COVID-19 Detection Using CXR Images Using Deep Learning

Abhishek Purnapatre[1], Pratik Patil[2], Tanmay Shinde[3], Nehal Chaudhari[4], Prof. Rupali T. Umbre[5]
[1, 2, 3, 4] Dept of Information Technology
[1, 2, 3, 4] JSPM'S Rajarshi Shahu College Of Engineering,Tathawade,Pune,India.

Abstract- *According to the world health organization, the coronavirus epidemic threatens the world's health system every day. Health resources in most countries are either insufficient or not fairly shared. There are various problems such as the number of health personnel, the number of beds, or the number of intensive care units. Using limited resources at the optimum level is the key to the country's health systems to overcome this epidemic. Disease detection is an important factor in preventing the epidemic. The higher the success, the more controlled the spread of the virus. Whether the person has a virus or not is usually done by the PCR test. In addition to the PCR method, chest x-ray images can be classified with deep learning methods. Deep learning methods have become popular in academic studies by processing multi-layered images in one go and by defining manually entered parameters in machine learning. This popularity reflected positively on limited health datasets. In this study, it was aimed to detect the disease of people whose x-rays were taken for suspected COVID-19. In such COVID-19 studies, a binary classification has generally been made. The data set includeschestx-raysofpatientswithCOVID-19,viralpneumonia, and healthy patients. Before the classification process, the data augmentation method was applied to the data set. These three groups have been classified through multi-class classification deep learning models.*

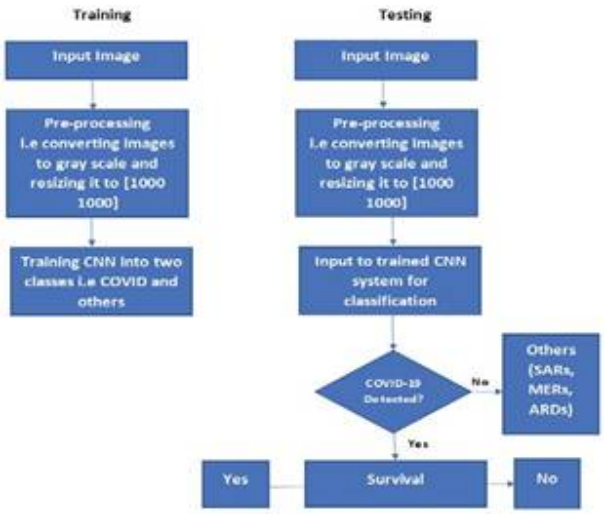
Keywords- Covid-19, chest X-ray images, convolution neural network, deep learning.

I. INTRODUCTION

In this study, it is aimed to classification healthy patients, COVID-19, and viral pneumonia cases. Deep learning is a machine learning method. It allows us to train artificial intelligence to predict outputs with a given data set. Both supervised and unsupervised learning can be used to train artificial intelligence. Deep learning is used invoice and face recognition, disease detection, defense, and security areas. The word deep in deep learning represents artificial neural networks. Artificial neural networks are inspired by the human brain. Just like the human brain, it consists of neurons. The difference between them is the amount and speed of learning. In other words, data set and processing power are needed to train artificial neural networks.

1)Robust Technique to Detect COVID-19 using Chest X-ray Images

The data used in this study is compiled using two publicly available datasets. The first dataset they used was provided in [10] In this database there are a total of 320 chest X-ray images out of which 259 images are of MERS, SARS and ARDS (pneumonia cases). These images are collected and extracted from different websites and online publications. The rest of 3347 images they used is from [4] for better DL classifier training and performance. Hence total of 3606 images are used in this study. 80% of images are used for training and 20% of images are used for testing the system. 3347 out of 3606 images are related to other diseases i.e MERS / SARS / ARDS / SARS-Cov-2 / without chest disease. The remained 259 are of COVID positive patient.
The goal of this study is based on four main objectives related to the answers of the following question:



1. Do patient is infected with COVID-19 or not?

The main methodology of this study is depicted in Figure 1. According to this, they first collected the images of CXR and arranged the data for detection of COVID-19 and

others. In initial stages they converted all the images into grayscale afterwards arranged it into two classes i.e. COVID-19 and other diseases (i.e. SARS, MERS and ARDS). If COVID-19 is detected then it is further classified to find the survival rate of patients. For this they used DL convolution neural network (CNN) with seven layers

