Detection of Covid-19 from Chest X-Ray Images Using Resnet50, Inception V3 Deep Learning models

[1]A PUNITHA, [2] D. NAVANITHA, [3] G S SRAVANTHI, [4] R.P.SHANTHI RANI

[1,2,3] Assistant Professor, Dept. of Computer Science and Engineering, CMR Engineering College, Hyderabad

[4] Assistant Professor, Dept. of Artificial Intelligence and Machine Learning, CMR Engineering College, Hyderabad

Abstract - The novel coronavirus 2019 (COVID-2019), which first appeared in Wuhan city of China in December 2019, spread rapidly around the world and increasing day by day and became a pandemic. There are a limited number of COVID-19 test kits available in hospitals due to the increasing cases daily. Therefore, it is necessary to implement an automatic detection system as a quick alternative diagnosis option to prevent COVID-19 spreading among people. Application of Deep learning techniques coupled with radiological images can be helpful for the accurate detection of COVID 19.Several deep learning architecture are deployed for the detection of COVID-19 such as ResNet, Inception, Googlenet etc.

In this paper, a model has been proposed for the detection of COVID infected patients using patient chest X-ray image (radiograph). We separate the dataset into two sets training and validation datasets for training the machine first and then use for detection. The databases used in most cases are about 50-100 x-ray images of both infected subjects with COVID-19 and normal subjects too. The proposed AI based approaches in the literature for detection of COVID-19 shows promising results such VGG19 with 97% of accuracy, ResNET with 96%, ResNet50 with 95% of accuracy, and InceptionV3 with 95%.

Keywords - COVID-19, ResNet, ResNet50, InceptionV3

1. Introduction

In December 2019, COVID-19 pandemic emerged in Wuhan china. It became a global health problem in a very short time because of its contact-transferred behaviours 1,2. COVID-19 is caused by a virus known as Severe Acute Respiratory Syndrome Coronavirus 2 or also SARS-Cov-23. A large family of these viruses are causing different kinds of diseases such as cold, Middle East Respiratory Syndrome and Severe Acute Respiratory Syndrome. The new addition to the coronavirus family COVID-19 was discovered in2019, which has never been detected in humans. COVID-19 is known as zoon tic disease because it is transferred from animal to human such as bat4, similar to SARS-CoV virus which is contaminated through cat and MERS-CoV from dromedary5. It is presumed that respiratory transmission and physical contact is caused to spread COVID-19 rapidly.

According to the world economic forum, people with no symptoms or mild symptoms are the main reason for spreading the epidemic. one out of four subjects shows no symptoms of COVID-19, although he/she is suffering from the disease6. Nearly 82% of the total infected subjects show mild or no symptoms and the remaining are in critical conditions7. Approximately

1,359,010 cases have been registered and 75,901 of them died and 293,454 were recovered till 7th April 2010. In the current status of infected subjects, the statistics provide the probability that 95% has the chances of recovery and 5% is the mortality rate as shown in figure 28. Dyspnea, fever and cough are symptoms of the infection while in more critical situations, this infection can lead to pneumonia, septic shock, SARS, organ failure and death4. Comparative studies show that the infection is widely spread in men as compared to women due to higher exposure and no deaths are reported in the children's age range of 0-9years 7,8. Respiratory problems spread faster in the subjects having pneumonia caused by COVID-19 as compared to healthy subjects.

In the developing countries, the pandemic spread very rapidly even after taking the precautionary measure. From march 19 till 31st march the number of infected subjects increased exponentially and the demand for intensive care units increased in parallel8.

The Chinese government publishes the updated guidelines, COVID-19 could be diagnosed by gene sequencing through blood samples and Reverse Transcript or Polymerase Chain Reaction (RT-PCR) is a specific indicator. The process gene sequencing using RT-PCR is time consuming and the subject should be hospitalized immediately. Thus, considering the fact, the subjects tested positive with COVID-19 may have pneumonia and can easily be indicated by using an automatic system for detection of COVID-19, and the concerned department should consider immediate isolation and treatment for the subjects. A subject with a critical stage, should suffer permanent lung damage, if not died. According to the world health organization, COVID-19 creates holes in the lungs similar to SARS which is not recoverable4 for detection of pneumonia; the technique of computed tomography of the chest is also useful. Artificial Intelligence based systems for automatic detection of COVID-19 can be helpful in monitoring, quantifying and distinguishing contact free subjective communication.

