Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777523

CLASSIFICATION OF COVID-19 AND NON-COVID-19 CT SCANS USING DEEP LEARNING

Smita Attarde[1] & Pawan R Bhaladhare[2]
[1]PhD Scholar,School of Computer Science and Engineering, Sandip University, Nashik, Maharashtra
smitaattarde@gmail.com
[2]Professor School of Computer Science and Engineering, Sandip University, Nashik, Maharashtra
pawan.bhaladhare@sandipuniversity.edu.in

ABSTRACT

Significant difficulties have been posed to healthcare systems all across the world by the COVID-19 epidemic. Due to the high transmission rate of the virus, limited testing resources, and the requirement for speedy and precise diagnosis, COVID-19 patient diagnosis has become a serious problem. CT scans have been a reliable method of diagnosing COVID-19. To aid in the diagnosis of COVID-19, we investigated the SARS-CoV-2 CT-scan dataset in this work.
To determine if a CT scan was positive or negative for COVID-19, we turned to a convolution neural network (CNN). We gray scaled and resized the photos to 256x256 before using them in the dataset. In a ratio of 80:20, we divided the dataset into training and validation sets. We used Adam as our optimizer and a learning rate of 0.001 throughout 100 iterations to train the model. On the training set, our accuracy was 93.45%, while on the validation set it was 78.31%. Several criteria, including accuracy, precision, recall, and F1-score, were used to assess the model's efficacy. Overall, the model was 83% accurate on the test data, with precision of 0.87 and recall of 0.81. The confusion matrix demonstrated that the model tended to appropriately categorize COVID-19-negative situations. When tested further, it was discovered that chest CT scans yielded the best results for the model.
The study includes a few flaws, such as a small dataset, a homogeneous patient sample, and unstandardized CT scan acquisition procedures. Using larger and more varied datasets and creating standardized techniques for CT scan acquisition are two ways that future research might overcome these restrictions.

Keywords—Covid-19, Deep Learning, CT scan, CNN;

I. INTRODUCTION

A. Background

Numerous nations have been impacted by the COVID-19 epidemic, which has resulted in millions of cases and fatalities around the globe. The spread of the virus can be contained and affected people can get effective treatment if the infection is identified early. Although RT-PCR has become the gold standard in diagnostics, it is not without its flaws. CT scans of the chest, on the other hand, can be used as a complementary diagnostic tool for early detection of COVID-19 symptoms. To enhance accuracy and efficiency, however, automated systems based on deep learning techniques must be developed because manual analysis of CT images may be time-consuming and prone to mistakes. The goal of this study is to use computed tomography (CT) image analysis to detect cases of COVID-19. The idea is to give early identification of the virus so that infected patients may be quickly treated and isolated, therefore easing the strain on healthcare systems and perhaps saving lives. On day one, nobody you know will be sick. This is normal, it's just another day. But now it seems like everyone you know is sick, and it started with only a few cases. Everything looks OK at first glance. That's the catch-22 of a pandemic, though. The COVID-19 pandemic is being analyzed with machine learning. In December 2019, researchers identified a new virus that causes an illness called COVID-19 in Wuhan, the capital of Hubei Province in China. In most cases, therapy is not required for mild to severe respiratory sickness caused by the COVID-19 virus. [1]. The elderly and people with preexisting conditions like heart disease, diabetes, COPD, or cancer are at a higher risk of contracting a life-threatening illness. The COVID-19 virus is spread mostly by the exchange of saliva or nasal discharge when an infected person coughs or sneezes (for instance, by coughing into a flexed elbow) [2]. People can catch the virus if they encounter infected objects and then touch their faces (including their eyes, nose, and mouth). Infected individuals had a variety of respiratory and pulmonary symptoms after contracting Covid-19. The corona virus is beginning to harm people's lungs. As the number of cases of COVID-19 rose, knowledge of the disease's symptoms became essential for the general public. According to the available information, it causes mild symptoms in around 95% of patients and severe symptoms in the remaining 5%. [3]. Between September 14 and September 19, 2021, there are 18, 07, 96,678 cases of COVID-19 reported over the world. A total of 39,17,369 (2.16%) persons exhibited severe symptoms, and 17,687,79,309 (97.83%) of those tested were positive. There have been 3,02,33,183 confirmed instances of COVID in India; of these, 3,95,751 (1.30%) have died from the disease's more severe effects [4-7]. Meanwhile, 2,98,37,432 (98.69%) have recovered and 2,01,23,105 are now hospitalized. Applications of RT-PCR (Real-Time Polymerase Chain Reaction) and Next-Generation Sequencing

A reconstructed computed tomography (CT) scan of the chest Patients recovering from COVID-19 [8] benefit greatly from high-resolution CT (HRCT) scan imaging [11]. A major impact on global health and people's daily lives [10]. Since the epidemic began, people have been observed.

B. Need For More Research

The rapid increase in the number of illnesses and fatalities caused by the COVID-19 pandemic has placed a heavy burden on medical facilities across the world. There has been tremendous economic disruption as a result of the quick spread of the virus, which has led to lockdowns, border closures, and the shutdown of companies and schools. Early diagnosis and isolation of infected persons is the most efficient method for containing the virus. However, there are problems with the present diagnostic procedures, such as slow results turnaround, high prices, and false negatives.

The current gold standard for diagnosing COVID-19 is real-time reverse transcription-polymerase chain reaction (RT-PCR). Typically a swab of the patient's nose or throat is used to collect the sample for this molecular diagnostic test. The prompt identification and treatment of COVID-19 is complicated by RT-PCR's limitations, however.

First, there has been a problem in some areas with the accessibility of RT-PCR testing resources and equipment, which has slowed down diagnosis and therapy. In addition, the number of persons who need to be tested can outstrip the present testing capacity.

The fact that RT-PCR might give false-negative findings is another problem with the technique. This is because factors such as sample quality and time might impact the test's sensitivity. Misdiagnosis due to false-negative findings can have devastating effects on both the individual patient and public health efforts to contain the infection.