The results of image processing and feature analysing COVID-19 are shown in Figure 3 which gives 91.67% accuracy as also depicted in Figure 6. Afterwards they further classified the data to find the survival ratio for which they again trained the CNN for feature extraction as shown in Figure 4(a). Figure 4(b) presents the confusion matrix. The accuracy of this classifier is 100% as depicted in Figure 7. The green color in Figure 3 and Figure 4(a) shows the images correctly predicted and red color shows the false detection. According to our study, the histogram results in Figure 5(a) and 5(b) emphasizing the age and sex of people infected mostly by this virus.

2) COVID-19 Detection Using Deep Learning Methods

The chest X-rays of the patients contain information that determines whether the person is COVID-19. In this article, chest x-rays of the patients were classified as COVID-19, healthy and viral pneumonia. The data set includes chest Xray images of 219 COVID-19 patients, 219 healthy patients and 219 viral pneumonia patients. This means a total of 657 images. Figure 1 shows sample chest X-ray images from the data set.

InceptionV3 is a type of convolutional neural network model. It was developed by Christian Szegedy. It consists of multiple convolution and maximum pooling steps. It consists of a 42-layer deep neural network. [15]. In the last layer, it involves a fully connected neural network. It has 78.1% accuracy rate on ImageNet dataset. The model requires 299x299x3 image size

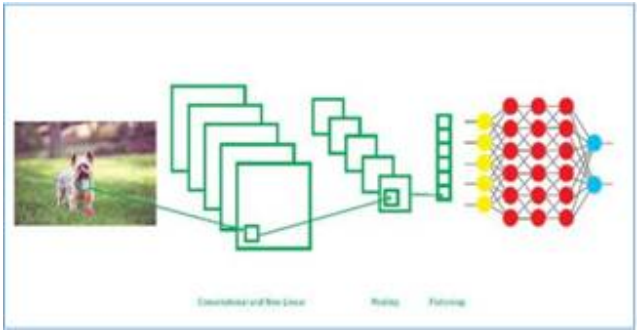
Table 2 shows the confusion matrix and other evaluation criteria of VGG16 model. The VGG16 model's accuracy rate is 93%.

3)A comparison of transfer learning approaches.

One of the critical factors behind the rapid spread of COVID-19 pandemic is a lengthy clinical testing time. The imaging tool, such as Chest X-ray (CXR), can speed up the identification process. Therefore, our objective is to develop an automated CAD system for the detection of COVID-19 samples fromhealthy and pneumonia cases using CXR images. Due to the scarcity of the COVID-19 benchmark dataset, they have employed deep transfer learning techniques, where they examined 15 different pre-trained CNN models to find the most suitable one for this task.

A total of 860 images (260 COVID-19 cases, 300 healthy and 300 pneumonia cases) have been employed to investigate the performance of the proposed algorithm, where 70% images of each class are accepted for training, 15% is used for validation, and rest is for testing. It is observed that the VGG19 obtains the highest classification accuracy of 89.3% with an average precision, recall, and F1 score of 0.90, 0.89, 0.90, respectively.

This study demonstrates the effectiveness of deep transfer learning techniques for the identification of COVID-19 cases using CXR images.



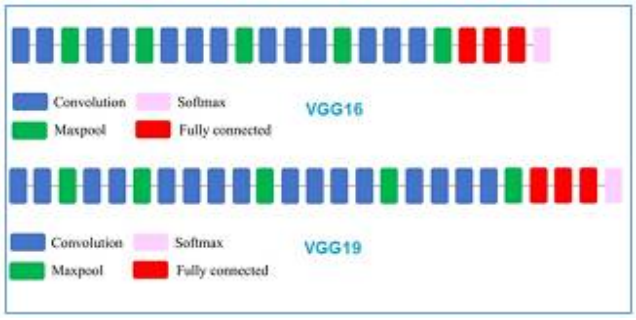
4) flat and hierarchical classification scenarios

