Machine learning-based automatic detection of novel coronavirus (COVID-19) disease

Anuja Bhargava[1] [iD] · Atul Bansal[1] · Vishal Goyal[1]

Abstract
The pandemic was announced by the world health organization coronavirus (COVID-19) universal health dilemma. Any scientific appliance which contributes expeditious detection of coronavirus with a huge recognition rate may be excessively fruitful to doctors. In this environment, innovative automation like deep learning, machine learning, image processing and medical image like chest radiography (CXR), computed tomography (CT) has been refined promising solution contrary to COVID-19. Currently, a reverse transcription-polymerase chain reaction (RT-PCR) test has been used to detect the coronavirus. Due to the moratorium period is high on results tested and huge false negative estimates, substitute solutions are desired. Thus, an automated machine learning-based algorithm is proposed for the detection of COVID-19 and the grading of nine different datasets. This research impacts the grant of image processing and machine learning to expeditious and definite coronavirus detection using CXR and CT medical imaging. This results in early detection, diagnosis, and cure for the accomplishment of COVID-19 as early as possible. Firstly, images are preprocessed by normalization to enhance the quality of the image and removing of noise. Secondly, segmentation of images is done by fuzzy c-means clustering. Then various features namely, statistical, textural, histogram of gradients, and discrete wavelet transform are extracted (92) and selected from the feature vector by principle component analysis. Lastly, k-NN, SRC, ANN, and SVM are used to make decisions for normal, pneumonia, COVID-19 positive patients. The performance of the system has been validated by the k (5) fold cross-validation technique. The proposed algorithm achieves 91.70% (k-Nearest Neighbor), 94.40% (Sparse Representation Classifier), 96.16% (Artificial Neural Network), and

✉ Anuja Bhargava
   anuja1012@gmail.com

   Atul Bansal
   atul.bansal@gla.ac.in

   Vishal Goyal
   vishal.goyal@gla.ac.in

[1] GLA University, Mathura, India

99.14% (Support Vector Machine) for COVID detection. The proposed results show feature combination and selection improves the performance in 14.34 s with machine learning and image processing techniques. Among k-NN, SRC, ANN, and SVM classifiers, SVM shows more efficient results that are promising and comparable with the literature. The proposed approach results in an improved recognition rate as compared to the literature review. Therefore, the algorithm proposed shows immense potential to benefit the radiologist for their findings. Also, fruitful in prior virus diagnosis and discriminate pneumonia between COVID-19 and other pandemics.

1 Introduction

The COVID-19 (Novel Coronavirus) appears as a crown for the microscope to cosmic so named as coronavirus. It emerges from SARS-COV-2 (a disorder of respiratory shown in Fig. 1) and an epidemic in November 2019 [17]. It composes of 14 secured debris which specifically connects with "human angiotensin-converting enzyme 2" and is associated with the "Coronaviradae Family". Figure 1 shows the organized hierarchically view of different pneumonia originated by micro-organisms. The COVID-19 is identified as a blustery respiratory disease. Globally, the economy is at high risk as COVID-19 affects distinct fields like production agriculture, supply chain, tourism, transport, insurance, and industries [58]. In the year 2020, due to COVID-19, the lowest economy is reported by the "Organization for Economic Cooperation and Development" [41].

Wuhan reports the very first wave of disease with syndrome-like nausea, cough, fever, fatigue, and lung infiltrates, dyspnoea, and shortness of breath. The "Public Health Emergency of International Concern" is claimed by World Health Organisation on the last day of January 2020 and realized as a pandemic in mid of March 2020 [16]. Now, the world reports a second wave of disease with symptoms like fever, persistent and dry cough, muscular pain, loss of taste and smell, headache, upset stomach, discoloration of toes and fingers, sore throat, and rashes. Figure 2 shows the top 5 country geographical dissemination and stats of reported death, recovered cases around the globe.
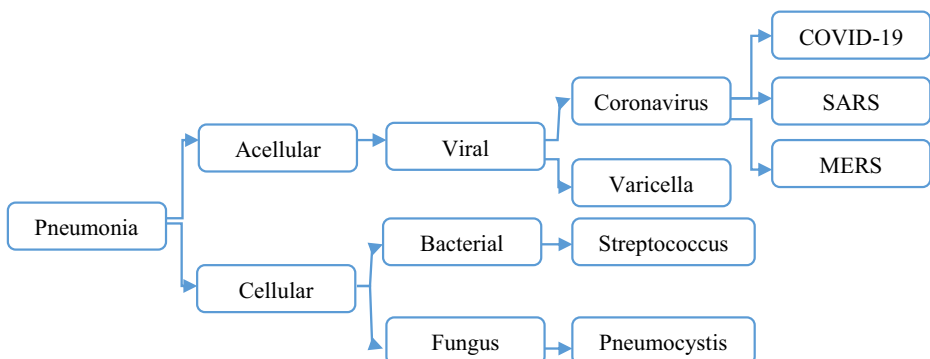

Fig. 1 Hierarchical structure of COVID-19

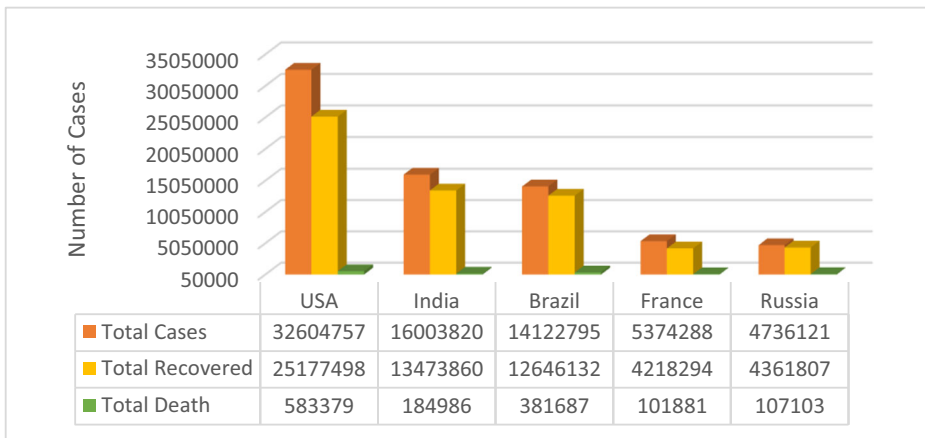|	USA	India	Brazil	France	Russia
Total Cases	32604757	16003820	14122795	5374288	4736121
Total Recovered	25177498	13473860	12646132	4218294	4361807
Total Death	583379	184986	381687	101881	107103

Fig. 2 Report of COVID 19 around the globe [62]

Also, in many of the cases, symptoms are hard to admit so one of the efficient ways of protection is lockdown that imposes emotional impose on peoples mind and health and influence the Gross Domestic Product of the country. "Richard Baldwin a professor of International Economics at the graduate institute in Genena said- This virus is as economically contagious as it is medically contagious" [49].

In early 2020, COVID-19 treatment diagnosis incorporates specific antibody detection and nucleic acid testing [32]. The initial screening recognizes the window phase for detection of antibody hence it is not convenient. The nucleic acid testing incorporates to recognize predominately a bacteria in urine or blood organism species but it is the time taken, operating concern with low recognition rate. In April 2020, the RT-PCR test (Reverse Transcription Polymerase Chain Reaction) a standard gold shield method is used on respiratory of COVID-19 patients [60]. This method exposes nucleoids from the nasopharyngeal swab, tracheal aspirate, and oropharyngeal swab. The patient shielding is done by assemble specimens from the throat or nose and bringing about diagnosis in the laboratories. However, the procedure is complicated, time-consuming, and bounded by swabs, test reagents and kits lead to challenges because of the huge amount of patients. Moreover, the RT-PCR test report reveals that sensitivity is not at all high enough for COVID-19 detection.

