SEGMENTATION OF CHEST CT IMAGES FOR RADIOMICS FEATURES BASED COVID-19 ANALYSIS USING DEEP LEARNING TECHNIQUES

**[1] Dr. V Anantha Natarajan, [2]Dr. Syed Jahangir Badashah, [3]Srilakshmi CH,
[4] Dr.M Sunil Kumar, [5]Jangam Bala Srinivasa Rao**

[1]Professor of CSE,Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous),
Tirupati,Andhra Pradesh,India.
vananthanatarajan@vidyanikethan.edu
[2]Professor, Department of ECE,
Sreenidhi Institute Science and Technology Autonomous Institution
Hyderabad, TS, India
syd_jahangir@yahoo.co.in
[3]Assistant Professor, R.M.D Engineering College
Tamil Nadu-601206
sricsbs@gmail.com
[4]Professor of CSE,Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous),
Tirupati,Andhra Pradesh,India.
sunilmalchi1@gmail.com
[5]Asst Professor, SRK Institute of Technology,Vijayawada, AP, India.
srinivasmm777@gmail.com

ABSTRACT

In recent past the COVID-19 pandemic cases surged in global level, and cause several issues and challenges for the physicians in diagnosis and treatment process. This research aims at presenting an efficient and intelligent support tool for the doctors in diagnosis and analysis of the Chest CT images of the infected patients. In general the COVID-19 affects the normal functioning of the lungs by damaging the small air sacs known as alveoli. During the COVID period the Chest CT were the prominently used screening mechanism before certain critical and emergency surgeries and treatment in ICU instead of the regular RT-PCR procedure which is time consuming. Usually the CT image analysis is done manually which is prone to error. It is hard for a physician to analyze hundreds of CT slices. Also the COVID-19 infections has more similarity with other pulmonary diseases including pneumonia and others. This research study attempts to detect the COVID-19 infections in chest CT images by segmentation the lung region in the CT images and then analyzing using radiomics features. The radiomic features aids in discriminating the normal non-transparent regions in the chest, GGO's, other consolidations, and few high intensity region including the blood vessels. The radiomics features are classified using a back propagation neural network. This automated analysis shall help to reduce the subjective variability and improve the overall efficiency of the diagnosis process when compared to the manual or other similar diagnosis process.

Keywords: COVID 19, chest CT, radiomics, GGO's, classification, neural network

1. INTRODUCTION

Coronavirus (COVID-19) disease is an infectious disorder caused by SARS-COV-2 virus whichis a recently discovered. This coronavirus is one of the families of RNA viruses that causes remarkable pathogens in animals and humans. People who are infected with this virus will experience modest to moderate illness in the respiratory tract. It may give serious impacts to people with a medical illness like cardiovascular disease, diabetes, chronic respiratory disease, and cancer. There are two main broad categories in the COVID-19 diagnostic approach. They are chest/ thoracic radiography and laboratory-based approach. At recent times it was proved that the COVID-19 can be diagnosed from chest CT images even for patient for whom the reverse transcriptase Polymerase Chain Reaction (RT-PCR) test results were negative. In case of RT-PCR, false negative test results are attributed by multiple factors including the low level of viral RNA in the collected sample, insufficiency of cellular components, and improper derivation of acid from samples.

In past only a limited set of reports showing the reliability and efficiency of the chest CT images in COVID 19 diagnosis in case of asymptomatic patients was available. Based on the few literatures [1, 2] it can be believed that chest CT includes essential features to detect the asymptomatic infections for which RT-PCR result were negative. Another CT scan named thoracic CT scan plays a major part in the diagnosis of COVID-19. For detecting COVID-19 thoracic CT is used as a primary tool that achieves high sensitivity in diagnosis. The CT scan requires X-rays to transmit via thoracic region of the patient, which can be detected by a radiation detectors, and in dominant images, they are reconstructed. Early diagnosis of the lung infection due to the viral disease may help to prevent patients requiring intensive care where the disease has worse outcomes.

