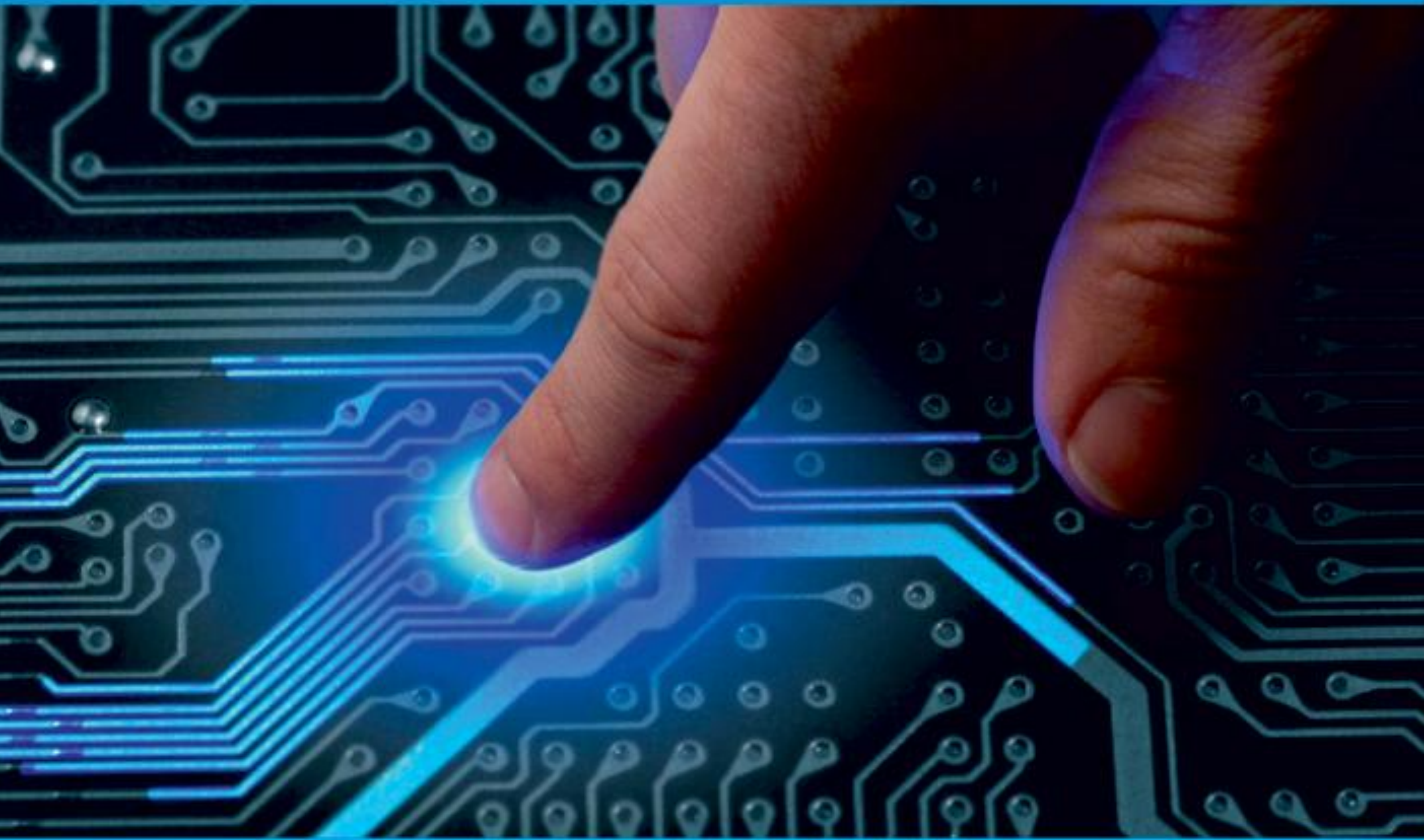
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Special Issue 2, March 2023

