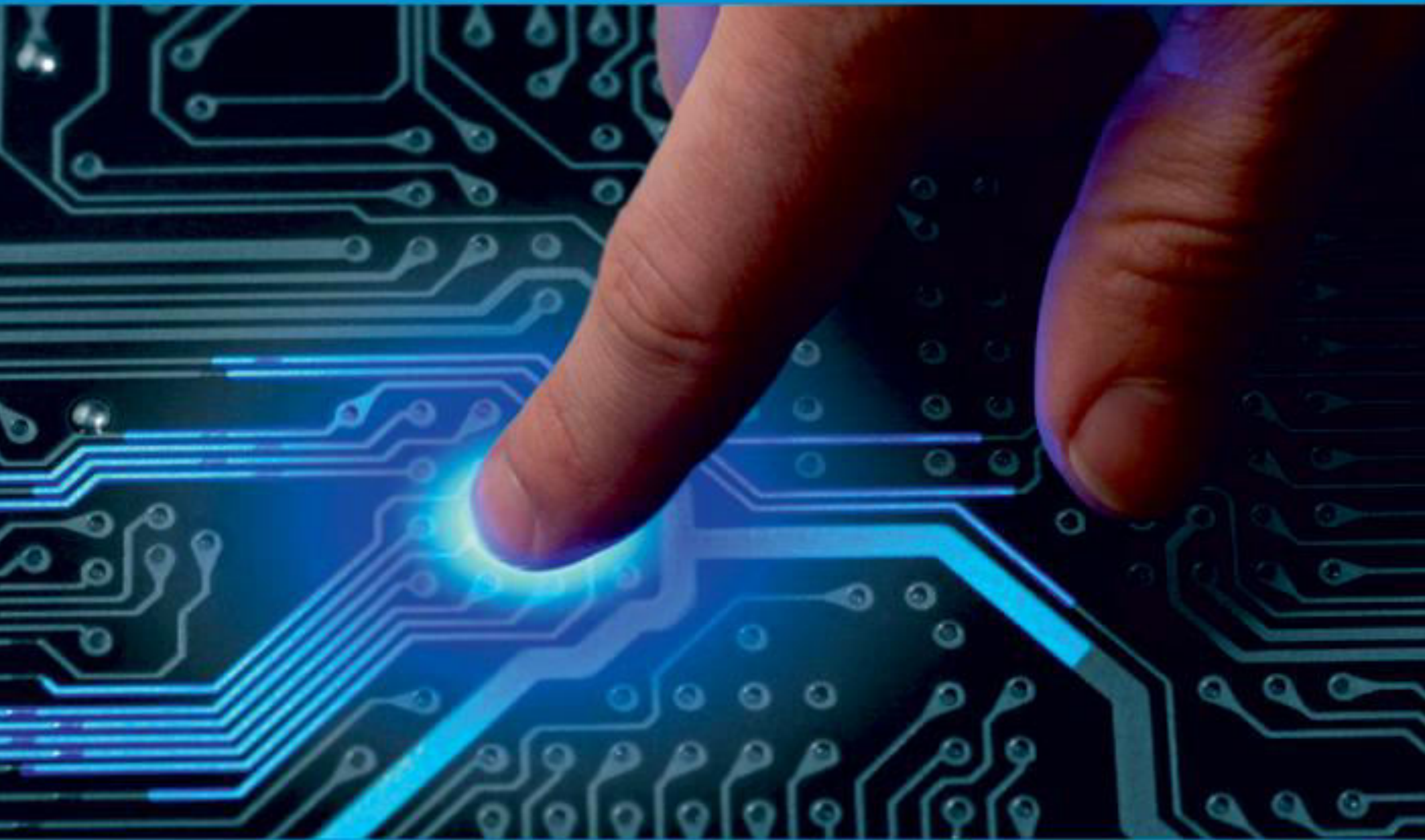
IJIRCCE

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



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005137|

Covid 19 Lung Infection Detection System Using Machine Learning

Akanksha Dhaybar, Prachi Chakranarayan, Shubhangi Gonte, Aaditya Chouhan,

Prof. Ashwini Kamble

Department of Computer Engineering, K J College of Engineering & Management Research, Pune, India