Artificial Intelligence (AI) was has a vital part in computer aided diagnosing SARS-COV-V2 viral infection and its associated side effects. AI has been utilized in bio-medical studies since its prominence and has become broadly considered because of its high detection accuracy and low misclassification. During the COVID-19 diagnostic phase, AI can be used to identify patterns in CT images. AI can be used in the process of segmenting the focused regions (ROI) and capturing refine components in chest tomographic image, features representing the abnormalitieswere extracted in an automated fashion from the CT images and used for further analysis. Many recent studies have proposed AI based approaches specifically that detects COVID-19 and can also distinguish it from other lung problems and community pneumonia.
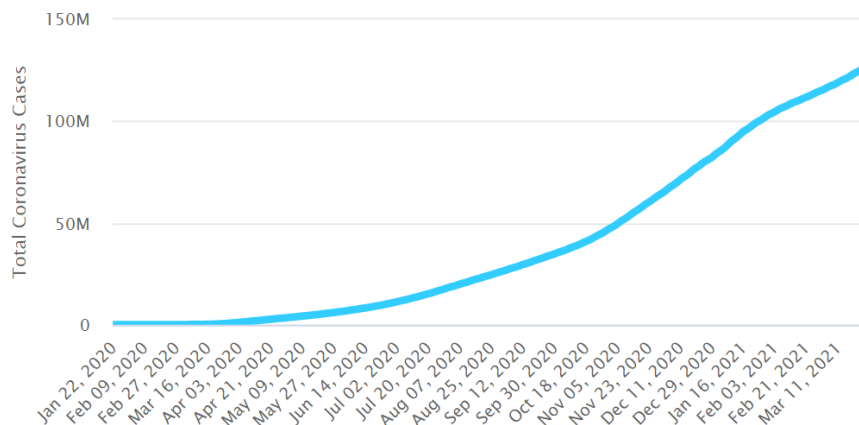

Fig. COVID-19 cases reported every two week by the WHO Region (Source: [3])

Machine Learning algorithms (ML) which are considered to be a subset of the AI were used in COVID-19 diagnosis by giving them the knowledge on how to acquire the necessary features from the biomedical images and take an accurate decision based on the extracted features. Deep Learning techniques are

717

subset of machine learning algorithm that focus on architecting and training deep neural architectures with the strong ability to uncover hidden patterns from the complex data. Deep Learning (DL) algorithms including belief network (DBN), Deep Neural architecture (DNN), and convolutional network (CNN) were employed in bio-medical image processing during the recent past and the corresponding results are better when analyzed along with their predecessors called as the machine learning algorithms. To control the COVID 19 outbreak, diagnosis and treatment process the role of AI especially machine learning and deep learning algorithms are increasing day by day. In a recent research, deep learning algorithms were employed to predict novel combinations in drug discovery for treating SARS-Cov2 by examining their protein related genetics. The large volume of COVID-19 data and COVID 19 related data collected from the social networks were used to build complex deep neural architectures that can predict the disease outbreak in a particular community, transmission path of the disease, and their impact. Deep learning algorithms were also used in combination with computer vision techniques for monitoring the crowd movement in public places and detect the infected person[30].

An efficient automated diagnosis approach can assist the physicians in controlling and preventing the pandemic disease from spreading. It is also obvious that a regular lab based testingmethodsconsumes more time and a chain of well-equipped laboratories are not sufficient to cater in high workload conditions. Thus bio-medical imaging especially CT images have become fast and reliable method in diagnosingCOVID-19 in critical and emergency situations. In general the CT image based COVID19 analysis focuses on estimating the volume of the ground glass opacity (GGO), consolidations, and gray pattern present in the images. Based on the estimations the severity or the level of infection can be determined [29].

Many recent research articles have been published for the segmentation of chest CT images for diagnosing COVID19. In majority of the approaches the lung segmentation from the chest CT images becomes a preliminary and essential activity. The lung region has to be detected for further feature extraction based on which Covid-19 severity analysis has to be conducted. Object detection approaches fails to detect the lung region because of its complex shape. A radiologist has to analyze a huge volume of CT scan images which is laborious and the likelihood of inaccurate detections is higher. Hence automated segmentation of lung nodules or the region of interest (RoI) is essential which takes advantage of the essential features available in the high-resolution CT images. This automated segmentation does aids the physician in the diagnosis process by looking in to the inner details of the ROI instead of the whole CT images. The basic and fundamental image processing techniques including morphological operations, active shape model and clustering techniques were employed for the segmentation of lung area. Segmentation based on hand crafted features was not accurate as like the end-to-end learning architecture. The majority of the biomedical image segmentation utilizes the encoder-decoder architecture such as U-Net or its variants. One of the major issues in lung segmentation is the utilization of various datasets for the evaluation of the performance of a segmentation model. Benchmark models are not available for considering as a primary reference for segmentation [28].

