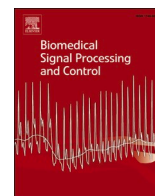




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Effectiveness evaluation of different feature extraction methods for classification of covid-19 from computed tomography images: A high accuracy classification study

Farid Al-Areqi^a, Mehmet Zeki Konyar^{b,*}

^a Department of Biomedical Engineering, Kocaeli University, 41001 Kocaeli, Turkey

^b Department of Software Engineering, Kocaeli University, 41001 Kocaeli, Turkey

ARTICLE INFO

Keywords:

Covid-19
CT images
Diagnosis
Features
Machine learning

ABSTRACT

Rapid diagnosis of the Covid-19 disease is the best way to prevent infection. In this paper, it is proposed to use machine learning methods to aid diagnoses quickly Covid-19 and focused on effect of several features on classification accuracy. In the proposed method 746 axial computed tomography (CT) images of the lung; 349 Covid-19 (positives) and 397 non-Covid-19 (negative) are used. Gray-level texture, shape and first order statistical features were extracted from the images. The feature vector for model training is constructed with one feature group or combination of more than one group. We then classified with Support Vector Machine, Random Forest, k-nearest neighbor and XGBoost classifier models. The hyperparameter of the models were controlled by the tuning test. Experimental results obtained with 10-fold cross-validation. The results of cross-validation verified with the additionally independent test. The best overall accuracy was 98.65% with first order statistics features classified with XGBoost. In the gray level features, the best individual results given by GLSZM as 81.25%, and the best combination result is with GLDM, GLRLM and GLSZM features as 85.52%. An important finding of this paper is that, for Covid-19 classification, the shape and first order statistics features are more valuable than gray level features. The proposed results compared with the literature studies under some Covid-19 dataset for accuracy, precision, sensitivity and F1-score metrics. Also, the literature studies which used the different Covid-19 dataset were compared with the proposed study. Our results have the significant superiority when compared with the literature studies.

1. Introduction

The world is witnessing a health crisis with the emergence of the Corona virus, Covid-19, in 2019 which invaded all countries of the world without exception, as it confused all off them. All countries were racing to discover the appropriate vaccine for the epidemic, and to discover and develop faster, more efficient and more reliable ways and tests to detect the virus. Conventional methods like polymerase chain reaction (PCR) and Enzyme-linked immunosorbent assay (ELISA) have shown risk of given false-positives or false-negatives and they are time and money consuming. On the other hand, lateral flow assays have shown fast and low-cost diagnosis but they give poor sensitivity [1,2]. Medical images have been known to be sources of information that used in diagnosis and treatment health problems including respiratory problems with fast time of results. However, non-expert radiologists

encounter difficulties in the detection and diagnosing process [3]. Artificial intelligence techniques have proved to be useful for radiologists in detection and classification of medical images by discovering the hidden patterns and learn the difference between the normal condition and the abnormal condition. Current studies in the scope have used many ways for Covid-19 diagnosis. They have used different datasets obtained from different medical imaging modalities such as X-ray and CT scan. They also have used different artificial intelligence techniques such as machine learning algorithms.

1.1. Literature review

Ardakani et al. [4] proposed a CAD system to discriminate COVID-19 from non-COVID-19 pneumonia patients. They used a data set consists of 612 CT images, 306 for COVID-19 and 306 non-COVID-19 pneumonia

* Corresponding author.

E-mail address: mzeki.konyar@kocaeli.edu.tr (M.Z. Konyar).

<https://doi.org/10.1016/j.bspc.2022.103662>

Received 4 December 2021; Received in revised form 18 February 2022; Accepted 19 March 2022

Available online 25 March 2022

1746-8094/© 2022 Elsevier Ltd. All rights reserved.

patients. The extracted 20 radiological features and fed them in five classifiers Decision Tree (DT), k- nearest neighbor (KNN), Naive Bayes, Support Vector Machine (SVM) and Ensemble. The achieved the best result using the proposed ensemble classifier with Accuracy of 91.94%. Al-Karawi et al. [5] proposed schemes to automatically analyze the COVID-19 in CT images. The data set they used contains 275 positive and 195 negative COVID-19 cases. They have used Gabor filter on the Fast Fourier Transform of the CT images. Classification process was done using SVM and they obtained 95.37% Accuracy. Barstugan et al. [6] in their study used 150 CT images and extracted 4 different patch regions (16x16, 32x32, 48x48, 64x64). The patches used to extract radiomics features such as FOS, GLCM, GLRLM and GLSZM and the have been classified using SVM. The best Accuracy the got was 99.64% with 10-fold and DWT (Discrete Wavelet Transform) features. Dey et al. [7] used 400 CT images 200 for COVID-19 and 200 for non-COVID-19 patients. They proposed a system that segments the COVID-19 infected regions and then extract features from those regions. For classification they implemented four classifiers; Random Forest, k-Nearest Neighbors (KNN), Support Vector Machine and Decision Tree. They achieved an accuracy of 87% with KNN classifier.