ABSTRACT*:* The coronavirus disease (COVID-19) is rapidly spreading all over the world, and has infected more than 1,966,000 people in more than 200 countries and territories as of October, 2020. Detecting COVID-19 at early stage is essential to deliver proper healthcare to the patients and also to protect the uninfected population. To this end, we develop a framework to automatically diagnose COVID-19 from the community acquired pneumonia (CAP) in chest computed tomography (CT). In particular, we propose a Noise Robust Segmentation approach with a recurrent neural network (RNN) to focus on the infection regions in lungs when making decisions of diagnoses. Note that there exists imbalanced distribution of the sizes of the infection regions between COVID-19 and CAP, partially due to fast progress of COVID-19 after symptom onset. Our framework is evaluated upon the largest multi-centre CT data for COVID-19 from hospitals.

KEYWORDS : COVID-19, Deep Learning, Recurrent neural network, noisy label, segmentation, pneumonia

I. INTRODUCTION

The coronavirus disease 2019 (COVID-19) has become a global pandemic since the beginning of 2020. The disease has been regarded as a Public Health Emergency of International Concern (PHEIC) by the World Health Organization (WHO) and the end of January 2020. Up to April 10, 2020, there have been more than 1.5 million cases of COVID19 reported globally, with more than 92 thousand deaths. The most common symptoms of COVID-19 patients include fever, cough and shortness of breath, and the patients typically suffer from pneumonia. Computed Tomography (CT) imaging plays a critical role for detection of manifestations in the lung associated with COVID-19, where segmentation of the infection lesions from CT scans is important for quantitative measurement of the disease progression in accurate diagnosis and follow-up assessment. As manual segmentation of the lesions from 3D volumes is labour-intensive, time-consuming and suffers from inter- and intra-observer variabilities, automatic segmentation of the lesions is highly desirable in clinic practice.

Despite its importance for diagnosis and treatment decisions, automatic segmentation of COVID-19 pneumonia lesions from CT scans is challenging due to several reasons. First, the infection lesions have a variety of complex appearances such as Ground-Glass Opacity (GGO), reticulation, consolidation and others. Second, the sizes and positions of the pneumonia lesions vary largely at different stages of the infection and among different patients. In addition, the lesions have irregular shapes and ambiguous boundaries, and some lesion patterns such as GGO have a low contrast with surrounding regions.The goal of this work is three-fold. First, to deal with noisy annotations for training RNNs to segment COVID19 pneumonia lesions, we propose a novel noise-robust Dice loss function, which is a combination and generalization of MAE loss that is robust against noisy labels and Dice loss that is insensitive to foreground background imbalance. Second, we propose a novel noise-robust learning framework based on self-ensembling of RNNs where an Exponential Moving Average (EMA, a.k.a. teacher) of a model is used to guide a standard model (a.k.a. student) to improve the robustness. Differently from previous self ensembling methods for semi-supervised learning and domain adaptation, we propose two adaptive mechanisms to better deal with noisy labels: adaptive teacher that suppresses the contribution of the student to EMA when the latter has a large training loss, and adaptive student that learns from the teacher only when the teacher outperforms the student. Thirdly, to better deal with the complex lesions, we propose a novel COVID-19 Pneumonia Lesion segmentation network (COPLE-Net) that uses a combination of max-pooling and average pooling to reduce information loss during down sampling, and employs bridge layers to alleviate the semantic gap between features in the encoder and decoder.

II. LITERATURE REVIEW

1. In summary [1], this experiment examined three common types of label noise in medical image datasets, as well as the relative effectiveness of several approaches for mitigating label noise's negative impact. Label

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005137|

noise in medical imaging has a variety of sources, statistics, and strengths, and this study demonstrates that the effects of label noise should be carefully analysed when training deep learning algorithms. This necessitates additional research and the development of robust models and training algorithms.

