Classification of lung infected Corona Virus X Ray Images using Deep Learning CNN Model

Yogesh Kumar Gupta, Saroj Agrawal, Aman Mittal

1. Department of Computer Science, Banasthali Vidyapith, Rajasthan, gyogesh@banasthali.in
2. Banasthali Vidyapith, Rajasthan. sarojagr708@gmail.com
3. IIITKanpur, amanmitl@iit.ac.in

Abstract. All Start In this examination as learning, a dataset of COVID-19 lung contaminated X-beam pictures with common infection pneumonia, insisted Covid-19 sickness, and ordinary occurrences were utilized for the programmed acknowledgment of the Corona virus infirmity. The purpose of the investigation is to the comparable portrayal of Covid-19 utilizing Convolution Neural Network and assesses the precision proposed model convolutional neural framework over the progressing years for clinical picture characterization. Corona Virus lung infected X-Ray Images accessible by Kaggle a complete 590 images, which classified in 2 classes: typical and Covid-19. To survey the hypothesis accuracy of the models. The plan models end up being productive appeared differently in relation to made by (Arpan Mangal et al, 2020), (A. A. Saraiva et al., 2020), (Ioannis D. Apostolopoulos et al, 2020), (Ali Narin et al. 2020), which gained accuracy 90.5, 95.30, 96.78, 98 independently and the current work had an ordinary precision of 99.54.

1. Introduction

As we know mostly all the lung diseases are infectious and can be transmitted from one person to another. Thus making Lung diseases very deadly as they can spread across the globe in very minimal time. In the Light of the recent event COVID19 is the most recent lung disease that we are encountered with and since then entire world is fighting against it day and night. Our lung plays a vital role in respiration process and its major role is to provide oxygen and remove carbon dioxide, with the increase in air pollution our respiratory system comes in contact with many hazardous elements which are not good and can cause major lung diseases. The virus is spread with people in close contact via droplets produced by sneezing, coughing, talking etc. The droplets remain on nearby surfaces thus the surfaces come in contact with other people and hence infecting them. That's why people are focusing on maintaining social distancing, wearing masks and sanitising hands regularly.

It is seen that COVID19 is more risky to persons having weak immunity particularly children and old-citizens. Same goes with the persons who are having diseases such as heart disease, diabetes, asthma etc. COVID19 is not only a danger for community health but it has also seen to generate discrimination in the infected countries. The first case of COVID19 came in December 2019 in China (Wuhan city) and was observed that it came from a fish later it spread across the world.

In this paper, our essential centre is to distinguish COVID19 from lung x-beam pictures. With the fast multiply of COVID19 and the expanding request of conclusion analysts are headed to discover more clever determination measure. The classification consists of two parts, first part is the training and the second part is the validation and testing the model on the unknown images.
We divided this paper into 7 sections; section 2 is to show the previous studies in this field. In section 3 explained the steps of the algorithm. In section 5 and 6 presented model results and conclusion.

2. Literature Review

Many past researches are being bone on detecting COVID19 from radiological images. We have shown all such studies in the below table 1

