Covid-19 Detection Using X-Ray Images

Major Project I

Submitted By

Lakshya Khera(9917103014)

Gaurav Parihar(9917103015)

Shubham Tak(9917103031)

Under the Supervision of

Dr. Arti Jain

Department of CSE/IT

Jaypee Institute Of Information Technology University, Noida

October 2020

ACKNOWLEDGEMENT

We would highly like to place on record our deep sense of gratitude to Dr. Arti Jain, Assistant Professor at, Jaypee Institute of Information Technology, Noida for his generous guidance and constant supervision as well as for providing necessary information regarding the project and also his support in completing this work. I express my sincere gratitude to faculty, Dept. of Computer Science, and project evaluators for their stimulating guidance, continuous encouragement and supervision throughout the course of present work. I also extend my thanks to my college for providing adequate resources for the project, my group members and seniors for their insightful comments and constructive suggestions to improve the quality of this project work.

Signature(s) of Students

Lakshya Khera (9917103014)

Gaurav Parihar (9917103015)

Shubham Tak (9917103031)

Signature(s) of Supervisor:

Dr. Arti Jain
Assistant Professor
(Department of CSE & IT,JIIT)

ABSTRACT

The novel coronavirus (COVID-19), known as SARS-CoV-2 previously known as 2019-nCoV causes an illness which affects the respiratory system mainly. It has caused a huge impact on the lives of human lives, health and global economy. The critical problem the world is facing apart from the invention of a vaccine is to detect the presence of coronavirus in the human body in the early stage and prevent the situation from getting worse. To detect the presence of coronavirus the need of tools for diagnosis increased at the fast pace. Recent studies and radiologists specify that the physical examination of the body shows the chest radiographs were found abnormal. Studies also state that chest X-Rays are insensitive in the early stages of the disease. Further if quarantined the patient's X-Rays often reveals changes in the lungs. So, the above findings could be a green signal to make advancement in the field of computer vision, machine learning and deep learning as X-Ray images could help to detect the presence of Coronavirus. Computer Vision is a scientific field which completely focuses on how computers gain a high level of understanding from digitally created images and videos. Computer based analysis for the images have been an issue in the field of medical image analysis. Recent advancement in the field of machine learning and especially deep learning helps to identify and classify patterns in medical images. The proposed model is built keeping in mind to provide accurate results whether a person is Covid-19 positive or negative.

Table of Contents

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In this epidemic of COVID-19, people will continue to suffer until people receive the vaccine. A lot of scientific research is being done continuously. So, until people receive vaccines they should be tested at a correct time so that they can be cured and prevent the spread of coronavirus from that person. Test analysis is often performed in automated, high-throughput, medical laboratories by medical laboratory scientists. Test samples can be obtained by various methods, including a nasopharyngeal swab, sputum (coughed up material), throat swabs, deep airway material collected via suction catheter or saliva.

Reverse transcription polymerase chain reaction (RT-PCR) first uses reverse transcription to obtain DNA, followed by PCR to amplify that DNA, creating enough to be analyzed. RT-PCR can thereby detect SARS-CoV-2, which contains only RNA. The RT-PCR process generally requires a few hours.

This complete process from collection, testing and making of reports takes about a day. The situation is even worse at places that are not equipped with testing laboratories and the biggest problem is contamination of the equipment with which samples are taken.

After going through many articles the world demands a fast, reliable and secure way of testing. Here comes the role of Computer Vision, Machine Learning and especially Deep Learning as it is described as the state-of-the-art for all analysis of medical imaging and plays an important role. Recent studies and radiologists specify that the physical examination of the body shows the chest radiographs were found abnormal as it infects the respiratory system mainly. Studies also state that chest X-Rays are insensitive in the early stages of the disease. Further if quarantined the patient's X-Rays often reveals changes in the lungs. So, the above findings could be a green signal to make advancement in the field of computer vision, machine learning and deep learning as X-Ray images could help to detect the presence of Coronavirus.

Recent advancement in the field of machine learning and especially deep learning helps to identify and classify patterns in medical images. Similarly X-Rays were used to detect COVID-19 but, the models are struggling to cope up with the accuracy, more false positive rates and mandatory characteristics.

So, we will be introducing a new model with better accuracy and which will try to fulfill all the flaws and the model will finally be merged with flask api for making it available for users so that

they can upload their chest X-Ray and automatically find whether he/she is covid positive or negative. The scalability doesn't end here; our application can be integrated with the systems available in the laboratory. This complete process is not costly as all the laboratories are equipped with the mobile x-ray machines which can produce x-ray instantly and our application will classify the image into positive or negative COVID. The presence of patchy and/or confluent, band-like ground glass opacity or consolidation in a peripheral and mid-to-lower lung zone distribution on a chest radiograph. Studies of chest also shows that lung shadowing goes to 69%. Our model will be built keeping all the facts above.