2. On the first layer of the CNN model, the most primitive [2] building blocks that comprise the images are located; these building blocks correspond to the motifs. By applying filters to the images, the CNN detects these motifs. Each filter is composed of pixels that have the same shape as the corresponding motif. The first layer filters in this example correspond to the letters of the alphabet. Each filter is sequentially shifted to each location in the image, and the degree to which the image's local properties match the filter at each location is measured, a process known as convolution. Convolution produces a new array (or new image) called a featuremap as a result of this process. The degree to which the filter matches each local region in the original image is quantified by feature maps. If there are N first layer filters, the convolutional process generates N 2D feature maps.
3. The purpose of this study [3] was to evaluate a quantitative CT Image Parameter, defined as the percentage of lung opacification (QCT-PLO), that was automatically calculated using a deep learning tool. We evaluated QCT-PLO in covid -19 patients at baseline and on follow-up scans, with an emphasis on cross-sectional and longitudinal differences in patients with varying degrees of clinical severity.
4. The diagnosis of 2019-nCoV pneumonia [4] was made based on epidemiologic characteristics, clinical manifestations, chest images, and laboratory findings. After three days of treatment with interferon inhalation, the patient's clinical condition deteriorated, with progressive pulmonary opacities discovered on repeat chest CT.
5. This article [5] proposes that a deep learning model can accurately detect and distinguish COVID -19 from community-acquired pneumonia and other lung diseases.
6. The author [6] used this pipeline to compare the evolution of two confirmed COVID-19 cases from Wuhan, China, who received similar supportive therapy. Figure 1 depicts the favourable evolution of a 48-year-old woman imaged at four time points over a 16-day period, whereas Figure 2 depicts the disease progression of a 44-year-old man over a 12-day period, particularly between the second and third studies.
7. Authors present UNet++, a new, more powerful architecture for medical image segmentation, in this paper [7]. This architecture is essentially a densely supervised encoder-decoder network, with the encoder and decoder sub-networks connected via a series of nested, dense skip paths.
8. A cluster of patients with pneumonia of unknown origin was linked to a seafood wholesale market in Wuhan, China, in December 2019[8]. Unbiased sequencing of samples from pneumonia patients identified a previously unknown betacoronavirus. We isolated a novel coronavirus, 2019-nCoV, from human airway epithelial cells. This virus formed a new clade within the subgenus sarbecovirus, Orthocoronavirinae subfamily. In contrast to MERS-CoV and SARS-CoV, 2019-nCoV is the seventh member of the human coronavirus family.
9. This study [9] describes the same population genetic dynamic as the SARS 2003 epidemic, emphasising the critical need for the development of effective molecular surveillance strategies for Betacoronavirus in animals and Rhinolophus in the bat family.
10. This article [10] discusses how artificial intelligence (AI) can be used to provide safe, accurate, and efficient imaging solutions for COVID-19 applications. COVID-19 covers the entire pipeline of AI-enabled imaging applications, including intelligent imaging platforms, clinical diagnosis, and pioneering research. Two imaging modalities, X-ray and CT, are used to demonstrate the efficacy of AI-assisted medical imaging in the diagnosis of COVID-19.

III. PROPOSED SYSTEM

Despite several recent studies on automatic segmentation of COVID-19 pneumonia lesions from CT scans, previous work has relied heavily on off-the-shelf models such as U-Net and a standard training procedure that ignores the presence of noisy labels. The purpose of this work is to develop a more advanced SVM model for the challenging segmentation task and to attempt to overcome the effect of noisy annotations on segmentation performance.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005137|

