Automatic Diagnosis of Covid-19 Related Pneumonia from CXR and CT-Scan Images

Naresh Kumar
Department of Computer Science & Engineering
Maharaja Surajmal Institute of Technology
New Delhi, India
narsumsaini@gmail.com

Manish Gupta
Department of Computer Science & Engineering
Moradabad Institute of Technology
Moradabad, India
manishymca2007@gmail.com

Adeel Hashmi
Department of Computer Science & Engineering
Maharaja Surajmal Institute of Technology
New Delhi, India
ashashmi10@gmail.com

Ankit Kundu
Department of Computer Science & Engineering
Maharaja Surajmal Institute of Technology
New Delhi, India
ankitkundu93@gmail.com

Abstract—Covid-19 is a highly infectious disease that spreads extremely fast and is transmitted through indirect or direct contact. The scientists have categorized the Covid-19 cases into five different types: severe, critical, asymptomatic, moderate, and mild. Up to May 2021 more than 133.2 million peoples have been infected and almost 2.9 million people have lost their lives from Covid-19. To diagnose Covid-19, practitioners use RT-PCR tests that suffer from many False Positive (FP) and False Negative (FN) results while they take a long time. One solution to this is the conduction of a greater number of tests simultaneously to improve the True Positive (TP) ratio. However, CT-scan and X-ray images can also be used for early detection of Covid-19 related pneumonia. By the use of modern deep learning techniques, accuracy of more than 95% can be achieved. We used eight CNN (CovNet)-based deep learning models, namely ResNet152v2, Xception, InceptionResNetv2, ResNet50, Inceptionv3, NASNetLarge, DenseNet201, and VGG-16 for both X-rays and CT-scans to diagnose pneumonia. The achieved comparative results show that the proposed models are able to differentiate the Covid-19 positive cases.

Keywords—artificial intelligence; covid-19 detection; convolutional neural networks; deep learning

I. INTRODUCTION

Covid-19 epidemic has caused millions of deaths worldwide [1-4]. The detection of corona virus positive cases is done through RT-PCR tests which take approximately 2 days to give results [5-7]. It is very crucial to diagnose the positive cases at early stages and quarantine them instantly. The infection can also be identified by pneumonia in Chest X-Ray (CXR) or CT-Scan (CTS) images of the patients, which is a much faster process [8-10]. These images are analyzed by the experts to locate the infection, which may take time due to the large number of patients. An automated and accurate artificial intelligence-based system to locate the infection in the images can considerably speed up the process.

This work aims to help identify the Covid-19 pneumonia from CXR and CTS images. A deep learning approach has been used to build the models, i.e. Convolutional neural Networks (CovNets) including ResNet152v2, Xception, InceptionResNetv2, ResNet50, Inceptionv3, NASNetLarge, DenseNet201, and VGG-16 for the classification of the images as corona virus positive or not. The models were trained with a dataset containing CXR and CTS images of covid and non-covid patients. Comparative analysis of all the models is performed using various metrics to determine the best approach.

In the recent studies, researchers have observed the bilateral relation of CXR and CTS images with covid-19. They have observed the imaging patterns on the CXR and CTS of the lungs for the diagnosis of nCov (novel coronavirus) infection [11]. Also, it was found that with the help of deep learning techniques it is possible to achieve much higher image classification accuracy from CXR and CTS images in the diagnostic process of corona virus. It has been observed that 4% of the patients had False Negative (FN) corona result when tested with the RT-PCR kit. The study in [12] was related to the sensitivity of the real time polymerase chain reaction and CTSs during diagnosis of nCov were acquired. The authors observed that the RT-PCR is less accurate and also its sensitivity is much lower than that of the CTS. Deep learning techniques provide much better results in the diagnosis process of acute pneumonia in CSRs [13] or CTSs [14, 15]. In [14], strong deep learning-based models have been used for the classification of CTS images. The authors have used the imaging patterns of the CTS images of lungs to differentiate the covid positive from the covid negative cases. However, their model achieved sensitivity under 90% and was unable to reach much accurate results. The authors in [15] built an automated system for the diagnosis and quantification of Sars-CoV-2 based on artificial intelligence. Their automated system generates an output giving the quantitative opacity and 3-
dimensional output for opacities. Their model is more efficient against pixel spacing. In [16], a transfer learning-based system was developed for the diagnosis of Covid-19 from CXRs. From the above discussion and from [17, 18], it has been concluded that CXR and CTS images can be utilized for the early diagnosis of nCov positive cases.