CHAPTER 2: PROBLEM STATEMENT

The novel coronavirus (COVID-19), known as SARS-CoV-2 previously known as 2019-nCoV causes an illness which affects the respiratory system mainly. It has caused a huge impact on the lives of human lives, health and global economy.

A lot of scientific research is being done continuously. So, until people receive vaccines they should be tested at a correct time so that they can be cured and prevent the spread of coronavirus from that person. Test analysis is often performed in automated, high-throughput, medical laboratories by medical laboratory scientists. Test samples can be obtained by various methods, including a nasopharyngeal swab, sputum (coughed up material), throat swabs, deep airway material collected via suction catheter or saliva.

So, to solve the above problem many models and research are being done everyday. When talking in terms of deep learning models, there are many models which build till now but, they all had flaws of not having a large size of dataset and more false positive rate.

Fig 2.1: Complete worldwide cases till 13 October.

Fig 2.2: Complete worldwide death cases till 13 October.

CHAPTER 3: LITERATURE SURVEY

Title: Deep Learning in Medical Image Analysis

Author: Dinggang Shen,Guorong Wu,Heung-Il Suk

Summary:

The computer based analysis for better interpreting images have been in the medical imaging field. In the view of the image interpretation, recent advancement in the field of machine learning and especially deep learning has a support hand to identify and classify patterns in medical images. The hierarchical feature representations from the data, instead of handcrafted features mostly designed based on domain-specific knowledge, lies at the core of the advances.

Deep Learning has proved to be the state-of-the-art in the field of image analysis in the terms of advancement of medical advancement. The decade has witnessed the importance of medical imaging e.g. computed tomography (CT), magnetic resonance (MR), positron emission tomography (PET), mammography, ultrasound, X-ray, and so on, for the early detection, diagnosis, and treatment of diseases. In real life the analysis of the medical images are done by the experts specially radiologists and physicians.

This paper has given overview of various deep learning techniques by giving explanation of DEEP MODELS like:

1. Feed-forward neural networks, Deep Models.
2. Unsupervised feature representation learning-Stacked Auto-Encoder, Deep Belief Network, Deep Boltzmann Machine.

3. Fine-tuning deep models for target tasks.
4. Convolutional neural networks.
5. Reducing overfitting

The explanation gives a great overview how deep learning has advantage. It attracts researchers to work on this field of improving medical imaging to investigate CT, MRI, PET, and X-Ray, etc. The introduction about how deep learning can help in medical imaging for localization, cell structures detection, tissue segmentation, computer aided detection and diagnosis/prognosis.

Title: COVID-19 detection from chest X-Ray images using Deep Learning and Convolutional Neural Networks

Author: Antonios Makris**,** Ioannis Kontopoulos, Konstantinos Tserpes

Summary:

This paper tells about the state-of-the-art that pre-trained convolutional neural networks were evaluated as of their ability to detect infected patients from chest X-Ray images. Due to the small number of samples, they used transfer learning, which transfers knowledge extracted by pre-trained models to the models to the model to be trained. A dataset consisting of 336 X-Ray images was created by mixing available X-ray images from patients with confirmed COVID-19 disease, common bacterial pneumonia and healthy individuals. The experimental results demonstrate that the classification performance can reach an accuracy of 95% for the best two models.

The proposed model in this paper incorporates a large number of CNN architectures in an attempt to not only distinguish X-Rays between COVID-19 patients and people without the disease, but also find the pneumonia patients. The model is completely a classifier for respiratory diseases.

A brief description about Deep learning approaches for COVID-19 detection based on image classification is introduced. The proposed CNN model is based on pre-trained transfer models (ResNet50, InceptionV3 and Inception-ResNetV2). In terms of getting high accuracy from the small dataset sample of X-ray images a transfer learning technique is applied by employing the ImageNet dataset. The result after training and testing was that ResNet50 was found to be superior.

The ResNet18 pre-trained ImageNet network is used and the results showed an accuracy of 95.12% on CXR images. A deep convolutional neural network called COVID-Net showed a high sensitivity (87.1%) and a precision of 96.4% for COVID-19 cases.