The diagnosis appliance can be epidemiological history, positive pathogenic testing, clinical symptoms analysis, positive radiographic images (Chest Radiography (CXR/X-Ray), Computed Tomography (CT)). The CXR and CT are mostly used due to their uninventive essence that records the lung's digital images. The radiographic images are extensively available at hospitals and have agile turnaround space. These images have a lack of experienced radio-logistic that contribute to the accurate and timely interpretation of images [38]. The computed tomography imaging estimate patient with different symptoms and has small misdiagnosis rate, highly available and sensitivity is high. The images incorporate "crazy paving pattern and ground glass pique shadows" with an increase in lesion density [24]. The characteristic of COVID-19 images, features are shown in Fig. 3.

The CT images are extensively preferred to CXR images due to the lung 3D view of images. These images can be converted into 2D images as a sagittal axial coronal view for improved diagnosis. April 2021 reports COVID-19 has rapid rise and formerly spread amount the globe with limited doctors, hospitals, oxygen cylinder, and inadequate RT PCR test kit and
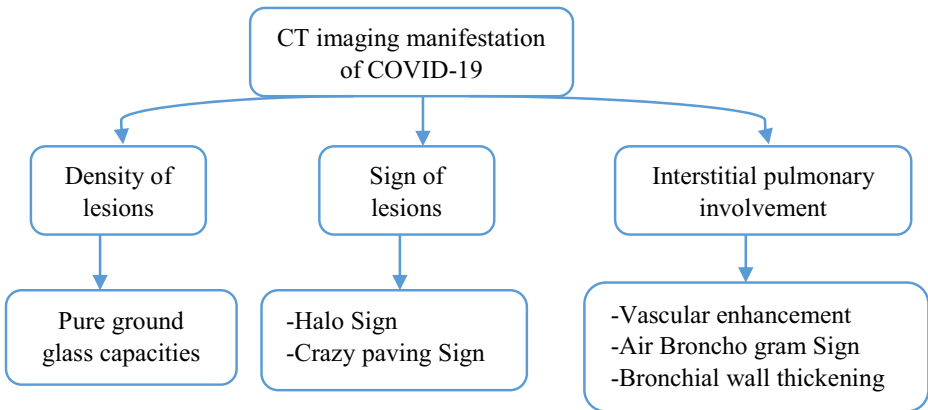
Fig. 3 CT image of features

its cost and test result wait time show the machine learning approaches importance. With the acute hike in COVID-19 patients, the current resources of medical as well as capabilities of diagnosis are inadequate. Also the cross-infection risk increase in hospital staff as well as corona warriors. "According to the National Health Commission of the people's republic of China (2020), Diagnosis and protocol of COVID-19 CT detection is not only one of the diagnostic standards for COVID-19 but also has an important significance in the treatment of COVID-19" [44].

Presently, CT and CXR images are estimated by highly qualified doctors and feasible treatments are recommended and implemented to shorten the level of infection. The individual detection from images will be tedious and time-consuming. Hence an automated machine learning-based process is needed that assists the doctor to examine the patient. This process will lower the ample amount of burden for diagnosis, optimizing medical resources, saving doctors energy, and improving diagnosis efficiency.

2 Related work

Since the whole world is affected by coronavirus COVID-19 since December 2019 distinct studies have been started to diagnosis the disease. Last year, several studies have been done using deep learning, artificial learning, and neural network for automatic detection of COVID-19 [11]. In [36] utilizes a model of Efficient Net to detect COVID-19 using X-Ray images achieves 93.90%accuracy. In [53] presented an algorithm using harmony-search optimization, Otsu thresholding, segmentation region of interest for disease detection on CT scan images. In [50] proposed nCOVnet method of deep learning neural network to detect COVID-19 for X-Ray images. In [65] introduce an early detection of COVID-19 using three-dimensional deep learning with an 86.70% recognition rate. Here, a classification model based on local attention and the Bayesian function is used. In [13] proposed a pre-processing and deep-learning-based automatic detection of COVID-19. The dataset is pertained with a deep convolutional network and achieves 98.30% accuracy. In [39] presented architecture based on detail-oriented capsule networks for COVID-19 disease detection. This model achieves 96.10% accuracy for CT Scan images.

In [4] proposed network model NABLA-N for segmentation and inception residual recurrent CNN for COVID-19 detection and achieves 98.78%, 84.67% recognition rate for CT Scan, X-Ray respectively. In [6] utilizes transfer learning with the architecture model of convolutional neural networks and yield 96.78% accuracy. In [31] presents architecture based on ResNet101. Here, U-Net is used to segment the lungs of X-Ray images with a confidence score for the classification of COVID-19 patients and achieves a 98.00% recognition rate. In [35] proposed conditional generative adversarial network with AlexNet, VGGNet16, GoogleNet, and ResNet50 deep convolutional neural network model achieves 93.33% accuracy.

In [43] utilizes ResNet50, Inception VS, Inception ResNetV2 based COVID-19 detection. Here, 5-fold cross-validation is used and yield 98.00% highest accuracy for RESNet50. In [28] proposed CGAN and architecture based on deep learning for ResNet8 achieves 98.70% accuracy. In [30] presents deep learning using CNN for both CXR and CT images and achieves 92.00% accuracy. In [66] propose an algorithm that adjusts the global intensity extracts for block feature for disease detection and classification is done using a deep convolutional neural network. In [40] utilize ResNet18, SqueezeNet, DebseNetF1 to detect COVID-19 an X-Ray image. In [3] presents a framework based on capsule network to identify COVID-19 with COVID-CAPS and achieves 95.70% accuracy. In [52].

It can be seen from Table 1 that automatic detection of COVID-19 is done using deep learning techniques mostly. Hence machine learning-based automatic detection of Novel Coronavirus (COVID-19) disease adds innovation to the existing technique.

This paper incorporates different methods that differentiate among normal, COVID-19, pneumonia patients using machine learning techniques. The proposed algorithm utilizes pre-processing, segmentation, feature extraction, classification methods of image processing. The remaining paper consists of methods and materials in Section 3, experimental results in Section 4, and conclusion in Section 5.

3 Materials and methods

The proposed approach aims to set up an arrangement that detects COVID-19 (for a patient with COVID-19), Pneumonia (for the patient with non-COVID-19 cases of pneumonia), and Normal (for healthy patients) among the numerous CT scan and CXR images. The individual detection from images will be tedious and time-consuming. Hence an automated machine learning-based process with rigors pre-processing, segmentation, and feature extraction is needed that assist the doctor to examine the patient. This process will lower the ample amount of burden for diagnosis, optimizing medical resources, saving doctors energy, and improving diagnosis efficiency. Figure 4 shows the proposed approach performance of the system. Each step with used algorithm is explained in the following sub-section.

3.1 Image acquisition

Our proposed algorithm uses multiple images i.e. Chest Radiography (CXR/X-Ray), Computed Tomography (CT scan) images. The database contains normal, pneumonia, COVID positive images obtained from different datasets available on the different sources. Table 2 shows the database used by the proposed algorithm. Some of the database samples are shown in Fig. 5.

