Images

As these collection of databases are available for open source contribution they were the best choice for building the X-Ray Dataset. Also they are freely accessible to each person who want to use it. The photos obtained are combined, and exactly similar images from the dataset are deleted using Image Hashing method. This method generates a hash value based on the contents of an image that identifies an input image uniquely. The dataset contains more than 460 images of COVID chest and 1266 Normal-CXR. Sadly, the number of COVID-CXR is much smaller than required to train an accurate and useful learning model. Therefore, we have created a GAN model that can be used to develop new COVID-CXR and Normal-CXR.

3.2 B. Auxiliary Classifier Generative Adversarial Network

The ACGAN was used by Abdul et al [1] to make their own CovidGAN. It is difficult for GANs to produce high-resolution images from unbalanced form of data. They used this GAN in their research because to generate better image quality this model rather than focusing on only the specific images it focuses on external knowledge. To produce an image for a specific class, along with the label of that class, a random matrix of points generated through random function is given to the generator. An image and a class label are given to the discriminator, determining whether the image is correct or not. The main advantage of a ACGAN is that the discriminator of this GAN tries to output specific image's class label and not to receive it as an input unlike in most other GANs. It helps to control the overfitting part and assists the production of high-resolution images. In this not only it produces better images but learns features of that image that is not dependent on given image's class. The generator uses the label c for a class and the random noise to produce a sample and then the discriminator D produces a probability distribution. This distribution is between labels and the sources. D maximizes Ls + Lc, and G maximizes Lc - Ls.

- Lc = E[log P(C = c— Xreal)] + E[log P(C = c— Xfake)] (1)
- Ls = E[log P(S = real— Xreal) + E[log P(S = fake— Xfake)] (2)

3.3 Cyclic GANs

Tahmina Zebin et al [5]Used Cyclic GANs for image augmentation. As the dataset were unbalanced CyclicGANs were helpful in creating new data because it uses 2 Generators and Discriminators and then learns characteristics from one image sample data which is further used to create new data. In this a type of Cycle is formed as first the image is generated by first Generator G1 which hen acts as an input to another Generator G2 and that resembles the original image.

3.4 Model Architecture

Most of the previous researchers have used VGG16, ResNet50, InceptionV3, and Xception pre-trained models which are backboned by a fully connected layer. They replaced the pretrained models final classifier with their classification layer of two classes (COVID-19 positive and COVID-19 negative samples), each with a ReLU and ELUs activation to supply the ultimate output.

All these models have already achieved high accuracy on the Standard ImageNet Dataset. The ResNet50 model consists of 48 Convolution layers. It takes

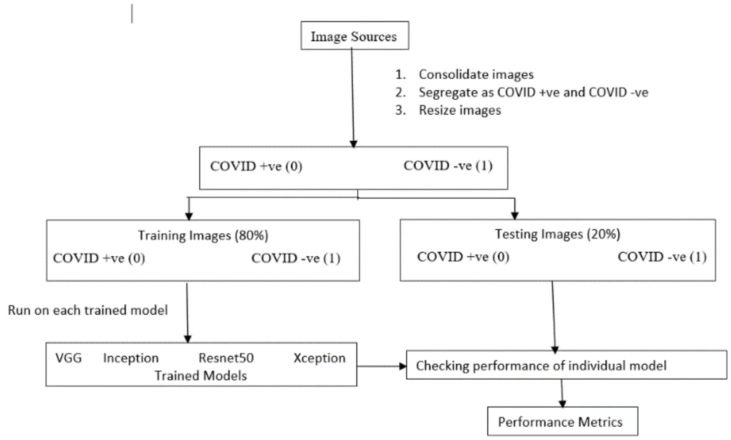

Fig. 3. Methodology

224 × 224 size of the image. It uses CONV2D, Batch Normalization and Max-Pooling as a combined layer and then uses it extensively. At the End Global Average Pooling is used, and then softmax is attached to find COVID or Non-COVID instead of labelling it into ten classes as in ImageNet. Also, batch normalization is used in reasonable amounts which helps to reduce overfitting. RMSProp Optimizer is used in the case of InceptionV3 Model. The Xception model is 71 layers deep. It uses Depth wise Separable Convolution. Convolution size is d × d × 1 instead of d × d × c where d is the filter size, and c is the channel. While training and testing, for calculation the losses, binary cross-entropy was used as the function because output was binary that is COVID or Non-COVID.