The results demonstrated a strong correlation between estimated uncertainty in prediction and classification accuracy, thus enabling false predictions identification. Finally, after the evaluation of the performance of five pre-trained CNN networks regarding the detection of COVID-19 from CXR. The results showed that VGG19 and MobileNetv2 achieved the higher accuracy, 93.48% and 92.85% respectively.

Title: Unveiling COVID-19 from Chest X-ray with deep learning: a hurdles race with small data

Author: Enzo Tartaglione, Carlo Alberto Barbano, Claudio Berzovini, Marco Calandri, and Marco Grangetto.

Summary:

In this work, possible obstacles have been highlighted in successfully training a deep model, ranging from the proper choice of the architecture to-be-trained to handling removable biases in medical datasets. The model consists of certain phases. First pre-processing of chest images and lung segmentation will help in removing any bias present in the data.Deep model will be Pre-trained on the feature extractor.

Extensive experiments show that extracting a “COVID” feature from CXR isn't a simple task. Such a problem should be addressed very carefully: it's very easy to misinterpret excellent results on test-data, still showing poor generalization on new data within the same domain. Such a test can be performed thanks to the possibility of using CORDA, a larger dataset comprising COVID cases. Of course, the quantity of available data is still limited but allowed us to find some promising seminal classification results. The ongoing collection and sharing of an outsized amount of CXR data is the only way to further investigate if promising CNN results can aid within the fight to COVID pandemic.

Title: Efficient Pneumonia Detection in Chest X Ray Images Using Deep Transfer Learning

Author: Mohammad Farukh Hashm, Satyarth Katiyar ,Avinash G Keskar 3 ,Neeraj Dhanraj BokdeandZong Woo Geem

Summary:

This paper described how CT scans can be utilised to diagnose Covid-19 in early stages, the paper also describes best comparative study between multiple implementations. As the dataset was very less so using data augmentation techniques, a larger dataset was created. The chest x-ray images

were passed through the process of scan line optimization such that it eliminates all the other body parts.

Multiple existing models are taken into consideration including Dense CNN, GoogleNet121, ResNet etc.

Following architecture was discussed:

- **Transfer learning:** In transfer learning, a model that is trained for a particular task is employed as the starting point for solving another problem.
- **Data Augmentation:** Duplicating dataset to make sure the model doesn't overfit.
- **Fine tuning the architecture:** SGD(Stochastic Gradient Optimizer) was used to provide better generalisation and models were trained for 25 epochs.

Critical analysis:

- On multiple implementations, it was turned that all models had 99% or more accuracy and 0.03% except the Xception model.
- DenseNet121 turned out to be best among, but using a single model wasn't able to solve the problem completely.
- A weighted matrix was used with multiple models, resulting in each model contributing to 97% accuracy with 0.087% loss.
- Weighted matrix results outweighed others in confusion matrix.
- Following analysis shows that using a deep learning model can provide faster diagnosis.

CHAPTER 4: PROPOSED SOLUTION

Keeping in mind the problem the world is facing we proposed a solution based on deep learning architectures. We will try to implement all available pre-trained models. Our proposed solution has the following features:

1. After going through the various model implementations of COVID-19 detection and analysing the drawbacks we worked on increasing the dataset and we collected 940 X-Rays of both the classes.
2. We will implement the pre-trained models like VGG, ResNet, Inception V3 and Xception and finally observe the overall accuracy of all these models and keep a model with highest accuracy for further prediction of results.
3. To make the working smooth and keeping in mind the scalability of the application we will finally integrate the final model with the UI build on flask.
4. The above features are enough to make the application complete package reliable, smooth and the one with accurate solutions.

CHAPTER 5: DESIGN

The proposed work is carried out in four main steps:

- Raw samples collection(X-Ray images)
- Pre-processing (includes labeling)
- Feature extraction and
- Training & Testing

The below figures gives the overview of the model building process and the steps followed throughout the process and how the complete application will look.

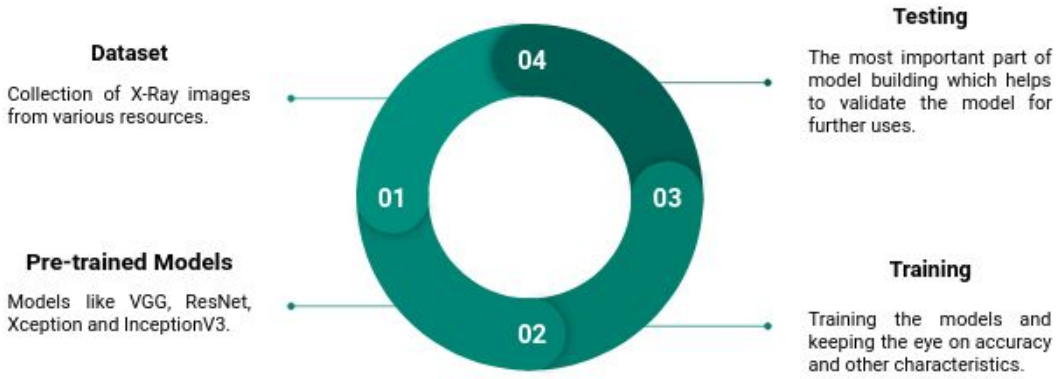

Fig 5.1: Complete model building life cycle of COVID-19 detection model.