Table 1 Literature review based on COVID-19 detection

Authors	Image Type	Data Size	Method
Ozturk et al. [48]	X-Ray	125 COVID-19+, 500 Finding 500 Pneumonia	Dark COVID Net
Hemdan et al. [21]	X-Ray	25 COVID-19+, 25 COVID-19-	COVIDX-Net
Barstugan et al. [10]	CT	53 COVID-19+, 97 COVID-19-	DWT+SVM
Wang et al. [34]	X-Ray	358 COVID-19+, 5538 COVID-19,8066 Pneumonia	COVID-Net
Maghdid et al. [37]	X-Ray, CT	203 COVID-19+, 153 COVID-19-	Deep Learning
Ghoshal et al. [18]	X-Ray	1583 Normal, 2786 Bacterial, 1504 COVID-, 68 COVID+	Bayesian CNN
Abbs et al. [1]	X-Ray	116 COVID-19+, 80 COVID-19-	Deep Transfer Learning, PCA
Farooq and Hafeez [15]	X-Ray	650 with non COVID-19, 1203 COVID-19-	ResNet5
Singh et al. [55]	CT		Multi-objective differential evolution CNN
Soares et al. [56]	CT	1252 COVID-19+, 1230 COVID-19-	Deep Learning Classification
Yang et al. [67]	CT	349 COVID-19+, 397 COVID-19-	Multi-task learning, Self-supervised learning
Wei et al. [33]	CT	305 COVID-19+, 1498 Pneumonia	3D ResNet18
Hu et al. [22]	CT	150 COVID-19+, 300 Pneumonia	Weakly Supervised
Wu et al. [64]	CT	400 COVID-19+, 350 COVID-19-	Deep Learning
Sun et al. [57]	CT	1495 COVID-19+, 1027 Pneumonia	Adaptive feature selection
Jaiswal et al. [23]	CT	1262 COVID-19+, 1230 COVID-19-	DenseNet201 based deep transfer learning
Abraham et al. [2]	X-Ray	453 COVID-19+, 497 COVID-19-	Bayesian Classifier
Altan & Karsau [5]	X-Ray	1262 COVID-19+, 1230 COVID-19-, 1614 Pneumonia	2D Curvelet transform
Nour et al. [45]	X-Ray	219 COVID-19+, 1341 COVID-19-, 1345 Pneumonia	Deep features, Bayesian optimization
Kassani et al. [29]	X-Ray, CT	117 X-Ray, 20 CT COVID-19+, 117 X-Ray, 20 CT COVID 19-	MobileNet, DenseNet, ResNet
Ardakani et al. [7]	CT	306 COVID-19+, 306 COVID-19-	k-NN, SVM
Zhou et al. [68]	CT	500 COVID-19+, 500 COVID-19-	Ensemble deep learning model
Gupta et al. [19]	X-Ray	219 COVID-19+, 1341 COVID-19-, 1345 Pneumonia	Integrated Stacking Insta CovNet19 Model
Asian et al. [9]	X-Ray	361 COVID-19+, 365 COVID-19-, 362 Pneumonia	CNN based transfer learning
Luz et al. [36]	X-Ray	8066 Normal, 5521 Pneumonia, 183 COVID+	Deep learning model
Rajinikanth et al. [53]	CT	Benchmark Database	Optimization and Thresholding
Panwar et al. [50]	X-Ray	192 COVID-19+, 145 COVID-19-	nCOVnet
Xu et al. [65]	CT	219 COVID-19+, 224 COVID-19-, 175 Normal	Deep learning Technique
Chowdhury et al. [13]	X-Ray	1341 Normal, 1345 Pneumonia, 190 COVID+	Deep Convolutional Neural Network
Mobiny et al. [39]	CT	349 COVID-19+, 397 COVID-19-	Detail-oriented capsule network
Alom et al. [4]	X-Ray, CT	1375 Normal, 3875 Pneumonia	Inception residual recurrent CNN
Aposlolopoulos & Mpesiana [6]	X-Ray	1008 Normal, 1414 Pneumonia, 448 COVID+	Deep learning

Table 1 (continued)

Authors	Image Type	Data Size	Method
Kusakunniran et al. [31]	X-Ray	5218 Normal, 200 Pneumonia, 142 COVID+	ResNet101
Loey & Khalifa [35]	CT	11,202 COVID-19+, 12,312 COVID-19-	Deep Convolutional Neural Network
Narin et al. [43]	X-Ray	50 Normal, 50 COVID+	ResNet50
Karakanis & Leontidis [28]	X-Ray	415 Normal, 420 Pneumonia, 420 COVID+	Deephear Model
Khurana et al. [30]	X-Ray	5760 COVID-19+, 1885 COVID-19-	Convolutional Neural Network
Yan et al. [66]	CT	861 COVID-19+, 20,797 COVID-19-	Deep Convolutional Neural Network
Minaree et al. [40]	X-Ray	536 COVID-19+, 5000 COVID-19-	DenseNet 121, SquuezeNet
Afshar et al. [3]	X-Ray	94,323 COVID-19+, 112,120 Abnormal	Convolutional Neural Network
Pereira et al. [52]	X-Ray	2000 Normal, 1374 Pneumonia, 288 COVID+	Convolutional Neural Network

3.2 Image pre-processing

Images obtained by multiple techniques might consist of distinct buzz that reduces the quality of the image. Consequently, it cannot contribute ample data for image processing. Hence, intensity values are adjusted or color mapping is done for image enhancement. It strengthens the contrast by choosing a dynamic range in the image. Typically, RGB colored images are obtained from digital cameras which is non-linear and brightness information is not possible to separate from color. For this purpose, grayscale images (converted from RGB) are used that dwindle the processing time. The conversion took place by RGB (average values) intensity values [14].

$$G_{Int.} = \frac{1}{3}\ (R + G + B) \tag{1}$$

where G_{Int} is the intensity value of grayscale, R is linear red, G is linear green and B is linear blue.

The image enhancement (normalization) is also done to improve the contrast of the image in the range of 0 to 255. Here, nine different datasets are used of different sizes. Hence, all the images are resized to 256 X 256-pixel images.

3.3 Image segmentation

A digital image separation towards definite areas is required after pre-processing of the image [51]. Segmentation segregates the defected part easily using thresholding. An analysis of image-appropriate segmentation is necessary. An enormous segmentation will reduce the performance of the classifier. The proposed approach uses fuzzy c-means segmentation.

A fuzzy partitioning uses multiple membership values, where each partition value permeates alter groups [8]. The intention is to reduce dissimilarity values for the centroid cluster. Mathematically, matrix for membership value
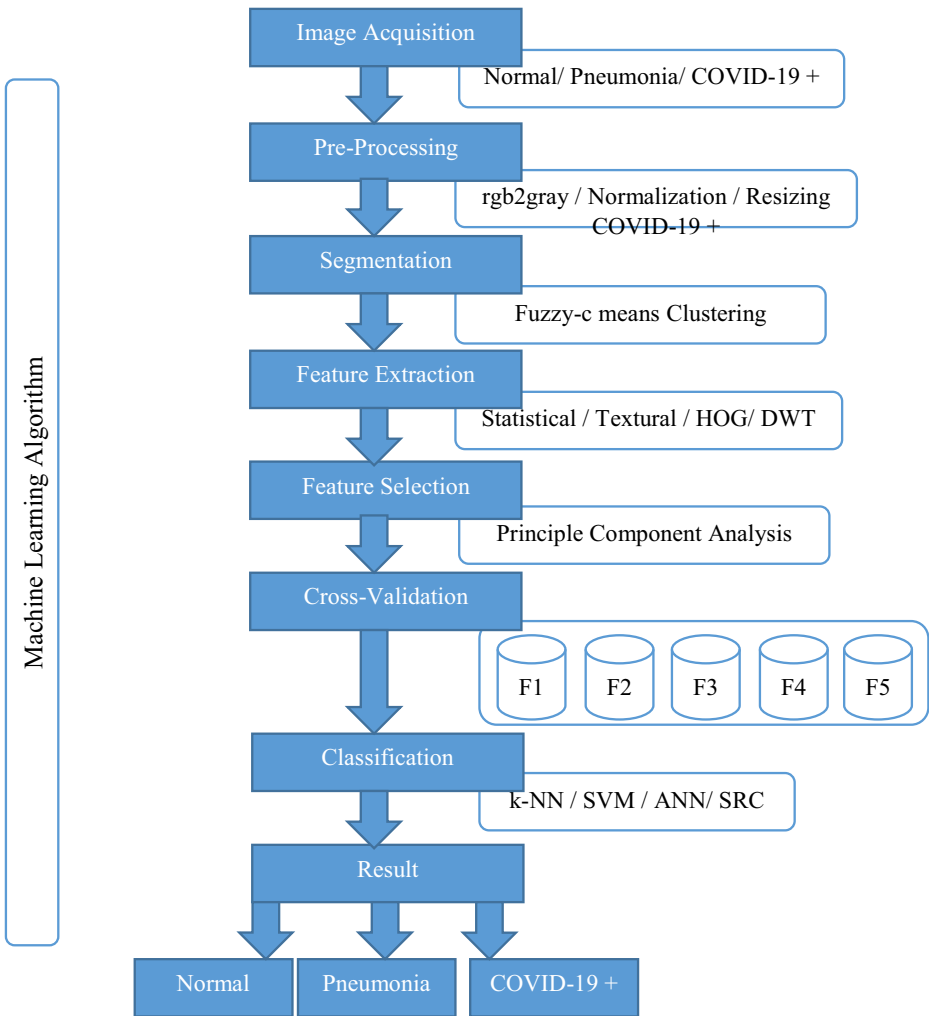
Image Acquisition

