International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Deep Learning System for COVID-19 Diagnostic and Predictive Analysis

S. Sivasankara Rao[*1], E.Madhusudhana Reddy[2], Shashi Bhushan Tyagi[3]

[*1] Research Scholar, CSE, Shri JJTUniversity, Associate Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

[*2] Professor & Principal, Department of CSE, Bhoj Reddy Engineering College for Women, Hyderabad, Telangana, India

[*3] Professor, Department of CSE, Shri JJTUniversity, Rajasthan, India

ARTICLEINFO

Article History:

Accepted: 10 Feb 2023
Published: 23 Feb 2023

Publication Issue
Volume 10, Issue 1
January-February-2023

Page Number
402-407

ABSTRACT

Beginning in early 2020, COVID-19, a health emergency and existential threat to society, began to spread globally. The current healthcare approach to combat COVID-19 has the potential to be considerably improved by automated lung infection detection utilising computed tomography (CT) images. To segment contaminated areas from CT slices, however, there are a number of difficulties, including low intensity and a wide range of infectious features when comparing infected and healthy tissues. Additionally, it is not feasible to collect a big amount of data quickly, which inhibits the deep model from being trained. To overcome these difficulties, a novel COVID-19 Lung Infection Deep Network Segmentation is proposed to autonomously separate unhealthy areas from slices of a chest CT image. This study provides a segmentation technique for Ground Glass Opacity or ROI identification in CT images caused by corona viruses. The area of interest was categorised down to the pixel level using a modified Unet model structure. Instead of the time consuming RT-PCR test, CT scans can be utilised to diagnose COVID-19. Using this segmentation method, doctors were able to diagnose COVID-19 more quickly, precisely, and consistently.

Keywords: CT scan, RT-PCR, ROI identification, UNet

I. INTRODUCTION

A health and financial crisis was caused by the Covid-19 (Coronavirus) virus's global spread. It is also highly transmissible and infectious. It also comes in different waves. One cannot anticipate when it will end. The lungs are largely impacted by the second wave of COVID-19, which is brought on by the Delta variation. The present healthcare strategy to reduce COVID-19 has the potential to be considerably improved by automated lung infection diagnosis utilizing computed tomography (CT) images. Many obstacles, such as the vast range of infection features and the diffuse contrast between infected and healthy tissues, make it difficult to separate ill regions from healthy ones in CT slices. The suggested CNN+ Unet-based solution addresses these problems. To mark the issue of enormous data requirements, a semi-supervised technique was then adopted. The proposed Unet-like model enhances the predicted region of interest's receptive field, allowing for the collection of more data and enhancing the model's capacity to identify edges. The performance of the recommended model performs well due to its quantitative results when compared to the fundamental Unet technique and other cutting-edge models. Because CT screening provides advantages over X-rays, such as providing a threedimensional image of the lung, it is the best technique. According to recent studies, CT slices can reveal infection related symptoms like Ground Glass Opacity (GGO) in the early stages and lung accumulation in the late stages. It would be beneficial to qualitatively describe COVID-19 and monitor longitudinal changes in CT slices in order to combat the infection successfully. However, manually identifying lung infections requires a significant amount of time and work. Additionally, radiologists' annotation of infections is a highly subjective endeavour that frequently takes into account their own biases and clinical experiences. We suggest a special COVID-19 Lung Infection Segmentation Deep Network (Inf-Net) for CT slices to address the aforementioned problems. [1].When diagnosing a lung infection, medical professionals must first roughly pinpoint the diseased area in order to appropriately extract its contour from the other features. The danger of dying increases with lung infections. Only a small portion of doctors are capable of early detection during a pandemic, which could lower the likelihood of a death.