Liu et al. [8] in their study used 61 COVID-19 and 27 general pneumonia CT images and extracted 34 statistical texture features. They used Ensemble of bagged tree and compare it with four different classifiers including Linear Regression, Support Vector Machine, Decision Tree and k-Nearest Neighbors (KNN). They used Ensemble of bagged tree and compare it with four different classifiers including Linear Regression, Support Vector Machine, Decision Tree and k-Nearest Neighbors (KNN). The best classification accuracy the obtained was 94.16% with Ensemble of bagged tree classifier. Özkaya et al. [9] used the same data set of [6] and divided it into two Subsets (16x16 Subset-1 and 32x32 Subset-2). After that they extracted features with convolutional neural network architecture and classified them using SVM. They got an accuracy of 98.27% with Subset-2. Kassani et al. [10] have proposed a method to extract features using different deep learning pre-trained networks. They used 274 images as data set; 117 X-ray and 20 CT positive and 117 X-ray and 20 CT negative. For classification, Random Forest, XGBoost, Decision Tree, AdaBoost, LightGBM and Bagging classifier were used. The accuracy is 99% on features extracted with DenseNet121 and classified with Bagging tree classifier. Shi et al. [11] in their study used CT images; 1658 COVID-19 and 1027 bacterial pneumonia as negative class. They segmented the infection regions and extracted radiomics and handcrafted features. Random Forest & LightGBM-based classification method was proposed and compared with other methods including Support Vector Machine, Linear Regression, Neural Network and Random Forest. The best performance was obtained from the proposed method with accuracy of 89.4% on the handcrafted features. Zheng et al. [12] proposed to use 3D deep convolutional neural network (DeCoVNet) for detection of Covid-19 from CT images. Enrolled 540 patients in their study, 313 of them are Covid-19 and 229 patients are without Covid-19. The network is pre-trained with a simple 2D UNet in a unified manner. Detect Covid-19 by varying the thresholds and best accuracy is 90.8%. Xu et al [13] proposed a deep learning-based classification of three class included 618 CT images with 219 Covid-19, 175 healthy people and 224 influenza a viral pneumonia (IAVP). 3D Convolutional neural network (CNN) segmentation method was used with transfer learning model. Their transfer learning model is structured based on ResNet-18-based traditional classification and the location-attention. They classify Covid-19, IAVP and healthy case as three class with 86.7% accuracy rate. In another deep learning-based classification of three class Song et al.'s work [14] included 88 patients with Covid-19, 100 patients with infected with bacteria pneumonia and 86 patients with healthy people. They used pre-trained ResNet50 model based deep learning model named Details Relation Extraction neural network and achieve 93% accuracy rate for three-class classification. Wang et al. [15] used a CNN model included used transfer learning with pre-trained model named GoogleNet

inception network model to classify Covid-19 patients. Their classification accuracy is 89.5% for 325 COVID-19 positive and 740 COVID-19 negative patients. Alsharman et al. [16] used a pre-trained model named GoogleNet based CNN classifier and achieve 82.14% accuracy rate. Their work contains 463 non-COVID-19 images and 349 COVID-19 CT images. Table 1 shows a summary of the literature review in the scope in terms of publication year, dataset detail, preprocessing, classification method and results.

In the literature, few methods used the same dataset which is also used in our proposed method. Saeedi et al. [28] used a web service with number of well-known deep neural network architectures. Pham [29] presents results for several pretrained CNNs. Sakagianni [30] evaluated an automated machine learning performance. Sakagianni et.al evaluated an automated machine learning performance. Elaziz et al. [31] used hybridization of swarm-based algorithms with deep learning algorithms. Madhawi et al. [32] analyzed publicly available convolutional neural network models. Shaik et al. [33] combines the strength of multiple deep neural network architectures. Cruz [34] combine the 2-stage transfer learning and an existing ensemble methods based model. Polsinelli et al. [35] proposed the SqueezeNet based a light CNN design for the efficient classification of Covid-19.

1.2. Key contribution of the proposed method

The CT images are in gray level, but existing papers on Covid-19 classification did not discuss enough the effect of gray-level feature sets for these images. However, proposed method focused on the effect of gray level based, shape based, intensity based statistical features on Covid-19 classification for the following reasons: Gray level-based texture features provide the most meaningful information for problems where tissue heterogeneity is important. Because texture-based features can capture spatial relationships between neighboring pixels. Shape-based features describe the geometry of tissue and are useful in the sense that they have high discrimination for problems such as tumor malignancy prediction. Intensity-based first-order features are easy to calculate and have the potential to differentiate various tissues such as benign and malignant.