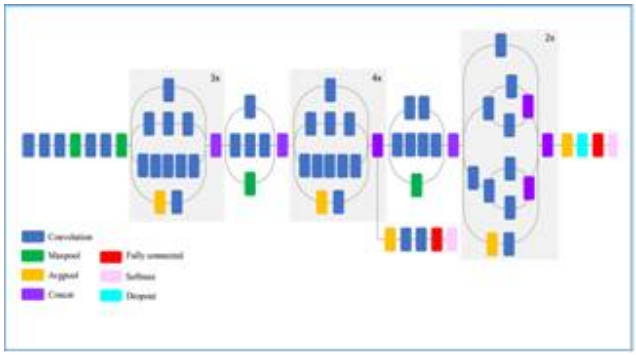
In order to achieve the objectives, they have proposed a classification schema considering the following perspectives: i) a multi-class classification; ii) hierarchical classification, since pneumonia can be structured as a hierarchy. Given the natural data imbalance in this domain, they also proposed the use of resampling algorithms in the schema in order to re-balance the classes distribution. They observed that, texture is one of the main visual attributes of CXR images, our classification schema extract features using some well-known texture descriptors and also using a pre-trained CNN model. They also explored early and late fusion techniques in the schema in order to leverage the strength of multiple texture descriptors and base classifiers at once.

To evaluate the approach, they composed a database, named RYDLS-20, containing CXR images of pneumonia caused by different pathogens as well as CXR images of healthy lungs. The classes distribution follows a real-world scenario in which some pathogens are more common than others.

The proposed approach tested in RYDLS-20 achieved a macro-avg F1-Score of 0.65 using a multi-class approach and a F1-Score of 0.89 for the COVID-19 identification in the hierarchical classification scenario.

As far as they know, the top identification rate obtained in this paper is the best nominal rate obtained for COVID-19 identification in an unbalanced environment with more than three classes. They must also highlight the novel proposed hierarchical classification approach for this task, which considers the types of pneumonia caused by the different pathogens and lead us to the best COVID-19 recognition rate obtained here.



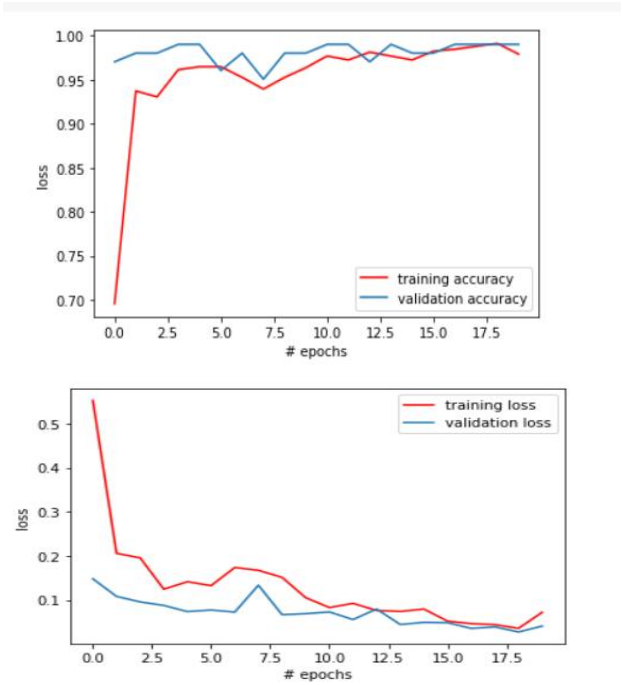


II. CONCLUSİON

In the case of parallel CNN ,we are getting promising accuracy of 97% .

At the 20 epochs we are getting maximum accuracy.

In neural network architecture we are using keras API to implement this neural network and we are using tensor flow as a back end. Input image size will be 100*100 pixel.



Parallel CNN is most the successful model that has 97% accuracy rate. COVID-19 patients, healthy patients, and viral pneumonia cases are classified successfully by this model. InceptionV3 is the most unsuccessful method for the dataset. This survey emphasizes the importance of the COVID-19 outbreak. When the number of x-ray images, ultrasound videos increases day by day the deep learning methods will more be used day by day in the health sector. The results of the studies in the literature show the importance of deep learning in combating the COVID-19 outbreak [17]. In future studies, the success ratio can be increased by strengthening the dataset. Lungtomo graphy can be used in addition to chest radiographs. By developing different deep learning models, success ratio and performance can be increased.