A. Architecture

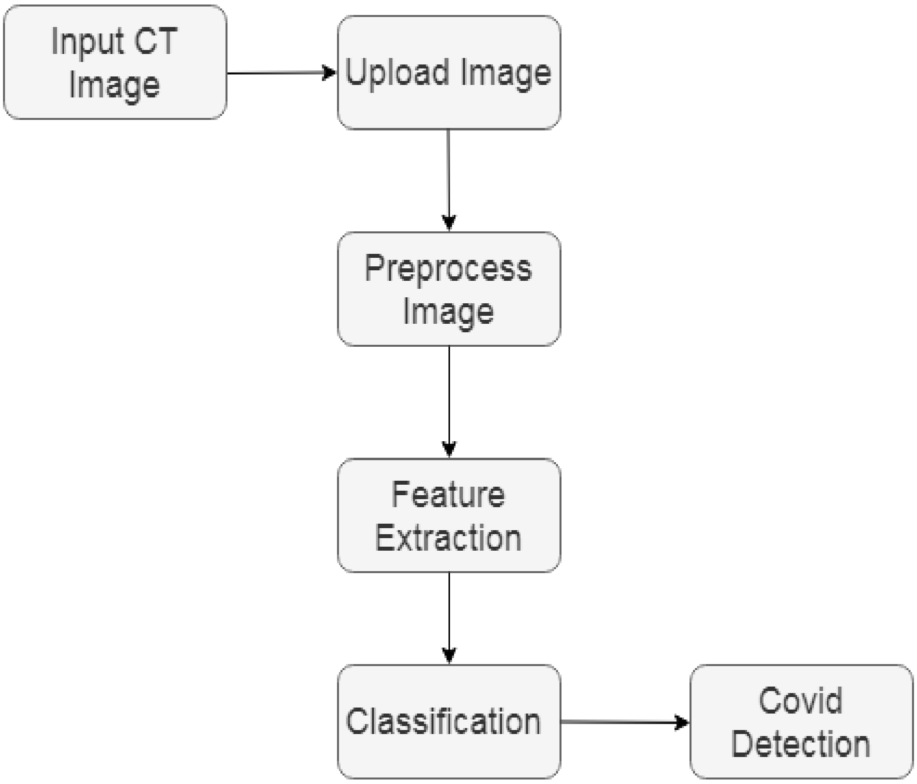


Fig 1. System Architecture

1. Input Image:

Fig 2.Home Page

2. Image Pre-processing:

In this step we will applying the image pre-processing methods like grey scale conversion, image noise normal.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005137|

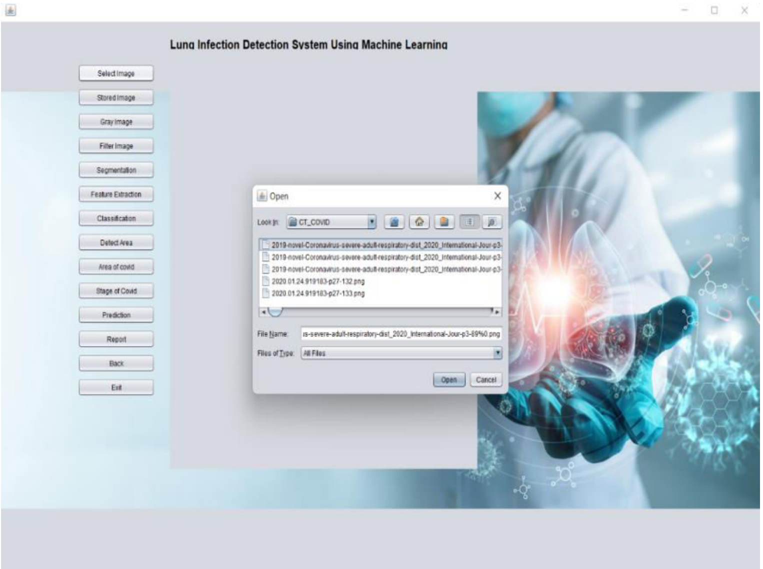


Fig 3. Selection of CT Images

3. Image Feature Extraction:
In this step we will applying the image pixel extraction methods to remove the image features from image.

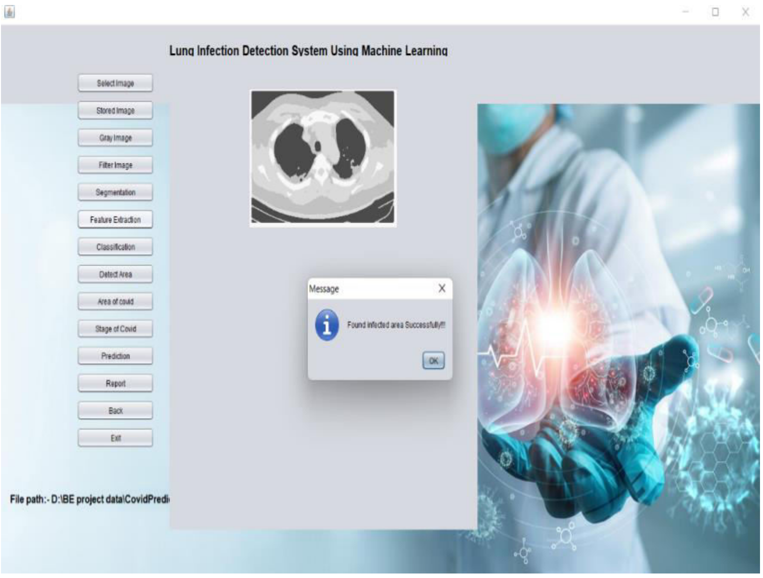


Fig 4. Feature Extraction

4. Image Classification:
In this stage we will applying the picture classification methods to distinguish the contaminated region and safe area from features.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005137|