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

| International Conference on Recent Innovations in Engineering and Technology (ICRIET'23)|

| Sharadchandra Pawar College of Engineering, Pune, India |

|| Volume 11, Special Issue 2, March 2023 ||

Using CT images, Deep Learning Allows Correct Detection of New Coronavirus (COVID-19)

Pritesh Bhor[1], Sujit Waman[2], Shivam Rain[3], Bhushan Pawar[4], Dr.Monika Rokade[5]

Student, Dept. of Computer Engineering, Sharadchandra Pawar College of Engineering, Otur, India[1]

Student, Dept. of Computer Engineering, Sharadchandra Pawar College of Engineering, Otur, India [2]

Student, Dept. of Computer Engineering, Sharadchandra Pawar College of Engineering, Otur, India [3]

Student, Dept. of Computer Engineering, Sharadchandra Pawar College of Engineering, Otur, India [4]

Professor, Dept. of Computer Engineering, Sharadchandra Pawar College of Engineering, Otur, India [5]

ABSTRACT: A novel coronavirus (COVID-19) has recently appeared as an acute respiratory illness, causing a worldwide pneumonia outbreak. Since the COVID-19 virus spreads rapidly over the world, computed tomography (CT) has become critical for timely diagnosis. As a result, it is critical to create an accurate computer-aided method to help clinicians detect COVID-19-infected individuals using CT imaging. For comparison and modelling, we obtained chest CT scans from 88 patients diagnosed with COVID-19 from hospitals in two Chinese provinces, 100 patients infected with bacteria pneumonia, and 86 healthy people.A deep learning-based CT diagnosis method was created based on the data to identify patients with COVID-19. With an AUC of 0.95, recall (sensitivity) of 0.96, and precision of 0.79, the testing findings demonstrated that our model could accurately distinguish COVID-19 patients from bacteria pneumonia patients. When three types of CT scans were combined, our model achieved a recall of 0.93 and a precision of 0.86 for distinguishing COVID-19 patients from others. Furthermore, our model was able to extract key lesion traits, particularly ground-glass opacity (GGO), which is visually useful for doctor-assisted diagnosis.

KEYWORDS: Deep learning, COVID-19, pneumonia detection, and unsupervised learning

I. INTRODUCTION

A number of pneumonia cases with an undetermined cause were recorded in late December 2019 [1]. Deep sequencing of samples from their lower respiratory tracts identified one novel coronavirus type that resembled the severe acute respiratory syndrome coronavirus (SARS-CoV) [3], which is now known as the 2019 novel coronavirus (COVID-19) by the World Health Organization. We still don't fully understand the human disease's range. Infection symptoms are generally non-specific and include dyspnea, fever, coughing, respiratory symptoms, and viral pneumonia [1]. Due to a lack of nucleic acid detection boxes and low detection rates in the epidemic area, the diagnosis has become a rising issue in major hospitals with the daily growth in the number of newly diagnosed and suspected cases [8]. Therefore, radiography and computed tomography (CT) techniques have become crucial components in the early identification and diagnosis of COVID-19 [9], [10], [11]. High false positive rates were nevertheless a result of the excessive patient load and the paltry number of radiologists [11]. In order to reliably confirm suspected instances, screen patients, and carry out virus surveillance, advanced computer-aided lung CT diagnosis tools are urgently required.

Computer vision techniques, initially used for categorising general images, have been used to medical images, particularly CT images, as a result of the quick growth of artificial intelligence [21]. Convolutional neural networks (CNN) [27] are a popular feed-forward artificial neural network among the currently employed techniques, and many models built on them have showed tremendous promise in collecting feature representations [13]. Because CT pictures are fine-grained, have low interclass variances, and are challenging to discriminate, these generic models do not adequately classify CT images [28].

For comparison and modelling purposes, we have collected chest CT scans from hospitals in two Chinese provinces of 88 patients with COVID-19, 100 patients with bacterial pneumonia, and 86 healthy individuals. We have developed a new deep learning architecture called DRENet to efficiently capture the minute variations in medical images. We integrated the pre-trained ResNet50 with the FPN network to first identify probable lesion sites at various sizes for a particular CT imaging slice. ResNet50 was once more used to extract local features at each region and relational features between regions in accordance with the regions that were found.