II. RELATED WORK

The diagnosis of lung issues frequently uses CT imaging. We suggest a new Deep Network (Inf-Net) for COVID-19 Lung Infection Segmentation for CT Slices to address the former problems [1]. Recent performances of numerous compositions have been well welcomed. These methods usually employ an extract feature classifier for nodule segmentation in lung CT. We segment the COVID-19 infection zones into sub regions in our work to track and assess the disease's evolution. Despite being able to discover the anomalous region, the (unsupervised)anomaly detection/ segmentation was unable to determine whether it was connected to COVID-19.The semi-supervised model, which is better suited for evaluating COVID-19 based on the weakly labelled data, could separate the target region from other anomalies. Semi-supervised learning (SSL) seeks to enhance model performance with a big amount of unlabeled data and a small amount of labelled data. The SSL technique is increasingly used for deep neural network training. In order to reduce the loss on the labelled data, these techniques frequently combine a supervised loss applied to the labelled data and an unsupervised loss applied to either the unlabelled data alone or both the labelled and unlabelled data. The use of a cross-entropy loss, which is calculated using the unlabeled data with fictional labels. and is regarded as an additional supervision loss, is therefore advised.

III. PROPOSED SYSTEM

In medical image processing, both segmentation and classification studies can be performed. The purpose of segmentation is to identify Regions of Interest (ROI) such as organs or medical abnormalities by labeling each pixel in a medical image. On the other hand, medical image classification aims to label a complete image into predefined classes, and for this, it first divides the image into specified sub-domains. Thus, only the ROI is examined in medical images divided into sub-areas, and classification is made according to these areas. However, ROI areas cannot always be obtained in medical images. In such a case, first ROI should be segmented and then the classification algorithm should be applied. In the COVID-19 classification dataset (Dataset-2) used in the study, there are no ROI areas as masked image areas. For this reason, first of all, a successful segmentation algorithm based on deep learning was proposed by using Dataset-1 that has ground truth masked images. After the success of the proposed segmentation algorithm has been proven on Dataset-1, the resulting segmentation model has been used on the Dataset-2 to generate mask which are then used for classification. So, ROI of lung was segmented and

According to the classification results, images could be classified more accurately in a short time compared to the literature. Since both segmentation and classification were made within the scope of the study, the literature studies and results in these two fields were examined separately and comparisons were made with the results. Our proposed method can be summarized in two main stages (Figure 1) as segmentation and classification. The data was preprocessed before feeding it into the model. After preprocessing, the data has been fed into a deep learning model to obtain a binary mask of lesions in COVID19 CT images in Stage-1. For this purpose; UNet, UNet++, SegNet and PsPNet were used both separately and as hybrids with GAN, to automatically segment infected areas from chest CT slices. The resulting binary mask was then fed into Stage2, where a CNN, PatchCNN and CapsNet models were used to predict whether or not a patient has COVID-19 or not. The binary masks that contain COVID-19 have white areas on a black background while those without COVID-19 are completely black.

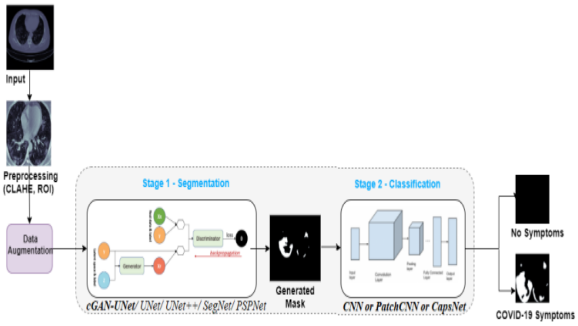

Figure 1. Steps of the proposed method

3.1 ALGORITHM

1) Acquiring images from Dataset-1

2) Extracting the image slices from original images as follows: x ∈ D: {xi}i=1 to 2112

3) Converting the images to gray-scale, resizing to 256x256 pixels and normalizing with min max scaling. (xpre = Preprocessing(x))

4) Applying CLAHE to enhance the images (xCL = CLAHE(xpre))

5) Extracting ROI by drawing contours over image and then cropping out the closed boundary ($xROI$ = ROI(xCL))

6) Applying data augmentation strategies using standard methods like rotation, shearing, zooming and horizontal flipping

7) Splitting data into 80% for training and 20% for testing.

8) Constructing the proposed cGAN-UNet model, training the model using the training dataset to get segmentation masks that will be used as input to our classifiers.

9) Evaluating, testing and comparing cGAN-UNet segmentation model to other baseline models.

10) Using trained model to generate segmentation masks on Dataset-2 and converting them to binary masks

11) Training the CNN / PatchCNN/ CapsNet model using generated binary masks

13) Evaluating, testing and comparing the performance of all classifiers.