Normal/ Pneumonia/ COVID-19 +

Pre-Processing

rgb2gray / Normalization / Resizing
~~COVID-19 +~~

Segmentation

Fuzzy-c means Clustering

Feature Extraction

Statistical / Textural / HOG/ DWT

Feature Selection

Principle Component Analysis

Cross-Validation

F1   F2   F3   F4   F5

Classification

k-NN / SVM / ANN/ SRC

Result

Normal   Pneumonia   COVID-19 +

Machine Learning Algorithm

Fig. 4 Proposed Approach for Detection and Grading COVID-19 disease

Table 2 Characteristics of the database used

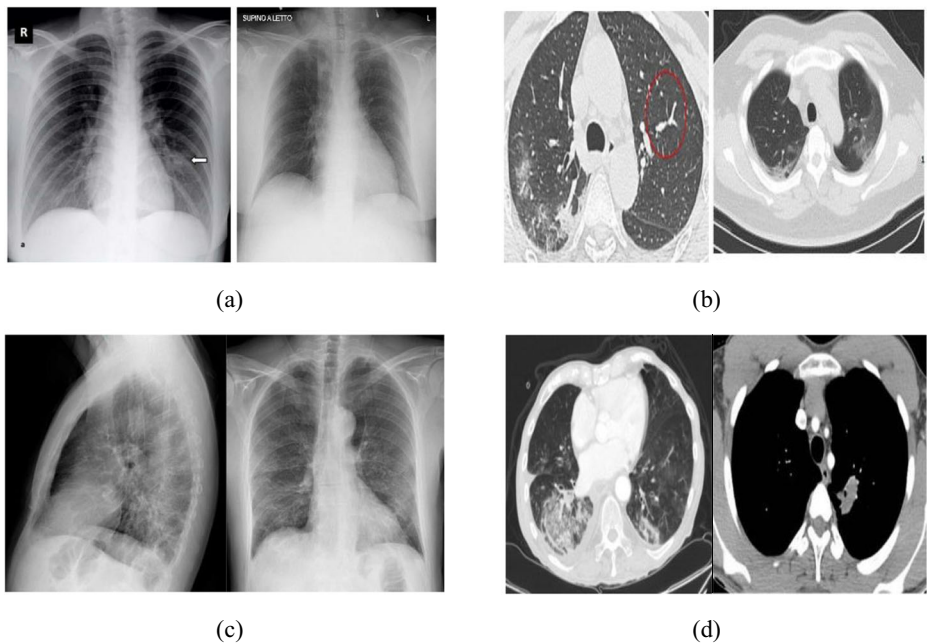| S.No.	Authors	Image Type	No. of COVID+	No. of N/P
1	Zhao et al. [67]	CT	349	397
2	Ozturk et al. [47]	X-Ray	125	500 P
3	Soares et al. [56]	CT	1252	1230
4	Joseph Chohen [27]	X-Ray	288	1374P/2000 N
5	Zhou et al. [69]	CT	2933	2500
6	Joseph [26]	X-Ray	25	25 N
7	Wang & Wong [59]	X-Ray	183	8066 N/ 5521P
8	Mooney [42]	X-Ray	–	1341 N/1345P
9	Siram [54]	X-Ray	2000	–

Fig. 5 Some of the samples of (a) COVID positive X-Ray (b) COVID positive CT (c) X-Ray Pneumonia (d) Normal CT database images

$$\sum_{x=1}^{n} Z_{xy} = 1 \qquad \forall y = 1, 2, \ldots, m \tag{2}$$

$$C(Z, n_1, n_2, \ldots, n_n) = \sum_{x=1}^{n} C_x = \sum_{x=1}^{n} \cdot \sum_{y=1}^{m} Z_{xy}^{a} \cdot d_{xy}^{2} \tag{3}$$

where Z_{xy} is a membership for y^{th} neighbor to x^{th} class, n_x is centroid for the cluster, n is the number of clusters and d_{xy} is the Euclidean distance between x and y.

In real life, the exact division is not feasible due to the existence of noise. Fuzzy c-means clustering is a natural type of clustering which is adequate to analyze data into multiple clusters accordingly and precisely.

3.4 Feature extraction

After segmenting the image, further analysis is done by estimating the features. These features contain image interpretation, forms feature vectors which is used to classify input dataset. These enhance the recognition rate. The proposed approach uses statistical, textural, histogram of gradients and discrete wavelet transform [46] to substantiate the accuracy. SURF (Speeded Up Robust Features) assesses 92 multiple features using a toolbox of image processing. The proposed approach uses the consolidation of 92 features shown in Fig. 6.
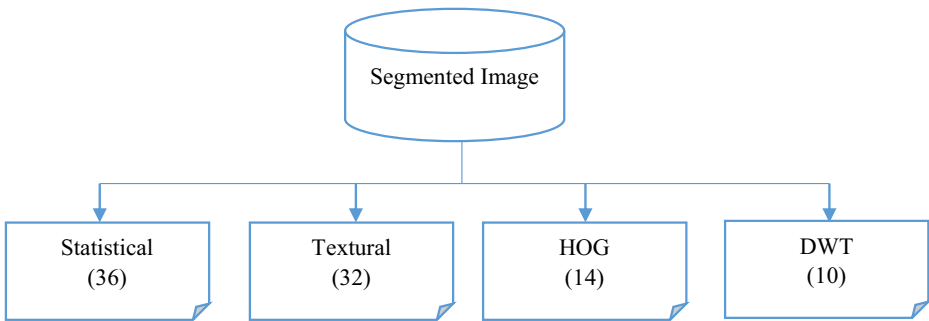
Fig. 6 Combination of Feature Set Used

3.4.1 Statistical features

The statistical features incorporate variance, smoothness, standard deviation, mean, RMS, kurtosis, Inverse difference moment and skewness represents the first-order probability measure. The proposed approach uses the histogram intensity of pixels for the defected part.

3.4.2 Textural features

The textural features incorporate energy, contrast, homogeneity, entropy, correlation represents the second-order probability measure. The grayscale distribution represents periodicity of the image texture for factors i.e. sequential, sequence created is not random and uniform area.

3.4.3 Histogram of oriented gradient

The histogram of oriented gradient incorporates detection and extraction of gradient and its direction of features. The features are calculated by taking the derivate of A and B.

3.4.4 Discrete wavelet transform

The discrete wavelet transform incorporates data compression and removal of noise from the frequency domain. This includes LL (low-low), LH (low-high), HL (high-low) and HH (high-high) components decomposed by wavelet. The highest energy and information contains in the LL component.

The combination of statistical, textural, histogram of gradients, and discrete wavelet transform incorporate the grading (normal/ pneumonia/ COVID positive) of the dataset.