II. MATERIALS AND METHODS

For this study, which assessed de-identified data and had no possible danger to patients, our institutional review board waived written informed consent.

1.Data Gathering

The Renmin Hospital of Wuhan University and two associated institutions (the Third Affiliated Hospital and Sun Yat-Sen Memorial Hospital) of the Sun Yat-sen University in Guangzhou supported the study with trustworthy resources. We acquired CT images of 88 COVID-19-infected patients in total, including 76 and 12 patients from the Third Affiliated Hospital and Renmin Hospital of Wuhan University, respectively. The individual laboratories extracted and calculated nucleic acids from the nasopharyngeal swabs of all COVID-19 patients. In the experiments, the novel coronavirus nucleocapsid protein gene (nCoV-NP) and the novel coronavirus open reading coding frame lab (nCoV ORFlab) sequences were compared with the viral nucleic acid sequences discovered by fluorescence RT-PCR. Only patients with positive nucleic acid readings and chest HRCT images that lacked major artefacts or missing images were chosen as patients. We also obtained CT scans of the chest from 86 healthy people and 100 patients with bacterial pneumonia from Sun Yat-Sen Memorial Hospital and Renmin Hospital of Wuhan University for comparison's sake.

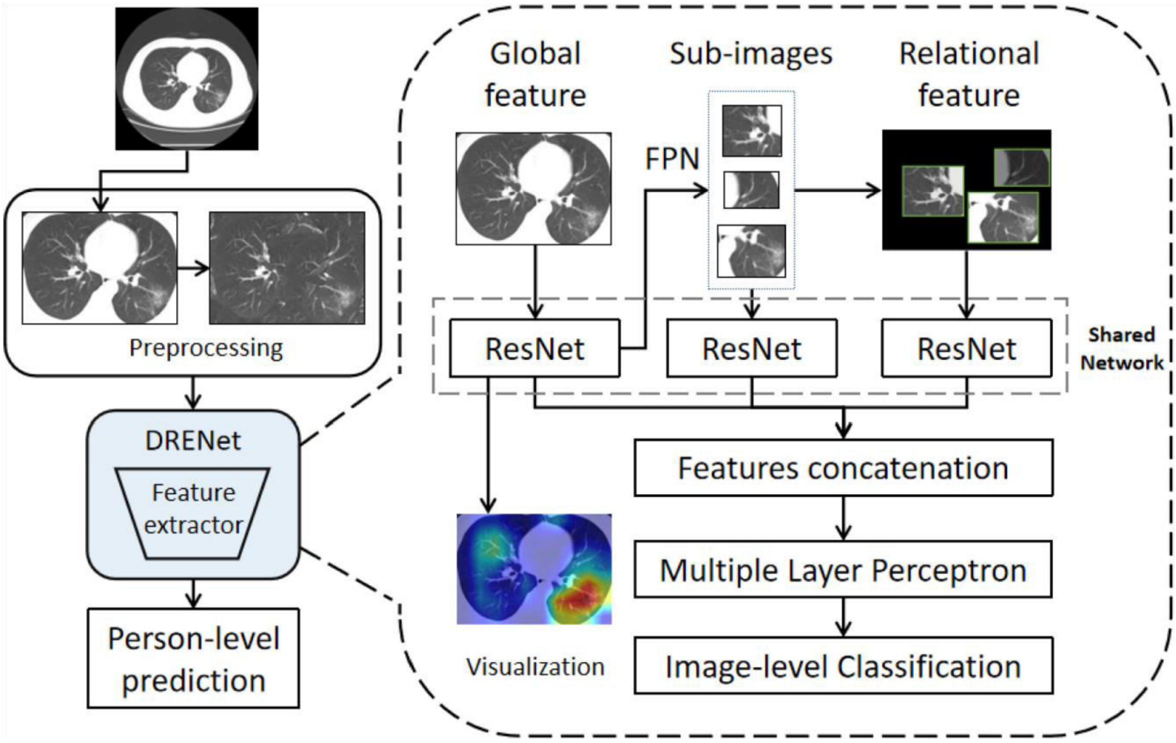
2.DATA PREPROCESSING

While there could be more than 200 photos in a single 3D CT scan of a patient, and neighbouring images are often rather similar, we simply chose 15 typical images at random, or images index [i*ntot/15, I = 0..14] with ntot as the total number of images. The calculation might be accelerated by removing unnecessary images, which would also lessen the effect of hospitals having differing quantities of scanned photos. Since many CT pictures had partial lung structures, we further identified lung structures using the OpenCV package[12] and eliminated any images with lung sections that made up less than 50% of the overall image. The fact that human lung outlines varies significantly is a challenge since deep learning models may overfit the characteristics of lung contour. To address this issue, we We merely filled the image's empty space with its rotating lung regions. Eventually, we included 86 healthy individuals with 708 slices, 100 individuals with bacterial pneumonia and 88 COVID-19 patients with 777 CT pictures in this investigation.