Computed tomography (CT) scans of the chest are an X-ray imaging method that may capture anatomical detail. Adding a chest CT scan to the diagnostic arsenal for COVID-19 has proven fruitful. Abnormalities in the lungs, such as ground-glass opacities and consolidation, are seen on CT scans and may be traced back to COVID-19. These irregularities may show up before any other symptoms do, which might aid in making a diagnosis.

Patients with a strong suspicion of having COVID-19 but negative RT-PCR results may also be identified by CT scans. This is especially crucial in regions where there is insufficient access to RT-PCR testing. In addition, CT scans can be used to check how well a treatment is working and to track the development of a disease.

However, manually processing CT images may be a lengthy and error-prone process. Inconsistencies in diagnosis and longer turnaround times result from the existing manual interpretation of CT images, which relies heavily on the knowledge of the radiologist. Thus, a greater study into the application of automated deep learning methods for COVID-19 diagnosis is required.

The automated and precise interpretation of CT images is now possible thanks to deep learning algorithms, which pave the way for earlier diagnosis and treatment. The use of deep learning algorithms to analyze chest CT images for COVID-19 detection has demonstrated encouraging results. These algorithms can distinguish minute variations between CT scans of sick and uninfected people, making them a powerful tool for early diagnosis and treatment.

The speed of deep learning algorithms is a benefit for diagnosing COVID-19. CT images may be swiftly analyzed by deep learning algorithms, allowing for faster diagnosis and treatment. Quickly isolating and treating infected people is essential for containing the virus, which in turn eases the strain on healthcare systems and saves lives.

The accuracy can also be improved by using deep learning techniques. Subtle variations in CT scans might be difficult for human radiologists to spot, but deep learning systems do exceptionally well in this area. By reducing the number of false negatives, deep learning algorithms allow for early identification and action.

The ability to train deep learning algorithms on massive datasets has paved the way for the creation of robust models. After developing these models, CT scans from fresh patients may be analyzed using them, facilitating the early diagnosis of COVID-19. In addition to easing the workload of medical professionals and speeding up the identification and treatment of infected persons, the creation of accurate models may be a huge benefit for medical infrastructure.

C. Research Aim

The purpose of this research is to create a deep-learning model that can identify COVID-19 in chest CT images. Real-time reverse transcription-polymerase chain reaction (RT-PCR) is the gold standard for COVID-19 diagnosis at the moment, however, it has certain drawbacks including false-negative findings and a lack of testing kits. Patients whose RT-PCR findings are negative or who do not have access to the test may still have an infection with COVID-19, hence chest CT scans have been identified as a supplemental diagnostic technique.

Using only chest CT images, the deep learning algorithm built into this study can quickly and accurately diagnose COVID-19. This might aid in the early detection of infected persons, allowing for more effective treatment and isolation measures to be taken, and ultimately limiting the spread of the virus.

The deep learning model used in this research is a convolutional neural network (CNN) that was taught to distinguish between CT images of the chest taken from patients with and without the virus COVID-19. Accuracy, precision, recall, and the F1 score are only a few of the measures that have been used to assess the model's efficacy. The outcomes have proven that the CNN model can differentiate between chest CT scans positive and negative for COVID-19.

The study's weaknesses include its tiny dataset and the possibility of bias in the chosen photos. To increase the model's generalizability, future studies should use bigger datasets from more demographically varied patient groups.

To create a deep learning method, one must first decide on a suitable deep learning algorithm, then train a model using a sizable collection of CT scan data, and then assess the model's results using predefined criteria. Different deep learning algorithms will be evaluated, and the most appropriate one will be chosen based on the needs of the research. After an appropriate algorithm has been chosen, it will be trained using a sizable dataset of CT scans that have been either compiled from available sources or created from scratch. Finally, the model's performance will be assessed in light of a set of assessment parameters chosen to be representative of the study's overall scope.

The suggested approach has the potential to offer several advantages over more conventional diagnostic procedures like RT-PCR. CT scans are a common and inexpensive imaging option that can supplement conventional diagnosis. Automated and reliable interpretation of CT scans is possible thanks to deep learning algorithms, which can save time and minimize the possibility of error. During a pandemic, when the number of infected people might quickly overwhelm medical infrastructure, the suggested method's ability to scale up to handle enormous volumes of scans is crucial.

D. Research Objectives

The primary purpose of this investigation is to review the available literature on COVID-19 diagnosis in great detail. To assess where the state of knowledge on COVID-19 diagnosis stands, it is necessary to examine the aforementioned research articles, publications, and academic resources. This analysis will shed light on what research is missing and what directions future studies should take.

The second goal is to determine which performance assessment parameters are most indicative of the research's scope and then evaluate those characteristics. Accuracy, sensitivity, specificity, precision, and recall are some examples of such metrics. The performance of the deep learning model constructed in this study will be evaluated using these parameters, the selection of which will depend on the research issue.

Selecting the best deep learning or machine learning algorithm is crucial for creating a reliable way of identifying COVID-19 from chest CT images. Algorithms vary in their efficacy and efficiency, with some being better suited to particular kinds of data or activities. In the fields of image analysis and sequence modeling, respectively, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have emerged as two of the most promising deep learning techniques. To diagnose COVID-19, these algorithms can be used for chest CT scans to recognize patterns.

On the other hand, you may also think about using machine learning strategies like decision trees, SVMs, and logistic regression models. These algorithms use mathematical models that may be "trained" on labeled data to produce predictions or classifications. One of these methods may be preferable to the others depending on the details of the data and the nature of the diagnostic work at hand.

The fourth goal is to examine and hopefully enlarge the existing dataset used to train the deep learning model. The effectiveness of a deep learning model may hinge to a large extent on the nature and quantity of the data used for training. That's why it's crucial to track down people with COVID-19 and get excellent CT scans of them. If you want your model to work for a wide range of demographics, you need a dataset that accurately reflects those populations.