IV. RESULTS

In this study, we proposed a novel fully automatic DL system using raw chest CT image to help COVID-19 diagnostic and prognostic analysis. To let the DL system mine lung features automatically without involving any time-consuming human annotation, we used a two-step transfer learning strategy. Firstly, we collected 4106 lung cancer patients with both CT image and EGFR gene sequencing. Through training in this large CT-EGFR dataset, the DL system learned hierarchical lung features that can reflect the associations between chest CT image and micro-level lung functional abnormality. Afterwards, we collected a large multi-regional COVID-19 dataset (n=1266) from six cities or provinces to train and validate the diagnostic and prognostic performance of the DL system. The good diagnostic and prognostic performance of the DL system illustrates that DL could be helpful in the epidemic control of COVID-19 without adding much cost. Given a suspected patient, CT scanning can be acquired within minutes. Afterwards, this DL system can be applied to predict the probability the patient has COVID-19. If the patient is diagnosed as COVID-19, the DL system also predicts their prognostic situation simultaneously, which can be used to find potential high-risk patients who need urgent medical resources and special care. More importantly, this DL system is fast and does not require human-assisted image annotation, which increases its clinical value and becomes more robust. For a

typical chest CT scan of a patient, the DL system takes less than 10 s for prognostic and diagnostic prediction. During building and training the DL system, we did not involve any human annotation to tell the system where the inflammatory area was. However, the DL system managed to automatically discover the important features that are strongly associated with COVID-19. In figure 4, we visualised the DL discovered suspicious lung areas that were used by the DL system for inference. These DL discovered suspicious lung areas had high overlap with the actual inflammatory areas that are used by radiologists for diagnosis. In previous studies, some radiological features such as ground-glass opacities, crazy-paving pattern and bilateral involvement have been reported to be important for diagnosing CVOID-19 [7]. In the DL discovered suspicious lung areas, we also observed these radiological features. This demonstrates that the high-dimensional features mined by the DL system can probably reflect these reported radiological finding.

Recently, deep learning methods with different processes and models were reported to diagnose COVID-19 using CT images. These methods can be classified into three types. 1) Using manually or automatically segmented lesions for diagnosis. WANG et al. used manually annotated lesions as ROI, and a modified ResNet34 model combined with decision tree and AdaBoost classifiers was used to diagnose COVID-19. To avoid time-consuming lesion annotation by radiologists, automatic lesion segmentation models were used in further studies. Afterwards, 3-dimensional CNN models such as 3DResNet were used to diagnose COVID-19 using the lesion images. 2) Using 2-dimensional lung image slices to train the DL model. Since lesions can be distributed in many locations in lungs, and automatic lesion segmentation may not guarantee very high precision. More studies used the whole lung image slices for analysis. In the study by SONG et al., a feature pyramid network using ResNet50 as the backbone was used to analyse 2-dimensional image slices of the whole lung area. Similarly, JIN et al. used DeepLabv1 and LI et al. used U-Net to segment lung from CT images, and then used the 2-dimensional ResNet model to analyse image slices of lung area. 3) Using a 3-dimensional DL model to analyse whole lung in CT images. To consider 3-dimensional information of the whole lung, ZHENG et al. used the 3DResNet model to analyse the 3-dimensional lung area in CT images. Compared with these studies, our study has three main differences. 1) We used the 3-dimensional bounding box of the whole lung as ROI instead of only using lesions or segmented lung fields. Since lesion segmentation may not guarantee a very high accuracy, inaccurate lesion segmentation may cause information loss. Compared with segmenting lung lesion, lung segmentation is easier, and analysing the whole lung can mine more information. However, different with the methods using only the segmented lung areas , we used the 3-dimensional bounding box of lung as ROI. In figure S2, we illustrated the lung segmentation results. In most situations, the lung segmentation method generated good results.

However, for some patients with severe symptoms and consolidation lesions, the performance of the lung segmentation method may be affected. Consequently, we used the 3-dimensional bounding box of the segmented lung mask as ROI, which ensures the lung-ROI covering the complete lung area. Combined with the non-lung area suppression strategy, the lung-ROI can reserve complete lung area, and suppress images outside lung area. 2) We used a large auxiliary dataset including chest CT images of 4106 patients to pre-train the proposed COVID-19Net, making it learn lung features. Many existed studies used DL models pre-trained in ImageNet dataset, this may increase the generalisation ability of the DL model. However, natural images in the ImageNet dataset have large difference to chest CT images. Consequently, using a chest CT dataset for auxiliary training (pre-training) enables the DL model learn features that are more specific to chest CT images. 3) Most studies used a small dataset and randomly selected data for validation. To assess the generalisation ability of the deep learning model, we used a large dataset and two independent validation sets from different regions.