As far is known this paper is one of the first paper that evaluate the effect of both individual gray level based, shape based, intensity based statistical feature extraction methods and their combination on classification of Covid-19 disease.

Literature methods used complex methods and models. However, it is easy to implement the features by means of proposed method. Also, the proposed method focused accurate classification, and quick detection with machine learning methods than the traditional test methods.

The most important requirement of medical classification models is classification accuracy. The proposed method achieves superior accuracy when compared with most of the literature methods. Also, the precision and sensitivity values of the proposed method were compatible with the accuracy values.

This paper organized as follows; in Section 2 detail of dataset, feature extraction techniques, machine learning algorithms and cross validation presents. Section 3 shows experimental results. Discussion and conclusion is given in Section 4.

2. Materials & methods

In this section the overview of the proposed method explained which is given in Fig. 1. At first, we extracted several features from the CT images with different feature extraction methods. Proposed method aims to investigate the effect of different features on classification of COVID-19 disease. Several features, such as gray level-based texture features, shape features and luminance based first order statistical extracted. Each feature set has different effect on classification accuracy. The extracted features are used individually or combine in a vector group to test their effects on the classification results. After the

Table 1
Summary of the literature studies.

Authors	Year	Dataset	Pre-processing	Classification method	Results
Ardakani et al. [4]	2020	612 CT images (306 COVID-19 and 306 non-COVID-19)	20 Radiological features extraction	DT, KNN, Naive Bayes, SVM and Ensemble.	Accuracy of 91.94% with Ensemble
Al-Karawi et al. [5]	2020	470 CT images (275 positive and 195 negative)	Features extraction (FFT-Gabor)	SVM	95.37% Accuracy
Barstugan et al. [6]	2020	150 CT images	Patch regions cropping and features extraction	SVM	99.64% Accuracy
Dey et al. [7]	2020	400 CT images (200 normal and 200 COVID-19)	Segmentation and feature extraction	RF, KNN, SVM, and DT	Accuracy of 87% with KNN
Liu et al. [8]	2020	88 CT images (61 COVID-19 and 27 general pneumonia)	Delineation of ROIs and feature extraction	DT, SVM, LR, KNN and Ensemble of bagged tree	94.16% Accuracy with EBT
Özkaya et al. [9]	2020	150 CT images	Deep learning-based feature extraction	SVM	98.27%
Kassani et al. [10]	2021	274 CT images (117 X-ray and 20 CT positive and 117 X-ray and 20 CT negative)	Deep learning-based feature extraction	DT, RF, XGBoost, AdaBoost, Bagging, LightGBM	99.00% accuracy on features extracted by DenseNet121 with Bagging
Shi et al. [11]	2021	2,685 CT images (1658 COVID-19 and 1027 bacterial pneumonia)	Segmentation and feature extraction	SVM, LR, NN and RF & LightGBM-based proposed method	89.4% Accuracy
Zheng et al. [12]	2020	540 CT images (313 Covid-19 and 229 without Covid-19)	Deep learning-based feature extraction	2D UNet, 3D CNN	90.8%. Accuracy
Xu et al [13]	2020	618 CT images (219 Covid-19, 224 IAVP, 175 healthy people)	Deep learning-based feature extraction	3D CNN, ResNet18, the location-attention	86.7% Accuracy
Song et al. [14]	2021	274 CT images (88 Covid-19, 100 infected with bacteria pneumonia and 86 healthy people)	Deep learning-based feature extraction	ResNet50, Details Relation Extraction neural network	93% Accuracy
Wang et al. [15]	2021	1065 CT images (325 COVID-19 positive and 740 COVID-19 negative)	Deep learning-based feature extraction	CNN, GoogleNet inception network	89.5% Accuracy
Alsharman et al. [16]	2020	812 CT images (349 COVID-19 and 463 non-COVID-19)	Deep learning-based feature extraction	GoogleNet based CNN	82.14% Accuracy