3.4.5 Feature selection

The system is not feasible by using all 92 features simultaneously which leads to degradation of the performance [20]. Thus, a technique principle component analysis for the reduction of features is used to select a subdivision of 30 features among 92 features. The advantage of using the feature reduction technique is low energy space and less estimation of time. "The central idea of PCA is to reduce the dimensionality of a dataset consisting of a larger number of interrelated variables while retaining as much as possible of the variation present in the dataset. This is achieved by transforming to a new set of variables, the principal components,

which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables" [25]. The proposed approach selects a combination of 30 features using PCA. Figure 7 shows the algorithm steps on the dataset image.

3.5 Classification

The fundamental giat for quality evaluation of vegetables and fruits is classification. Here, appropriate information is extracted from the processed image. The proposed approach incorporates the following classification techniques used for detecting and grading vegetables and fruits.

3.5.1 K-NN (k-nearest neighbour)

The measurement of distinct metric closeness and value assignment to further convenient category in nearest neighbour are measure by k-NN. It measures the gap in input and trained data values using "Euclidean Distance". Mathematically,

$$ED = \sqrt{\sum_{m=1}^{N} (x_m - y_m)^2} \tag{4}$$

where x = $x_1, x_2, \ldots, x_m$ & y = $y_1, y_2, \ldots, y_m$ are the points of input trained dataset. Figure 8 shows the example of the k-NN classifier.

3.5.2 Sparse representative classifier (SRC)

SRC (non-parametric learning method) is primarily popularized for investigation in the field of recognition of face [63]. SRC consist of a dictionary having a label on each class for training samples that elect the label on test classes. SRC follows the following steps to label the test class.
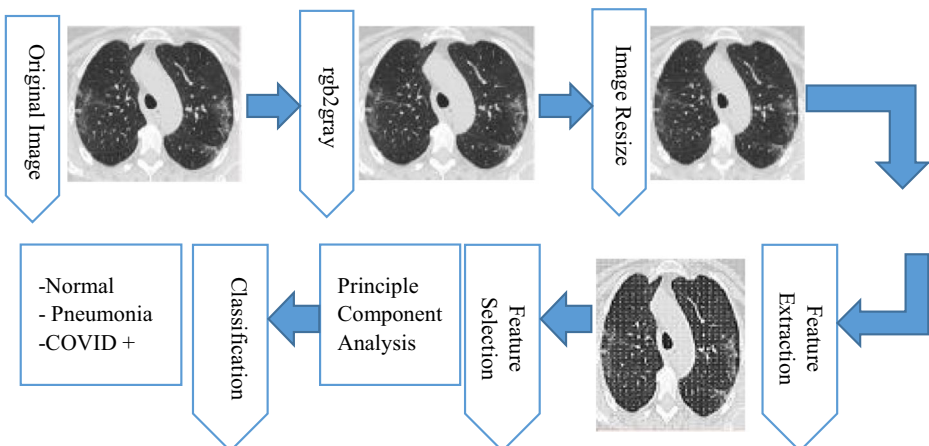

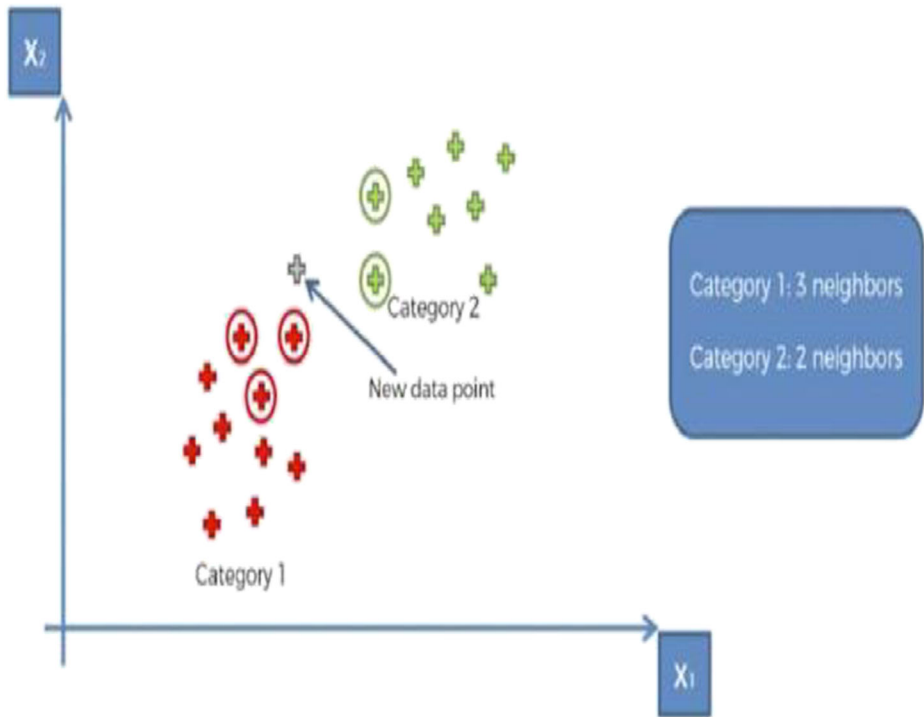**Fig. 7** Overview of the algorithm on dataset image

Fig. 8 Example of k-NN

Step 1. Given train, sample forms a dictionary for n classes.

$$Z = [Z_1, Z_2, Z_3, ..., Z_n] \in R^{x*y} \tag{5}$$

Step 2. Initially, consider tolerance $\varepsilon > 0$ error and test sample $A \in R^x$.
   Step 3. The dictionary values are normalized.
   Step 4. Determine the minimization.

$$\min s.t. \ A = ZB \tag{6}$$

Step 5. Calculate the residue.

$$S_i(A) = \|A - Z\epsilon_i B\| \qquad for \ i = 1, 2, 3, ..., n \tag{7}$$

Step 6. Form test sample.

$$Identity \ (A) = min_i S_i(A) \tag{8}$$

3.5.3 Support vector machine (SVM)

A machine learning algorithm (SVM) incorporates classification for linear as well as non-linear data. SVM is a supervised learning algorithm that modifies the hyperplane for dividing the two classes with maximal margin. The biasing is presented by ordering the samples

randomly. Figure 9 shows a solid line for the hyperplane and hollow lines for training. The linear regression is used to bias the samples.

$$C(k) = sgn\left(\sum_{i=1}^{N} d_i^* \, c_i + e^*\right) \qquad (9)$$

where d_i^* is multiplier values for Lagrange, e* is the threshold values, c_i is 1 or $-$ 1.

In the case of linear data, mapping into higher dimensional space using kernel Hilbert Space [12] is done resulting in linear data. Mathematically,

$$L\left(y, y'\right) = \left\langle \varphi(y), \varphi\left(y'\right) \right\rangle \qquad (10)$$

Here, y is the input vector and L is Mercer's condition,

$$L\left(y, y'\right) = \sum_{n}^{\infty} a_n \varphi_n(y)\varphi_n\left(y'\right) \qquad (11)$$

$$\iint L\left(y, y'\right) g(y) g\left(y'\right) dy \, dy' \qquad (12)$$

Above mentioned values are known as kernel tricks resulting from the inner value product.

3.6 Artificial neural network (ANN)

An ANN consist of weights, threshold values with transfer function united connected through neurons [61]. If the mathematical structure is not classifying the object then ANN results in leading choice conception capacity. The classification of agricultural objects using ANN results in persistent performance. The input with biasing of weighted sum results in neuron output. The majority of neurons are considered feature vectors with an activation function.