A deep learning technique is also developed for extraction of graphical characteristics of COVID-19 from CT images to provide quick and precise diagnosis as compared to pathogenic testing and save the critical time10. COVID-19 belongs to the same family of SARS-CoV and MERS-CoV, Scientific evidence supports the possibility to detect SARS-CoV and MERS-CoV using chest x-ray and CT images. Researchers have used the techniques of features extraction and data mining to identify the pneumonia caused by MERS-CoV and SARS-CoV11. X-ray machines are normally used to scan the body for detection of fractured bones, tumors, pneumonia, and lung infections while CT scanning is a Little advanced and more sophisticated system to examine different body parts, tissues and organs more clearly. Using x-ray images is a bit cheap and easier as compared to CT. While wrong detection may lead epidemic worse than expected12. In this article, we will discuss the existing artificial intelligence-based system for the detection of COVID-19 and the challenges these systems are facing.

Considering the challenges related to pandemic COVID-19, Artificial Intelligence (AI) can provide sophisticated solutions. The human knowledge, intelligence, and creativity along with the updated technology, its possible to beat the problems. The COVID-19 challenges are somehow exposing the drawback related to AI. The existing form of AI, in the form of machine learning and deep learning is trying to identify different pattern in the training databases. AI can provide sufficient results just in case having enough data for training and testing different systems with several approaches.

In the existing AI techniques, all the approaches are detecting COVID-19 by using chest X-ray images database. Among the challenges AI is facing a problem with aspects like accuracy, reliability. The regular methods used for detection and identification of COVID-19 are Nasal

Swab test where swab is taken from the nasal cavity of human being and being tested whether it has virus or not. This system takes 24-48 hours of time duration to get the results. Another recent development is Rapid kits through which COVID-19 detection is done by taking blood samples and this takes 3-4 hours of time duration but the accuracy of these kits is 60% and it is unreliable.

2. Literature Survey

Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020,**Authors** Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., Yan, P., and Chowell, G. , The initial cluster of severe pneumonia cases that triggered the COVID-19 epidemic was identified in Wuhan, China in December 2019. While early cases of the disease were linked to a wet market, human-to-human transmission has driven the rapid spread of the virus throughout China. The Chinese government has implemented containment strategies of city-wide lockdowns, screening at airports and train stations, and isolation of suspected patients; however, the cumulative case count keeps growing every day. The ongoing outbreak presents a challenge for modelers, as limited data are available on the early growth trajectory, and the epidemiological characteristics of the novel corona virus are yet to be fully elucidated.

They use phenomenological models that have been validated during previous outbreaks to generate and assess short-term forecasts of the cumulative number of confirmed reported cases in Hubei province, the epicenter of the epidemic, and for the overall trajectory in China, excluding the province of Hubei. We collect daily reported cumulative confirmed cases for the 2019-nCoV outbreak for each Chinese province from the National Health Commission of China. Here, we provide 5-, 10-, and 15-day forecasts for five consecutive days, February 5th through February 9th, with quantified uncertainty based on a generalized logistic growth model, the Richards growth model, and a sub-epidemic wave model.

Our most recent forecasts reported here, based on data up until February 9, 2020, largely agree across the three models presented and suggest an average range of 7409–7496 additional confirmed cases in Hubei and 1128–1929 additional cases in other provinces within the next five days. Models also predict an average total cumulative case count between 37,415 and 38,028 in Hubei and 11,588–13,499 in other provinces by February 24, 2020.

Mean estimates and uncertainty bounds for both Hubei and other provinces have remained relatively stable in the last three reporting dates (February 7th – 9th). We also observe that each of the models predicts that the epidemic has reached saturation in both Hubei and other provinces. Our findings suggest that the containment strategies implemented in China are successfully reducing transmission and that the epidemic growth has slowed in recent days.

Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan, Background: COVID-19 appeared in Wuhan, China in December 2019, and since then it has immediately become a serious public health problem worldwide. No specific medicine against COVID-19 has been found until now. However, mortality risk in patients could potentially be predicted before they transmit to critically ill. Methods: We screened the electronic records of 2,799 patients admitted in Tongji Hospital from January 10th to February 18th, 2020. There were 375 discharged patients including 201 survivors. We built a prognostic prediction model based on XGBoost machine learning algorithm and then tested 29 patients (included 3 patients from other hospital) who were cleared after February 19th. Results: The mean age of the 375 patients was

58.83 years old with 58.7% of males. Fever was the most common initial symptom (49.9%), followed by cough (13.9%), fatigue (3.7%), and dyspnea (2.1%). Our model identified three key clinical features, i.e., lactic dehydrogenase (LDH), lymphocyte and High-sensitivity C-reactive protein (hs-CRP), from a pool of more than 300 features. The clinical route is simple to check and can precisely and quickly assess the risk of death. Therefore, it is of great clinical significance. Conclusion: The three indices-based prognostic prediction model we built is able to predict the mortality risk, and present a clinical route to the recognition of critical cases from severe cases. It can help doctors with early identification and intervention, thus potentially reducing mortality.

First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020, A novel corona virus (severe acute respiratory syndrome corona virus 2, SARS-CoV-2) causing a cluster of respiratory infections (corona virus disease 2019, COVID-19) in Wuhan, China, was identified on 7 January 2020. The epidemic quickly disseminated from Wuhan and as at 12 February 2020, 45,179 cases have been confirmed in 25 countries, including 1,116 deaths. Strengthened surveillance was implemented in France on 10 January 2020 in order to identify imported cases early and prevent secondary transmission. Three categories of risk exposure and follow-up procedure were defined for contacts. Three cases of COVID-19 were confirmed on 24 January, the first cases in Europe. Contact tracing was immediately initiated. Five contacts were evaluated as at low risk of exposure and 18 at moderate/high risk. As at 12 February 2020, two cases have been discharged and the third one remains symptomatic with a persistent cough, and no secondary transmission has been identified. Effective collaboration between all parties involved in the surveillance and response to emerging threats is required to detect imported cases early and to implement adequate control measures.

A recent cluster of pneumonia cases in Wuhan, China, was caused by a novel beta coronavirus, the 2019 novel coronavirus (2019-nCoV). We report the epidemiological, clinical, laboratory, and radiological characteristics and treatment and clinical outcomes of these patients. Methods: All patients with suspected 2019-nCoV were admitted to a designated hospital in Wuhan. We prospectively collected and analyzed data on patients with laboratory-confirmed 2019-nCoV infection by real-time RT-PCR and next-generation sequencing. Data were obtained with standardized data collection forms shared by the International Severe Acute Respiratory and Emerging Infection Consortium from electronic medical records. Researchers also directly communicated with patients or their families to ascertain epidemiological and symptom data.

Outcomes were also compared between patients who had been admitted to the intensive care unit (ICU) and those who had not. Findings: By Jan 2, 2020, 41 admitted hospital patients had been identified as having laboratory-confirmed 2019-nCoV infection. Most of the infected patients were men (30 [73%] of 41); less than half had underlying diseases (13 [32%]), including diabetes (eight [20%]), hypertension (six [15%]), and cardiovascular disease (six [15%]).

Median age was 49·0 years (IQR 41·0-58·0). 27 (66%) of 41 patients had been exposed to Huanan seafood market. One family cluster was found. Common symptoms at onset of illness were fever (40 [98%] of 41 patients), cough (31 [76%]), and myalgia or fatigue (18 [44%]); less common symptoms were sputum production (11 [28%] of 39), headache (three [8%] of 38), haemoptysis (two [5%] of 39), and diarrhoea (one [3%] of 38). Dyspnoea developed in 22 (55%) of 40 patients (median time from illness onset to dyspnoea 8·0 days [IQR 5·0-13·0]). 26 (63%) of 41 patients had lymphopenia. All 41 patients had pneumonia with abnormal findings on chest

CT. Complications included acute respiratory distress syndrome (12 [29%]), ANaemia (six [15%]), acute cardiac injury (five [12%]) and secondary infection (four [10%]). 13 (32%) patients were admitted to an ICU and six (15%) died.