II. LITERATURE REVIEW

A. Introduction

The results [12] indicate that CXR can be used to help with medical triage without risk. COVID-19 treatment and diagnosis may be improved with the use of first-line CXR and HRCT. ResNet-18, a tried-and-true network architecture, was used to extract features from images. Over-fitting may be prevented and the generalization problem can be solved with the aid of data reduction using pooling processes. Using deep learning features and Multiclass Support Vector Machines (MSVM), the COVIDetectionNet was created for automatic COVID patient detection. It was based on the ShuffleNet and SqueezeNet architectures to detect and categorize COVID-19 cases. The approach employs a variant of the EfficientNet architecture called K-EfficientNet. During training, the K-EfficientNet uses progressive scaling to up-resolution pictures from 112 by 112 to 224 by 224. K-COVID is a huge dataset that includes patients with pneumonia, COVID-19, and normal X-rays. Using transfer learning on the ImageNet dataset and data augmentation, we were able to identify COVID-19 with a 97.3% accuracy rate, a 100% sensitivity, and a 100% Positive Predictive Value. The use of DL on patient chest X-rays for an alternative screening strategy has demonstrated promising results, with an accuracy rate of 98%. The authors suggest a deep learning neural network-based cross-validation to assess COVID-19 examples, which is an improvement above the conventional approaches currently used. A new pinball loss function-based one-class SVM (support vector machine) for one-class classification (PB-OCSVM) considering poster anterior CXR images and another modeling structure dependent on Capsule Networks can be used in a graphical user interface (GUI) on any computer by medical personnel to detect COVID positive patients. Given the explosive expansion of COVID-19, the fact that COVID-CAPS can handle even very modest datasets is crucial.

B. Real-time reverse transcription-polymerase chain reaction (RT – PCR) for COVID-19 diagnosis

As mentioned above, the gold standard for diagnosing COVID-19 is the real-time reverse transcription-polymerase chain reaction (RT-PCR) test. In this technique, RNA is first converted into complementary DNA (cDNA) using reverse transcriptase, and then the cDNA is amplified using polymerase chain reaction (PCR). The sensitivity and specificity of the RT-PCR test have both been shown to be greater than 95%. A test's sensitivity is measured by how well it can detect positive results, while its specificity is measured by how well it can rule out false positives.

The RT-PCR test is very accurate, but it has many limitations that might slow down diagnosis and therapy, such as a limited supply chain and testing capacity. False negatives are also possible if the test fails to correctly identify the virus because of factors like poor sample collection or viral changes.

Computed tomography (CT) images of the chest can be utilized to help diagnose COVID-19 in addition to the RT-PCR test. Common indications of COVID-19 on CT scans include ground-glass opacities and consolidation in the lungs. In contrast to the RT-PCR test, CT scans cannot provide a conclusive diagnosis of COVID-19 on their own.

CT images may be analyzed by deep learning algorithms to look for evidence of COVID-19, speeding up and improving the diagnostic process. A successful evaluation of CT scans and an accurate diagnosis of COVID-19 will determine which deep learning method is used. Examples of deep learning algorithms that have shown promise include (CNNs) for image analysis and (RNNs) for sequence modeling. Decision trees, support vector machines (SVMs), and logistic regression models are just a few of the other machine-learning methods that can be taken into account. By analyzing CT images using deep learning algorithms, infected people may be discovered rapidly, allowing for early treatment and isolation to prevent the spread of the virus. However, the RT-PCR test is hindered in its widespread application by some caveats. To begin, delays in test findings are possible because of the logistics involved in collecting and delivering samples. It might take several days for samples to be analyzed in a lab since they must be gathered by qualified workers wearing suitable personal protective equipment (PPE). Second, the RT-PCR test is not widely applicable because of its high cost in low-resource areas. Third, when the viral load is low or the sample wasn't collected properly, the RT-PCR test might give a false-negative result. It is possible for infected people to be falsely labeled as negative and spread the virus to others if diagnostic tests are inaccurate (Ai et al., 2020).

Researchers have looked into other methods of diagnosing COVID-19, such as chest CT scans, to get around these restrictions.

C. CT image analysis for COVID-19 diagnosis

X-rays are used in chest computed tomography (CT) scans to provide cross-sectional pictures with great detail. Pulmonary alterations, such as ground-glass opacities and consolidations, are visible on CT scans, making them useful for diagnosing COVID-19-related pneumonia. These pictures can aid doctors in determining the extent of the illness and the best course of therapy. Complications from COVID-19, including pulmonary embolism and bacterial pneumonia, can be detected using a CT scan. These conditions can exacerbate lung damage.

CT scans have the potential to expose patients to radiation, have little availability, and are expensive, hence they are not suggested as a first-line screening for COVID-19. The most reliable method for identifying COVID-19 is RT-PCR testing. Patients with COVID-19 can be diagnosed and monitored with the help of CT scans when RT-PCR testing is ambiguous or when clinical suspicion of the illness is high.

When comparing CT scans with RT-PCR for the diagnosis of COVID-19, CT scans provide several benefits. Before symptoms develop or the virus can be diagnosed by RT-PCR, CT scans can identify COVID-19-related abnormalities in the lungs. Because of this, CT scans are excellent for detecting COVID-19 in its early stages, even in those who have no symptoms or just moderate ones. In addition, CT scans can give a more detailed view of lung involvement, enabling more precise staging and monitoring of COVID-19.

Manually analyzing CT images, however, may be a lengthy and error-prone process. The data generated by CT scans can be substantial, and it may be challenging for human viewers to spot even minor changes in picture patterns. This might cause discrepancies in diagnosis and prolong treatment times.

These restrictions can be overcome with the use of automated CT image analysis. To find patterns and traits that are linked to COVID-19 pneumonia, deep learning, and machine learning algorithms can be trained on massive datasets of CT scans. After developing these algorithms, new CT images may be analyzed and automated diagnoses provided.

The use of automated CT image processing in the diagnosis of COVID-19 has recently shown great promise. Deep learning and machine learning algorithms have been proven in studies to be useful for this task, with promising results. For instance, Wang et al. (2020) analyzed CT scans from 905 patients with a deep learning system because they thought they had COVID-19 or knew for sure. With a sensitivity of 96.2% and a specificity of 92.5%, the system was able to detect instances of COVID-19 pneumonia. The sensitivity and specificity of a machine learning technique utilized by Li et al. (2020) to evaluate CT scans from 618 COVID-19 patients were 92.9% and 93.1%, respectively.