Radiomics is a non-invasive machine learning technology that involves extracting more number of beat representational features from bio-medical images in an efficient manner, thus transforming the multi-dimensional bio medical images in to features that can best describe the nature and components (that are not visible to the naked eye) present in the images. Radiomics' potential advantages in enhancing diagnosis process, and the corresponding accuracy in predicting the various categories of carcinoma including lung, rectal, and other non-neoplastic disorders have been emphasized [10-13]. There is currently less volume of data to justify the effectiveness of radiomics features extracted from the chest CT in identifying COVID-19 pneumonia quickly and effectively [27].

This study attempts to classify CT images of COVID 19 infected person from the CT images of normal and infected with other common lung infections after segmenting the lung region from the CT images.

718

For lung segmentation UNET architecture was used and later radiomics features (texture and shape) were extracted from the CT image slides. For classification a neural network trained with back propagation algorithm was employed[25][26].

2. LITERATURE SURVEY

As part of computer aided diagnosis the automatic delineation of different infected organs in bio medical images has become one of the important steps. In automatic organ delineation the deep learning techniques were gaining more importance especially the semantic segmentation architecture like UNet. This section shall discuss the various deep learning models used for semantic segmentation of lungs and other organs of interest from CT images. In a research article [4], the advantage of using dilated convolutional layers which captured the global information from the input images was analyzed. The dilated convolutional layers were capable of efficiently learning the anatomy of the organ and thus the overall delineation accuracy was increased. The major challenge in lung segmentation from medical images is that the lung disease changes the density of the tissue causing variations in the intensity of the CT images. Therefore the segmentation techniques based on intensity are not reliable. In [5] graph search was used to automatically segment the pathological lung from the CT image [32]. The graph search method includes anobjective function derived from combined intensity, gradient, smooth boundary and the meta-information of the rib region. With an objective to support radiotherapy treatment planning, a deep neural architecture based methods was proposed for segmenting multiple organs from chest CT images using an adversarial network termed as UNet-GAN [6]. A set of U-Net model were used as generators and a fully convolutional network was used as discriminator. For a given input image the generator network yields a segmentation map based on the knowledge gained during the training process. During the model training it learns how to map a CT image to multi-organ segmented image in an end-to-end fashion [19][20][22].

In a recent study the UNET architecture was used to segment lungs from the Electrical Impedance Tomography (EIT) images [7]. Even without prior knowledge the trained U-Net was capable of segmenting the conductivity changes in the EIT images. In an earlier attempt, using thresholding techniques the lungs segmentation in each CT slice was implemented wherein the segmentation was improved by separating the right and left parts of the lung with respect to the anterior junction line [8]. The bronchi and trachea were eliminated and the diaphragm region was suppressed during the post-processing stages[23]. To accurately segment lungs from the Low Dose tomographic (LDCT) images the challenges caused by the presence of arteries, veins, and bronchi which appears closer to the chest tissues must be addressed. Hence the segmentation approach should consider the spatial relationship between the other components of the lungs for preserving the information and details available in the lung region. In [9] a two stage segmentation process was adopted; initially the Linear Combination of Discrete Gaussians (LCDG) was used for segmentation of lungs from LDCT and later using Markov-Gibbs-random-field model (MGRF) the segmentation was refined[21][24].

Radiomics features, which include efficient and high dimension representational features [18]that are derived from the quantitative analysis of grayscale image pixels, statisticaland textural information may be used to forecast and analyze disease growth.

3. METHODOLOGY

The CT images are pre-processed using a median filter to remove the salt and pepper noise and preserve the edges. A 3x3 sized square shape filter is used as a median filter. In general the mean filter is used for removing noise present in the bio-medical images but it tends to blur the images and overall affects the Lung segmentation and GGO detection. To overcome such issue median filter was utilized and also the noise is not a serious issue in CT images and the median filter is employed to remove the spike points and smooth the image.

