IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

VOLUME : 5 ISSUE : 12 Print / Issue Publication Date: 13-Jul-2021

ISSN : 2455-2143

DOI : 10.33564/IJEAST.2021.v05i12.042

Indexed In

Google SCHOLAR · OPEN ACCESS · doi · Crossref

WWW.IJEAST.COM

editor@ijeast.com

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 5, Issue 12, ISSN No. 2455-2143, Pages 242-246
Published Online April 2021 in IJEAST (http://www.ijeast.com)

ANALYZING CHEST X-RAY TO DIFFERENTIATE SYMPTOMS OF COVID-19 THROUGH ML APPROACH

Deepali R. Deshpande
Department of IT
Vishwakarma Institute of Technology, Pune, India

Raj L. Shah
Department of IT
Vishwakarma Institute of Technology, Pune, India

Anish N. Shaha
Department of IT
Vishwakarma Institute of Technology, Pune, India

***Abstract*— The motive behind the project is to build a machine learning model for detection of Covid-19. Using this model, it is possible to classify images of chest x-rays into normal patients, pneumatic patients, and covid-19 positive patients. This CNN based model will help drastically to save time constraints among the patients. Instead of relying on limited RT-PCR kits, just a simple chest x-ray can help us determine health of the patient. Not only we get immediate results, but we can also practice social distancing norms more effectively.**

***Keywords*— Chest x-rays, CNN-model, Data-Augmentation, Gray scaling, Image Classification.**

I. INTRODUCTION

The idea struck when the cases of covid-19 rose exponentially throughout the globe. Limited availability of test centers and lack of skilled manpower not only delayed results of the tests but also made the testing process slow and inefficient. There was an immediate call for an alternate solution for testing purposes. We came up with the idea of detecting covid-19 with the help of chest x-rays, one simple parameter that would not crowd the test centers and won't delay the results, our major aim is to reduce the time lag and provide quick results so that the patient can seek the appropriate diagnosis further. We build a machine learning CNN based model, that takes in a x-ray image as an input, the model grayscales it according to its requirements, and then is passed through multiple convolutional layers which have been tested to the accuracy of approximately 95% and in the end we get the result thus reducing the time lag by great extent.

II. LITERATURE REVIEW

COVID-19 began with unknown causes of pneumonia in Wuhan and has rapidly become a pandemic. The symptoms of SARS-CoV-2 virus were unknown to the world. But as per study went on, the idea about different symptoms got clarified. Firstly, real-time reverse transcription-polymerase chain reaction (RT-PCR) technique was used, and it is still in use because of its accuracy. But in this method, samples were taken through naso-pharyngeal cavity which makes direct physical contact between two persons which increases chances of spread. Also, there are limited numbers of COVID-19 test kits available in hospitals due to the increasing cases daily. So, scientists and researchers started research and finally found a method using X-Ray images as well as CT images of lungs. But again, the problem of shortage radiologists came into picture. So, to make testing automated and to decrease the physical contact, Artificial Intelligence (AI) was introduced.

While glancing through this technique, three different convolutional neural network (CNN) based models (ResNet50, InceptionV3 and Inception-ResNetV2) have been proposed for the detection of coronavirus pneumonia infected patient using chest X-ray radiographs. ROC analyses and confusion matrices by these three models are given and analyzed using 5-fold cross validation. Considering the performance results obtained, it is seen that the pre-trained ResNet50 model provides the highest classification performance with 98% accuracy among other two proposed models (97% accuracy for InceptionV3 and 87% accuracy for Inception-ResNetV2).

A lot of models are there in working conditions which are using Convolutional Neural Network (CNN) to distinguish whether the person is suffering through COVID or not. Generally, one trained model is available in the market that is for detecting COVID. But we are going to train one more model for pneumonia. We have taken databases from *kaggle.com* which consists of lung X-Ray images of COVID-19 patients, pneumatic patients, and normal uninfected humans.

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 5, Issue 12, ISSN No. 2455-2143, Pages 242-246
Published Online April 2021 in IJEAST (http://www.ijeast.com)

III. METHODOLOGY/WORKING

1. ***Decision Tree:***

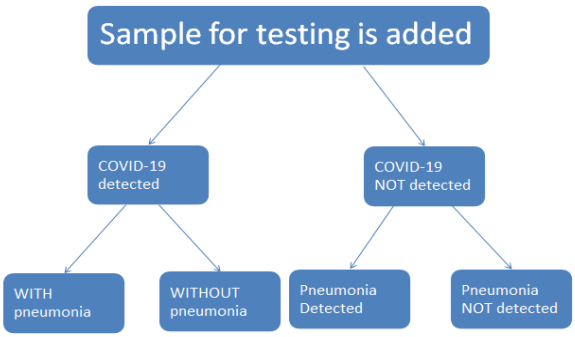

Fig. 1: Flowchart of working process.