| Author Name                  | Paper Title                                                                 | Algorithm                                      | Observation                                                                 | Accuracy  |
|------------------------------|------------------------------------------------------------------------------|------------------------------------------------|----------------------------------------------------------------------------|-----------|
| Ioannis D. Apostolopoulos1, Tzani A, Mpesiana2 | Covid-19 automatic detection from X-ray images utilizing transfer learning with convolutional neural networks | Deep learning with CNNs                        | The aim is to evaluate the performance of CNN architectures for medical image classification. | 96.78%    |
| Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U. Rajendra Acharya | Automated detection of COVID-19 cases using deep neural networks with X-ray images | The DarkNet model in deep neural networks       | to detect and classify COVID-19 cases from X-ray images.                   | 98.08%    |
| Ali Narin1, Ceren Kaya2, Ziyin Panuk2 | Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks | three different convolutional neural network based models (ResNet50, InceptionV3 and Inception-ResNetV2) | the detection of coronavirus pneumonia infected patient using chest X-ray radiographs. | 98%       |
| A. A. Saraiva, N. M. Fonseca Ferreira | Classification of Images of Childhood Pneumonia using Convolutional Neural Networks | CNN Model using K-fold                          | Classify images in normal and pneumonia.                                           | 95.30%    |
| Nour Eldeen Mahmoud Khalifa et al | Detection of Coronavirus (COVID-19) Associated Pneumonia based on Generative Adversarial Networks and a Fine-Tuned Deep Transfer Learning Model using Chest X-ray Dataset | CNN Model using K-fold                          | a pneumonia chest x-ray detection based on generative adversarial networks (GAN) with a fine-tuned deep transfer learning | 99%       |
| Halgurd S. Maghdid, Aras T. Asaad | Diagnosing COVID-19 Pneumonia from X-ray and CT Images using Deep Learning and Transfer Learning Algorithms | deep learning and transfer learning algorithms, CNN and modified pre-trained AlexNet model | COVID-19 Detection                                                              | 98%       |
| Srinivas Atreya                | DISTINGUISHING COVID-19 USING CHEST X-RAYS | Uniform Manifold Approximation and Projection(UMAP) uses manifold learning techniques | separate out the Normal, Pneumonia and COVID-19 chest X-Ray images | Not Analyzed |
| Ho Yuen Frank Wong, Hui Yin Sonia Lam, Ambrose Ho-Tung Fong | Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients | RT-PCR Method                                  | finding the severity of CXR                                                     | Not Analyzed |
| Sohaib Asif, Yi Wenhui         | Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Networks | Deep convolutional neural networks (DCNN)       | detection of coronavirus pneumonia infected patients using chest X-ray radiates gives a classification accuracy | 98%       |

3. Approach to Our Model

3.1 Dataset Description

The dataset we utilized contains 590 X-beam (JPEG) pictures with two classifications: Normal and COVID-19. The X-beam pictures are taken from Kaggle.com. All pictures in dataset are physically examined so as to evacuate inferior quality pictures, just as being characterized by expert doctors to affirm the ground truth and miss-characterization.
• Training Dataset – My training dataset comprises of 370 images with 130 COVID X-Ray images and 240 Normal X-Ray images.
• Validation Dataset - My validation dataset comprises of 220 images with 82 COVID X-beam pictures and 138 Normal X-beam images.
• Labels – COVID19 : 0, Normal : 1

3.1.1. COVID-19 and Normal. Radiologists utilise a wide range of highlights in a x-beam picture to identify the COVID19. It is seen that ground glass design (GGO) is usually found in COVID19 patients. The GGO examples are distinctive in the later phases of the infection when contrasted with its beginning phases. In radiography, murkiness is utilised to speak to aspiratory combination which compares to the whitish territory, as appeared in the Figure 1 (b), the X-beam picture of a COVID19 tainted individual and Figure 2 (a) shows the x-beam of a typical patient.

![Figure 1. Normal Images](image1.png)  ![Figure 2. COVID Images](image2.png)

3.2 Input Structure of the Model
In Figure 3 essential design of our model is spoken to, which comprises of the image in the CNN network (which is utilized for categorization of pictures). Right off the bat all the pictures are resized to (224x224) continued by standardization with the goal that all the values lie between 0 to 1.

![Figure 3. Input Structure](image3.png)

3.3 Training composition of the Model
Our training composition has many steps; the first step is the division of images into training and validation data sets than all the images are resized and normalized. The second step is the training of the images. Than the model validation is done through the validation data set.

3.4 Proposed Convolution Neural Network (CNN) Model
Every CNN has two parts first is the feature extraction part and the second one is the classification part. Feature Extraction part consists of convolutional layers and pooling layers while the categorization fully connected layers.