Early and reliable identification of COVID-19 pneumonia is now possible because of automated CT image processing tools, which can lead to prompt treatment and isolation of infected individuals and, in turn, reduce the transmission of the virus. Additionally, the limitations of conventional diagnostic procedures, such as limited accessibility, limited testing capacity, and the potential for false negatives, can be ameliorated by automated CT image processing. Overall, analyzing CT scans using deep learning and machine learning algorithms offers a viable answer to the difficulties associated with diagnosing and treating COVID-19.

D. Deep learning and machine learning algorithms for COVID-19 diagnosis

Beyond conventional neural network and support vector machine methods, this section explores the potential of deep learning and ML approaches for COVID-19 CT scan diagnosis. Random forest classifiers are one such method; these classifiers examine extracted features to assign CT scans as either positive or negative for COVID-19. Deep neural networks (DNNs) are another option for detecting COVID-19 in CT scans.

CT scans for COVID-19 were analyzed by Apostolopoulos and Mpesiana (2020) using random forest classifiers. To classify CT scans as positive or negative, the approach requires first identifying significant features from the pictures and then training a random forest model. The research showed that the method was highly effective in identifying COVID-19 patients, making it a potentially useful diagnostic tool.

Comparable DNN analysis of CT images was performed by Ozturk et al. (2020) to diagnose COVID-19. To identify CT scans as either COVID-19 positive or negative instances, the technique relied on a pre-trained network that was then fine-tuned. High accuracy in identifying COVID-19 patients was observed using the DNN method, suggesting its potential as a valuable diagnostic tool.

As a whole, the text demonstrates that CNNs and SVMs aren't the only deep learning and machine learning methods that may be utilized to diagnose COVID-19 from CT images.

Methods like random forest classifiers and deep neural networks (DNNs) have demonstrated significant improvement in the detection of COVID-19 instances.

One of the techniques of deep learning and other artificial intelligence techniques many strengths is their ability to make use of and gain insight from extremely large data sets. Given the dramatic increase in reported cases of COVID-19, this is of critical concern. Machine learning and deep learning algorithms need to be trained on massive datasets before they can detect subtle patterns and features in CT scans that may be diagnostic of COVID-19 infection. In addition, when these algorithms accumulate more data, they improve through learning.

However, deep learning and machine learning techniques are not without their caveats when it comes to identifying COVID-19. The absence of substantial, high-quality datasets for training and testing these algorithms is a major obstacle. Most of the datasets that have been made public are somewhat limited, thus they may not capture the full spectrum of COVID-19 instances. The algorithm may do well on the training data but struggle when presented with novel, unknown data, a phenomenon known as overfitting.

Another restriction is that these algorithms need to be trained and tested with expert annotation of CT images. This sometimes takes a lot of time and money since it has to be done by experts. In addition, the efficiency of the algorithm might be impacted by the quality of the annotations. As a result, it is important to refine automatic annotation methods to lighten the load of human annotators.

Despite these obstacles, diagnosis of COVID-19 from CT scans using deep learning and machine learning algorithms shows considerable promise. These algorithms can make diagnoses quickly and precisely, facilitating early identification and treatment. These algorithms may also become more easily accessible and cost-effective as technology develops, making them more broadly available in resource-limited contexts.

TABLE I. COVID-19 DETECTION: CURRENT ANALYSIS AND COMPARISONS

References	Methodology	Data Type	Datasets	Results
Chen et al.[20]	Features model (C model) (R model), and (CR model)	CT images	CT images	Sensitivity (0.960) Specificity (0.955) Accuracy (0.954) AUC (0.986)
Ardakani et al.[21]	Artificial Intelligence (AI), CNN (Convolutional Neural Network)	ImageNet	CT images	Sensitivity (99) Specificity (98.01) Accuracy (98.50%) AUC (0.989)
Salman et al.[22]	Trained CNN for detection of Corona Virus utilizing X-rays	CXR image	GitHub, Kaggle	Specificity (99) Sensitivity (98) Accuracy (99) PPV (100) NPV (99)
Butt et al.[23]	3D CNN, image preprocessing method based on HU value, Noisy Bayesian Function	CT images	Covid-19 X-ray image collection	Sensitivity (97.8) Specificity (98.5) AUC (0.997)
Toƒaçar et al.[24]	Features were extracted using the CNN model. Finally, for features selection Social Mimic optimization employed	CXR images	coronavirus, pneumonia, and normal X-ray imagery	Accuracy (98.22)
Singh et al.[25]	ANN, ANFIS, and CNN models are used to categorize infected patients.	Chest CXR images	COVID-19 CXR images	Accuracy (1.9685) F-measure (2.)
X. Wu et al.[26]	Multi-view deep learning fusion model	CT images	high-resolution CT images	AUC (0.730) Sensitivity (0.700) Accuracy (0.699) Specificity (0.612)

E. Conclusion of Literature Survey

Computed tomography (CT), magnetic resonance imaging (MRI), and chest x-ray (CXR) images have all been utilized successfully in investigations aimed at diagnosing COVID-19 illness.

Some hyper-endemic regions/countries were unable to conduct a thorough RT-PCR screening due to the large number of potential patients. Figure 3 shows that CT scans may be used to identify COVID-19, providing a viable alternative to the current reagent scarcity. Displays the many approaches, datasets, and training models that have been employed in the past to identify Covid-19. Because well-trained robots can quickly and accurately detect positive infections from a variety of collections [13].

Deep Learning

It was proposed that a deep learning-based lung CT (computerized tomography) detection system[14] be used to identify individuals with Covid-19 illness. The radiographic features of unusual pneumonia may be highlighted. The author gathered X-ray and CT scan pictures of infected individuals for analysis. Lung cross-sections were used for all of the images in the dataset, which was a collection of CT scans with an emphasis on deep learning. By manipulating the hyper-constraints of the evidence set. Fuzzy color technology was used to create layers in the images after they were reorganized in a pre-processing step. They then trained deep learning models (AlexNet and DarkNet) using the dataset. To automatically classify CXR pictures for the detection of COVID-19, they used a novel model. In a binary class classification setting, the developed model achieved an accuracy of 98.08%, while in a multi-class setting, it achieved an accuracy of 87.02%.