The model will first decide if the person is suffering through COVID or not.
According to the result, the machine will move for the pneumonia detection model and will produce the final report. This is how checking for both pneumonia and corona virus is done.

2. ***Dataset:***

Image data set which consisted of chest x-ray images was downloaded from www.kaggle.com .

3. ***Data Pre-processing:***

Detection of covid-19 using ML approach is an image classification problem. Accuracy of any machine learning model significantly depends on data cleaning or data pre-processing. Data is removed or modified that is incorrect, incomplete, irrelevant, duplicated or improperly formatted.

Data cleaning helps ensure that information always matches the correct field. Several methods can be adopted for cleaning data as per the user's requirement. We have used resizing and grey scaling in our model.
Resizing images is a critical pre-processing step because machine learning models can be trained faster on smaller images. Moreover, many deep learning model architectures require that images to be of same size whereas raw collected images may vary in size. So, images were resized in fixed size 64x64 pixel images. Note: Resizing doesn't affect the quality of image.
Also, many irrelevant chest x-ray images (like images of side view of chest) whose features were not exactly contributing to the training of our model were removed manually.
Many chest x-ray images had blue tints and red marks. Color images are complicated and harder to process. So grey scaling was performed on the image data set. Grey scaled black and white images are comparatively easier to process and are sufficient for many tasks.

4. ***Data Augmentation:***

Accuracy of any machine learning or deep learning model significantly depends on training data set. The performance of the model often improves with available data. If we have limited number of images, then our model will not be able to produce accurate predictions. Initially we had around 60,000 images in our training data set, but it was not sufficient. To increase accuracy from this limited data we implemented data augmentation technique. Data augmentation is a technique to artificially create hey a new data set from the existing data set. It creates diversity in the existing data set. There are various data augmentation techniques for example cropping, padding, flipping etc. In our model we have used horizontal flipping that is mirror images of the existing data set to increase the training data.

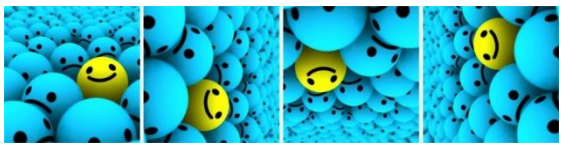

Fig. 2: Data Augmentation

5. ***Encoding Categorical Column***

The performance of a machine learning model not only depends on the model and the hyperparameters but also on how we process and feed different types of variables to the model. Since most machine learning models only accept numerical variables, pre-processing the categorical variables becomes a necessary step. We need to convert these categorical variables to numbers such that the model can understand and extract valuable information.

One hot encoding:
We use this categorical data encoding technique when the features are nominal (do not have any order). In one hot encoding, for each level of a categorical feature, we create a new variable. Each category is mapped with a binary variable containing either 0 or 1. Here, 0 represents the absence, and 1 represents the presence of that category.

For example, in our model we have encoded:
1. COVID-19 Patients: 1 0 0
2. Pneumatic Patients: 0 1 0
3. Healthy Patients: 0 0 1

6. ***Creating the model***

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 5, Issue 12, ISSN No. 2455-2143, Pages 242-246
Published Online April 2021 in IJEAST (http://www.ijeast.com)

We build a sequential model, From the definition of **Keras** documentation the Sequential model is a **linear stack of layers**. You can create a Sequential model by passing a list of layer instances to the constructor.

7. *Description of Layers*

1. Convolutional Layer

Convolutional layers are the major building blocks used in convolutional neural networks. A convolution is the simple application of a filter to an input that results in an activation.

In our project we have used two convolutional layers. First layer has 32 filters whereas second layer has 64 filters.

Repeated application of the same filter to an input results in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input, such as an image. The innovation of convolutional neural networks is the ability to automatically learn many filters in parallel specific to a training dataset under the constraints of a specific predictive modelling problem, such as image classification. The result is *highly* specific features that can be detected anywhere on input image.

2. Max Pooling 2-D Layer

The above two convolutional layers are separated by one max pooling layer. Max pooling operation for 2D spatial data. Down samples the input representation by taking the maximum value over the window defined by pool_size for each dimension along the feature's axis. The window is shifted by strides in each dimension.

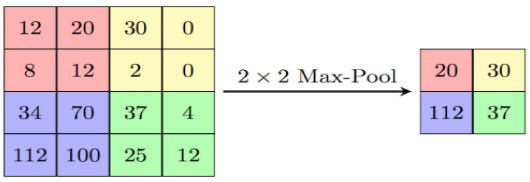

Fig. 3: Max Pooling

3. Feature Selection

Feature selection in machine learning refers to the process of choosing the most relevant features in our data to give to our model. By limiting the number of features we use (rather than just feeding the model the unmodified data), we can often speed up training and improve accuracy, or both.