719

Given a set of images $X \in \mathbb{R}^{n_r, n_c}$ and respective binary mask $Y \in \{0,1\}^{n_r, n_c}$ the objective is to build a mathematical model M using the encoder-decoder architecture and conduct a set of experiments to detect the optimal set of hyper-parameters. The model is trained on the given set of images so that the overall loss is reduced and the generalization ability is tested by using a validation set. In overall the objective of the training process is to detect a point in the weight space such that for a given input X_{test} the model yields an output $\hat{Y}_{test}$ which is equivalentalmost to the expected or target output Y_{test}. The testing will be more reliable when the test samples considered for the performance analysis obeys the following condition; $X_{test} \cap X = \emptyset$.

To standardise the image scale in the slice bicubic resampling was used, which results in a pixel size of 0.5 mm x 0.5 mm. Using the actual CT image, and the smoothed image, intensity based statistical features, gray-level co-occurrence matrix (GLCM) features, and gray-level run-length matrix (GLRLM) features were generated. Based on the parameters calculated in the training set, the radiomic characteristics are normalised to z-scores.

3.1 Modeling

Radiomics features are strongly related with the characterization of patients with and without COVID-19 and the same has been exploited in this research. When compare to the time consuming manual process in segmenting the abnormal or the infections on CT images, deep learning-based techniques have proved to be efficient.Also our experimental results indicate that using radiomics to diagnose COVID is a reliable option.The accuracy of the analysis of infected images in this study was assured by the absence of substantial statistical variations in radiomics characteristics acquired from the manual segmentation of lung region amongst radiologists.

The back-propagation neural network (BPNN) classifier was utilized in this study to test the efficiency of radiomics characteristics for the categorization of COVID 19. The output of the network nodes were sent to differentiable activation function, and the use of backpropagation algorithm to update the weights connecting the internal network nodes after everyepoch in the training process, helped the BPNN to effectively discriminate the two classes of input. BPNN has strong ability to construct complicated decision surfaces in a high dimensional space of the features, the BPNN was chosen as a classifier.There's even evidence that the output of a BPNN is close tothe results of Bayesian posterior probabilities, according to some research. This is significant because for a given distribution of feature data, a Bayesian classifier delivers the highest accuracy (i.e., low mis-classification). It is impossible to anticipate the performance of a BPNN a priori, as it is with other non-parametric techniques to pattern categorization. Additionally, the volume of samples required for training, the number of nodes in the hidden layer, and the learning parameter are also BPNN parameters that must be selected.

Based on Baum and Haussler's [17] work, a limit (m) on the volume of training samples required to produce a given level of performance over a set of test samples having the same probability distribution as the samples used for training can be set. It can be predicted that a fraction equal to $1 - \frac{\epsilon}{2}$ test samples will be classified correctly when m training samples were used to train a neural network with N nodes and W parameters. The value of m can be defined mathematically as

$$m \geq O\left(\frac{W}{\epsilon} \log \frac{N}{\epsilon}\right)$$

The number of training samples was kept around 10 times the number of weights in the network to keep the error % less than 10% when classifying the test samples. More than 10000 training samples are required for a typical network with three layers, each with 100 nodes. Creating so many photos is simply

720

not feasible. Fortunately, as several studies have actually proved, this bound does not prohibit the possibility of constructing a good classifier with fewer training examples. New training data samples were developed from existing training photos using data augmentation techniques, and experiments were done on the dataset with and without augmentation.

Theoretical foundations for determining the number of hidden nodes to use in a single hidden layer network are lacking. The only general strategy for optimizing this value is to run the network with various numbers of hidden nodes and choose the optimal one. To correct the values of the network's hyperparameters, a cross validation approach was used. Based on the related radiomics investigations [14–16], the segmented pictures were randomly separated into training and test datasets with an 80:20 ratio. Images from the test datasets were utilized as a validation dataset after the classification model was obtained from the training dataset. This allowed the method provided in this work to be independently verified.

4. EXPERIMENTS AND RESULTS
The following section presents result achieved by the proposed approach which includes COVID-19 dataset, experimental setup, quantitative results, qualitative results, failure analysis and discussion.