• The Convolutional Layer: This layer helps in extracting features from the input image which afterwards is helpful in making predictions. We give image matrix and a filter which is smaller than the image matrix as an input to the layer. Filters are a set of weights which are learnt by our model to detect specific features. The filter is moved on every part of the picture,
left to right to check a particular feature. We basically perform dot product between the small parts of input matrix and the filter. The learning of filter weights is done by training process.

- The Pooling Layer: It is utilized to reduce the parameters of the network and quantity of boundaries of our model in this way helps in diminishing the computational expense. This is generally put after every convolutional layer. For the simple understanding it just simply condenses the feature maps. Thus all the operations afterwards are applied on these condensed feature maps. There are mainly 3 types of pooling max-pooling, average pooling and global pooling. For our model we used max-pooling (most preferred type of pooling), in this it selects the max element from every small parts of the feature map which are covered by filter, as shown in figure 4.

- Fully Connected Layers: It is used for the classification, before giving input to the FC layers we convert our feature matrix into a vector using a flatten layer. Output of the FC layer is the 1-D matrix with the scores of all the classes in our model using a sigmoid activation function ranging from 0 to 1.

3.4.1. Binary Cross Entropy. The loss function we use is the cross-entropy loss function it helps in making a decision about the presentation of our model, with the yield being likelihood esteem between 0 to 1.

- \( \text{Loss} = - (y \log (y^\prime) + (1-y) \log(1-y)) \)  \( y = \text{original value}, y^\prime = \text{predicted value} \)

3.4.2 Convolutional Neural Network Architecture: In Figure 5 is the CNN architecture designed in 11 layers, seven feature extraction layers and four classification layers, the network input receives a 224x224 pixel image after normalization, used RELU activation function for each convolutional layer and a kernel of size 3x3, also gradually increased the amount of the features to be extracted by convolutional layers as go deep into the model. In this used max-pooling layer helps in dogging the problem of over fitting. To reduce the training time we also use dropout layers. At the end we used sigmoid function for the classification output.
4. Measuring Parameters of the Evaluation

4.1 Confusion Matrix
The confusion matrix contains true positives, false positives, true negatives, false negatives.
- **True +VE**: People are COVID-19 according our model.
- **False -VE**: People are actually with COVID-19 but categorized as Normal our model.
- **False Positive**: People are actually Normal, but categorized as COVID-19, according our model.
- **True -VE**: Peoples are really Normal and categorized as Normal according our model.

4.2 ROC Curve
The (ROC) Receiver Operating Characteristic curve is plotted involving true positives and false positives by varying thresholds. It is used to visualise the performance for our classification problem. The AUC (area under the curve) score of ROC curve lies between 0 to 1. More the AUC score means more our model is accurate. It is basically a curve which shows relationship between test accuracy and the probability of miss-classification as shown in Figure 6.

4.3 Precision-Recall plot
It is a plot between precision and corresponding recall values as shown in figure 7. We can also calculate an AUC score for this plot. Similar to ROC more the AUC score of the p-r plot more accurate our model.

4.4 Classification report
Classification report consists of all the other features used to measure a model such as precision (for all the classes), recall (for all the classes)), F-1 score, accuracy. Accuracy - Accuracy shows how correct are the predictions made by our model as compared to round truth.
Precision - Precision is a measure that mentions to us what extent of patients that we analyzed as having COVID19, really had Covid19. Precision = T.P. / (T.P. + F.P.)
Recall - Recall is a measure that mentions to us what extent of patients that really had COVID19 was analyzed by the calculation as having COVID19. Recall = T.P. / (T.P. + F.N.)
F-1 score - It is the harmonic mean of precision and recall, more F-1 score means more accuracy.

Table 2: Categorize Report

|    | Precision | Recall | F-1 score | Support |
|----|-----------|--------|-----------|---------|
| 0  | 0.99      | 1.00   | 0.99      | 82      |
| 1  | 1.00      | 0.99   | 1.00      | 138     |
| Accuracy | | 0.95 | 1.00 | 220 |
| Macro Avg | 0.99 | 1.00 | 1.00 | 220 |
| Weighted Avg | 1.00 | 1.00 | 1.00 | 220 |

5. Result
In this area will be introduced the outcomes got by our characterization model and the measurements we depicted previously. Figure 6, 7, respectively shows ROC curve, precision-recall plot. Table 2 shows the classification report of our model.
6. Conclusion
In this research Convolutional Neural Networks is utilised for building our arrangement model. We have additionally contrasted our model with the related past examinations (Shown under section 2 of Literature Review) and it comes out to be classification model has achieved a highest accuracy of 99.54 percent.