4. Activation Function

Activation functions are an extremely important feature of the artificial neural networks. They basically decide whether a neuron should be activated or not. Whether the information that the neuron is receiving is relevant for the given information or should it be ignored. The activation function is the nonlinear transformation that we do over the input signal. This transformed output is then sent to the next layer of neurons as input. We have used ReLU. ReLU is the most widely used activation function while designing networks today. The ReLU function is nonlinear, which means we can easily backpropagate the errors and have multiple layers of neurons being activated by the ReLU function.

The main advantage of using the ReLU function over other activation functions is that it does not activate all the neurons at the same time. If you look at the ReLU function if the input is negative, it will convert it to zero and the neuron does not get activated. This means that at a time only a few neurons are activated making the network sparse making it efficient and easy for computation.

5. Dropout Layer

The Dropout layer is a mask that nullifies the contribution of some neurons towards the next layer and leaves unmodified all others. We can apply a Dropout layer to the input vector, in which case it nullifies some of its features; but we can also apply it to a hidden layer, in which case it nullifies some hidden neurons.

Dropout layers are important in training CNNs because they prevent overfitting on the training data. If they aren't present, the first batch of training samples influences the learning in a disproportionately high manner. This, in turn, would prevent the learning of features that appear only in later samples or batches.

6. Dense Layers

We have used three dense layers. A densely connected layer provides learning features from all the combinations of the features of the previous layer, whereas a convolutional layer relies on consistent features with a small repetitive field.

After working the model for all the above layers, the model is tested and trained to predict labels for images of the x-rays and classify them into covid-19, pneumatic or healthy patient.

IV. Application Output

We have created a web application for ease of users. This web application was developed using Django. Django is a high-level Python web framework that enables rapid development of secure and maintainable websites. With the help of this

application, uploaded images of chest x-rays can be easily classified into one of the three specified categories.

Following are some snapshots of the web application:

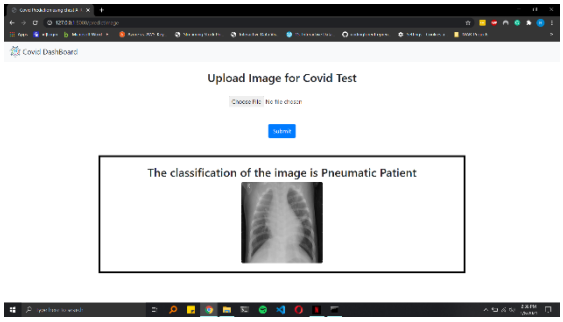

Fig. 4: Pneumatic Patient Classification

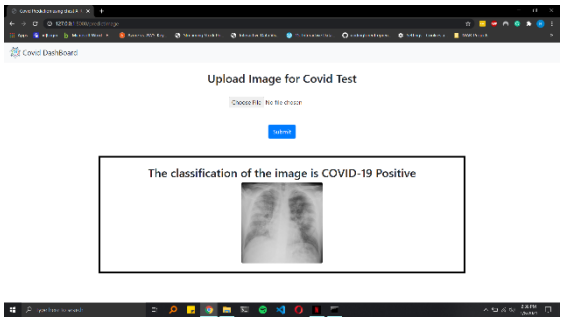

Fig. 5: COVID-19 Patient Classification

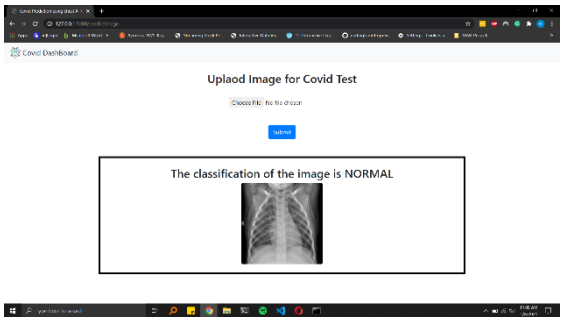

Fig. 6: Normal Patient Classification

V. Future Scope

This model can prove to be extremely useful in near future considering this to be an efficient and hassle-free technique to detect the disease. We have built a web-based application using Django to deploy the model for public use.
This model and be worked upon and can also be deployed as android or iOS application for mobile phones making the model more and more easy to access, thus making the overall experience highly user friendly.

VI. Conclusion

Our project "Analyzing chest x-rays to differentiate the symptoms of covid-19 using Machine learning approach" can be considered a great aid to the society in times of this pandemic. The aim was to build something quicker and faster to detect the disease and reduce crowding in the test centers and we all think this can be achieved with this project. The CNN based model has been trained to the accuracy of 95% and we continue to work on improving this further so that the model achieves its maximum potential.