4.1 Dataset
A total of 2482 CT scans are accessible in a public SARS-CoV-2 CT scan dataset, which includes 1252 CT scans that are positive for SARS-CoV-2 infection (COVID-19) and 1230 CT scans for patients who are not infected with SARS-CoV-2. The information was gathered from genuine patients in hospitals in Sao Paulo, Brazil. There were no statistically significant variations in demographic data between male and female patients($p < 0.05$).The ratio of persons classified as early, middle, or late in the SARS-CoV-2 infection based on the time between the onset of symptoms and the admission CT scan. Because the quantity of training samples is crucial for a model to be effectively trained, we used a data augmentation technique to increase the number of training samples.

4.2 Experimental setup and results
The pre-processed images are segmented using the trained U-NET architecture and the radiomics features were extracted from the segmented lung region. Data augmentation is also employed during training to improve network robustness, in which images are randomly and minimally scaled, rotated, skewed, warped, noise is added, and image signal strength is marginally altered. The initial parameter values for neural networks, such as weight, bias, and learning rate, are used to train them.It begins leaning with a starting value, and the weight is updated with each iteration. The training of a neural network takes a long time and has a complicated structure.

Because sensitivity and specificity measures are highly dependent on the threshold value used to modify the model's performance, evaluating a model only on the basis of these metrics is problematic. There are two types of faults to consider when working with a binary classification problem: false positive rate and false negative rate.A better option to assess a model is to employ measures that integrate FPR and FNR to provide a comprehensive view of the model.A common strategy is to use the Receiver operating characteristic (ROC), which depicts the true positive rate (y-axis) as a function of the false positive rate (x-axis) for different thresholds ranging from zero to one.
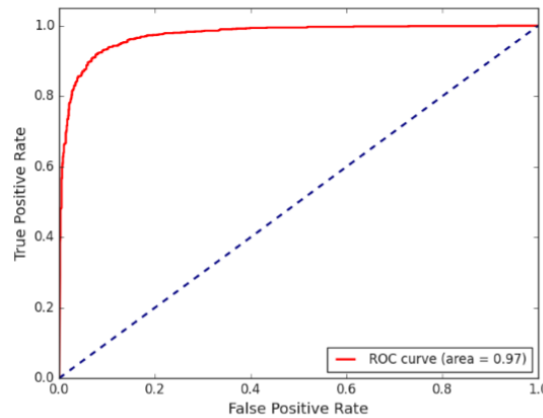
Fig 4.2(a): Reciever Operating Characteristic (ROC) of Neural Classifier

The receiver operating characteristic (ROC) curve was used to assess the radiomics, effectiveness of CT images, and models' classification abilities. The accuracy, sensitivity, and specificity of the area under the curve (AUC) were calculated. To assess the models calibration and clinical benefits, calibration curves and decision curve analysis (DCA) curves were created. The confusion matrix, specificity, sensitivity, and accuracy in the training and test cohorts were calculated at the cut-off value in the training set at the maximum of the index of each model for ROC analysis, and the confusion matrix, specificity, sensitivity, and accuracy in the training and test cohorts were derived at this cut-off value. The respective optimism adjusted AUC and average optimism indicate that the model built from the given feature has a high level of reliability. The performance of the classifier was calculated with respect to sensitivity, specificity, and accuracy. In additional, the ROC curve was employed to assess performance. ROC analysis was used to evaluate the classifier's ability to distinguish between genuine COVID images and false-positive candidates. Based on the output values from the classification model, the areas under the ROC Curve values were calculated.

Table 4.2(a): Dataset training/testing with and without data augmentation

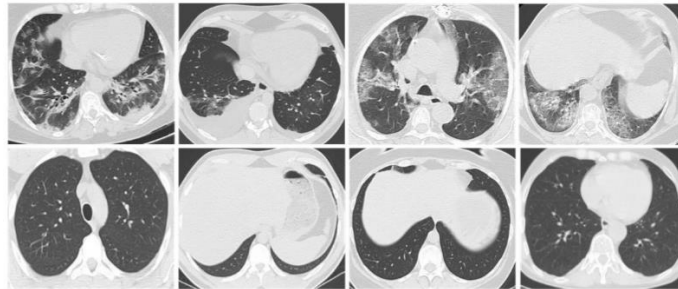| DATA	NO AUGMENTATION		WITH AUGMENTATION	
	COVID-19	Non-COVID-19	COVID-19	Non-COVID-19
Training	760	710	3040	2850
Test	492	520	504	530
total	1252	1230	3544	3380