In [19], the authors designed an automated ensemble system for the initial stage diagnosis of nCov from CXRs. They used two popular open-source datasets of CXR images for their research work. The first dataset was used for binary classification (i.e. Covid-Healthy and Covid-Positive) and the second dataset was used for multi-class classification (i.e. Covid-Healthy, Covid-Positive, Pneumonia-Positive, and Tuberculosis-Positive respectively). They used the pre-trained CovNet combination of EfficientNet, GoogLeNet, and Xception in their model. The results were summated and an absolute or relative voting was carried out to obtain the final decision. Comparative analysis of their proposed model was done with some other CovNets such as ResNet-152v2, DenseNet-201, InceptionResNetv2, and VGG-16 to assess the capabilities of the model. Their model outshined all the above-mentioned CovNets in both binary and multi-class classification and achieved decent accuracy rates of 99.21% and 98.95% for multi-class classification and binary classification respectively. The authors of [20] proposed a metaheuristic-based automated system for the early diagnosis of nCov positive patients from CXR images. They have used CNN for their experiments along with SPEA-2 for the proper tuning of the model’s parameters. They resized the dimensions of the images in the datasets to 259×259×3, which is the optimal size of the input image for the model. They computed the probability score of each class and the one with the maximum value was taken as the final. The performance of the proposed model was compared to VGG19, VGG16, ResNet50, AlexNet, ResNet-34, InceptionNet, Xception, DenseNet 201, and GoogLeNet in terms of accuracy, F1 score, sensitivity, recall, and area under the ROC curve. Their proposed metaheuristic-based model performed well in multi-class classification and achieved accuracy of 99.97% and 99.13% on training and testing sets respectively.

In [21] the authors used a multi-class dataset and deep CovNet and Xception model for their image classification model. The performance was evaluated with SVM, random forest, back propagation network, adaptive neuro-fuzzy inference systems, CovNet, VGGNet, ResNet50, AlexNet, etc. They achieved accuracy of 97.40%. The authors in [22] proposed an automated system for pre-screening of covid, heart illness, and diabetes using machine learning models. For portability, they built an android application for early prediction or severity diagnosis at home without having to travel to clinics. They used three binary datasets, one for each considered disease. For the covid dataset they arranged the data in two groups, one group containing the data points regarding the countries having less than 10,000 cases and the other one for countries with more than 10,000 cases and marked them as 0 and 1 respectively. They applied the logistic regression on the training set in all the datasets. They also applied the cross-validation technique to get the most optimized result from the model. They claimed that real-time diagnosis could be possible with the usage of an android application, also it would help in tweaking or dropping updates in the future. The proposed model was compare with other ML models such as LR, J48, KNN, SVM, ANN, RF, and GB. Their model achieved an accuracy rate of 95%, 94.20%, and 87% on diabetes, covid, and heart disease datasets respectively. The authors in [23] developed an automated system named Dark CovidNet for the early screening of Covid-19 from CXRs. Their model gave a decent accuracy of 98%, but it was trained on a small dataset of 125 CXR images.

II. METHODOLOGY

A. Dataset

Open source datasets from GitHub repository and Kaggle, were used to train the above-mentioned models. The dataset in [24] is a 740MB dataset source containing 7252 CXR images (3612 CXR images of Covid-19 patients and more than 3600 non-covid CXR images), whereas [25] is a dataset containing 1000 CXR images of Covid-19 and healthy people (495 Covid-19 positive and 505 Covid-19 negative images), as shown in Figures 1-2.

Fig. 1. Visualization of a few non-covid CXR images.

Fig. 2. Visualization of a few covid CXR images.

Regarding CTS images, [26] is a 231MB dataset consisting of 2481 CTS images (out of which 1252 images are of Covid-19 patients and 1229 are from healthy people), whereas [27] is a dataset used to train the models on chest CTS images. It consists of 750 images of Covid-19 patients and healthy people (353 images of Covid-positive patients and 397 of healthy people). The utilization ratio of the complete dataset is 80:20, i.e. 80% of the complete dataset is used to train the model and the rest 20% is used for testing. Figure 3 depicts a few Computed Tomography (CT) images from the dataset of the people infected from nCov 2019, while Figure 4 illustrates the visualization of a few CT images of healthy people present in the dataset used to train the models.

B. CovNet (CNN) Overview

CovNet has very good performance in diagnostics, agriculture, and industry [28-33]. Figure 5 illustrates a general
architecture of a CNN having an input layer, two convolution layers, each followed by a pooling layer, followed by two consecutive fully-connected layers, and at the end there is a SoftMax activation function [34, 35].

The utilized architecture is depicted in Table I. The model consists of convolution layers that extract possible characteristics from medical scans and feed them to the input layer, which is then followed by the pooling layer. The average pooling strategy of 2×2 is being used in this study, which minimizes or divides the feature mapping by 2 and transmits the output to the hidden or middle layer, which is masked behind the SoftMax activation function and comprises of pooling and convolution layers. The activation function aids the middle layer in selecting the best input based on the weights. The data processed by the middle layer are sent as input for prediction to the dense or fully-connected layer. Because it delivers the final classification or predictions, the dense layer is also known as the output layer.