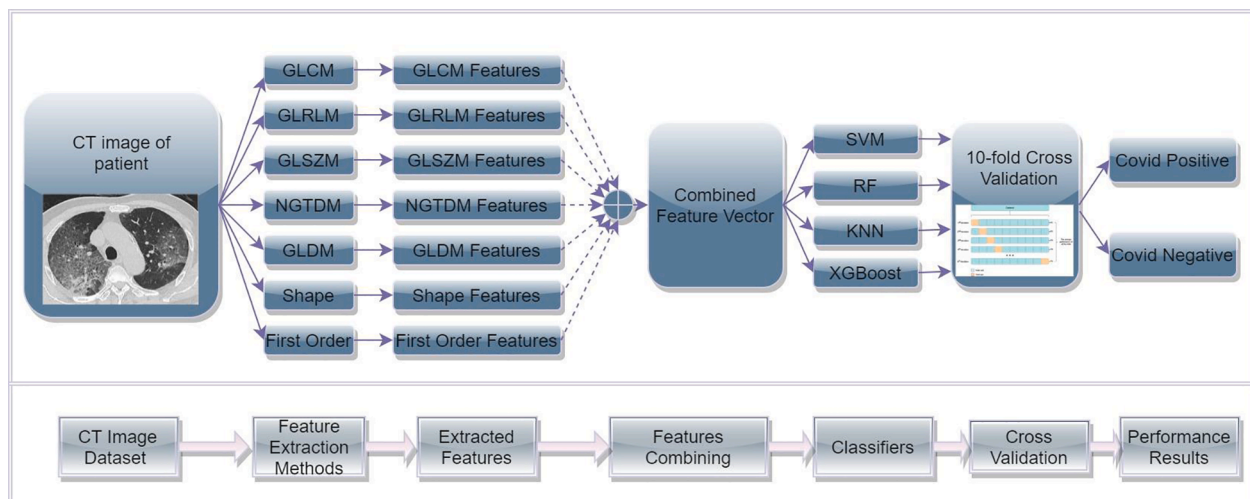


Fig. 1. Overview of the proposed method.

extracting the features and prepare feature vectors, dataset was split into training and test sets. We fed training set into machine learning algorithms to train them to classify the features with most known classifiers. For evaluation, we implemented 10-fold cross validation to evaluate the model performance on the test set. The evaluation process was conducted 10 times with different test set. At the end of all stage the performance results were calculated based on classifier output.

2.1. Dataset

The dataset used in this study has been approved by a senior radiologist in Tongji Hospital, Wuhan, China. It is available for public in [17]. The dataset contains 746 axial CT images of the lung; 349 images of patients with Covid-19 and 397 images of non-Covid-19 patients. The Images were collected from several COVID-19-related papers from

sources such as medRxiv, bioRxiv, NEJM, JAMA and Lancet and were captured by different CT modalities. Thus, the images in the data set are of different dimensions, aspect ratio and gray levels. Yang et al. [18] were collected the dataset from several public sources. So only some of Covid-19 positive patients have gender and age information. 137 of them have gender information and 86 male and 51 female. 169 of them have age information and 1–21, 22–41, 42–61, and 62–81 years old patients' numbers are 11, 59, 45 and 53 respectively. More details about the data set are available in [18]. Fig. 2 shows samples of the images in the data set for both COVID-19 and non-COVID-19 classes.

2.2. Features Extraction Techniques

Works on the numerical properties of radiological images using artificial intelligence methods is a rapidly growing field of research. The

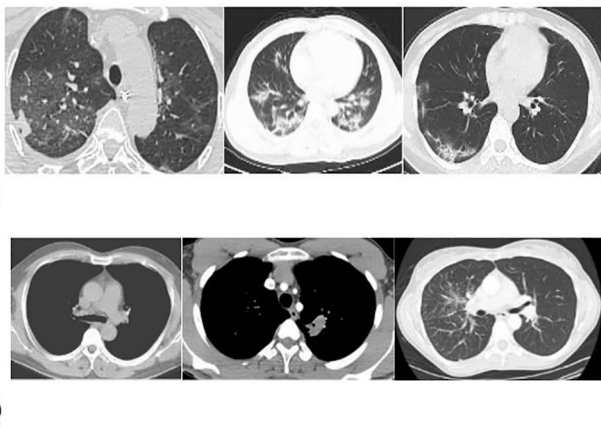


Fig. 2. Samples of the CT images dataset. a) images of Covid-19 infected patients, b) images of non-Covid-19 patients.

first step in this study was to extract features from the data. Pyradiomics library of Python [19] was one of the important tool for extract radiologic features from medical images. Pyradiomics library contains shape and texture-based method used to extract statistical features [20]. Texture-based features facilitate the detection of regions with different features in the tissue by detecting the relationships and characteristic features between pixels [21]. The texture-based method uses matrices to extract the texture feature and then the statistical features can be derived from those matrices. Various matrices such as Gray Level Co-occurrence (GLCM), Gray Level Run-Length Matrix (GLRLM), Neighborhood Gray Tone Difference Matrix (NGTDM), Gray-Level Zone Length Matrix (GLZLM) and Gray Level Dependence Matrix (GLDM) are used to extract texture-based features. Fig. 3 shows how the three common matrices are extracted from the images.