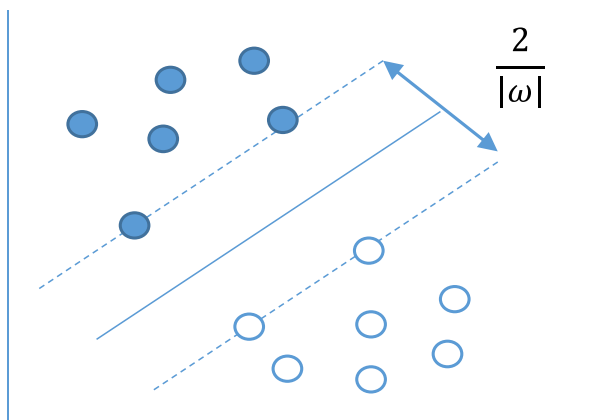$$f(net) = \tanh(x.net{-}y) + z \qquad (13)$$

Fig. 9 Linearly separable 2D hyperplane

The selection of the classifier is based on the complexity of each architecture. LR incorporates the decision based on the relationship between dependent and independent variables. A learning method SRC based on predications using training samples dictionary label the test samples. The neurons and their structure store all the information in ANN. SVM is an appropriate, favored classifier that results in accurate efficacy in multiple fields of medical classification. In the proposed approach, k-NN, SRC, SVM, and ANN are implemented using Math work.

4 Evaluation

The k-fold cross-validation process is used to evaluate the decision. Here, the whole dataset is divide into K subset and leaving one dataset at a time with repeating k time's results invalidation. A 5-fold cross-validation approach is shown in Fig. 10. The following parameters are considered for the performance prediction of the system. Accuracy measures the commutation of TP and TN values. Sensitivity measures truly classified undesirable objects. Specificity measures truly classified desirable objects.

To prohibit the model overfitting and establishment of predicted results, the dataset was divided into 75% for training, 25% for testing. The training samples calibrate the parameters using LR, SRC, ANN, and SVM.

$$Accuracy\ (\%) = \frac{True\ positives + True\ negatives}{Total}\ X\ 100\% \tag{14}$$

$$Sensitivity\ (\%) = \frac{True\ positives}{Total\ defect}\ X\ 100\% \tag{15}$$

$$Specificity\ (\%) = \frac{True\ negatives}{Total\ sound}\ X\ 100\% \tag{16}$$

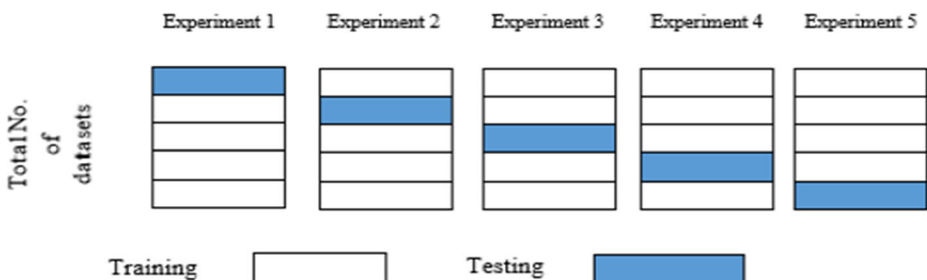$$FPR = 1 - Specificity \tag{17}$$

Fig. 10 A five-fold cross-validation

$$FNR = 1-Sensitivity \tag{18}$$

5 Results and discussions

Our proposed algorithm utilizes an Intel Core i5 processor with 64 bit, 4 GB RAM under operating system Window 7 (MATLAB 15). The proposed algorithm is sorted on nine distinct datasets of COVID-19 CXR and CT images. In this algorithm to detect the situation of the patient a combination of statistical, textural, histogram of gradients, and discrete wavelet transform features are extracted i.e. 92 features, and major features are selected using PCA i.e. 30 features. The implementation for COVID-19 detection using distinct classification techniques (k-NN, SRC, ANN, and SVM) to inspect particular classification sufficiency and to affirm whichever classifier is best for disease detection using selected features with 5 fold cross-validation method. The parameters accuracy, sensitivity, specificity, FPR, and FNR are used for classification performance with execution time. The proposed results for the corona-virus detection of the different datasets are represented in Table 3. Figure 11 shows the detection of the patient situations in the form of accuracy. The maximum recognition rate of 99.14% using SVM is attained by the system. It can be found that SVM is superior among all other classifiers used and is best for COVID-19 disease detection.

The segmentation impact has been studied and it has been found that 96.25% are correctly segmented in the proposed algorithm. The correctly segmented images result in an increase of 99.86% from 99.14% recognition for the proposed algorithm. Yet, these results indicate a minimal increase but this shows that the proposed algorithm recognition rate might be increased by using an enhanced segmentation approach. Note that, feature extraction with selection plays a crucial role resulting in high recognition rate contrast with various literary techniques. Hence, the algorithm proposed with proper features selected is capable of detecting the type and grading as normal, pneumonia, or COVID positive for various datasets of CT and CXR images.

Figures 11 show obtained accuracy on the x-axis and classification technique on the y-axis. For each plot, the SVM performance is much better than another classifier. The healthy kernels classification indicates specificity and sensitivity whose larger values indicate the healthy classifier class and low values indicate errors. The proposed algorithm errors are indicated by FPR & FNR with the least and most values of 2.60% & 11.30% for FPR and 7.14% & 9.31% for FNR respectively.

Table 3 Performance detection of the type of vegetable or fruit

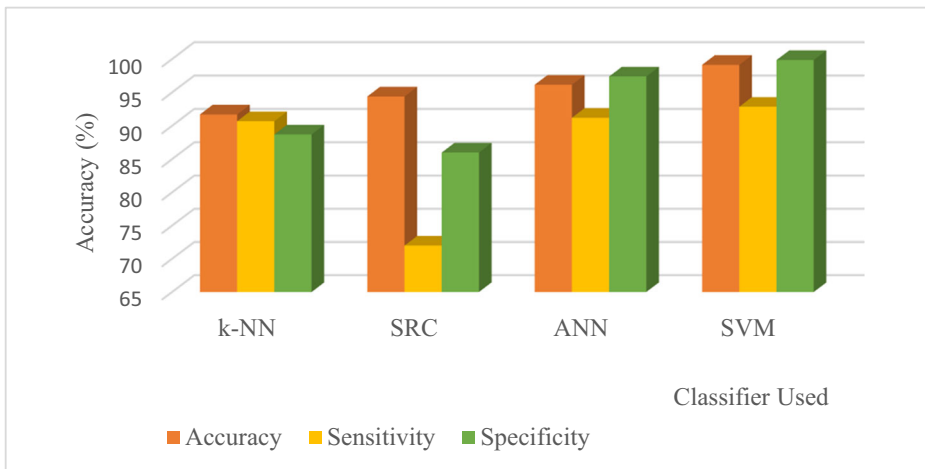| S.No.	Classifier	Accuracy	Sensitivity	Specificity	FPR	FNR	Time (s)
1	k-NN	91.70	90.69	88.70	11.30	9.31	39.98
2	SRC	94.40	72.00	86.00	14.00	28.00	33.46
3	ANN	96.16	91.20	97.40	2.60	8.80	25.87
4	SVM	**99.14**	92.86	99.86	0.14	7.14	14.34

Fig. 11 Various parameters achieved for detection of the type of vegetable or fruit

5.1 Comparison

Among the literature reviewed in recent years for coronavirus detection and grading few of them gain our attention as analyzed in Table 4. The system effectiveness is established with a large data sample. Also, the comparison of the average time taken for diagnosis of COVID-19 is shown in Table 5. The comparison between proposed algorithms with literature indicates a high and improved recognition rate with nine different datasets. Therefore, our proposed approach with selected essential features and relevant classifier results in improved accuracy.

5.2 Complexity

The complexity of the algorithm plays a crucial role in the overall system. The proposed algorithm shows the order of $O\ (a\ (b2\ +\ b*c\ +\ d))$ complexity where a is total images, b*b is the dimension of the image, b*c is the dimension of segmented image and d is features selected. Also, the execution time for the proposed algorithm including detection and grading is shown in Fig. 12.

