Covid-19 detection,analysis using machine learning

Najmul Uddin
BSc in CSE
student of AIUB
Dhaka,Bangladesh
dbnnoman1998536@gmail.com

Mahzabin Mostary
BSc in CSE
student of AIUB
Dhaka,Bangladesh
mahzabinmostary@gmail.com

Hassan KH Salar
BSc in CSE
student of AIUB
Dhaka,Bangladesh
khsalarhassan@gmail.com

***Abstract*—The novel coronavirus 2019 (COVID-19) is considered a major health challenge worldwide due to the rapid transmission of the virus to humans, leading to an increase in the number of people infected with the virus and death. Therefore early detection of COVID-19 is very important to control the spread of the epidemic and reduce mortality. The real-time reverse transcription-polymerase chain reaction, a key diagnostic method for coronavirus infection, has a high negative rate of diagnosis during the first phase of the disease. At the same time, the manifestations of COVID-19, as evidenced by medical imaging modalities such as computed tomography (CT), radiograph (X-ray), and ultrasound imaging, indicate unique features that are different from those of healthy cases or other forms of pneumonia. Mechanical learning (ML) applications for COVID-19 diagnosis, diagnosis, and clinical severity based on medical assumptions have received considerable attention. Here, we review the latest ML advances in the detection of COVID-19 with a special focus on ML models using CT and X-ray images published in high-quality journals, including a discussion of outstanding clinical imaging features in patients with COVID-19. Deep Learning algorithms, especially convolutional neural networks, have been widely used in image classification and classification of patients with COVID-19 and many ML modules have yielded predictable results using data sets with limited sample sizes.**

Keywords— in-depth reading, COVID-19 diagnosis, CT, x-Ray

I. Introduction

Since the end of 2019, the world is facing a pandemic caused by a highly contagious novel coronavirus called acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The newly discovered virus is called COVID-19, which has spread rapidly in 222 countries and territories, resulting in more than 102 million confirmed cases, and 2.2 million deaths on January 31, 2021.COVID-19 patients present with cough, fever, dyspnea, fatigue, and myalgia. However, hemoptysis, chest pain, sputum production, rhinorrhea, headache, sore throat and abdominal cramps are rare symptoms of COVID-19 Although people with the virus have mild or no symptoms, a large number of patients progress rapidly to respiratory failure great high risk of death, especially in the elderly, as well as in people with chronic illnesses, for example, chronic respiratory or cardiovascular disease, diabetes and cancer. Early and accurate diagnosis of COVID-19 has the potential to control the spread of the epidemic and reduce mortality. Reverse-transcription polymerase chain reaction (RT-PCR) of the nose and nasopharyngeal swabs is considered an important clinical diagnostic tool for SARS-CoV-2 infection. However, there are some shortcomings, including limited sensitivity to the early stages of the disease, a time-consuming process, and a lack of RT-PCR kits. At the same time, medical imaging has shown a positive effect on the diagnosis, especially with computed tomography (CT) which can be considered as a key and reliable tool for diagnosing and monitoring COVID-19.In recent years, multidisciplinary (ML) methods have been used successfully in the field of health and medical care to address challenges such as accurate diagnosis and predictor of disease outcomes. Therefore, researchers have made efforts to use ML techniques to help reduce the COVID-19 epidemic. This review aims to discuss the various ML algorithms developed to develop COVID-19 diagnostic and prognostic systems based on clinical modeling. Our contributions to this review include this review provides a list of publicly available data sets for training and testing of ML models. The data sets include collections of COVID-19 clinical chest images, another common pneumonia, and healthy controls. We are evaluating the manifestation of CT in different stages of the disease, and reporting on the main findings of COVID-19 using X-ray images.

II. Motivation

We focus on analyzing advanced ML models for COVID-19 diagnostic tests using imaging techniques (CT scan, X-ray, and ultrasound imaging), and briefly describe the normal operation of the image-based COVID-19 diagnostic system. We present the current ML techniques used in the COVID-19 study for image classification and classification functions, and discuss the general limitations of ML methods in detecting COVID-19 infections. One of our review manuscripts is organized as follows: In Section 2, we list the most popular open source data sets of medical imaging data from healthy samples, COVID-19, and other patients with pneumonia. We then present the highlights of medical imaging during SARS-CoV-2 infection, and highlight recent research that has used ML algorithms in the classification and diagnosis of COVID-19. Areas that have used CT-based, X-ray-based, and ultrasound-based diagnostics are emphasized in Phase 3 while the next Phase 4 introduces COVID-19 complexity studies and mortality predictions using medical imaging. The manuscript provides a brief overview of the limitations of ML COVID-19 diagnostic methods based on imaging.