2.2.1. Gray Level Co-occurrence Matrix (GLCM)

GLCM is a method to distinguish the texture of an image by calculating the occurrence of pairs of pixels with a specific value in a specific direction (Fig. 3). This method used to extract 24 statistical feature which are: Autocorrelation, Cluster Tendency, Cluster Shade, Cluster

Prominence, Contrast, Correlation, Difference Variance, Difference Average, Difference Entropy, Informational Measure of Correlation-1, Informational Measure of Correlation-2, Inverse Variance, Inverse Difference Normalized, Inverse Difference, Inverse Difference Moment Normalized, Inverse Difference Moment, Joint Average, Joint Entropy, Joint Energy, Maximal Correlation Coefficient, Maximum Probability, Sum Average, Sum of Squares, Sum Entropy.

2.2.2. Gray Level Run Length Matrix (GLRLM)

GLRLM is a method to defines the length/number of consecutive pixels (run) as shown in Fig. 3, which gray level value are the same. This matrix used to extract 16 statistical features as follows: Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Low Gray Level Run Emphasis, Short Run Emphasis, Long Run Low Gray Level Emphasis, Long Run High Gray Level Emphasis, Gray Level Variance, High Gray Level Run Emphasis, Long Run Emphasis, Run Entropy, Run Length Non-Uniformity, Run Length Non-Uniformity Normalized, Run Percentage, Run Variance, Short Run Low Gray Level Emphasis, Short Run High Gray Level Emphasis.

2.2.3. Gray Level Size Zone matrix (GLSZM)

GLSZM is a method to quantifies several pixels which values are have the same gray level (zone) shown as Fig. 3. The statistical features extracted by this method are 16 and they are: Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Zone Percentage, Gray Level Variance, Low Gray Level Zone Emphasis, Size-Zone Non-Uniformity, Size-Zone Non-Uniformity Normalized, High Gray Level Zone Emphasis, Large Area Emphasis, Large Area High Gray Level Emphasis, Large Area Low Gray Level Emphasis, Small Area High Gray Level Emphasis, Small Area Low Gray Level Emphasis, Zone Entropy, Zone Variance, Small Area Emphasis.

2.2.4. Neighboring Gray Tone Difference Matrix (NGTDM)

NGTDM is a method that measures the difference between a gray level value and its neighbors gray values average in a specific distance. This method can be used to extract 5 statistical features: Busyness, Coarseness, Complexity, Contrast, Strength.

2.2.5. Gray Level Dependence Matrix (GLDM)

GLDM is a method used to measure the number of connected pixels within the distance connected to a center pixel. GLDM is used to extract 14 statistical features and they are: Small Dependence Emphasis, Large Dependence Emphasis, Small Dependence Low Gray Level Emphasis, Gray Level Non-Uniformity, Dependence Non-Uniformity, Low Gray Level Emphasis, High Gray Level Emphasis, Dependence Entropy, Dependence Non-Uniformity Normalized, Gray Level Variance, Dependence Variance, Large Dependence High Gray Level Emphasis, Large Dependence Low Gray Level Emphasis, Small Dependence High Gray Level Emphasis.

2.2.6. First order statistics

First-order statistics is the smallest pixel value within the pixel values histogram of an image. The most important advantages of these features are that they are easy to obtain and allow easy differentiation of tissues containing benign and malignant tumors. The first order statistic features are: Energy, Kurtosis, 10th percentile, 90th percentile, Entropy, Robust Mean Absolute Deviation, Interquartile Range, Maximum, Mean, Mean Absolute Deviation, Median, Minimum, Range, Root Mean Squared, Skewness, Standard Deviation, Total Energy, Variance, Uniformity.

2.2.7. Shape features

Shape features use descriptive properties of the image's two-dimensional size and shape They are effective on the classification of tumor or infected area. Shape features are: 2D diameter, Elongation, Surface ratio, Maximum Mesh Surface, Minor Axis Length, Major Axis Length, Perimeter, Perimeter to Pixel Surface, Spherical Disproportion, Sphericity.

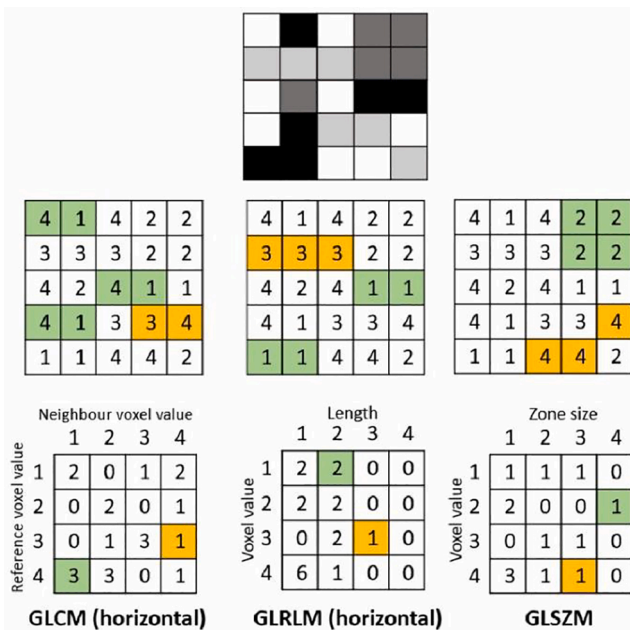


Fig. 3. An example of the calculations of GLCM, GLRLM and GLSZM matrices [22].