DRENet

The DRENet was developed based on the pre-trained ResNet50 that has been demonstrated to be robust to detect objects in images [13]. To further extract the top-K information from each image, we added the Feature Pyramid Network (FPN) [14]. An attention module is connected to learn the importance based on the extracted details [15].Top-K sub-images are then cropped from the original image based on the regions that have been detected. By multiplying the matching regions of the top-K sub-images with their confidence ratings, the scaled original image was created at the same time. The areas excluded from the top-K photos were set to zero. These K sub-images were sent into ResNet50 together with the freshly created image to extract the appropriate characteristics.

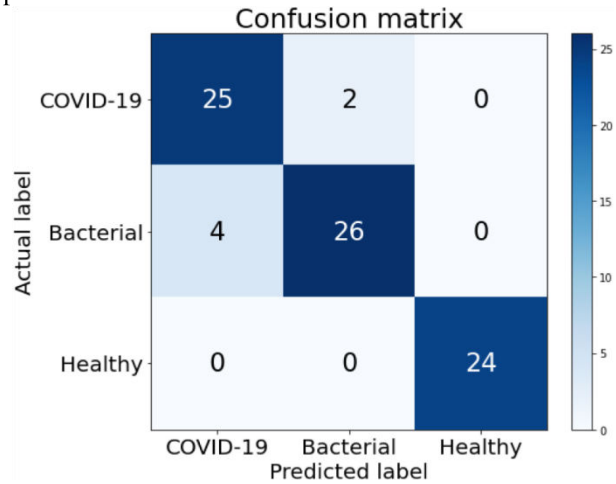
In order to predict the image-level state, we finally combined the features obtained from the original image, the top-K sub-images, and the created image into a 1-D vector. Our model, which is based on FPN, can not only give the doctor a person-level prediction score but also interpret the prognosis by weighing each pixel of original images.



The proposed training architecture including (1) Preprocessing: the CT images are first preprocessed to remove boundary regions around the lungs, and to further fill the blank regions around two lungs with its rotational lung areas (for model training only). (2) Image-level classifications by DRENet: each CT image is input to the pretrained ResNet50 to extract global features, from which FPN module is employed to identify top-K lesion regions. According to the top-K regions, the shared ResNet is utilized again to extract local features within the sub-images and relational features between the sub-images. These features are concatenated with the learned global features to input into MLP for the image-level prediction. (3) Person-level prediction: the image-level predictions will be aggregated for person-level prediction

Aggregation

For each patient, the predictions were made on each image slice, and the image-level scored results were aggregated for personlevel prediction. Here, the mean pooling was used to integrate the image-level scores into the morbidity of each person as the personlevel prediction.