Fig 4.2(b):The first row of samples belongs to the covid-19 cases, whereas the second row belongs to the non-COVID instances.
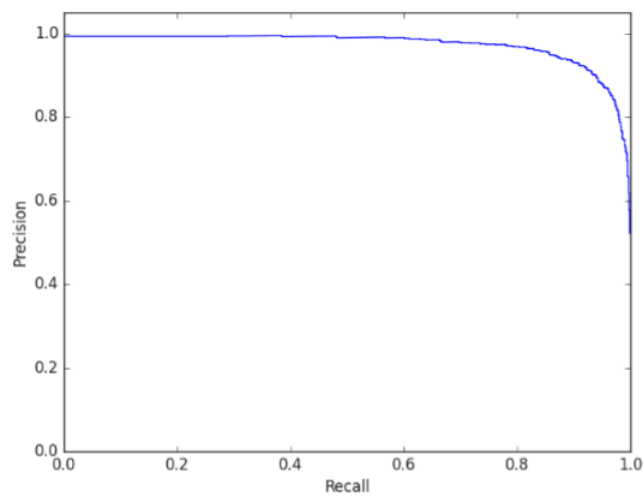

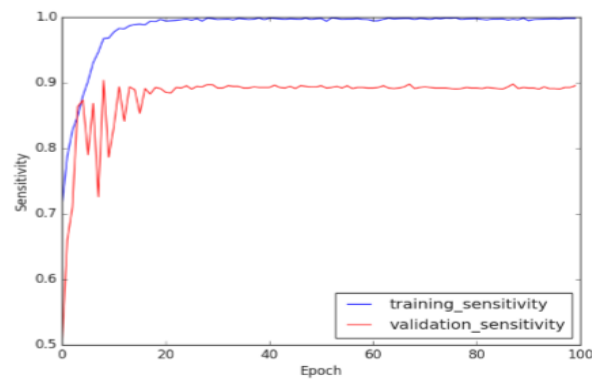**Fig 4.2(c): The test set's precision-recall curve**

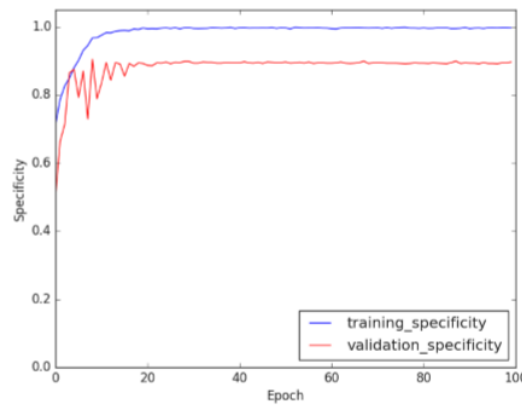**Fig 4.2(d): Model sensitivity during training and validation**

Fig 4.2(e): Model specificity during training and validation

Standard performance metrics, such as specificity, sensitivity, accuracy, and receiver operating characteristic curve, can be used to assess a classifier's performance (ROC curves). ROC curves are created by plotting the sensitivity and false positive rate for various threshold values. The area under the ROC curves summarizes the classifier's performance. The following formula can be used to calculate these statistics.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad , \qquad\qquad \text{Eq.4.2 (a)}$$
$$\text{Specificity} = \frac{TN}{TN+FP} \quad , \qquad\qquad \text{Eq.4.2 (b)}$$
$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad , \qquad \text{Eq.4.3 (c)}$$

True positive, negative, and false positive and negative labels are denoted by the letters TP, TN, FP, and FN.

Table 4.2(b): Summary of scores of sensitivity and specificity for various threshold values.