2.3. Machine learning algorithms

After the feature extraction, the next step was train machine learning models with the extracted features and evaluate the models with the test dataset.

In this study we have used the most powerful and commonly used machine learning algorithms because of their robustness. Support Vector Machine (SVM), Random Forest (RF), K- nearest neighbor (KNN) and XGBoost are machine learning algorithms that are used for classification. The studies referenced in this paper and many similar machine learning studies in the literature use one or more classifier from SVM, RF, KNN or XGBoost, we also preferred similar classifiers. We aim to evaluate feature effects, so we did not need to focus in detail on the performances of different classifiers.

2.3.1. Support vector machine SVM

The binary classification algorithm SVM is plot the data in an n-dimensional plane according to the number of features. In the SVM each feature correspond to the value of a special location. SVM tries to separate the features into two classes by creating a line between them. That line is called Hyperplane (in 3-dimensional space or higher) and the Equation (1) is used to find that line.

$$\mathbf{w}^T \mathbf{x} + \mathbf{b} = 0 \tag{1}$$

where the weights vector is \mathbf{w}^T , the features vector is \mathbf{x} and bias is \mathbf{b} . The hyperplane maximizes the margins using points called support vectors. Fig. 4 shows a hyperplane for two type of features X_1 and X_2 in 2-dimensional space separating the red and the green points (features) into two classes.

2.3.2. k-Nearest neighbors

The supervised machine learning algorithm KNN is an algorithm that is simple and used for regression and classification problems. KNN classifies objects based on their proximity to the training examples in the feature space. The basic form of the KNN algorithm is when $K = 1$, which means that the test example will be classified according to its nearest neighbor in the training data [23]. Fig. 5 explains the concept of this algorithm.

The distance between two examples p and q can be calculated using their coordinates in a plane by one of the Formulas below:

$$\text{Euclidean Distance } d(p, q) = \sqrt{\sum_{i=1}^K (p_i - q_i)^2} \tag{2}$$

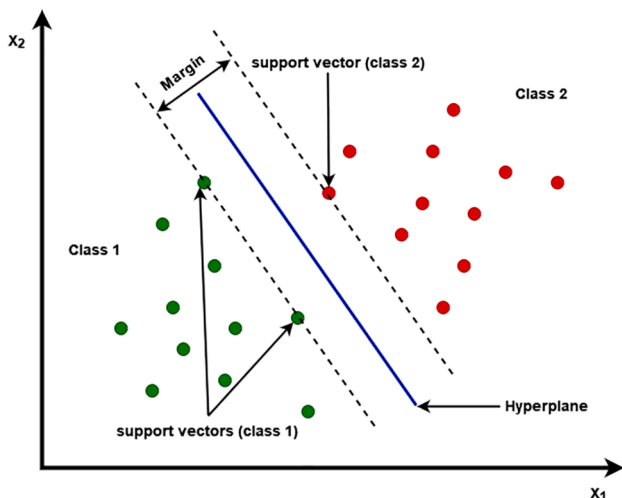


Fig. 4. SVM visualization in 2D.

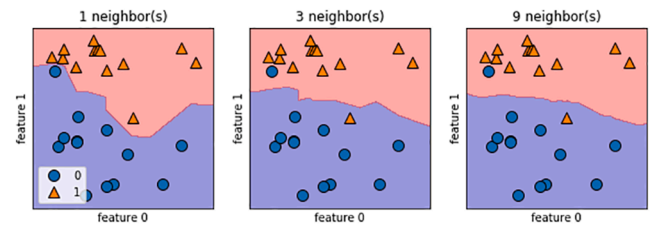


Fig. 5. Decision boundaries created by the nearest neighbors for different values of K [24].

$$\text{Manhattan Distance } d(p, q) = \sum_{i=1}^K |p_i - q_i| \tag{3}$$

$$\text{Minkowski Distance } d(p, q) = \left(\sum_{i=1}^K |p_i - q_i|^p \right)^{\frac{1}{p}} \tag{4}$$

2.3.3. Random Forest

Random forests are an fully automated machine learning techniques. It requires nearly no data preparation, or any modeling expertise. The basic building block of a random forest is the decision tree as classification and regression Trees. Models are constructed by splitting the data between different decision trees and taking the average of all decision tree predictions as the answer Fig. 6.