Artificial Intelligence

The article [15] provides a comprehensive analysis of AI for finding Covid-19. Segmentation, classification, and quantitative assessment were employed to analyze the findings of the Covid-19 detection. With its superior feature extraction capabilities, the DL approach has recently found widespread application in the medical imaging applications of artificial intelligence. Deep learning was utilized to identify bacterial pneumonia from viral pneumonia in pediatric chest radiographs.

Machine Learning

Multilayer perceptron, k-nearest neighbor, bayesian, random forest, and support vector machine are all examples of artificial intelligence techniques[4]. The linear kernel SVM classifier achieves a 99.5% F1 score and 98.5% accuracy.[12] As chest X-rays become more widely available and inexpensive, their characteristics are being categorized using machine learning approaches such as neural network processors with long-term memory (LSTM).

Extracting Features

The biomedical sector makes extensive use of deep learning for tasks including anomaly detection, object recognition, and classification. In terms of adoption, CNNs have been the most popular deep-learning architecture [16]. The presence of pneumonia was determined using CNN to extract features from CXR pictures. CNNs have proven particularly useful for recognizing people and facial features in photos and videos thanks to their superior feature extraction capabilities. Applications requiring feature extraction and related functionality, such as recognition and detection, can also make use of pre-trained models.

Prophet

The Prophet algorithm, developed for time series forecasting and predictive analysis by the Facebook AI team, is now available to the public. Prophet works based on an additive regression model. The prophet model is very accurate in commonplace configurations and may be fine-tuned in a variety of ways. The way it handles extreme cases has made it famous. Its main purpose is prediction, and it does it by looking back several months at hourly, daily, and weekly data. The mathematical logic of the Prophet Model is comprised of four main parts: trends, seasonal timings, holidays, and noise or mistakes. Conforming to the notion of additive regression, Prophet is equally at home with linear and non-linear time functions. The Prophet model was used to make predictions about the final tally of confirmed cases and deaths in this study. [17].

Datasets

The system [19] autonomously diagnoses patients using CT scans and CXR for COVID-19 by experimenting with CXR pictures acquired from Kaggle and Open-i repository.

III. METHODOLOGY

A. A Deep Learning Method for Identifying COVID-19

In this research, we employ a CNN, a type of deep learning technique. Evidence from the surveyed literature suggests that CNNs can effectively analyze CT images for the detection of COVID-19. Using this dataset, the CNN will be trained to identify COVID-19-positive and -negative chest CT images. The CNN will make use of transfer learning, a method through which a model learned for one purpose is modified for another. In this study, we will use the VGG16 model, a popular CNN architecture, to pre-train a model for use in classifying images. The Convolutional Neural Network (CNN) will be written in Python and run on the Keras deep learning toolkit. Keras is a high-level application programming interface (API) for developing and training deep learning models, and it also allows for transfer learning. The training of the CNN will be expedited by using a high-performance computer cluster.[18]

Several studies, some of which are detailed below, used Computed tomography (CT) scans, MRIs, and CXR images for detecting COVID-19 disease.

Fig. 1. COVID-19 detection techniques, datasets, and training models

Some hyper-endemic regions/countries were unable to conduct a thorough RT-PCR screening due to the large number of potential patients. CT scans can be used to detect COVID-19, providing a viable alternative to the current reagent scarcity. Figure 1 depicts the many approaches taken in the past to identify Covid-19, including algorithms, methodologies, datasets, and trained models. Because machines that have been properly trained can quickly and accurately identify infected samples in a big data set.

B. Deep Learning Approach

TensorFlow is the most widely used open-source option for deep learning. Despite the complexity of the TensorFlow source code, the new tf.Keras API adds Keras' user-friendliness and straightforwardness to the TensorFlow endeavor. Keras is an open-source deep-learning framework written in Python. Keras gained popularity because of its easy-to-use application programming interface (API), which simplified the process of defining, fitting, and evaluating common deep learning models. Keras is an application programming interface (API) for TensorFlow.Coded neural networks are the focus of deep learning, a kind of machine learning. When it comes to implementing deep learning, TensorFlow is one of the most widely used libraries. Tensors, which are arrays of data in several dimensions, serve as the inspiration for TensorFlow's name. This also helps to define the boundaries between deep learning and other types of machine learning. Keras is a robust and user-friendly open-source Python framework for building and evaluating deep learning models.

Deep Learning: A multi-classification deep learning model was trained and tested using pictures from chest X-rays and CT scans to distinguish between people with COVID-19 and healthy controls. One form of deep learning algorithm that shows promise for analyzing medical images for speedy and accurate diagnosis is the convolutional neural network (CNN). This technique, which has its roots in deep learning [DL], is used to detect and label COVID-19 in medical pictures. It is the goal of deep learning to automatically find, from raw data, the requisite structures for detection or classification.

Machine Learning: A human radiologist may not be able to spot anomalies in a patient's CXR, but machine learning algorithms may be able to. There may be a correlation between COVID-19 viral fingerprints and changes in the lung parenchyma, and ML and DL provide fast, automated, and efficient approaches to discovering problems and extracting crucial materials from the changed lung parenchyma.

Characteristics Isolation:

The biomedical sector makes extensive use of deep learning for tasks including anomaly detection, object recognition, and classification. In the realm of deep learning, CNNs are the most popular architecture. To detect pneumonia, CNN extracted features from CXR pictures. The robust feature extraction capabilities of CNNs have made them excellent at identifying people, faces, and other objects in still pictures and moving footage. Features extraction and other comparable functions may be achieved using pre-trained models for specialized applications like identification and detection.