Loss is calculated as:

– Loss=$y^i(-log(y_i)) + (1 - y^i) * -(log(1 - yi))(3)$

where yi is the ith vector in the model output, yi is the value. Binary loss entropy was used we have to classify between Covid Positive and Covid Negative.

SGD (Stochastic Gradient Descent) was used as the optimizer with learning rate kept as low as 0.0005. Its output is derivative of a set of parameters of input. It takes significantly shorter training time than old Gradient Descent. The path taken to reach the minima is usually noisier but efficient. The batch size was also kept low as 16. The model was trained for around 200 epochs for all four models and achieved high accuracy.

4 Result

Results obtained from training the deep learning models like Xception, VGG16, ResNet50 and Inception V3, were compared based on their accuracy, recall value and F1 score in order to choose the most efficient model that performs the best on the available dataset. First each model is trained with the available dataset. The accuracy score per epoch is elaborated in Fig6. The ResNet50 exhibits a very unstable accuracy in the testing data compared to the training data. Xception model performs very well in the training set but performs poorly in testing set. Only VGG16 and Inception V3 perform well both in training and testing data with Inception V3 showing a better performance compared to VGG16.

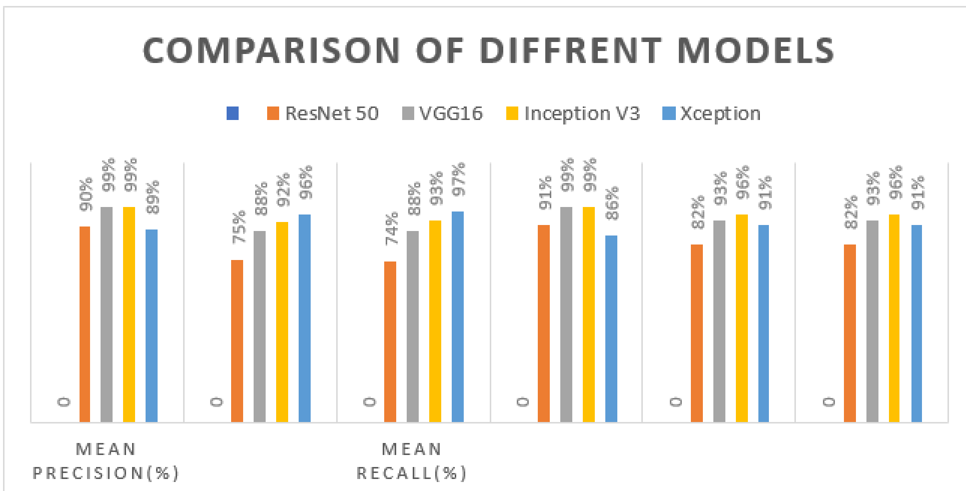

Fig. 4. Comparison of Different CNN Models.

To quantify and evaluate the CNN model's performance using synthetic data augmentation technology, recall (or sensitivity), F1-score, and specificity are used. Precision is the models ability to label a wrong sample as negative and correct sample as positive in other words, it asks the question "What proportion of positive identifications was actually correct?". Recall is the ratio of model's ability to identify all those with the problem [23] correctly in other worlds is asks the question, "What proportion of actual positives was identified correctly?".

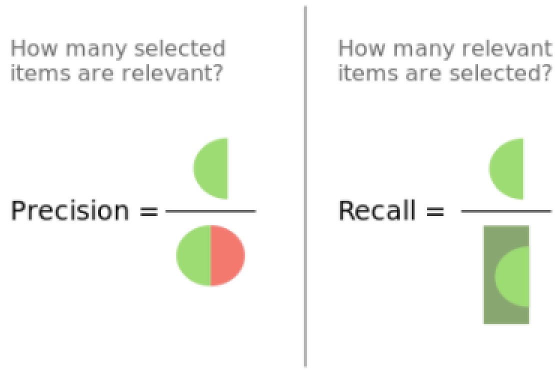

Fig. 5. Comparison between Precision and Recall

The weighted average of precision and recall is the F1-score. Here are the formulas of the measures :

$$\text{sensitivity} = \text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{F1score} = 2 * \left(\frac{\text{recall} * \text{precision}}{\text{recall} + \text{precision}}\right)$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

According to our requirement, we need a very high recall value on the cost of having a lower precision value. For example, it is necessary to identify every COVID positive patient on the cost of identifying a few negative cases as positive i.e. a person having COVID should never be misidentified.