References
[1]. Ioannis D. Apostolopoulos1 , Tzani A. Mpesiana, et al, 2020, “Covid-19: automatic detection from X-Ray images utilizing transfer learning with convolutional neural networks”, Physical and Engineering Sciences in Medicine 43:635–640.
[2]. Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U. Rajendra Acharya et.al. 2020, “Automated detection of COVID-19 cases using deep neural networks with X-Ray images”, Comput Biol Med.; 121: 103792, doi: 10.1016/j.compbiomed.2020.103792.
[3]. Ali Narin1, Ceren Kaya, Ziynet Pamuk et al, 2020, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-Ray Images and Deep Convolutional Neural Networks", arXiv preprint arXiv..
[4]. A. A. Saraiva, N. M. Fonseca Ferreira et al, 2020, "Classification of Images of Childhood Pneumonia using Convolutional Neural Networks", Computer Science, DOI: 10.5220/0007404301120119 Corpus ID: 88488645.
[5]. Nour Eldeen Mahmoud Khalifa et al; 2020, “Detection of Coronavirus (COVID-19) Associated Pneumonia based on Generative Adversarial Networks and a Fine-Tuned Deep Transfer Learning Model using Chest X-Ray Dataset”.
[6]. Halgurd S. Maghdid, Aras T. Asaad et al, 2020, “Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms”, Deep AI: The front page of A.I.
[7]. Srinivas Atreya et al 2020, “DISTINGUISHING COVID-19 USING CHEST X-RAYS”, https://www.analyticsinsight.net/distinguishing-covid-19-using-chest-x-rays.
[8]. Ho Yuen Frank Wong, Hiu Yin Sonia Lam Ambrose Ho-Tung Fong et al, 2020, "Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients"; https://doi.org/10.1148/radiol.2020201160.
[9]. Sohaib Asif, Yi Wenhui et al, 2020, “Classification of COVID-19 from Chest X-Ray images using Deep Convolutional Neural Networks”, doi: https://doi.org/10.1101/2020.05.01.20088211.
[10]. F. Shan+, Y. Gao+, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, and Y. Shi, 2020, “Lung infection quantification of covid-19 in ct images with deep learning,” arXiv preprint arXiv:2003.04655.
[11]. Zhang, Z. and Sabuncu, M. R., 2018, “Generalized cross entropy loss for training deep neural networks with noisy labels”, CoRR, abs/1805.0783.
[12]. Brzezinski, D. and Stefanowski, J., 2017, “Frequentical auc: properties of the area under the roc curve for data streams with concept drift. Knowledge and Information Systems”, 52 (2):531–562.
[13]. Cherian, E.K Mulholland, J.B Carlin, H. Ostensen, R. Amin, M. De Campo, D. Greenberg, R. Lagos, “Lucero, M.; Madhi, S.A. et al.; (2005); “Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies”, Bull. World Health Organ, 83, 353–359.
[14]. C Santosh, S. Antani; “Automated chest X-Ray screening: Can lung region symmetry help detect pulmonary abnormalities”, IEEE Transactions on Medical Imaging, 37 (5):1168–1177.
[15]. A. Krizhevsky, Sutskever, G.E Hinton, 3–6 December 2012, “Imagenet classification with deep convolutional neural networks”, In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Lake Tahoe, NV, USA.; pp. 1097–1105.
[16]. Zhu N, Zhang D, Wang W, Li X, Yang B, et al, 2020, “A novel coronavirus from patients with pneumonia in China, 2019”, N Engl J Med 382:727-733.