3. Proposed Methodology

Artificial Intelligence is most popular and renowned technology in present time which helps us to predict and to be prescient. By considering the drawbacks of existing systems, there is a need for automated detection system which reduces the time complexity and which can provide results more accurately and it should be completely reliable. Therefore, it is necessary to implement an automatic detection system as a quick alternative diagnosis option to prevent the rapid spread of COVID-19 by detecting and identifying at an early stage. In this study, an artificial neural network based model has been proposed for detection of COVID infected patients using patient chest x-ray images. The model uses ResNet50, Inception V3 and Inception ResNet V2 in-order to find out various accuracies given by each algorithm.

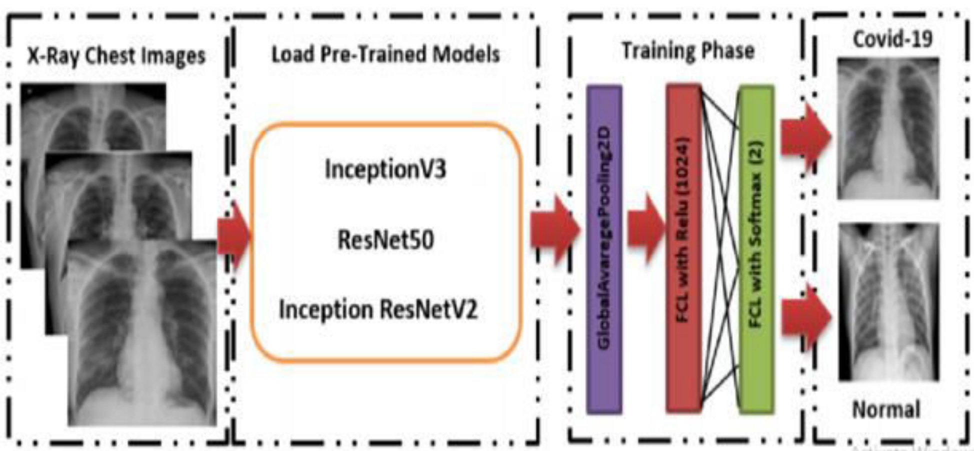

Fig.1 Architecture of proposed model

Artificial Intelligence based systems for automatic detection of COVID-19 can be helpful in monitoring, quantifying and distinguishing contact free subjective communication. A deep learning technique is also developed for extraction of graphical characteristics of COVID-19 from CT images to provide quick and precise diagnosis as compared to pathogenic testing and save the critical time. COVID-19 belongs to the same family of SARS-CoV and MERS-CoV, Scientific evidence supports the possibility to detect SARS-CoV and MERS-CoV using chest x-ray and CT images. Researchers have used the techniques of features extraction and data mining to identify the pneumonia caused by MERS-CoV and SARS-CoV. X-ray machines are normally used to scan the body for detection of fractured bones, tumors, pneumonia, and lung infections while CT scanning is a Little advanced and more sophisticated system to examine different body parts, tissues and organs more clearly. Using x-ray images is a bit cheap and easier as compared to CT.

A residual neural network (ResNet) is an artificial neural network (ANN) of a kind that builds on constructs known from pyramidal cells in the cerebral cortex. Residual neural networks do this by utilizing skip connections, or shortcuts to jump over some layers. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between. An additional weight matrix may be used to learn the skip weights; these models are known as HighwayNets. Models with several parallel skips are referred to as DenseNets. In the context of residual neural networks, a non-residual network may be described as a plain network.