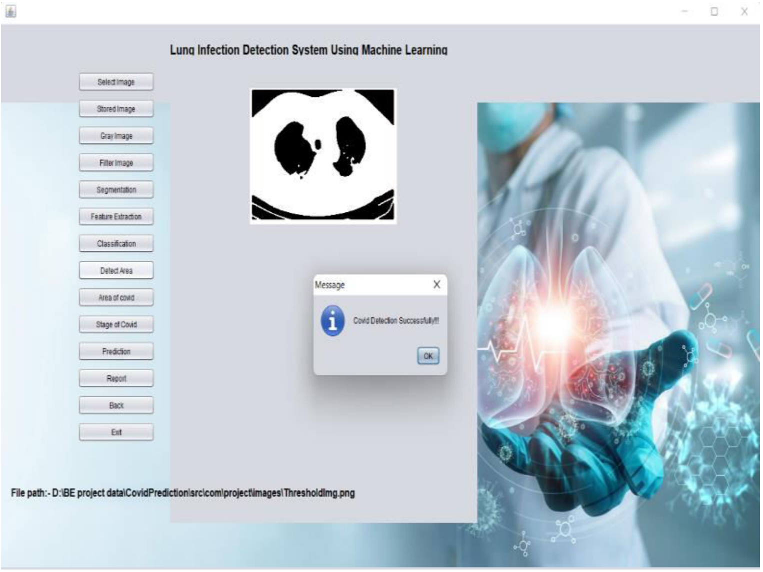


Fig 5. Classification and Covid Detection

5. Result:
In this step will show the final result detection result.

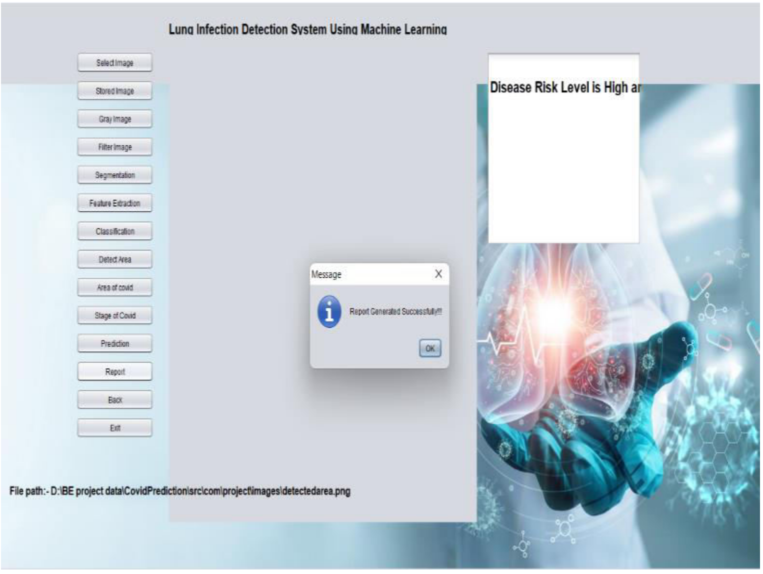


Fig 6. Result

B. Algorithm

1. Shape-Based Invariant Texture Index (SITI) for Feature Extraction
Feature Extraction image content identification.

Steps:

1. **Color feature** is one of the most widely used visual features in image retrieval, for its in variance with respect to image scaling, rotation, translation. In this work, an image is divided into four equal sized blocks and a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |
|| Volume 10, Issue 5, May 2022 ||
| DOI: 10.15680/IJIRCCE.2022.1005137|

centralized image with equal-size. For each block, a 9-D color moment is computed, thus the dimension of color comment for each image is 45. The 9-D color moment of an image segment is utilized, which contains values of mean, standard deviation and skewness of each channel in HSV color space.