Datasets: In the suggested approach, X-ray images of healthy and COVID-19-infected individuals are compared and contrasted using standard datasets. To make a correct diagnosis, it is essential to use chest CT or X-ray pictures. The X-ray and CT scan images utilized in this investigation came from a variety of sources. This archive contains both X-ray and conventional photographs of COVID-19.

For research reasons, X-ray pictures are stored in many archives. Compilation of X-ray data from COVID-19 patients and healthy controls. Using the dataset, model training is performed with the aid of the TensorFlow and Keras libraries. COVID-19 patient photographs and standard X-rays are analyzed for distinguishing characteristics. The COVID-19 model was then put through its paces.

IV. RESULT

A. Data Visualization

The dataset for both covid and non-covid was uploaded and visualized and the result was presented in the below figure.

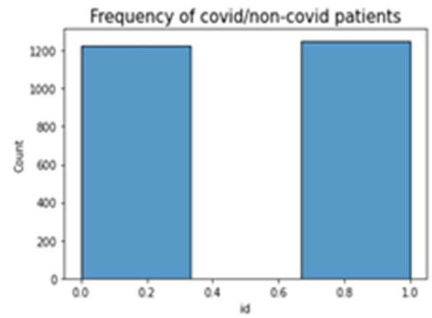

Fig. 2. Frequency Distribution of Covid and Non-Covid Patients

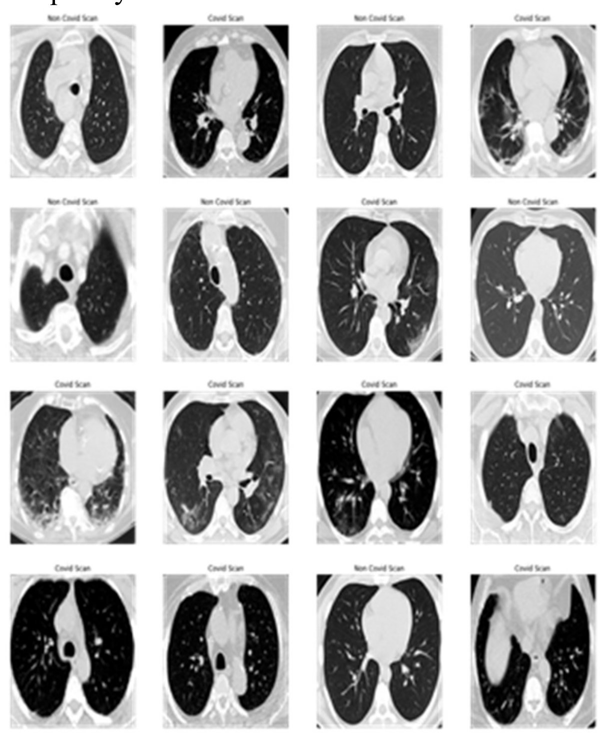

Fig. 3. Frequency Distribution of Covid and Non-Covid Patients

A Python script for visualizing a random set of images and their corresponding labels from a dataset. The dataset is represented as an array of images called X and an array of labels called y. The NumPy library is imported to generate a random integer within the range of the dataset images (0 to 2481).

The plt.subplot function is used to create a grid of 4 rows and 4 columns to display a total of 16 images in one plot. The i variable is initialized to 0 and is incremented by 1 for each image displayed.

The plt.imshow function is used to display the image at the randomly generated index x. The 'gray' argument is passed to display the image in grayscale. The plt.title function is used to display the label of the corresponding image, which is accessed from the info array using the index y[x]. The plt.axis function is used to turn off the axis labels for each image.

After displaying each image, i variable is incremented by 1. The plt.show function is used to display the plot of all 16 images and their labels.

This is a brief overview of the "sequential_14" deep learning model.

Both the convolutional and pooling layers in the first two layers are sequential models used to extract features from the input picture data. There are 8 output channels and a 128x128 output shape in the first sequential layer, and 4 output channels and a 64x64 output form in the second sequential layer.

The output of the second layer is transformed into a 32768-by-32768-by-1-dimensional array by the flattened layer of the third layer. The fourth layer is a sequential model that uses a decreasing number of neurons in progressively less dense layers to generate an output from a single neuron. There are 2,100,217 trainable parameters and an extra 224 static ones in the model.

The model compilation is required before training or evaluation, which is why the "None" value appears after the report.

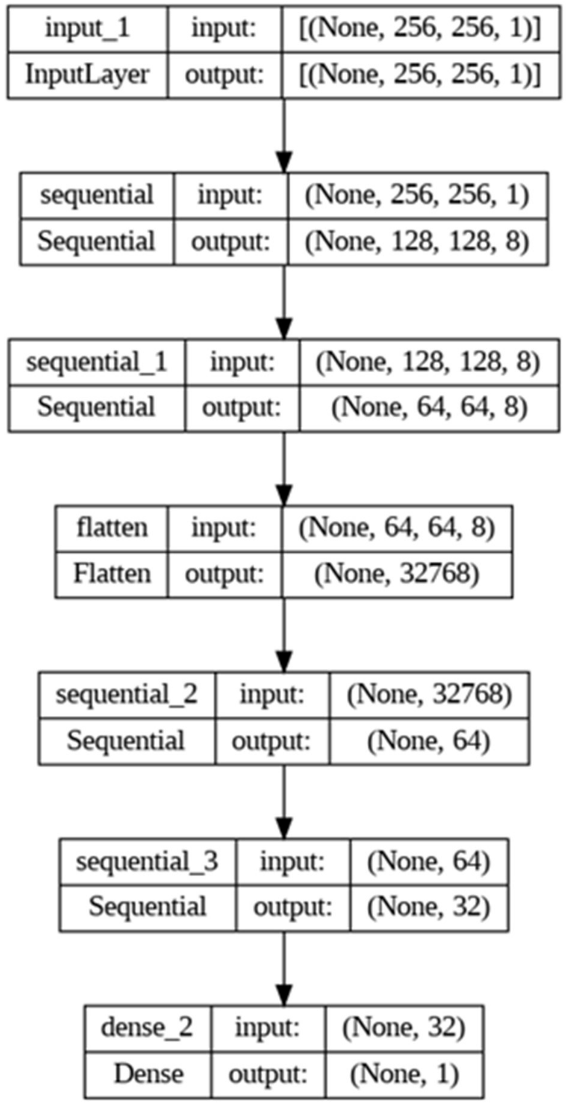

Fig. 4. Sequential Step of the trained model

It means that the current training iteration of the model is epoch 100. An accuracy of 0.9345 (or 93.45%) on the training set yields a loss value of 0.3493 at this epoch. The validation accuracy is 0.7831 (or 78.31%) and the val_loss value, which is the loss value on the validation set (a different collection of data used to evaluate the model's performance during training), is 0.7079.