III. METHODOLOGY

In this section, we present a detailed description of the proposed methodology designed for COVID-19 detection and the data used for the validation of the proposed model

A. Data processing steps

Our proposed forecast model consists of five main steps. The first step is data collection, followed by pre-data analysis that prepares chest X-ray images to extract the element from which the Histogram of oriented gradients (HOG) is extracted (step three). In the fourth step, the HOG element is trained by the dividers to detect the infected state of COVID-19. The final step is to evaluate the effectiveness of the three categories used. First we collect data or image randomly, then preprocessing then extract the feature for classification.

B. Data used

We have used image as a data for this research. we have collected more than 500 x-ray images of in general peoples then classify then by using CNN ,are the affected covid-19 or not?

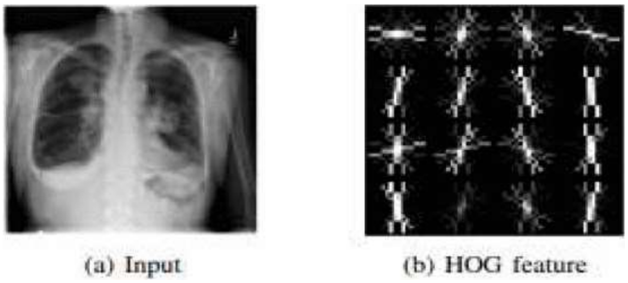

Figure 2:X-ray Image

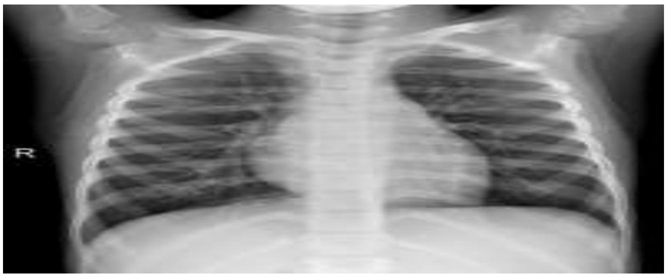

Figure 3: Covid-19 negative X-ray Image

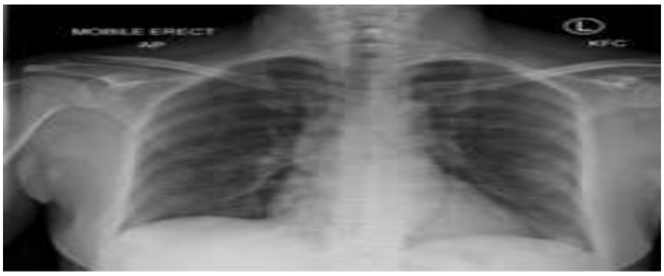

Figure 4: Covid-19 positive X-ray Image

In our study, we prepared three databases of chest X-ray images (COVID-19 patients, ordinary people, and patients with pneumonia).Contains 520 images (120 COVID-19 images, 200 pneumonia images and 200 healthy images). COVID-19 is a new disease, therefore, the number of COVID-19 chest x-ray images is limited. Data sets are collected from different medical. Then, the data sets were split by 80% training and 20% by segment divider. This steps shows an example of a chest X-ray image Data sets.