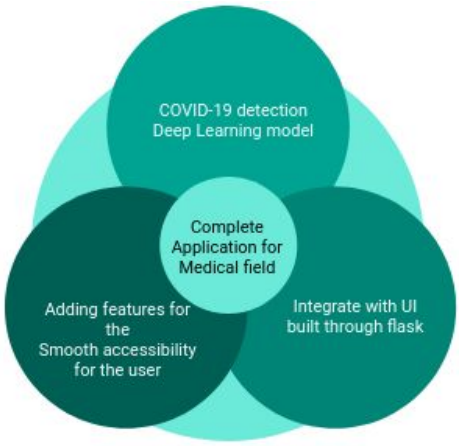

Fig 5.2:Integration of model file with flask based UI.

CHAPTER 6: TOOLS AND TECHNOLOGY USED

Here is a complete list of tools and technology which will be used throughout the project implementation.

- **Flask** : Flask is a micro web framework built on Python. It is designed to make getting started quick and easy, with the ability to scale up to complex applications.
- **Keras** : Keras is an open-source library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library.
- **Seaborn** : Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
- **Tensorflow** : TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks.
- **Matplotlib** : Matplotlib is used for data visualization.
- **Pandas** : Used for creation and manipulation of data and creating data frames.
- **Numpy** : Used for creation and handling the features in an array.
- **Sklearn** : Scikit-learn is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms

CHAPTER 7: IMPLEMENTATION

In order to detect features based on deep learning techniques. Our first step was to choose the best architecture to act on. So from the various research it's important to choose the architecture with good value accuracy, more trainable parameters and maximum depth(layers) to learn new features and come out with the best outcome. From the observations made we choose faster R-CNN which is ensembled with core network architecture ResNet50 and Inception-Res-Net-V3, google's latest inception architecture to detect images and other pre-trained models such as VGG, Xception, etc.

We have implemented the model with ResNet50 and achieved the accuracy of about 82%.

Architecture Design of ResNet50:

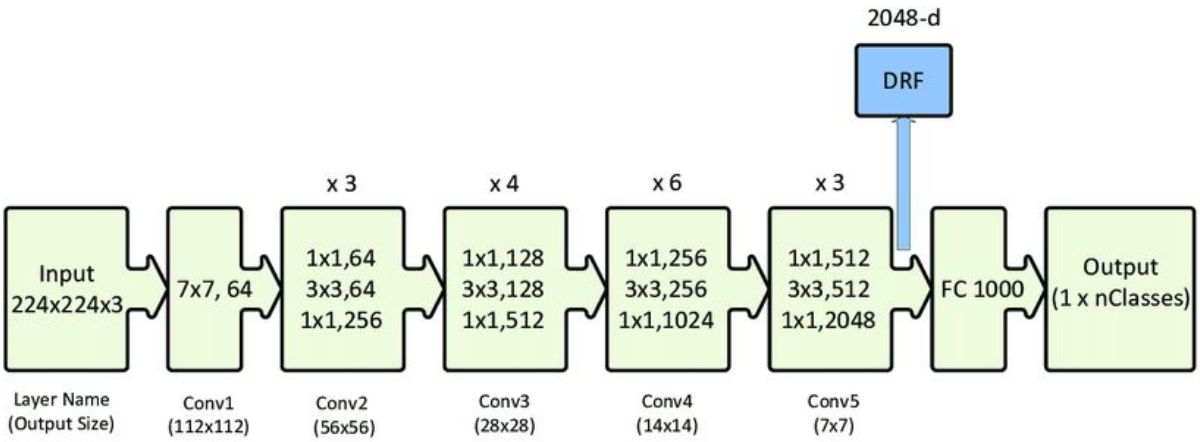

Fig 7.1: Complete Architecture Design of ResNet50.