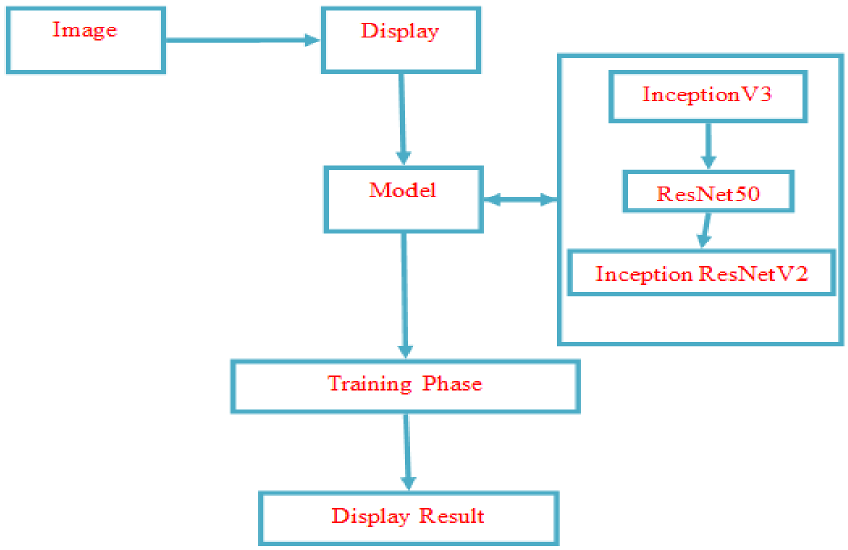

Fig.2 Methodology Implementation

X-ray images obtained from two different sources were used for the diagnosis of COVID-19. A COVID-19 X-ray image database was developed by Cohen JP using images from various open access sources. This database is constantly updated with images shared by researchers from different regions. At present, there are 127 X-ray images diagnosed with COVID-19 in the database. fig.3 shows a few COVID-19 cases obtained from the database and the findings of the experts.

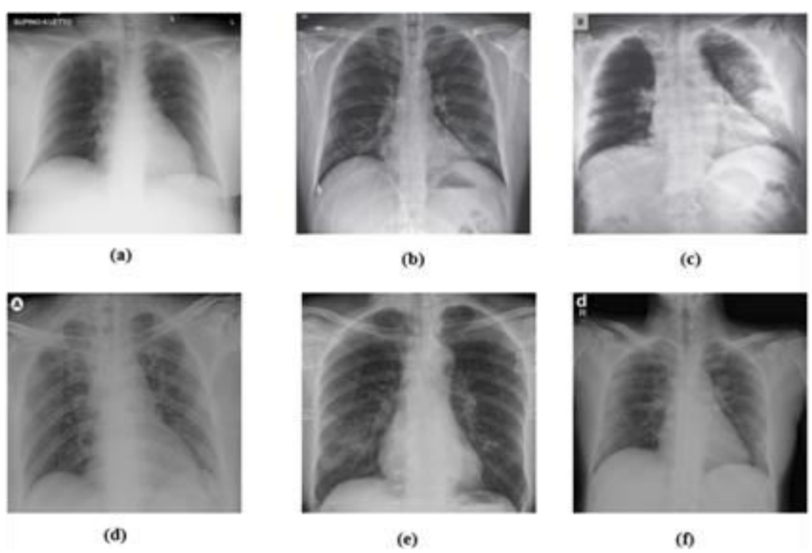

Fig.3 Scan images

A few COVID-19 cases and findings by dataset:

(a) Cardio-vasal shadow within the limits,

(b) Increasing left basilar opacity is visible, arousing concern about pneumonia,

(c) Progressive infiltrate and consolidation,

(d) Small consolidation in right upper lobe and ground-glass opacities in both lower lobes,

(e) Infection demonstrates right infrahilar airspace opacities, and

(f) Progression of prominent bilateral perihilar infiltration and ill-defined patchy opacities at bilateral lungs.

There are 43 female and 82 male cases in the database that were found to be positive. In this dataset, a complete metadata is not provided for all patients. The age information of 26 COVID-19 positive subjects is given, and the average age of these subjects is approximately 55 years. Also, the ChestX-ray8 database provided by Wang et al. [58] was used for normal and pneumonia images. In order to avoid the unbalanced data problem, we used 500 no-findings and 500 pneumonia class frontal chest X-ray images randomly from this database.

The advent of deep learning technology has revolutionized artificial intelligence. The word deep refers to the increase in the size of this network with the number of layers. The structure is named after convolution, a mathematical operator. A typical CNN structure has a convolution layer that extracts features from the input with the filters it applies, a pooling layer to reduce the size for computational performance, and a fully connected layer, which is a neural network. By combining one or more such layers, a CNN model is created, and its internal parameters are adjusted to accomplish a particular task, such as classification or object recognition.

4. Results & Discussion

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English words frequently whereas other languages use punctuation, and it has fewer syntactic constructions than other languages.

Numpy - An open-source library mainly useful for its N-dimensional array objects.

Pandas - A fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language. Python data analysis library, including structures such as dataframes.

Matplotlib - A comprehensive library for creating static, animated, and interactive visualizations in Python. 2D plotting library producing publication quality figures.

Scikit-learn - A free software machine learning library for the Python programming language used for data analysis and data mining tasks. Some features of Scikit-learn are Classification, regression and clustering.