2.3.4. XGBoost

The Gradient Boosting algorithm optimized with various modifications and XGBoost (eXtreme Gradient Boosting) classifier is obtained as a high-performance version. The most important features of the algorithm are that it can achieve high predictive power, prevent over-learning, manage empty data and do them quickly. It is cited as the best of the decision tree-based algorithms. The principle of the Gradient Boosting -Fig. 7- is to build multiple trees/models sequentially called weak learners [25]. Where each model tries to learn from the errors of the previous one making an improved and more accurate model.

2.4. Cross validation

In the machine learning algorithms evaluation of the methods are done with the statistical method named cross validation. The basic of this method is to divide the data into two parts for training and validation of the model. Typically, the training and validation sets are incremented sequentially so that every sample in the dataset can be found in the validation set. The k-fold cross validation is the most known form of cross validation. In this method the dataset is being split into k equally parts(folds) to train and validate the model. The training and validation process is being held through k iterations such that for each

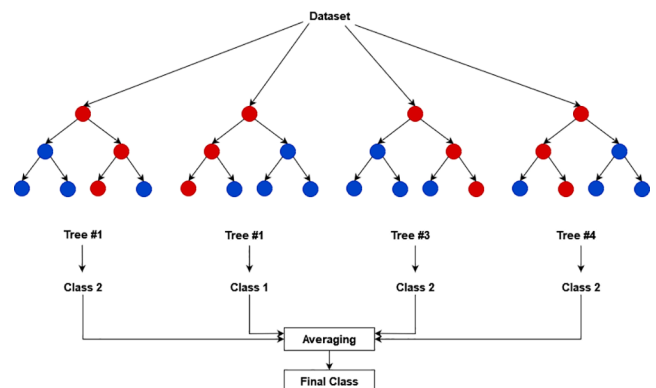


Fig. 6. Random forest model consists of four decision trees.

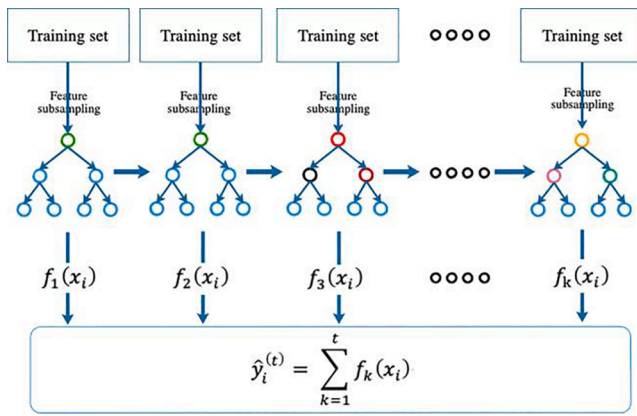


Fig. 7. The basic structure of XGBoost [24].

iteration the model is being validated using different fold. The overall performance is then taken by averaging the performance of all the folds. Fig. 8 shows how k-fold cross validation works.

3. Results

In this section the classification results of COVID-19 disease from CT images evaluated. All training and test results obtained with a Windows 10 operating system-based computer has Intel Core i5 8th generation processor and 8 GB RAM. Python 3.7.10 with open-source software packages Scikit-learn version 0.23.1 and Pyradiomics 3.0.1 were used. SVM, RF, KNN and XGBoost classifier are used for classification. These classifiers have parameters that can be tuned to control the training process called Hyperparameters. Several parameters controlled for each classification method. Also, the hyperparameters of each classification model was specified by a grid search with 10-fold cross-validation on the training data over the search spaces. The SVM results obtained with several parameters in the range. The RF, KNN and XGBoost classifiers achieve similar results for some parameters. Therefore, it was selected certain parameters from the specified range for these models. The specified hyperparameters given in the Table 2.

The test evaluation of the model is the essential part of any study. In our study the 10-fold Cross Validation technique was implemented to give realistic and more reliable results. The metrics that have been used in the evaluation process for the model performance were Accuracy, Precision, Sensitivity and F1 score. Accuracy is the number of correct predictions to the total number of predictions. Precision is the number of correct positive class predictions to the total positive class predictions. Sensitivity is the number of correct positive class predictions to the correct positive class predictions and the false negative class predictions. F1 score is the weighted average of Precision and Sensitivity. Accuracy

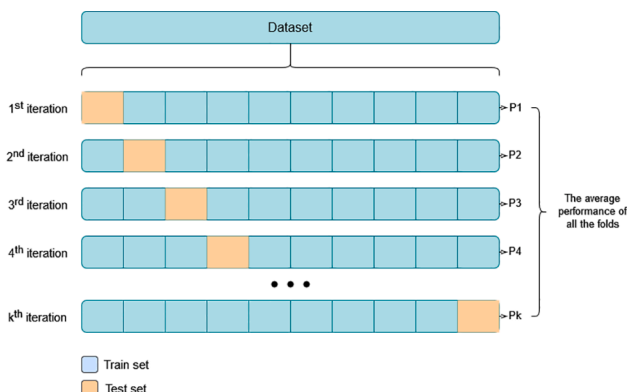


Fig. 8. The basic concept of k-fold cross validation. P stands for performance.