CHAPTER 8: RESULTS AND ANALYSIS

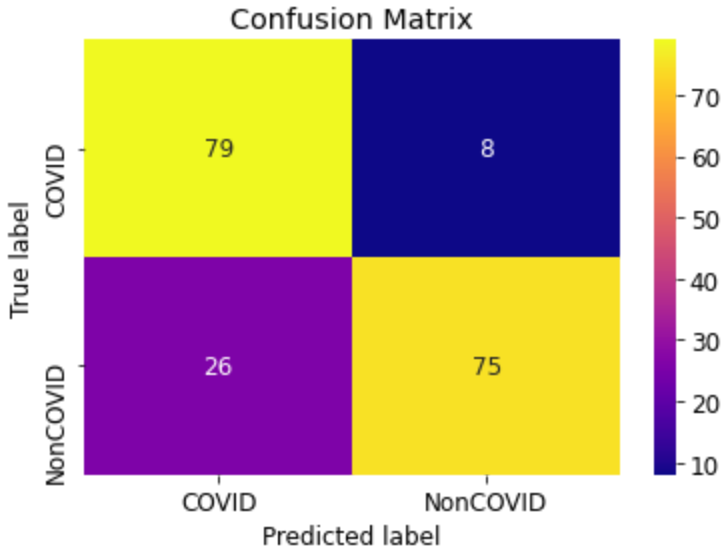

Fig 8.1: Confusion Matrix

	precision	recall	f1-score	support
0	0.75	0.91	0.82	87
1	0.90	0.74	0.82	101
accuracy			0.82	188
macro avg	0.83	0.83	0.82	188
weighted avg	0.83	0.82	0.82	188

Fig 8.2: Complete Classification Report

CHAPTER 9: CONCLUSION

After the complete observation of results we will try to implement other pre-trained models to obtain better and accurate results. If we finds flaws after all this we will make some hyper parameter tunings and try other algorithms too.

CHAPTER 10: REFERENCES

[1] Dinggang Shen,Guorong Wu,Heung-Il Suk " Deep Learning in Medical Image Analysis", Department of Radiology, University of North Carolina at Chapel Hill, Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea, June, 2017.

[2] Lawrence O. Hall, Rahul Paul, Dmitry B. Goldgof, and Gregory M. Goldg, "Finding COVID-19 from Chest X-rays using Deep Learning on a Small Dataset", Department of Computer Science and Engineering, University of South Florida, Dept. of Laboratory Medicine University of California, San Francisco, 2020.

[3]Muhammed Talo, Ozal Yildirima,Ulas Baran Baloglub, Galip Aydinc, U Rajendra Acharya, "Convolutional neural networks for multi-class brain disease detection using MRI images", Department of Computer Engineering, Munzur University, Tunceli, Turkey, f International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto, Japan, October, 2019.

[4]Antonios Makris, Ioannis Kontopoulos, Konstantinos Tserpes, "COVID-19 detection from chest X-Ray images using Deep Learning and Convolutional Neural Networks", Department of Informatics and Telematics Harokopio University of Athens Tavros, Greece, May,2020.

[5]Prabira Kumar Sethy, Santi Kumari Behera, "Detection of coronavirus Disease (COVID-19) based on Deep Features", International Journal of Mathematical, Engineering and Management Sciences, April,2020.

[6]Enzo Tartaglione, Carlo Alberto Barbano, Claudio Berzovini, Marco Calandri, and Marco Grangetto, "Unveiling COVID-19 from Chest X-ray with deep learning: a hurdles race with small data",University of Turin, Computer Science dept., Torino, Italy,University of Turin, Oncology Department, AOU San Luigi Gonzaga, Orbassano, Italy, April,2020.

[7]Arpan Mangal, Surya Kalia, Harish Rajgopal, Krithika Rangarajan, Vinay Namboodiri, Subhashis Banerjee, Chetan Arora," CovidAID: COVID-19 Detection Using Chest X-Ray", Indian Institute of Technology Delhi, Indian Institute of Technology Kanpur, All India Institute of Medical Sciences Delhi, University of Bath, UK, April,2020.

[8]Gianluca Maguolo , Loris Nanni,"A Critic Evaluation of Methods for COVID-19 Automatic Detection from X-Ray Images",2020.

[9]Mohamed Samir Boudrioua, "COVID-19 detection from chest X-ray images using CNNs models: Further evidence from Deep Transfer Learning",Ronin institute, Montclair, USA,2020.

[10]Asmaa Abbas, Mohammed M. Abdelsamea , Mohamed Medhat Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network",Mathematics Department, Faculty of Science, Assiut University, Assiut, Egypt,School of Computing and Digital Technology, Birmingham City University, Birmingham, UK,September, 2020.