Tensorflow - An open source library developed by Google which is used for large numerical computations and accepts data in the form of multi-dimensional arrays named as tensors. Tensorflow can be used sophisticatedly even while working with GPU’s

- Collect and sample an open-source database for Chest X-ray images of Covid-19 infected patients.
- Collect and sample an open-source database for Normal and healthy patient's chest x-ray images.
- Combine the above two datasets and make a one single dataset of both covid positive and negative chest x-rays.

- Train and test on CNN to automatically detect COVID-19 X-rays via the dataset we created.
- Since COVID-19 attacks the epithelial cells that line our respiratory tract, we can use X-rays to analyze the health of a patient's lungs.
- A drawback is that X-ray analysis requires a radiology expert and takes significant time — which is precious when people are sick around the world. Therefore, developing an automated analysis system is required to save medical professional's valuable time.
- Evaluate the results after processing the CNN model.

To load our data, we are giving paths to images in dataset directory. Then, for each imagePath, we are going to following processes:
Extract the class label (either covid or normal) from the path.
Load the image, and preprocess it by converting to RGB channel ordering, and resizing it to 224×224 pixels so that it is ready for our Convolutional Neural Network.
Update our data and labels lists respectively.
We then scale pixel intensities to the range [0, 1] and convert both our data and labels to NumPy array format.
Next we will one-hot encode our labels and create our training/testing splits:

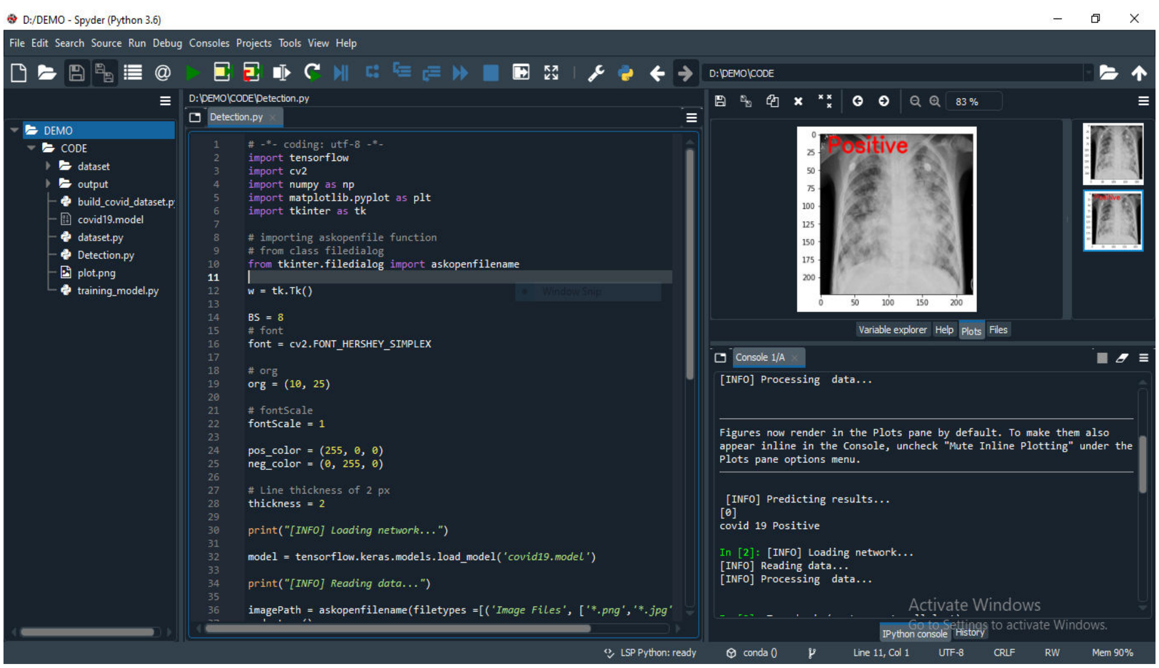

Fig.4 testing –Xray images (Positive)