Table 4 Comparative Analysis for COVID-19 detection and grading

S.No.	Authors	Image Acquisition	Segmentation	Feature Extraction	Classification	Accuracy
1	Chowdhury et al. [13]	2876	–	–	Deep Convolutional Neural Network	98.30%
2	Apostolopoulus et al. [6]	2870	–	–	Deep Learning	96.78%
3	Kusakumarni et al. [31]	5753	U-Net	–	ResNet101	98.00%
4	Narin et al. [43]	100	–	–	ResNet50	97.00%
5	Proposed	31,454	Fuzzy c-means	Statistical/ Textural/ HOG/DWT	k-NN/ SRC/ANN/SVM	99.14%

Table 5 Comparison of average time taken for diagnosis of COVID-19 detection

S.No.	Method Used	Time
1	RT-PCR	4–6 h
2	Radiologist	20–45 min
3	Proposed	1–1.5 min
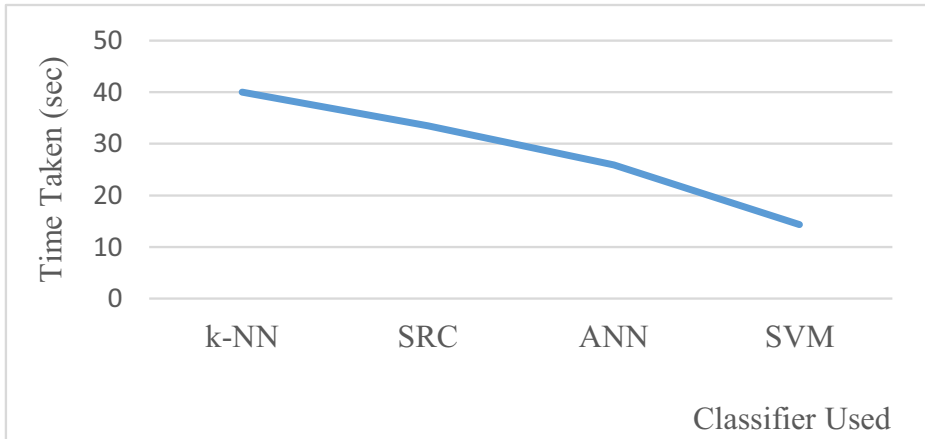
Fig. 12 Execution Time for different classifiers used for COVID-19 detection

5.3 Practical implementation

The proposed algorithm may be implemented in the medical diagnosis by the doctors. The system needs Processor P4 with 1.25 GHz and 512 M memory with an estimation time of 5 s/view. Though the estimation time can be reduced using powerful hardware with advanced software. In the proposed work, assessment is done on nine different CT and CXR based images which can be elongate for more medical imaging as per the demand of the medical industry.

6 Conclusion and future work

With the acute hike in COVID-19 patients, the current resources of medical as well as capabilities of diagnosis are inadequate. Also the cross-infection risk increase in hospital staff as well as corona warriors. Thus, detection of coronavirus became a crucial issue for research. In this work, an automated machine learning-based algorithm was presented and implemented for the detection and grading of CT and CXR images. The proposed algorithm incorporates firstly pre-processing normalization; then single object image was segmented through fuzzy c-means clustering; then multiple features statistical, textural, histogram of gradients and discrete wavelet transform (92) were extracted and essential features were selected by PCA (30) from vector space (feature) and then k-NN, SRC, ANN, and SVM classifiers were used to detect the image as normal, pneumonia and COVID positive.

The classification results show the feature selection importance and improved recognition rate of classification. The patient situation detection reports show that the proposed system classifies with the highest recognition rate of 99.14% SVM. The reported results with the proposed algorithm are compared with the literature and show the efficient results for SVM than k-NN, SRC, and ANN. The artificial intelligence-based acquisition, conventional class activation mapping, and supervised deep learning technique might be used in the future because of reduced computational time and cost. Also, the proposed approach may be used by doctors for the real-life environment.

Declarations

Conflict of interest The authors do not have any conflicts.

References

1. Abbas A, Abdelsamea M, Gaber M (2020) Classification of COVID-19 in chest X-ray images using detrac deep convolutional neural network. arXiv preprint arXiv:2003.13815
2. Abraham B, Nair M (2020) Computer-aided detection of COVID-19 from X-ray images using multi-CNN and Bayesnet classifier. Biocybern Biomed Eng 40:1436–1445
3. Afshara P, Shahin H, Farnoosh N, Anastasia O, Konstantinos N, Arash M (2020) COVID-CAPS: a capsule network-based framework for identification of COVID-19 cases from X-ray images
4. Alom Z, Rahman MMS, Nasrin S, Taha TM, Asari V (2020) Department of Electrical and Computer Engineering, University of Dayton, Dayton, OH, USA, "COVID MTNet: COVID-19 Detection with Multi Task Deep Learning Appraoches", 1–11.
5. Altan A, Karasu S (2020) Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique. Chaos Solitons Fractals 140:110071
6. Apostolopoulos I, Tzani A (2020) "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks" physical and engineering sciences in medicine
7. Ardakani A, Acharya U, Habibollahi S, Mohammadi A (2020) Covidiag: a clinical CAD system to diagnose COVID-19 pneumonia based on CT findings. Eur Radiol:1–10
8. Ashok V, Vinod DS (2014) Automatic quality evaluation of fruits using probabilistic neural network approach. International conference on contemporary computing and informatics (IC3I), 308-31.
9. Aslan MF, Unlersen K, Sabanci A, Durdu AF (2021) CNN-based transfer learning–BiLSTM network: a novel approach for COVID-19 infection detection. Appl Soft Comput 98:106912
10. Barstugan M, Ozkaya U, Ozturk S (2020) Coronavirus (covid-19) classification using ct images by machine learning methods, arXiv preprint arXiv: 2003.09424
11. Bhargava A, Bansal A (2021) Novel Coronavirus (COVID-19) Diagnosis using computer vision and artificial intelligence techniques: A Review. Multimedia Tools Appl 385:8
12. Chang, C.C., Lin, C.J. (2001). Libsvm: A Library for Support Vector Machines, via http://www.csie.ntu.edu.tw/cjlin/libsvm.
13. Chowdhury EH, Tawsifur R, Amith K, Rashid M, Abdul K, Zaid BM, Khandakar RI (2020) "Can AI help in screening Viral and COVID-19 pneumonia?" 1–12
14. Dorj UO, Lee M, Yum S (2017) An yield estimation in citrus orchards via fruit detection and counting using image processing. Comput Electron Agric 140:103–112
15. Farooq M, Hafeez A (2020) Covid-resnet: A deep learning framework for screening of covid19 from radiographs, arXiv preprint arXiv:2003. 14395
16. Fellous J, Sapiro G, Rossi A, Mayberg H, Ferrante M (2019) Explainable artificial intelligence for neuroscience: behavioral neurostimulation. Front Neurosci 13:1346