III.IMPLEMENTATION AND EVALUATION

In our experiments, we fixed K = 3 meaning to extract 3 sub-images from each input image. For practical purposes, we designed the models with two tasks: discriminating COVID-19-infected patients from the healthy people, and separating COVID-19 patients from bacterial patients and healthy controls. For each task, we employed the person-level split strategy following the LUNA16 competition [16] by using random splits of 60/10/30 percent for training, validation, and test sets, respectively. The training set was used to train models, and the validation set was used to optimize the hyperparameters for the best performance. The final optimal model was independently assessed on the test set. We computed the measured metrics as

DenseNet performed the worst with F1-score of 0.88. Fig. 4 shows the confusion matrix of the DRENet on the pneumonia three-class classification. We observed that the model can discriminate all healthy persons from pneumonia patients while make only six wrong predictions in classifying the bacterial pneumonia and viral pneumonia (COVID-19). When considering the separation of COVID-19 patients from others, the model achieved a recall of 0.93 and precision of 0.86.

By examining two COVID-19 patients who were wrongly predicted to be caused by bacteria, we found their respective predicted scores were 0.46 and 0.39 in the patient level. These scores were both marginally lower than the threshold of 0.5. As shown in Fig. 5, although the patient C has a score of 0.46 in the patient level, three most suspected slice images (the highest scores in the image level) all have scores greater than 0.5 in the image level. Therefore, DRENet could still assist doctors to make correct decisions by visualizing key images in occasionally wrong cases. On the other hand, the patient D of bacterial pneumonia was wrongly predicted as the COVID-19 patient. The misjudgements might be caused by the GGO abnormalities appearing in the patient of bacterial pneumonia for unknown reasons, as shown in the rectangles.

IV.DISCUSSION

The outbreak of COVID-19 has resulted in a world-wide panic recently. There are no specific signs for the infected patients. Those who have fallen into ills were reported with the ground-glass opacity. Thus, CT becomes the most convenient and fastest diagnosis approach. However, difficulties were met due to lacks of quantified diagnosis criteria and of experienced radiologists. It is essential to develop a publicly available model to assist the diagnosis by artificial intelligence.

In this study, we have developed a new deep learning architecture to construct the prediction model based on 88 patients diagnosed with the COVID-19, 100 patients infected with bacteria pneumonia, and 86 healthy persons from three hospitals in China. Our models achieved a recall of 96 percent and a precision of 93 percent for COVID-19 patients against bacteria pneumonia patients. When training a model to separate three classes simultaneously, the tri-classification model achieved a recall of 93 percent and a precision of 86 percent for the COVID-19 patients. Due to the high infectivity of COVID-19 virus, a high recall is essential for correctly recognizing COVID-19 patients to prevent the spreading of virus. The model developed here is appropriate for assisting clinical diagnoses considering the 60–70 percent recall rate by nucleic acid examinations. More importantly, our method also provides interpretation of the prediction through recognized lesion regions and detailed heat map. Since clinical diagnosis is of low endurance to wrong predictions, such interpretation is essential for assisting doctors to make decisions.

Since the training data is still a small data, the model can't effectively solve the batch effect to make accurate predictions for other sources of data, a well known problem in AI diagnosis through medical images [CITE]. One possible way for this problem is to accumulate data from multiple hospitals. Alternatively, the model can be fine-tuned or retrained by using data from hospitals that need to use the model.

V.CONCLUSION

In conclusion, our study demonstrated the feasibility of a deep learning approach to assist doctors in detecting patients with COVID-19 and automatically to identify the potential lesions from CT images. With the accurate discrimination of the COVID-19 patients, the proposed system may enable a rapid and accurate identification of patients.

There are also several weakness points in our work. One major problem is the dataset. We accomplished the work at the very early stage of the COVID pandemic, where there are only very few samples to build such a deep learning prototype. Second, due to the issue of batch effect, our method achieved good results on the datasets from original source data collection hospital but was not able to predict the external data directly. Fortunately, it can be alleviated by finetuning the model with the images from target sources. In the future, we may include a pre-trained model with more image sources for more accurate diagnosis of COVID-19.