Despite the good performance of the DL system, this study has several limitations. Firstly, there are other prognostic end events such as death or admission to an intensive care unit, and they were not considered in this study. Secondly, the management of severe and mild COVID-19 are different, thereby, exploring the prognosis of COVID-19 in these two groups separately should be helpful. However, CT images of different slice thickness were included in this study. In the future, we will use a generative adversarial network to convert CT images of different slice thickness into CT images with a unified slice thickness, which may further improve the diagnostic performance of the DL system.

V. CONCLUSION AND FUTURE WORK

On test data, our trained model scored a tremendous benchmark. The segmentation of lung infection mask is predicted with the highest level of accuracy (91.05%). Future model architectural improvements could raise this threshold, which would revolutionise the healthcare sector.

The experiments show that the COVID-19 testing of our models is successful. In the future, we'll pay close attention to how serious COVID-19 is judged to be and work to extract more useful data from CT scans to fight the epidemic. The experiments show that the COVID-19 testing of our models is successful. In order to combat the outbreak, we'll continue to closely monitor COVID-19's severity and work to extract more useful information from CT scans. In order to identify pertinent elements in the CT scans and assist clinical practitioners in patient selection, additional explanatory investigations will be conducted on the models. This study will provide light on the COVID-19 detecting process. The study is still in the theoretical research stage, and the models have not yet been tested in actual clinical operations, despite the system's successful performance on publicly available datasets. Our technology is still in clinical testing situations and talk to medical experts to find out how they use it and what they think of the models. Consequently, in our future work, further enhance the models will be done.

VI. REFERENCES

[1]. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. JAMA 2020; 323: 1061–1069.

[2]. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 2020; 382: 1199–1207.

[3]. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med 2020; 8: 475–481.

[4]. Ji Y, Ma Z, Peppelenbosch MP, et al. Potential association between COVID-19 mortality and health-care resource availability. Lancet Glob Health 2020; 8: e480.

[5]. Xie X, Zhong Z, Zhao W, et al. Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. Radiology 2020; 296; E41–E45.

[6]. Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 2020; 296; E32–E40.

[7]. Zhang S, Li H, Huang S, et al. High-resolution computed tomography features of 17 cases of coronavirus disease 2019 in Sichuan province, China. Eur Respir J 2020; 55; 2000334.

[8]. Wang L, Gao Y-h, Zhang G-J. The clinical dynamics of 18 cases of COVID-19 outside of Wuhan, China. Eur Respir J 2020; 55; 2000398.

[9]. Shi H, Han X, Jiang N, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. Lancet Infect Dis 2020; 20; 425–434.

[10]. Lee EY, Ng M-Y, Khong P-L. COVID-19 pneumonia: what has CT taught us? Lancet Infect Dis 2020; 20: 384–385.

[11]. Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. Eur Respir J 2019; 53: 1800986.

[12]. Walsh SL, Calandriello L, Silva M, et al. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. Lancet Respir Med 2018; 6: 837–845.

[13]. Walsh SL, Humphries SM, Wells AU, et al. Imaging research in fibrotic lung disease; applying deep learning to unsolved problems. Lancet Respir Med 2020; in press [https://doi.org/10.1016/S2213-2600(20)30003-5].

[14]. Angelini E, Dahan S, Shah A. Unravelling machine learning: insights in respiratory medicine. Eur Respir J 2019; 54: 1901216.

[15]. Wang S, Zhou M, Liu Z, et al. Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. Med Image Anal 2017; 40: 172–183.

Cite this article as :

S. Sivasankara Rao, E.Madhusudhana Reddy, Shashi Bhushan Tyagi, "Deep Learning System for COVID-19 Diagnostic and Predictive Analysis", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 1, pp. 402-407, January-February 2023.

Journal URL : https://ijsrset.com/IJSRSET2310352