17. Ghoshal B and Tucker A (2020) Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection. arXiv:2003.10769.
18. Ghoshal B, Tucker A (2020) Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection, arXiv preprint arXiv: 2003.10769.
19. Gupta S, Gupta R, Katarya A (2020) Instacovnet-19: a deep learning classification model for the detection of COVID-19 patients using chest X-ray. Appl Soft Comput 106859
20. Hall JG, Rapanotti LL (2013) Beauty inSoftware engineering. Computer 46:85–87 https://engineering.purdue.edu/RVL/Database/IFW/database/index.html
21. Hemdan E, Shouman M, Karar E (2020) Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, arXiv preprint arXiv: 2003.11055.
22. Hu S, Gao Y, Niu Z, Jiang Y, Li L, Xiao X, Wang M, Fang E, Menpes S, Xia J (2020) Weakly supervised deep learning for covid-19 infection detection and classification from ct images. IEEE Access 8:118869–118883
23. Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M (2020) Classification of the COVID-19 infected patients using densenet201 based deep transfer learning. J Biomol Struct Dyn 39:1–8
24. Jiang X, Coffee M, Bari A et al (2020) Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. Computer Mater Continua 63(1):537–551
25. Jolliffe I (1986) Principal component analysis. R.R. Donnelley & Sons, Harrisonburg
26. Joseph E https://github.com/ieee8023/covid-chestxray-dataset.
27. Joseph P., Morrison P and Dao L 2020 COVID-19 image data collection. arXiv 2003.11597
28. Karakanis S, Georgios L (2021) Lightweight deep learning models for detecting COVID-19 from chest X-ray images. Comput Biol Med 130:104181
29. Kassani PH, Kassasni MJ, Wesolowski KA, Schneider R (2020) "Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: A machine learning-based approach", arXiv preprint arXiv: 2004.10641.
30. Khurana S., Rohan C., Bharti K. (2021) "Automated processing of social media content for radiologists: applied deep learning to radiological content on twitter during COVID-19 pandemic" emergency radiology.
31. Kusakunniran W, Sarattha K, Thanongchai S, Punyanuch B, Krittanat S, Trongtum T COVID-19 detection and heatmap generation in chest x-ray images J Med Imag 014001–12.
32. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, … Feng Z (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med 382(13):1199–1207
33. Li Y, Wei D, Chen J, Cao S, Zhou H, Zhu Y, Wu J, Lan L, Sun W, Qian T (2020) Efficient and effective training of COVID-19 classification networks with self-supervised dual-track learning to rank. IEEE J Biomed Health Inf 24:2787–2797
34. LINDA W (2020) A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images. J. Netw. Comput, Appl
35. Loey M, Nour E, Khalifa M (2020) "A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images
36. Luz Eduardo JS, Pedro L, Rodrigo S, Ludmila PS, Gladston JPM and David M, "Towards an Effective and Efficient Deep Learning Model for COVID-19 Patterns Detection in X-ray Images", 1–10.
37. Maghdid HS, Asaad AT, Ghafoor KZ, Sadiq A, Khan MK (2020) Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms, arXiv preprint arXiv: 2004.00038.
38. Mahmud MB, Carolyn CM, Danny CK, Michael JT, Pallav K, Bien ST (2020) Radiology department preparedness for covid-19: Radiology Scientific expert panel. Radiology:200988
39. Cicalese MAPA, Zare S, Yuan P, Abavisan MS, Wu CC, Ahuja J, de Groot PM, and Nguyen HV (2020) "Radiologist-Level COVID-19 Detection Using CT Scans with Detail-Oriented Capsule Networks". 1–10.
40. Minaee S, Rahele K, Milan S, Shakib Y, Ghazaleh J (2020) Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning
41. Miotto R, Wang F, Wang S, Jiang X, Dudley J (2018) Deep learning for healthcare: review, opportunities, and challenges. Brief Bioinform 19(6):1236–1246
42. Mooney P, (2018) Chest X-Ray Images (Pneumonia). Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
43. Narin A, Ceren K, Ziynet P (2020) Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks
44. National Health Commission of the People's Republic of China, Diagnosis and protocol of COVID-19, trial version 6, 2020, [Online]. Available: http: //www.gov.cn/zhengce/zhengceku/2020-02/19/content_5480948.htm.
45. Nour M, Cömert K, Polat Z (2020) A novel medical diagnosis model for COVID- 19 infection detection based on deep features and Bayesian optimization. Appl Soft Comput 106580

46. Ou X, Pan W, Xiao P (2014) Vivo skin capacitive imaging analysis by using grey level co-occurrence matrix (GLCM). Int J Pharm 460(2):28–32
47. Ozturk M, Talo EA, Yildirim UB, Baloglu O, Yildirim UR (2020) Acharya, automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 103792:103792
48. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya U (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. 103792.
49. Pandemic AE (2020) Coronavirus disease (covid-2019) r&d. https://foreignpolicy.com /2020/03/09/coronavirus-economic-pandemic-impact-recession/Lastac-cessed on Mar 2020
50. Panwar H, Gupta P, Mohammad Khubeb S, Ruben M, Vaishnavi S (2020) Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet Chaos. Solitons and Fractals 138(109944):1–8
51. Payne AB, Walsh KB, Subedi PP, Jarvis D (2013) Estimation of mango crop yield using image analysis - segmentation method. Comput Electron Agric 91:57–64
52. Pereira R, Diego B, Lucas O, Carlos N (2020) COVID-19 identification in CHEST x-ray images on flat and HIERARCHICAL classification scenarios"
53. Rajinikant V., Nilanjan D, Raj ANJ, Hassanien AE, K.C. Santosh, Raja NSM (2020) Harmony-Search and Otsu based System for Coronavirus Disease (COVID-19) Detection using Lung CT Scan Images. Appl Sci
54. S.I. S. o. M. a. I. Radiology. (2020). COVID-19 Database. Available: https://www.sirm.org/category/senza-categoria/covid-19/
55. Singh D, Kumar V, Kaur M (2020) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. Eur J Clin Microbiol Infect Dis 39:1–11
56. Soares E, Angelov P, Biaso S, Higa F, Kanda A (2020) SARS-CoV-2 CT-Scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification, medRxiv
57. Sun L, Mo F, Yan L, Xia F, Shan Z, Ding B, Song W, Gao W, Shao F (2020) Adaptive feature selection guided deep forest for covid- 19 classification with chest ct. IEEE J Biomed Health Inf 24:2798–2805
58. Tang L, Zhang L, Wang Y, Zeng X (2020) Severe COVID-19 pneumonia: assessing inflammation burden with volume-rendered chest CT. Radiol: Cardiothoracic Imag 2:e200044
59. Wang L and Wong A (2020) "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images," arXiv preprint arXiv:2003.09871
60. Wang, W., Xu, Y., Gao, R. Lu, K. Han, G. et al., "Detection of SARS-CoV-2 in different types of clinical specimens," Jama, 2020.
61. Wen X, Fang J, Diao M, Zhang C (2012) Artificial neural network modeling of dissolved oxygen in the Heihe River, northwestern China. Environ Monit Assess 185(5):4361–4371
62. Worldometer information, https://www.worldometers.info/coronavirus/#countries, Reported on 22 April 2021.
63. Wright J, Yang A (2009) Y., Ganesh, a., Sastry, S., S., and Yi, M. "robust face recognition via sparse representation," pattern analysis andMachine intelligence. IEEE Transactions on 31:210–227
64. Wu Y, Gao S, Mei J, Xu J, Fan D, Zhao W, Cheng M (2020) JCS: An explainable COVID-19 diagnosis system by joint classification and segmentation. arXiv preprint arXiv:2004.07054.
65. Xiaowei X, Xiangao J, Chunlian M, Peng D, Xukun L, Shuangzhi L, Liang Y, Qin N, Yanfei C (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 6:1122–1129
66. Yan Q, Bo W, Dong G, Chuan L, Wei Z, Jianhu S, Qinfeng S, Shuo J, Liang Z, Zheng Y (2020) COVID-19 Chest CT Image Segmentation – A Deep Convolutional Neural Network Solution" 1–10
67. Zhao J, Zhang Y, He X, Xie P (2020) COVID-CT-Dataset: a CT scan dataset about COVID-19, arXiv preprint arXiv:2003.13865.
68. Zhou H, Lu Z, Yang S, Qiu B, Huo Y, Dong T (2021) The ensemble deep learning model for novel COVID-19 on CT images. Appl Soft Comput 98:106885
69. Zhou HL, Lu FY, Hu S, Qiu CY, Wu T (2020) A model of high-dimensional feature reduction based on variable precision rough set and genetic algorithm in medical image. Math Probl Eng 2020:7653946, 18 pages. https://doi.org/10.1155/2020/7653946