VII. Acknowledgement

Each and every one of us is overwhelmingly thankful to all the teachers and mentors who helped us create this project. Not only the guidance, but the overall learning experience from the professors helped us learn and understand a lot of things.
Our guide Prof Dr. Deepali Deshpande Madam helped us through the thick and thins of our project, her opinions and guidance was the driving force to all of us. Our mentor Prof Mrs. Deepali Joshi madam's suggestions and useful technical insights help us solve our problems whenever we got stuck, madam's expertise in the field helped us learn a lot new things and bought us to successful completion of our project.

VIII. REFERENCE

[1] M. Irfan Uddin, Syed Atif Ali Shah, Mahmoud Ahmad Al-Khasawneh. (2020). A Novel Deep Convolutional Neural Network Model to Monitor People following Guidelines to Avoid COVID-19, Hindawi Journal of Sensors Volume 2020, (DOI Number 8856801).

[2] Mohd Zulfaezal Che Azemin, Radhiana Hassan, Mohd Adli Md Al. (2020). COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings, Hindawi Journal of Sensors Volume 2020, (DOI Number 8828855).

[3] Pedro Silva, Eduardo Luz, Guilherme Silva, Gladston Moreira, Rodrigo Silva, Diego Lucio. (2020). *COVID-19 Detection in CT Images with Deep Learning: A Voting-Based Scheme and Cross-Datasets Analysis. Informatics in Medicine Unlocked Volume 20, (*DOI Number 100427*).*

[4] Ahmed Sahlol, Dalia Yousri, Ahmed Ewees, Robertas Damasevicius, Mohamed Elaziz. (2020). COVID-19 Image Classification Using Deep Features and Fractional-order Marine Predators Algorithm. Scientific Reports Volume 10, (DOI Number 15364).

[5] Linda Wang, Zhong Qui Lin, Alexander Wong. (2020). COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-ray Images. Scientific Reports Volume 10, (DOI Number 19549).

[6] Dingding Wang, Jiaqing Mo, Gang Zhou, Liang Xu, Yajun Liu. (2020). An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. Plos One Research Article, (DOI Number 0242535).

[7] Hoon Ko, Heewon Chung, Wu Seong Kang, Kyung Won Kim, Youngbin Shin, Seung Ji Kang, Jae Hoon Lee, Young Jun Kim, Nan Yeol Kim, Hyunseok Jung, Jinseok Lee. (2020). COVID-19 Pneumonia Diagnosis Using a Simple 2D Deep Learning Framework With a Single Chest CT Image: Model Development and Validation, JMIR Publications Volume 22. (DOI Number 19569).

[8] Ali Narin , Ceren Kaya, Ziynet Pamuk. (2020). Department of Electrical and Electronics Engineering, Zonguldak Bulent Ecevit University, Turkey. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Network. (DOI Number 10849).

[9] Antonios Markis, Ioannis Kontopoulos. (2020). COVID-19 Detection from Chest X-Ray Images Using Deep Learning and Convolutional Neural Networks. (DOI Number 20110817) .

[10] Florentin Smarandache, Mohamed Loey, Nour Eldeen M. Khalifa. (2020). Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. (DOI Number 12040651).

[11] Ishan Manandhar, Ahmad Alimadadi, Sachin Aryal. (2020). Artificial intelligence and machine learning to fight COVID-19. Physiological Genomics Volume 52. (PubMed 32216577).

[12] Shuo Wang, Yunfei Zha, Weimin Li, Qingxia Wu, Xiaoh u Li, Meng Niu, Meiyun Wang, Xiaoming Qiu, Hongjun Li, He Yu, Wei Gong, Yan Bai, Li Li, Yongbei Zhu, Lius u Wang, Jie Tian. (2020). A Fully Automatic Deep Learning System for COVID-19 Diagnostic and Prognostic Analysis. European Respiratory Journal Volume 57 (DOI:10.1183/13993003.00775-2020).

[13] Boran Sekeroglu, Ilker Ozsahin. (2020). Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Network. National Library of Medicine. (PMID: 32948098).

IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

ABOUT IJEAST

International Journal of Engineering Applied Science and Technology (IJEAST) is a peer-reviewed, open access journal that publishes high-quality research papers in the field of Engineering, Applied Science and Technology.

IJEAST aims to provide a platform for researchers, academicians, and professionals to share their innovative ideas, research findings, and practical experiences with the global scientific community.

FOCUS AREAS

- Engineering
- Applied Science
- Technology
- Innovation & Development
- Interdisciplinary Studies

PEER REVIEWED
All submissions are rigorously peer reviewed to ensure quality.

OPEN ACCESS
Free and unrestricted access to research for all.

GLOBAL REACH
Connecting researchers and professionals worldwide.

TIMELY PUBLICATION
We ensure a swift and efficient publication process.

For more information, visit our website
www.ijeast.com

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

editor@ijeast.com
www.ijeast.com
India

2455-2143