Threshold	Sensitivity	Specificity	F1 score
0.1	0.85 ± 0.002	0.96 ± 0.001	0.90 ± 0.001
0.2	0.87 ± 0.002	0.95 ± 0.001	0.90 ± 0.001
0.3	0.89 ± 0.002	0.94 ± 0.001	0.91 ± 0.001
0.4	0.89 ± 0.002	0.93 ± 0.001	0.91 ± 0.001
0.5	0.90 ± 0.002	0.93 ± 0.001	0.91 ± 0.001
0.6	0.89 ± 0.002	0.93 ± 0.001	0.91 ± 0.001
0.7	0.89 ± 0.002	0.94 ± 0.001	0.91 ± 0.001
0.8	0.87 ± 0.002	0.95 ± 0.001	0.90 ± 0.001
0.9	0.85 ± 0.002	0.96 ± 0.001	0.90 ± 0.001

The number of COVID positive cases as determined by the RT-PCR report is TP, and the machine learning algorithm properly predicts these images as COVID cases.The number of non COVID cases but having infected by common flu/ fever are to be considered FP, however the machine learning model mistakenly predicts few of these cases as positive.The number of non-COVID cases as classified on the pathology report is TN, and the machine learning algorithm accurately predicts these modules as non

724

COVID cases. In the training set, the GGO (Ground Glass Opacity) was not found in the majority of slices. Features were retrieved using the already trained feature extraction block, and then a binary decision was made using fully linked dense layers, similar to the segmentation model.

5. CONCLUSION

This work was focused on reviewing the efficiency of the machine learning architectures for COVID-19 prediction from chest CT images. A neural network was trained on the RADIOMICS features extracted from the segmented images which shall assist the physician in COVID 19 diagnosis and treatment process. The model was trained on the public dataset and data augmentation techniques were followed to increase the number of training samples. A comprehensive experimental evaluation of model performance in terms of specificity, sensitivity, precision-recall curve, and ROC curve was performed. The segmentation model, UNET was trained and tested end-to-end with CT images and corresponding lung masks comprehensive images of early, medium, and final stages of COVID-19 disease.

REFERENCES

1. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology 2020; 296: E115–7.
2. Inui S, Fujikawa A, Jitsu M, Kunishima N, Watanabe S, Suzuki Y, Motoyuki J, et al. Chest CT findings in cases from the cruise ship diamond princess with Coronavirus disease (COVID-19). Radiology 2020; 2: e200110.
3. https://www.worldometers.info/coronavirus/worldwide-graphs/
4. Lei Geng, Siqi Zhang, Jun Tong & Zhitao Xiao (2019) Lung segmentation method with dilated convolution based on VGG-16 network, Computer Assisted Surgery, 24:sup2, 27-33
5. Hua, Panfang, Qi Song, Milan Sonka, Eric A. Hoffman, and Joseph M. Reinhardt. "Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm." In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 2072-2075.
6. Dong, Xue, Yang Lei, Tonghe Wang, Matthew Thomas, Leonardo Tang, Walter J. Curran, Tian Liu, and Xiaofeng Yang. "Automatic multiorgan segmentation in thorax CT images using U-net-GAN." *Medical physics* 46, no. 5 (2019): 2157-2168.
7. Ko, Yen-Fen, and Kuo-Sheng Cheng. "U-Net-based approach for automatic lung segmentation in electrical impedance tomography." *Physiological Measurement* 42, no. 2 (2021): 025002.
8. Armato III, Samuel G., and William F. Sensakovic. "Automated lung segmentation for thoracic CT: impact on computer-aided diagnosis1." Academic Radiology 11, no. 9 (2004): 1011-1021.
9. El-Ba, Ayman, Georgy Gimel'farb, Robert Falk, Trevor Holland, and Teresa Shaffer. "A new stochastic framework for accurate lung segmentation." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 322-330. Springer, Berlin, Heidelberg, 2008.
10. Guiot J, Vaidyanathan A, Deprez L, et al. Development and validation of an automated radiomic CT signature for detecting COVID-19. medRxiv. 2020.
11. Liu Z, Zhang XY, Shi YJ, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer.
12. Chen BT, Chen Z, Ye N, et al. Differentiating peripherally-located small cell lung cancer from non-small cell lung cancer using a CT radiomic approach.
13. Liu Y, Dong D, Zhang L, et al. Radiomics in multiple sclerosis and neuromyelitis optica spectrum disorder. Eur Radiol. 2019; 29:4670–7.