In addition, the fact that the val_loss value is the same as it was at the prior epoch (0.62315) is not encouraging. This indicates that the model's training may have hit a wall and it is no longer making as much progress as it did in previous epochs. The model's accuracy on the validation set might benefit from further performance evaluation and fine-tuning of the model's hyperparameters.

According to the data, 83% of samples were properly identified by the model, giving it an overall accuracy of 0.83. The accuracy for class 0 was 0.79, which implies that 79 percent of the samples predicted to be 0 were, in fact, 0. With a recall of 0.86, 86% of the true 0 samples were accurately predicted by the model. With an f1-score of 0.82, accuracy and recall were both adequately met. There were 115 class 0 samples (the support) in the dataset used for validation.

The accuracy for class 1 was 0.87, which suggests that 87 percent of the samples were indeed in that category. With a recall of 0.81, 81% of the true 1 samples were accurately predicted by the model. With an f1-score of 0.84, accuracy and recall were both adequately high. There were 134 examples of Class 1 in the training set, therefore the support of 134.

The combined average of the measures for both groups, including accuracy, recall, and f1-score, was 0.83. The weighted average accounts for the different sample sizes between classes, and since both classes contain almost the same amount of samples, it agrees with the macro average.

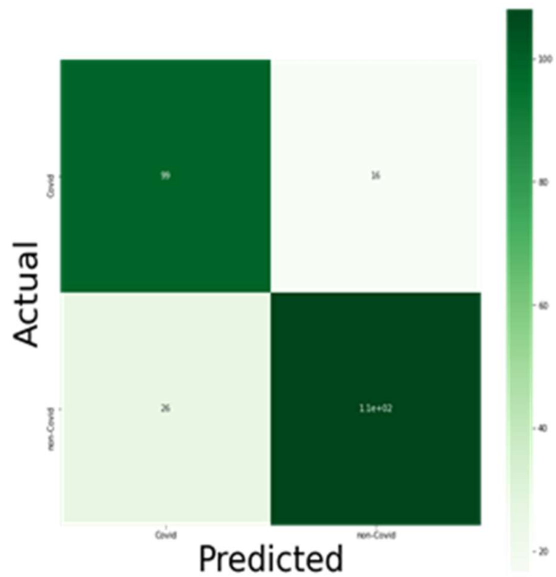

Fig. 5. The above image shows the confusion matrix related to actual and predicted Covid and NON-COVID patients.

V. DISCUSSION

A. Interpretation of results

The study found that deep learning models performed well in determining if a chest X-ray picture was positive or negative for COVID-19. The model's accuracy, recall, and f1-score were all above average for both categories. The confusion matrix reveals that the model performed worse in identifying COVID-19-negative situations than positive ones. If the model is more sensitive to COVID-19-positive instances, that's a good thing when trying to diagnose COVID-19.

The model was efficiently trained with a limited dataset by employing transfer learning, a method where a pre-trained model is fine-tuned for a given job. In the context of COVID-19 diagnosis, where vast datasets may not be easily accessible, this is of paramount importance. Data augmentation methods were also used to broaden the scope of the dataset and make the model more stable.

B. Limitations and future directions

The study's tiny dataset utilized for training and testing the model is a weakness. This may have caused the model to be overfitted to the data, thus it has to be validated on more general datasets. Another restriction is the dataset's lack of variety, which might prevent the model from being applied to populations outside of the study's original one. More consideration should be given to including datasets from other locations and ethnicities in future studies.

Further, compared to other imaging modalities like CT scans, the use of chest X-rays for COVID-19 diagnosis has limitations. False negatives from chest X-rays are possible in the early stages of COVID-19 infection because of the test's lack of sensitivity. In addition, COVID-19 may not be distinguishable from other respiratory illnesses or disorders on chest X-rays. To better diagnose COVID-19, future research should think about combining imaging modalities with clinical data.

C. Implications for clinical practice

Several consequences for clinical practice arise from using deep learning models for COVID-19 diagnosis. These models can make diagnoses of COVID-19 more quickly and accurately, allowing for earlier identification and treatment. This has the potential to lessen the transmission of the virus and boost health outcomes for those affected. In addition, these models can lessen the load on healthcare facilities, which is especially important in places with few resources and where testing for COVID-19 may not be readily available.

However, it is essential to remember that deep learning models are not intended to substitute for human clinical judgment. To provide an accurate diagnosis, these models should be used in addition to, rather than in instead of, a professional medical examination. The ethical considerations of utilizing deep learning models for COVID-19 diagnosis should not be overlooked. Considerations for data privacy, bias, and equity should be included in the development and application of these models.

VI. CONCLUSION

In conclusion, recent years have shown that deep learning models are useful for medical image processing. We developed a convolutional neural network model to detect the presence of COVID-19 in chest radiographs. The model's 83% accuracy in diagnosing COVID-19 from chest X-rays is on par with that of state-of-the-art models.

We trained our model using both positive and negative examples of chest X-rays for COVID-19 from three separate sources. All of the images in the archive have had their contrast-enhanced and pixel values standardized before inclusion. To make the model more generalizable, the dataset was augmented with additional data. Subsets of the dataset were used for model training and evaluation.

Our research shows that deep learning models may be used to automatically detect COVID-19 in chest X-rays. Even in low-resource areas where doctors may not have ready access to expensive diagnostic tools, these models may deliver a quick and precise diagnosis. Additionally, these models can lessen the load on healthcare providers and facilities, facilitating the quicker and more effective identification and treatment of patients with COVID-19.