2. **Edge Detection:** Most of the shape information of an image is enclosed in edges. So first we detect these edges in an image and by using these filters and then by enhancing those areas of image which contains edges, sharpness of the image will increase and image will become clearer.
   Canny Edge Detection:
   Canny edge detection is a technique to extract useful structural information from different vision objects and dramatically reduce the amount of data to be processed. It has been widely applied in various computer vision systems. Canny has found that the requirements for the application of edge detection on diverse vision systems are relatively similar. Thus, an edge detection solution to address these requirements can be implemented in a wide range of situations. The general criteria for edge detection include:
   1. Detection of edge with low error rate, which means that the detection should accurately catch as many edges shown in the image as possible.

   2. The edge point detected from the operator should accurately localize on the center of the edge.

   3.A given edge in the image should only be marked once, and where possible, image noise should not create false edges.

   The Process of Canny edge detection algorithm can be broken down to 5 different steps:
   1. Apply filter to smooth the image in order to remove the noise.
   2. Find the intensity gradients of the image.
   3. Apply non-maximum suppression to get rid of spurious response to edge detection.
   4. Apply double threshold to determine potential edges.
   5. Track edge by hysteresis: Finalize the detection of edges by suppressing all the other edges that are weak and not connected to strong edges.

3. Texture feature describes the structure arrangement of surfaces and their relationship to the environment, such as fruit skin, clouds, trees, and fabric. The texture feature in our method is described by hierarchical wavelet packet descriptor (HWVP). A 170- D HWVP descriptor is utilized by setting the decomposition level to be 3 and the wavelet packet basis to be DB2.

4. Support Vector Machine:
Support Vector Machine (SVM) is used to classify the fruit quality. SVM Support vector machines are mainly two class classifiers, linear or non-linear class boundaries. The idea behind SVM is to form a hyper plane in between the data sets to express which class it belongs to.The task is to train the machine with known data and then SVM find the optimal hyper plane which gives maximum distance to the nearest training data points of any class.

Steps:
Step 1: Read the test image features and trained features.
Step 2: Check the all test features of image and also get all train features.
Step 3: Consider the kernel.
Step 4: Train the SVM using both features and show the output.
Step 5: Classify an observation using a Trained SVM Classifier.

IV. RESULTS AND DISCUSSION

1) Positives and Negatives: Suppose there is a CT image t and the result class S. The output of the classifier is whether t belongs to S or not. A common way to evaluate the classifier's performance is to use true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). These metrics are defined as follows:

a) TP CT image of class S correctly classified as belonging to class S.
b) FP CT image not belonging to class S incorrectly classified as belonging to class S.
c) TN CT image not belonging to class S correctly classified as not belonging to class S.
d) FN CT image of class S incorrectly classified as not belonging to class S.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005137|

To measure the ability to detect result, we also import true positive rate (TPR) and false positive rate (FPR).
a) TPR is defined as the ratio of those positive CT image correctly classified as belonging to class positive to the total number of CT image in class positive, it can be calculated by

TPR=TP/(TP+FN)............... (1)

b) FPR is defined as the ratio of those negative CT image incorrectly classified as belonging to negative class S to the total number of negative CT image

FPR=FP/(FP+FN)............... (2)

2) Precision, Recall, and F-measure: By using precision, recall, and F-measure to evaluate per-class performance.\\
a) Precision is defined as the ratio of those CT image that truly belong class S to those identified as class S, it can be calculated by

Precision=TP/(TP+FP)............... (3)

b) Recall (which is also known as detection rate in the detection scenario) is defined as the ratio of those CT image correctly classified as belonging to class S to the total number of users in class S, it can be calculated by\\

Recall=TP/(TP+FN)........... (4)

c) F-measure is a combination of precision and recall, it is a widely adopt metric to evaluate per-class performance, it can be calculated by

F-measure=(2*Precision*Recall)/(Precision+Recall)....... (5)

As a result, F-measure, which is combination of precision and recall, decreased dramatically due the decrease of precision. We find that the F-measure of machine learning-based classifiers is quite low as there are much more negative CT image than positive CT image.