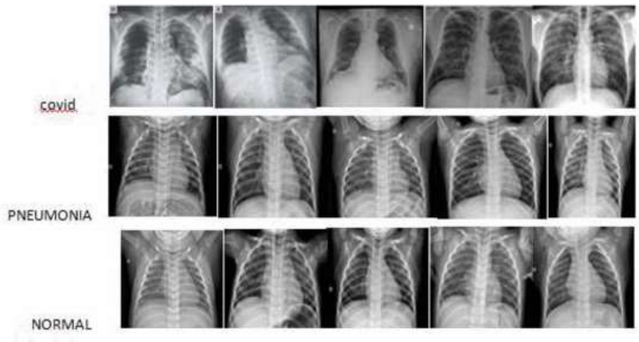

Figure 1: Chest X-ray image datasets

C. Methods

Training a CNN from scratch is tough since it needs a big amount of training data as well as a high level of skill in order to choose the right model architecture for successful convergence. Data is often sparse in medical applications, and expert annotation is costly. Deep CNN training necessitates a lot of computing and memory resources, hence it takes a long time. In the absence of sufficient data, transfer learning (TL) offers a viable option for fine-tuning a CNN that has previously been trained on a large batch of labeled pictures from a different category. This aids in training by accelerating convergence while reducing computing complexity. The early layers of a CNN often learn low-level visual properties that are useful for most vision applications.

After prior processing, the elements are extracted from the images using the Histogram of oriented gradients (HOG). The basic advantages define the shape and properties of the image container. Figure 3 shows an example of a HOG feature extracted from an x-ray image then we use CNN and different algorithm to find optimize accuracy. For categorical data, we use support vector machine

a) Vector Support Machine (SVM)

SVM is a supervised algorithm that is used for both data classification and retrieval challenges. It is often used for data classification problems. SVM uses the kernel strategy as a way to have a linear hyper-plane between classes. The advantages of SVM compared to other algorithms for its speed, efficiency, and accuracy. It is also considered an appropriate method for the division of small databases.

b) Convolutional neural networks (CNNs)

CNN are a subset of deep neural network that is used for in deep learning to analyze and visual images from raw data .We normally think of matrix multiplications but when we think about neural networks this isn't the case with convolutional neural networks. It makes use of a technique known as Convolution. Convolution is a mathematical operation on two functions that makes another third function that shows how one affects the different shape

IV. DECISION VISUALIZATION

In order to get an idea about the COVID-19 detection transparency, we employed the concept of Gradient Class Activation Map (Grad-CAM for detecting the regions where the model paid more attention during the classification. We use CNN for classification and visualization our images Here we uses x-Ray images for a dataset

Source code :https://github.com/dbnnomanbhai/Covid-19-Dataset-Analysis-using-Machine-Learning

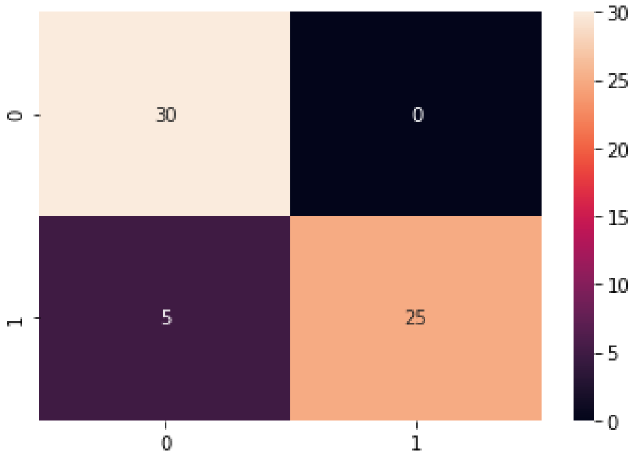

Figure 5: Confusion matrix

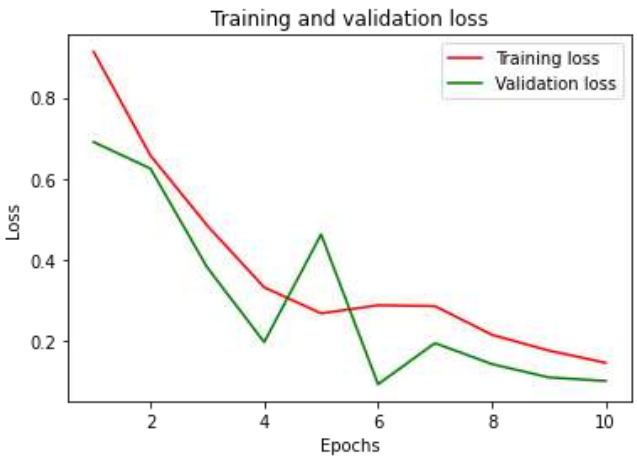

Figure 6: Training and Validation loss.