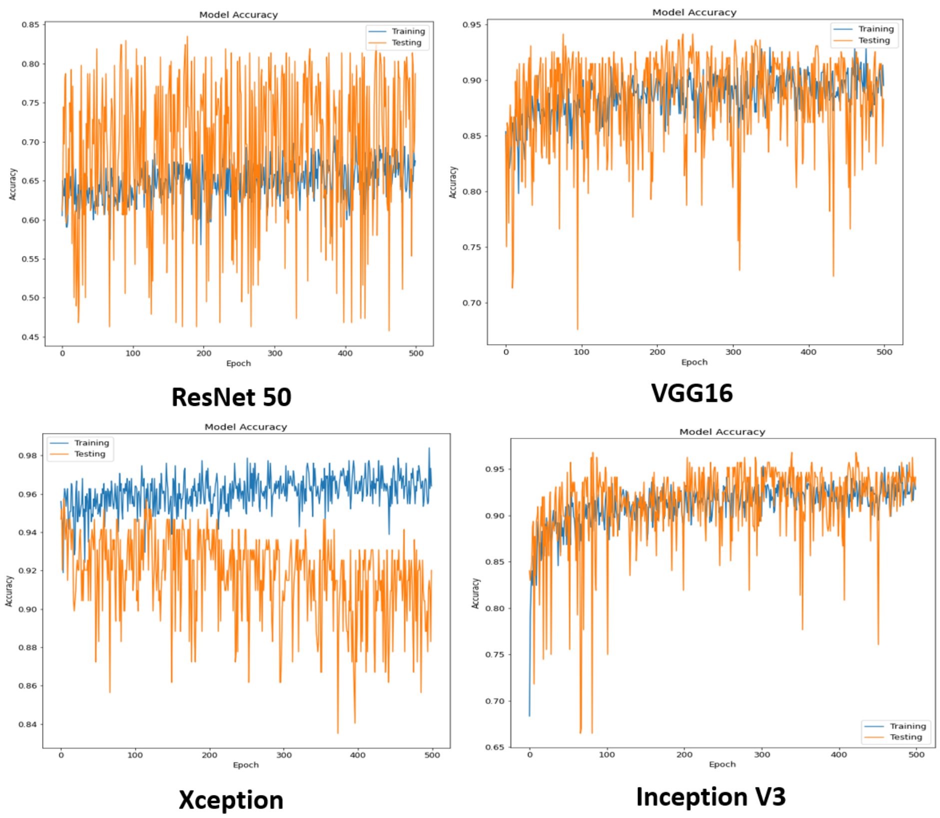

Fig. 6. Training and Testing for Different CNN Models.

VGG

Dataset	Class	Precision	Recall	F1-score	Support	Accuracy (%)	Sensitivity (%)	Specificity (%)
Actual Data (CNN-AD)	COVID-CXR	0.89	0.69	0.78	72	**85**	**69**	**95**
	Normal-CXR	0.84	0.95	0.89	120			
	Macro-average	0.87	0.82	0.84	192			
	weighted-average	0.86	0.85	0.85	192			
Actual data + Synthetic Augmentation (CNN-SA)	COVID-CXR	0.96	0.90	0.93	72	**95**	**90**	**97**
	Normal-CXR	0.94	0.97	0.96	120			
	Macro-average	0.95	0.94	0.94	192			
	weighted-average	0.95	0.95	0.95	192			

InceptionV3

Dataset	Class	Precision	Recall	F1-score	Support	Accuracy (%)	Sensitivity (%)	Specificity (%)
Actual Data (CNN-AD)	COVID-CXR	0.82	0.79	0.75	72	**93**	**69**	**95**
	Normal-CXR	0.93	0.96	0.84	120			
	Macro-average	0.94	0.84	0.85	192			
	weighted-average	0.95	0.82	0.95	192			
Actual data + Synthetic Augmentation (CNN-SA)	COVID-CXR	0.96	0.93	0.73	72	**95**	**90**	**97**
	Normal-CXR	0.94	0.96	0.86	120			
	Macro-average	0.95	0.97	0.99	192			
	weighted-average	0.95	0.93	0.96	192			

Resnet 50

Dataset	Class	Precision	Recall	F1-score	Support	Accuracy (%)	Sensitivity (%)	Specificity (%)
Actual Data (CNN-AD)	COVID-CXR	0.81	0.65	0.76	72	**82**	**69**	**95**
	Normal-CXR	0.83	0.95	0.89	120			
	Macro-average	0.87	0.84	0.86	192			
	weighted-average	0.86	0.85	0.84	192			
Actual data + Synthetic Augmentation (CNN-SA)	COVID-CXR	0.96	0.90	0.93	72	**95**	**90**	**97**
	Normal-CXR	0.94	0.97	0.94	120			
	Macro-average	0.95	0.94	0.94	192			
	weighted-average	0.94	0.95	0.95	192			