Though CNN achieves leading results in medical image analysis tasks, there is still scope for development. First of all, researchers can develop a “partly new” CNN model for the analysis of COVID-19 chest X-ray images by selecting the top CNN models and find their separate advantages and merge their “best parts” together to enhance the finial classification performance. Secondly, there is a scarcity of publicly accessible COVID-19 CXR image datasets. Therefore to

develop a publicly accessible database would be beneficial for future researchers.

REFERENCS

[1] A. L. Blum and P. Langley, “Selection of relevant features and examplesinmachinelearning,”Artif.Intell.,vol.97,no.1–2,pp.245– 271, Dec.1997.

[2] Y. Bengio, A. Courville, and P. Vincent, “Representation Learning: A Review and New Perspectives,” IEEE Trans. *Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828,2013.

[3] (https://github.com/ieee8023/covid-chestxray-dataset)covidchestx-ray dataset fromgithub (https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset)Pneumoniadatasetfrom kaggle.

[4] F. Rustam et al., "COVID-19 Future Forecasting Using Supervised Machine Learning Models," in IEEE Access, vol. 8, pp. 101489- 101499, 2020, doi:10.1109/ACCESS.2020.2997311.

[5] S. Roy et al., "Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound," in IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2676-2687, Aug. 2020, doi:10.1109/TMI.2020.2994459.

[6] COVID-19 identification in chest X-ray images on flat and hierarchical classificationscenarios(https://doi.org/10.1016/j.cmpb.2020.105532)

[7] Z. Long, R. Alharthi and A. E. Saddik, "NeedFull – a Tweet Analysis Platform to Study Human Needs During the COVID-19 Pandemic in New York State," in IEEE Access, vol. 8, pp. 136046-136055, 2020, doi:10.1109/ACCESS.2020.3011123.

[8] İ. Aydın, M. Sevi, U. M. Salur, “Makine Öğrenmesi Algoritmaları ile Sahte Twitter Hesapların Tespiti,” in International Conference on Artificial Intelligence and Data Processing,2018.

[9] G. Gündüz, İ. H. Cedimoğlu, “Derin Öğrenme Algoritmalarını Kullanarak Görüntüden Cinsiyet Tahmini,” Sakarya Unıversıty Journal Of Computer And Informatıon Scıences, 2, 9-17,2019.

[10] K. Shankar, V. Iyer, K. Iyer and A. Pandhare, "Intelligent Video Analytics(IVA) and Surveillance System using Machine Learning and Neural Networks," 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2020, pp. 623-627, doi: 10.1109/ICICT48043.2020.9112527.

[11] https://medium.com/@tuncerergin/convolutional-neural-network- convnet-yada-cnn-nedir-nasil-calisir-97a0f5d34cad/Convolutional Neural Network (ConvNet yada CNN) nedir, nasıl çalışır? 15 Agust 2020.

[12] M.M.Leonardo,T.J.Carvalho,E.Rezende,R.ZucchiandF.A. Faria, "Deep Feature-Based Classifiers for Fruit Fly Identification (Diptera: Tephritidae)," 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, 2018, pp. 41-47, doi: 10.1109/SIBGRAPI.2018.00012.

[13] S. Toraman, B. Daş, “Evrişimsel Sinir Ağları Kullanılarak Normal Ve Göğüs Kanseri Hücreleri İçeren Genomların Sınıflandırılması” DÜMF Mühendislik Dergisi, 11-1, 81-90,2020.

[14] https://www.researchgate.net/figure/Schematic-diagram-of-a-VGG16-and-b-VGG19-models_fig4_326421398/ Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery 15 Agust2020.

[15] A. P. Rahmathunneesa and K. V. Ahammed Muneer, "Performance Analysis of Pre-trained Deep Learning Networks for Brain Tumor Categorization," 2019 9th International Conference on Advances in ComputingandCommunication(ICACC),Kochi,India,2019 ,pp.253- 257, doi:10.1109/ICACC48162.2019.8986151.

[16] https://www.researchgate.net/publication/326421398_Very_Deep_Convolutional_Neural_Networks_for_Complex_Land_Cover_Mapping_ Using_Multispectral_Remote_Sensing_Imagery/ Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery 15 Agust2020.

[17] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio and S. K. Antani, "Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays," in IEEE Access, vol. 8, pp. 115041-115050, 2020, doi:10.1109/ACCESS.2020.3003810.