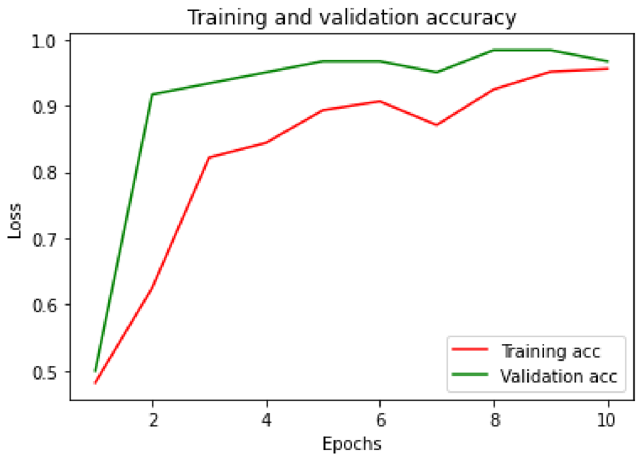

Figure 7: Training and Validation accuracy

V. RESULT

For testing, we randomly split the data twice and conducted the studies twice. For testing and training, the first split comprises 40 photos of 30 COVID-19 patients and 700 images of 502 additional pneumonia patients. The second split includes 50 photographs of 37 COVID-19 patients as well as 617 images of 616 additional pneumonia patients for testing and training. The overall performance is the average of two split. We get 95% accuracy within 7 times approaches.

VI. DISCUSSION AND CONCLUSION

Our X-ray-based model performs similarly to the CT-based screening technique. More importantly, our model only learns from 70 COVID-19 participants, or less than 5% of all COVID-19 subjects. As a result, the suggested model, which employs a chest X-ray, may be deemed a useful computer-aided diagnostic (CAD) tool for low-cost and quick COVID-19 screening. Our model, however, has significant flaws, including missing 4% of COVID-19 instances and a false positive rate of over 36%. Our future efforts will be directed on further lowering the false negative rate and, if possible, the false positive rate. We'll also look into ways to use a chest X-ray to discern COVID-19 severity and therefore identify potentially severe instances for early therapy, which will need additional clinical diagnostic data.

VII. ACKNOWLEDGMENT

The authors express their gratitude to Dr. MM Manjurul Islam for his constructive comments, advice, and, most importantly, his inspiration in completing our research study. The authors appreciate the time and effort put in to gather and share COVID-19 chest X-ray pictures and the ChestX-ray14 dataset for research on low-cost, quick COVID-19 screening using X-rays

VIII. DECLARATION OF CONFLICTING INTERESTS

The authors state that the publishing of this paper does not include any conflicts of interest. For the research, writing, and/or publishing of this paper, Nazmul Uddin, Mahzabin Mostary and KH.Salar Hassan did not receive any funding.

IX. REFERENCES

[1] Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O. and Acharya, U.R., 2020. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in biology and medicine*, *121*, p.103792 .

[2] Ismael, A.M. and Şengür, A., 2021. Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Systems with Applications*, *164*, p.114054.

[3] Hussain, E., Hasan, M., Rahman, M.A., Lee, I., Tamanna, T. and Parvez, M.Z., 2021. CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. *Chaos, Solitons & Fractals*, *142*, p.110495.

[4] Nayak, S.R., Nayak, D.R., Sinha, U., Arora, V. and Pachori, R.B., 2021. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. *Biomedical Signal Processing and Control*, *64*, p.102365.

[5] Jain, G., Mittal, D., Thakur, D. and Mittal, M.K., 2020. A deep learning approach to detect Covid-19 coronavirus with X-Ray images. *Biocybernetics and biomedical engineering*, *40*(4), pp.1391-1405.

[6] Panwar, H., Gupta, P.K., Siddiqui, M.K., Morales-Menendez, R. and Singh, V., 2020. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos, Solitons & Fractals*, *138*, p.109944.

[7] Zhang, J., Xie, Y., Li, Y., Shen, C. and Xia, Y., 2020. Covid-19 screening on chest x-ray images using deep learning based anomaly detection. *arXiv preprint arXiv:2003.12338*, *27*.