III. RESULTS

The proposed models were implemented using keras. All the tests were run on a Google Colaboratory console with 12GB RAM and a Tesla K80 GPU. The validity of the CovNets has been examined in two stages. In the first stage, classification reports of the 8 CovNet models were produced for 500 epochs for both CXR and CTS images. After that, confusion matrices were constructed. After a comparison analysis of the models, the best method was identified. Table II compares the results of the proposed model with 8 standard CovNet models on the CXR test dataset. The Inception V3 outperformed the other CovNets in terms of accuracy, peaking at 9%. The Xception model, on the other hand, yields nearly identical prediction figures to the Inception V3 network and scores 2% lower in terms of accuracy, while it also gives 97%, 87%, and 89% specificity, sensitivity, and f1-score respectively. In terms of accuracy, precision, recall, and f1-score, the ResNet 50 network has the lowest prediction figures of 83%, 76%, 92%, and 81% respectively. The VGG16 model produces some incredible results, with accuracy, precision, sensitivity, and f1-score values of 92%, 86%, 98%, and 92%. The rest of the models produced similar findings, with the exceptions of DenseNet 201 and ResNet 152 V2, which produce significantly better results than the other two models (InceptionResNet V2, NASNetLarge CovNet model) in all metrics. In terms of overall statistics, the proposed model produced the best end prediction values when all of the measures were taken into consideration.

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASNetLarge</td>
<td>0.89</td>
<td>0.83</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>DenseNet201</td>
<td>0.94</td>
<td>0.82</td>
<td>0.88</td>
<td>0.89</td>
</tr>
<tr>
<td>ResNet 152 V2</td>
<td>0.91</td>
<td>0.85</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>Inception V3</td>
<td>0.86</td>
<td>0.80</td>
<td>0.83</td>
<td>0.84</td>
</tr>
<tr>
<td>Inception ResNet V2</td>
<td>0.76</td>
<td>0.92</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>ResNet 50</td>
<td>0.97</td>
<td>0.87</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>VGG 16</td>
<td>0.87</td>
<td>0.98</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>Proposed Model</td>
<td>0.96</td>
<td>0.95</td>
<td>0.96</td>
<td>0.98</td>
</tr>
</tbody>
</table>

The statistics of the models in diagnosing nCov illness from CTSs are shown in Table III. All of the models were developed using a binary classification data sample. With accuracy, specificity, sensitivity, and f1-score values of 93%, 98%, 92%,
and 96% on the CT testing data sample, the Xception model performed exceptionally well and delivered highly competitive figures. The VGG 16 and Inception V3 both provide 92% accuracy, however the rest of the metrics for both neural networks are 92%, 92%, 97%, and 95% respectively, while NASNetLarge and ResNet produced the lowest prediction figures. The accuracy rates of DenseNet 201, InceptionResNet V2, and ResNet 152 V2, are all in the range of 80-90%. In terms of overall statistics, the proposed model produced the best possible end prediction values when all the measures were taken into consideration.

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASNetLarge</td>
<td>0.75</td>
<td>0.79</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>DenseNet 201</td>
<td>0.89</td>
<td>0.79</td>
<td>0.79</td>
<td>0.81</td>
</tr>
<tr>
<td>ResNet 152 V2</td>
<td>0.88</td>
<td>0.83</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>Inception ResNet V2</td>
<td>0.75</td>
<td>0.99</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td>ResNet 50</td>
<td>0.83</td>
<td>0.74</td>
<td>0.76</td>
<td>0.79</td>
</tr>
<tr>
<td>Xception</td>
<td>0.93</td>
<td>0.98</td>
<td>0.92</td>
<td>0.96</td>
</tr>
<tr>
<td>Inception V3</td>
<td>0.93</td>
<td>0.97</td>
<td>0.95</td>
<td>0.92</td>
</tr>
<tr>
<td>VGG 16</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>Proposed Model</td>
<td>0.97</td>
<td>1.00</td>
<td>0.96</td>
<td>0.97</td>
</tr>
</tbody>
</table>

IV. CONCLUSION AND FUTURE WORK

CXR and CTS images were employed in this study to overcome the lack of precision with real-time polymerase chain reaction tests for nCoV illness diagnosis. CXR and CTS reports were employed as medical imaging procedures. CXR and CTS reports suggest certain prospective patterns (pneumonia) and have bilateral effects with nCoV, according to a large number of research publications and detailed reviews [36, 37]. Manual testing of these medical photos on the other hand takes a long time and may culminate in diagnostic inaccuracy. Image categorization, speech recognition, and other applications benefit greatly from CovNets. All the CovNet models listed above have been correctly trained, and the parameters have been fine-tuned. The models were trained using a dataset with a substantial number of images. To assess the capabilities of all the aforementioned models, a thorough comparison was carried out using a variety of popular benchmarks such as accuracy, precision, recall, and F1-score. Inception V3 and xception showed the highest accuracy of 95% and 93% respectively whereas in CTSs, Xception Inception V3 exhibited the highest scores (96% and 92% respectively). Both the CovNet models, i.e. Inception V3 and xception surpassed all the other models.

REFERENCES


Kumar et al.: Automatic Diagnosis of Covid-19 Related Pneumonia from CXR and CT-Scan Images