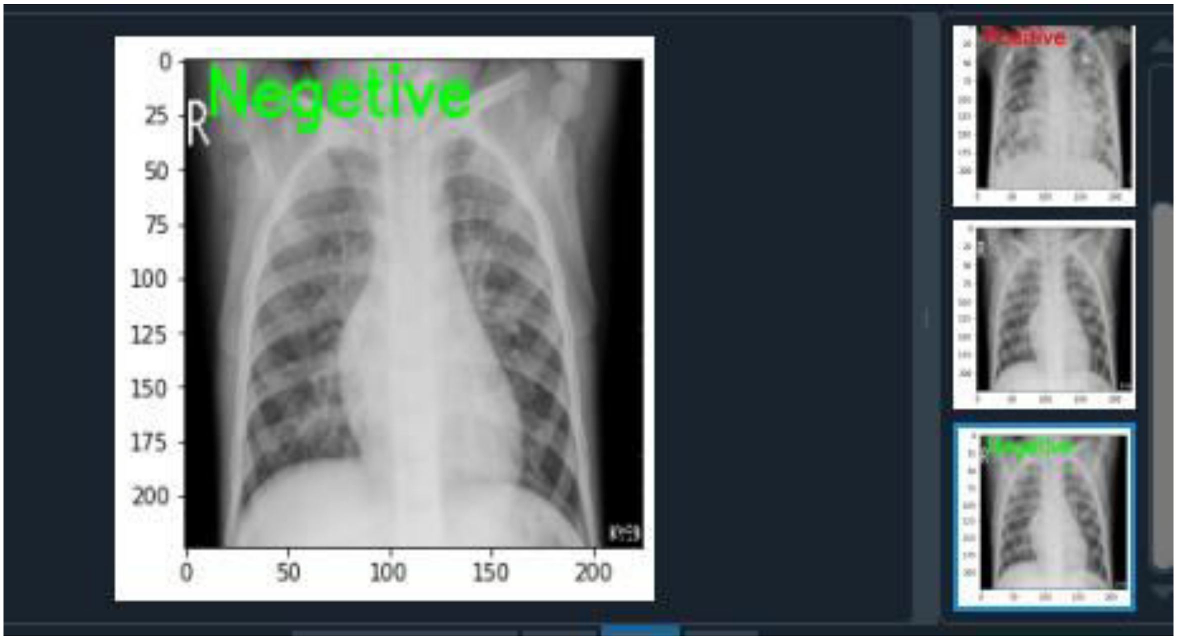

Fig.5 testing –Xray images (Negative)

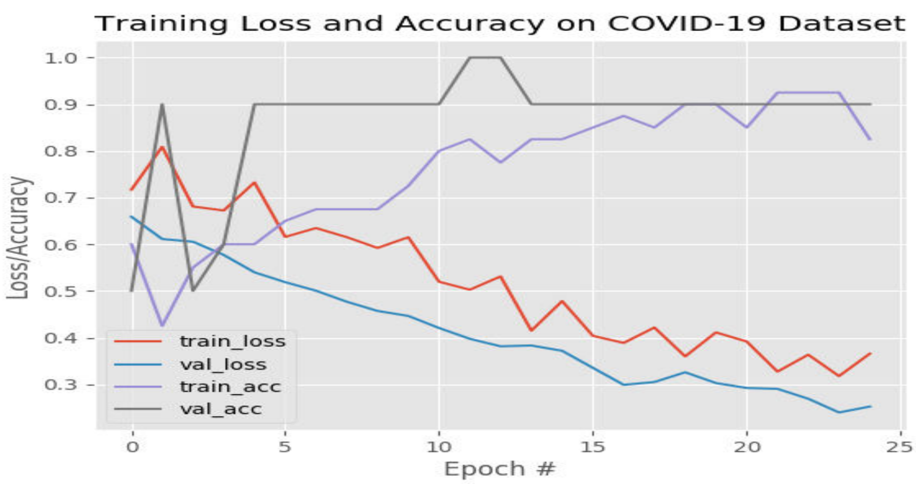

Fig.6 loss& accuracy of proposed Model

5. Conclusion

The proposed AI based approaches in the literature for detection of COVID-19 shows promising results such VGG19 with 97% of accuracy, ResNET with 96%, ResNet50 with 96% of accuracy, and InceptionV3 with 95%. The databases used in most cases are about 50-100 x-ray images of both infected subjects with COVID-19 and normal subjects too. All the proposed approaches used binary classification techniques.