14. J. Zhou, H. Tan, Y. Bai, J. Li, Q. Lu, R. Chen, M. Zhang, Q. Feng, M. Wang, Evaluating the HER-2 status of breast cancer using mammography radiomics features, Eur. J. Radiol. 121 (2019) 108718.

15. R.T. Leijenaar, M. Bogowicz, A. Jochems, F.J. Hoebers, F.W. Wesseling, S. H. Huang, B. Chan, J.N. Waldron, B. O'Sullivan, D. Rietveld, Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study, Br. J. Radiol. 91 (1086) (2018) 20170498.

16. H.-Y. Tsao, P.-Y. Chan, E.C.-Y. Su, Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms, BMC Bioinformatics 19 (9) (2018) 283.

17. E.B. Baum and D. Haussler, What size net gives valid generalization?, *Neural Computation*, vol. 1, no. 1, pp. 151-160, 1989.

18. V. Anantha, et al. "Prediction Of Soil Ph From Remote Sensing Data Using Gradient Boosted Regression Analysis." Journal of Pharmaceutical Negative Results (2022): 29-36.

19. Kumar, M. Sunil, et al. "Deep Convolution Neural Network Based solution for Detecting Plant Diseases." Journal of Pharmaceutical Negative Results (2022): 464-471.

20. Ganesh, D., et al. "Implementation of AI Pop Bots and its allied Applications for Designing Efficient Curriculum in Early Childhood Education." International Journal of Early Childhood 14.03: 2022.

21. Kumar, M. Sunil, et al. "APPLYING THE MODULAR ENCRYPTION STANDARD TO MOBILE CLOUD COMPUTING TO IMPROVE THE SAFETY OF HEALTH DATA." Journal of Pharmaceutical Negative Results (2022): 1911-1917.

22. Prasad, Tvs Gowtham, et al. "Cnn Based Pathway Control To Prevent Covid Spread Using Face Mask And Body Temperature Detection." Journal of Pharmaceutical Negative Results (2022): 1374-1381.1911-1917.

23. P. Sai Kiran. "Power aware virtual machine placement in IaaS cloud using discrete firefly algorithm." Applied Nanoscience (2022): 1-9.

24. Peneti, S., Sunil Kumar, M., Kallam, S., Patan, R., Bhaskar, V. and Ramachandran, M., 2021. BDN-GWMNN: internet of things (IoT) enabled secure smart city applications. Wireless Personal Communications, 119(3), pp.2469-2485.

25. Malchi, Sunil Kumar, et al. "A trust-based fuzzy neural network for smart data fusion in internet of things." Computers & Electrical Engineering 89 (2021): 106901.

26. Sangamithra, B., P. Neelima, and M. Sunil Kumar. "A memetic algorithm for multi objective vehicle routing problem with time windows." 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE). IEEE, 2017.

27. Sunil Kumar, M., and A. Rama Mohan Reddy. "An Efficient Approach for Evolution of Functional Requirements to Improve the Quality of Software Architecture." Artificial Intelligence and Evolutionary Computations in Engineering Systems. Springer, New Delhi, 2016. 775-792.

28. Kumar, T. P., & Kumar, M. S. (2021). Optimised Levenshtein centroid cross-layer defence for multi-hop cognitive radio networks. IET Communications, 15(2), 245-256.

29. Natarajan, V. Anantha, et al. "Segmentation of nuclei in histopathology images using fully convolutional deep neural architecture." 2020 International Conference on computing and information technology (ICCIT-1441). IEEE, 2020.

30. Ganesh, D., et al. "Extreme Learning Mechanism for Classification & Prediction of Soil Fertility index." Journal of Pharmaceutical Negative Results (2022): 37-43.

31. D.Ganesh,M.S Kumar., et al. "Implementation Of Convolutional Neural Networks For Detection Of Alzheimer's Disease",BioGecko, A Journal for New Zealand Herpetology, Vol 12 Issue 01 2023

32.   Huang Y Q, Liang C-H, He L, Tian J, Liang C-S, Chen X, Ma Z L and Liu Z Y 2016 Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancerJ. Clin. Oncol. 34 2157–64.