VI.ACKNOWLEDGMENTS

The work was supported in part by the National Key R&D Program of China under Grant 2018YFC1315405, the National Natural Science Foundation of China under Grants U1611261, 61772566, 81801132 and 81871332, Guangdong Frontier and Key Tech Innovation Program under Grants 2018B010109006 and 2019B020228001, the Natural Science Foundation of Guangdong, China under Grant 2019A1515012207, and Introducing Innovative and Entrepreneurial Teams under Grant 2016ZT06D211. Ying Song, Shuangjia Zheng, Liang Li, and Xiang Zhang are contributed equally.

REFERENCES

1. C. Huang et al., "Clinical features of patients infected with 2019 novel Coronavirus in Wuhan, China," Lancet, vol. 395, no. 10223, pp. 497–506, 2020.
2. WHO website WM, Accessed: Feb. 5, 2020. [Online]. Available: https:// www.who.int
3. R. Lu et al., "Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding," Lancet, vol. 395, 565–574, 2020.
4. Q. Li et al., "Early transmission dynamics in Wuhan, China, of novel Coronavirus–infected pneumonia," New Engl. J. Med., vol. 382, pp. 1199–1207, 2020.
5. D. S. Hui et al., "The continuing 2019-nCoV epidemic threat of novel Coronaviruses to global health–The latest 2019 novel Coronavirus outbreak in Wuhan, China," Int. J. Infect. Dis., vol. 91, pp. 264–266, 2020.
6. N. Zhu et al., "A novel Coronavirus from patients with pneumonia in China," New Engl. J. Med., vol. 382, pp. 727–733, 2020.
7. WHO reports, WHONC-ns, Assessed: Feb. 6th, 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/ situation-reports/
8. D. K. Chu et al., "Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of Pneumonia," Clin. Chem., vol. 66, pp. 549–555, 2020.
9. J. Lei, J. Li, X. Li, and X. Qi, "CT imaging of the 2019 novel coronavirus (2019-nCoV) Pneumonia," Radiology, vol. 295, 2020, Art. no. 200236.
10. H. Shi, X. Han, and C. Zheng, "Evolution of CT manifestations in a patient recovered from 2019 novel Coronavirus (2019-nCoV) Pneumonia in Wuhan, China," Radiology, vol. 295, 2020, Art. no. 200269.
11. Monika D.Rokade ,Dr.Yogesh kumar Sharma,"Deep and machine learning approaches for anomaly-based intrusion detection of imbalanced network traffic." IOSR Journal of Engineering (IOSR JEN), ISSN (e): 2250-3021, ISSN (p): 2278-8719
12. Monika D.Rokade ,Dr.Yogesh kumar Sharma"MLIDS: A Machine Learning Approach for Intrusion Detection for Real Time Network Dataset", 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), IEEE
13. Monika D.Rokade, Dr. Yogesh Kumar Sharma. (2020). Identification of Malicious Activity for Network Packet using Deep Learning. International Journal of Advanced Science and Technology, 29(9s), 2324 - 2331.
14. Sunil S.Khatal ,Dr.Yogesh kumar Sharma, "Health Care Patient Monitoring using IoT and Machine Learning.", IOSR Journal of Engineering (IOSR JEN), ISSN (e): 2250-3021, ISSN (p): 2278-8719
15. Sunil S.Khatal ,Dr.Yogesh kumar Sharma, "Data Hiding In Audio-Video Using Anti Forensics Technique ForAuthentication ", IJSRDV4I50349, Volume : 4, Issue : 5
16. Sunil S.Khatal Dr. Yogesh Kumar Sharma. (2020). Analyzing the role of Heart Disease Prediction System using IoT and Machine Learning. International Journal of Advanced Science and Technology, 29(9s), 2340 - 2346.
17. Dubois B, Padovani A, Scheltens P, Rossi A, DellAgnello G. Timely diagnosis for Alzheimers disease: a literature review on benefits and challenges. Journal of Alzheimer's disease. 2016 Jan 1;49(3):617-31.

19.Monika D. Rokade;Sunil S. Khatal"Detection of Malicious Activities and Connections for Network Security using Deep Learning",2022 IEEE Pune Section International Conference (PuneCon),Year: 2022 | Conference Paper | Publisher: IEEE

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379

doi crossref

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details

www.ijircce.com