To conclude the existing work, it is hard to fight COVID-19 because of its mysterious behavior, and unknown biological origin. We can try the precautionary measures and lessons learned from other public health outbreaks such as SARS-CoV and MERS-CoV. The awareness about wearing masks, social distancing, isolation, hygiene and quarantine can reduce the chances of spreading pandemic. Convalescent plasma is also considered as a potential therapy for COVID-19.

The future challenges for AI related to the detection of COVID-19 is: training and testing different deep learning architectures with huge databases with all kinds of chest x-ray images infected by different kinds of pneumonia. A multi- classification technique is required with data for more precise detection of COVID-19.

References

1. Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., Yan, P., and Chowell, G. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. Infectious Disease Modelling, 5 : 256-263, 2020.

2. Yan, L., Zhang, H. T., Xiao, Y., Wang, M., Guo, Y., Sun, C., Tang, X., Jing, L., Li, S., Zhang, M., Xiao, Y., Cao, H., Chen, Y., Ren, T., Jin, J., Wang, F., Xiao, Y., Huang, S., Tan, X., Huang, N., Jiao, B., Zhang, Y., Luo, A., Cao, Z., Xu, H., and Yuan, Y. Prediction of criticality in patients with severe Covid-19 infection using three 6/8 clinical features: a machine learning-based prognostic model with clinical data in Wuhan. medRxiv preprint doi: https://doi.org/10.1101/2020.02.27.20028027, 1-18, 2020.

3. Stoecklin, S.B., Rolland,P., Silue,Y., Mailles, A.,Campese, C.,Simondon, A.,Mechain, M.,Meurice, L.,Nguyen, M., Bassi C., Yamani, E., Behillil, S., Ismael,S., Nguyen, D., Malvy, D., Lescure, F. X., Georges, S., Lazarus, C., Tabaï, A., Stempfelet, M., Enouf, V., Coignard, B., Levy-Bruhl, D. and Team, I. First cases of corona virus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. Euro surveillance, 25(6) :2000094, 2020

4. https://www.who.int/health-topics/coronavirus [Last visit 01.04.2020].

5. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin Q., Wang, J., and Cao, B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, 395(10223): 497-506, 2020.

6. https://www.weforum.org/agenda/2020/03/people-with-mild-or-no-symptoms-could-be-spreading-covid-19/ [Last visited 01/04/2020]

7. https://www.nationalgeographic.com/science/2020/02/here-is-what-coronavirus-does-tothe-body/ [Last visited 01/04/2020]

8. https://www.worldometers.info/coronavirus/ [Last visited 01/04/2020]

9. Wang, Y., Hu, M., Li, Q., Zhang, X. P., Zhai, G., and Yao, N. Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner. arXiv preprint arXiv : 2002.05534, 1-6, 2020.

10. Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., Li, Y., Meng, X., and Xu, B. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv preprint doi: https://doi.org/10.1101/2020.02.14.20023028, 1-26, 2020.

11.Xie, X., Li, X., Wan, S., and Gong, Y.Mining X-ray images of SAR Patients. DataMining: Theory,Methodology, Techniques, and Applications, Williams, Graham J., Simoff, Simeon J. (Eds.), pp. 282-294, ISBN: 3540325476 , Springer-Verlag, Berlin, Heidelberg, 2006.

12. https://biodifferences.com/difference-between-x-ray-and-ct-scan.html
13. Narin, A., Kaya, C., Pamuk, Z. (2020). Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv preprint arXiv:2003.10849.
14. Sethy, P. K., Behera, S. K. (2020). Detection of coronavirus Disease (COVID-19) based on Deep Features.
15. Zhang, J., Xie, Y., Li, Y., Shen, C., Xia, Y. (2020). COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection. arXiv preprint arXiv:2003.12338.
16. Hemdan, E. E. D., Shouman, M. A., Karar, M. E. (2020). COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. arXiv preprint arXiv:2003.11055.
17. Apostolopoulos, I. D., Bessiana, T. (2020). Covid-19: Automatic detection from X-Ray images utilizing Transfer Learning with Convolutional Neural Networks. arXiv preprint arXiv:2003.11617.
18. https://github.com/ieee8023/covid-chestxray-dataset
19. Chest X-Ray Images (Pneumonia) https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
20. https://github.com/ieee8023/covid-chestxray-dataset