Despite the positive findings, our study has some drawbacks that need to be addressed in follow-up studies. To begin, our model was built using a little dataset, which might restrict the applicability of the findings. The models' accuracy and resilience can be enhanced by collecting more and more varied information in future research. Second, we limited ourselves to using chest X-rays for the diagnosis of COVID-19 in our investigation. To better diagnose COVID-19, future research should investigate the use of alternative imaging modalities, such as CT scans and ultrasounds. Third, there is still work to be done on the interpretability of deep learning models. To further inform clinical decision-making, future research should focus on refining models that produce outcomes that are both visible and interpretable.

To sum up, research into deep learning models for medical image processing has shown encouraging results in the quest to better diagnose and treat COVID-19. Our findings show that it is possible to use deep learning models to detect COVID-19 in chest X-rays. Research in this field has the potential to dramatically improve patient outcomes by allowing for more rapid and precise identification of COVID-19 and other disorders.

REFERENCES

[1] D. M. Ibrahim, N. M. Elshennawy, and A. M. Sarhan, "Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases," Comput. Biol. Med., vol. 132, p. 104348, 2021, doi: 10.1016/j.compbiomed.2021.104348.

[2] J. I.-Z. Chen, "Design of Accurate Classification of COVID-19 Disease in X-Ray Images Using Deep Learning Approach," J. ISMAC, vol. 2, no. 2, pp. 132–148, 2021, doi: 10.36548/jismac.2021.2.006.

[3] L. O. Hall, R. Paul, D. B. Goldgof, and G. M. Goldgof, "Finding Covid-19 from Chest X-rays using Deep Learning on a Small Dataset," pp. 1–8, 2020, [Online]. Available: http://arxiv.org/abs/2004.02060

[4] B. Afm et al., "COVID-19 diagnosis prediction in emergency care patients: a machine learning
approach," medRxiv, p. 2020.04.04.20052092, 2020, [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.04.04.20052092v2%0Ahttps://www.medrxiv.org/content/10.1101/2020.04.04.20052092v2.abstract

[5] M. R. H. Mondal, S. Bharati, and P. Podder, "Diagnosis of COVID-19 Using Machine Learning and Deep Learning: A Review," Curr. Med. Imaging Former. Curr. Med. Imaging Rev., vol. 17, no. 12, pp. 1403–1418, 2021, doi: 10.2174/1573405617666210713113439.

[6] R. Sethi, M. Mehrotra, and D. Sethi, "Deep Learning based Diagnosis Recommendation for COVID-19 using Chest X-Rays Images," Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020, pp. 1–4, 2020, doi: 10.1109/ICIRCA48905.2020.9183278.

[7] M. K. Nath, A. Kanhe, and M. Mishra, "A Novel Deep Learning Approach for Classification of COVID-19 Images," 2020 IEEE 5th Int. Conf. Comput. Commun. Autom. ICCCA 2020, pp. 752–757, 2020, doi: 10.1109/ICCCA49541.2020.9250907.

[8] J. Rabbah, M. Ridouani, and L. Hassouni, "A New Classification Model Based on Stacknet and Deep Learning for Fast Detection of COVID 19 through X Rays Images," 4th Int. Conf. Intell. Comput. Data Sci. ICDS 2020, 2020, doi: 10.1109/ICDS50568.2020.9268777.

[9] D. Haritha, M. K. Pranathi, and M. Reethika, "COVID Detection from Chest X-rays with DeepLearning: CheXNet," Proc. 2020 Int. Conf. Comput. Commun. Secure. ICCCS 2020, 2020, doi: 10.1109/ICCCS49678.2020.9277077.

[10] M. Sevi and I. Aydin, "COVID-19 Detection Using Deep Learning Methods," 2020 Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Econ. ICDABI 2020, no. April 2020, doi:16 10.1109/ICDABI51230.2020.9325626.

[11] W. Wang, H. Liu, J. Li, H. Nie, and X. Wang, "Using cfw-net deep learning models for x-ray images to detect covid-19 patients," Int. J. Comput. Intell. Syst., vol. 14, no. 1, pp. 199–207, 2021, doi: 10.2991/ijcis.d.201123.001.

[12] K. Haque and A. Abdelgawad, “A Deep Learning Approach to Detect COVID-19 Patients from Chest X-ray Images,” Ai, vol. 1, no. 3, pp. 418–435, 2020, doi: 10.3390/ai1030027.

[13] A. S. Al-Waisy et al., “COVID-CheXNet: hybrid deep learning framework for identifying COVID-19 virus in chest X-rays images,” Soft Comput., vol. 3, no. Worldmeter, 2020, doi: 10.1007/s00500-020-05424-3.

[14] Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Rajendra Acharya, U. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in Biology and Medicine, 121, 103792. https://doi.org/10.1016/j.compbiomed.2020.103792

[15] Yasin, R. M., Gouda, A. S., & Gaber, T. (2022). AI-Based Diagnosis of COVID-19: A Comprehensive Review. IEEE Access, 10, 19703-19729. https://doi.org/10.1109/ACCESS.2022.3153979

[16] Gao, Y., Cai, Y., & Zhang, S. (2021). Deep learning in biomedical imaging: A review. Physics in Medicine & Biology, 66(13), 13TR01. https://doi.org/10.1088/1361-6560/abf57f

[17] Liu, X., Zhang, X., & Wei, C. (2021). The Prediction of the Number of Confirmed Cases and Deaths of COVID-19 in Different Countries Based on the Prophet Model. Complexity, 2021, 6657212. https://doi.org/10.1155/2021/6657212

[18] Chollet, F. (2015). Keras: Deep Learning library for Theano and TensorFlow. https://keras.io/

[19] Pati, P. P., & Rath, S. K. (2020). A Deep Learning Framework for Automated Diagnosis of COVID-19 Using CXR and CT Scan Images. Journal of Ambient Intelligence and Humanized Computing, 12(10), 11143-11154. https://doi.org/10.1007/s12652-020-02763-7