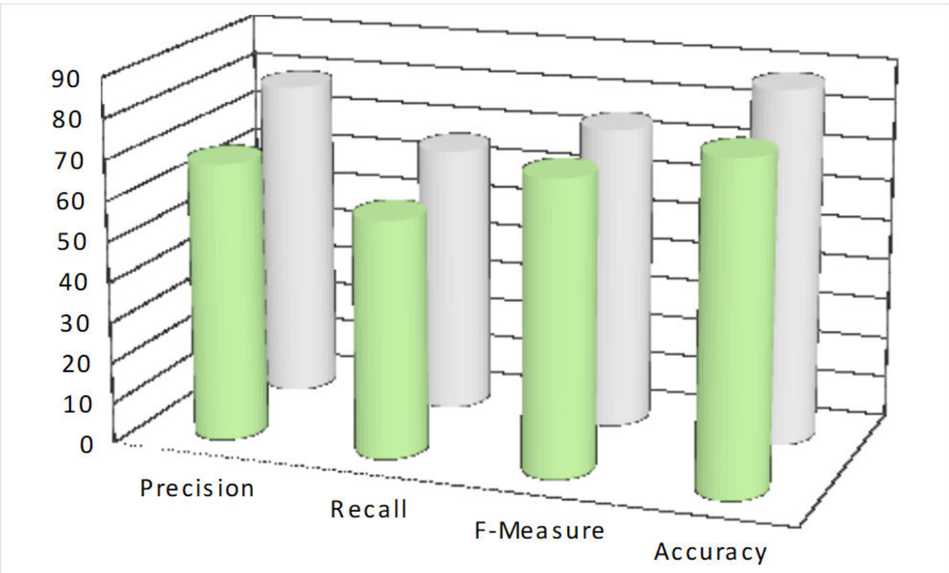


Fig 2. Classification result

	Existing System	Proposed System
Precision	68.45	78.70
Recall	58.44	65.64
F-Measure	72.11	74.31
Accuracy	80.29	87.26

V.CONCLUTION

COVID-19, it is important to get the diagnosis result at soon as possible. CT is shown as a powerful tool and could provide the chest scan results in several minutes. It is beneficial to develop an automatic diagnosis method based on chest CT to assist the COVID19 screening. In this study, we explore a deep-learning based method to perform automatic COVID-19 diagnosis from CAP in chest CT images. We evaluate our method by the largest multi-centre CT data in the world.

REFERENCES

[1] N. Zhu, D. Zhang, W.Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, and W. Tan, "A novel coronavirus from patients with pneumonia in China, 2019," N. Engl. J. Med., vol. 382, pp. 727–733, 2020.
[2] D. Benvenuto, M. Giovanetti, M. Salemi, M. Prosperi, C. De Flora, L. C. Junior Alcantara, S. Angeletti, and M. Ciccozzi, "The global spread of 2019-nCoV: A molecular evolutionary analysis," Pathog. Glob. Health, pp. 1–4, 2020.
[3] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, and D. Shen, "Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19," IEEE Rev. Biomed. Eng., vol. 3333, no. c, pp. 1–13, 2020.
[4] L. Huang, R. Han, T. Ai, P. Yu, H. Kang, Q. Tao, and L. Xia, "Serial quantitative chest CT assessment of COVID-19: Deep-learning approach," Radiol. Cardiothorac. Imaging, vol. 2, p. e200075, 2020.
[5] J. Lei, J. Li, X. Li, and X. Qi, "CT imaging of the 2019 novel coronavirus (2019-nCoV) pneumonia," Radiology, p. 200236, 2020
[6] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Gao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, and J. Xia, "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT," Radiology, p. 200905, 2020.
[7] Y. Cao, Z. Xu, J. Feng, C. Jin, X. Han, H. Wu, and H. Shi, "Longitudinal assessment of COVID-19 using a deep learning–based quantitative CT pipeline: Illustration of two cases," Radiol. Cardiothorac. Imaging, vol. 2, no. 2, p. e200082, 2020.
[8] D. Karimi, H. Dou, S. K. Warfield, and A. Gholipour, "Deep learning with noisy labels: exploring techniques and remedies in medical image analysis," arXiv:1912.02911, pp. 1–17, 2020.
[9] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," Annu. Rev. Biomed. Eng., vol. 19, no. 1, pp. 221–248, 2017.