Xception

Dataset	Class	Precision	Recall	F1-score	Support	Accuracy (%)	Sensitivity (%)	Specificity (%)
Actual Data (CNN-AD)	COVID-CXR	0.86	0.63	0.77	72	**91**	**69**	**95**
	Normal-CXR	0.74	0.95	0.89	120			
	Macro-average	0.87	0.82	0.84	192			
	weighted-average	0.86	0.83	0.85	192			
Actual data + Synthetic Augmentation (CNN-SA)	COVID-CXR	0.92	0.90	0.92	71	**95**	**90**	**97**
	Normal-CXR	0.93	0.92	0.92	121			
	Macro-average	0.93	0.92	0.91	191			
	weighted-average	0.93	0.91	0.92	193			

Fig. 7. Detailed results for different CNN Models with and without using GANs.

5 Conclusion and Future Scope

Timely recognition of patients with COVID-19 is vital for opting correct handling and also prevent widespread of virus. The results from the above research shows that an effective model can be trained using existing methodologies with slight modification and a sufficient amount of images in the dataset. Deep Leaning is an essential to achieve such a result. The proposed method doesn't have any clinical study to support its efficiency and reliability. Thus, right now it can't replace a medical diagnosis by an expert medical professional. Therefore a more thorough investigation and a model trained to a comparatively larger dataset is required. Under such scenarios, the work shows high probability of a precise, automated, quick and affordable method for the diagnosis. For future work, the amount of images of both the classes can be increased in the collection of images by adding more X-ray images of people already tested positive for COVID and also adding other diseases which affects lungs in a similar way as in COVID-19, thus making the approach more efficient and generic. This will allow doctors and medical professionals throughout the globe to carry out more extensive testing without the use of existing testing techniques, which are slow and can also act as a hotspot for the spread of virus. Furthermore, the proposed approach can be compared with techniques based on fine-tuning, and many other models can be trained and tested from scratch.

References

1. Abdul Waheed , Muskan Goyal , Deepak Gupta ,Ashish Khanna , Fadi Al- Turjman And Placido Pinherio "CovidGAN : Data Augmentation using Auxilliary Classifier GAN for Improved Covid-19 Detection" IEEE Access 2020
2. Elene Firmeza Ohata, Gabriel Maia Bezerra, João Victor Souza das Chagas, Aloísio Vieira Lira Neto, Adriano Bessa Albuquerque, Victor Hugso Costa de Albuquerque , "Automatic Detection of COVID-19 Infection Using Chest X-Ray Images Through Transfer Learning" IEEE 2020 Tahima Zebin , ShahadateRezvy "COVID-19 detection and disease progression visualization: Deep learning on chest X-rays for classification and coarse localization" SpringerLink 2020.
3. "Jiang Xiango, et.al "Towards an artificial intelligence framework for data-driven perdiction of coronavirus clinical severity." CMC(2020)
4. Irem Mertyuz ,Tolga Mertyuz , Beyda Tasar , Oguz Yakut "Covid-19 Diesease Diagnosis From Radiology Data with Deep Learning Algorithms" IEEE(2020)
5. A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," 2018 International Interdisciplinary PhD Workshop (IIPhDW), pp. 117–122, 2018.
6. A. Beers, J. M. Brown, K. Chang, J. P. Campbell, S. Ostmo, M. F. Chiang, and J. Kalpathy-Cramer, "High-resolution medical image synthesis using progressively grown generative adversarial networks," ArXiv, vol. abs/1805.03144, 2018.
7. Wu Fan et al. "A new coronavirus associated with human respiratory disease in china" Nature 579.7898(2020): 265-269
8. Parul Hooda , Kumar Akshi , Vikrant Dabas, "Text Classification algorithms for mining unstructured data , a SWOT analysis" . IJIT(2018)

9. Aouatif Najoua , Hassan Ouajji , Youssef Mnaoui "Analyzing COVID-19 Crisis in North America". IEEE Conference(2020)
10. A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks, " arXiv preprint arXiv: 2003.10849, 2020.
11. Milan S.Milvojevic , Ana Gavrovska "Long Short-Term Memory Forecasting for COVID19 Data"
12. Zizhan Tang "Adaptive Griup Testing Models for infection Detection of COVID-19" IEEE 2020