Table 2
The hyperparameter spaces of the machine learning models*.

Model	Hyperparameter Spaces
SVM	Kernel = { rbf , poly, sigmoid, linear} Regularization (C) = {10 ⁻² , 10 ⁻¹ , 10 ⁰ , 10 ¹ , 10 ² , 10 ³ }
RF	Gamma = {10 ⁻⁷ , 10 ⁻⁶ , 10 ⁻⁴ , 10 ⁻⁴ , 10 ⁻³ , 10 ⁻² } Min samples leaf = {1, 3, 5, 7} Min samples split = {2, 8, 10, 12} Number of estimators = {10, 50, 100, 500 , 1000}
KNN	Number of neighbors (K) = {3, 5, 7} Weights = { uniform , distance}
XGBoost	Distance metrics = { euclidean , manhattan, minkowski} Learning rate = {0.01, 0.1, 0.3 , 1.0} Max depth = {4, 6, 8, 10} Number of estimators = [10, 50, 100 , 500]

* The bold parameters were selected for overall results.

(ACC), Precision (PRE), Sensitivity (SENS) and F1 score that given in Equation 5–8 respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{7}$$

$$\text{F1 score} = \frac{2 * \text{Precision} * \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \tag{8}$$

where TP is abbreviation of true positive that the number of positive class examples predicted correctly. TN is abbreviation of true negative that the number of negative class examples predicted correctly. FP is abbreviation of false positive that the number of negative class examples predicted as positive. FN is abbreviation of false negative that the number of positive class examples predicted as negative.

Effect of different feature set was tested in the proposed method and performance results given in the tables. Here it is important to note that a lot of different feature combinations were tested. The mean results over 70% accuracy are given in this paper. Lower results are not enough meaningful in terms of classification accuracy. We extracted features with PyRadiomics package of Python. In this package, the value of each feature of the GLCM or GLRLM is calculated for each angle degree separately, after which the mean of these values is returned [26,27]. Therefore, our features are calculated for the mean of 0, 45, 90, 135 degrees. we also add trial and error by using different distance values for all feature extraction methods. We saw that these trials did not change the results significantly.

Table 3 shows classification results for GLCM features. The mean classification accuracy of GLCM low which is 75.24%. Best accuracy result obtained with XGBoost classifier as 77.9% and best F1 score obtained with KNN as 79.68%.

Table 4 shows classification results for GLRLM features. The mean classification accuracy of GLRLM is about 76.16%. Best accuracy and F1 score results are obtained with RF classifier as 79.35% and 80.9% respectively.

Table 3
Results of GLCM features.

GLCM	ACC	PRE	SENS	F1
SVM	72.66%	73.21%	75.98%	75.25%
RF	77%	79.57%	79.67%	79.68%
KNN	73.40%	74.50%	76.40%	74.10%
XGBoost	77.90%	79%	77.90%	78%

Table 4
Results of GLRLM features.

GLRLM	ACC	PRE	SENS	F1
SVM	72.60%	70.4%	71.42%	70.93%
RF	79.35%	79%	82.46%	80.9%
KNN	74.27%	75.6%	78.75%	77.96%
XGBoost	78.42%	79.42%	82.42%	78.3%

Table 5 shows classification results for GLSZM features. The mean classification accuracy is 78.45%. RF and XGBoost are achieve best results to classify GLSZM features. RF achieves 81.25% accuracy and 82.85 F1 score where XGBoost achieves 80.54% accuracy and 82.18 F1 score.

As a result of previous tables, each feature influences classification accuracy. Therefore, new feature vector is set with all the GLCM, GLRLM, GLSZM features. Table 6 shows the results for combination of these features. When we used the combined vector the mean classification accuracy is improved as 83.4% and the mean F1 is improved as about 83.77% for all classifiers. RF classifier shows best results with 85.52% and 86% for accuracy and F1 score respectively.

When the GLDM features used instead of GLCM features the feature vector is set as combination of GLDM, GLRLM, GLSZM features. Results of this combination is given in Table 7. According to Table 7 the total classification results with SVM, RF and XGBoost are decreased only KNN results are increased. But the mean values for all metrics are decreased when compared with feature vector in the Table 6. The mean accuracy is 81.9% for all classifier in Table 7 where the best accuracy is 83.24% for RF classifier.

In Fig. 9, a confusion matrix of a simple test data for RF classifier is given. The test data size is 150 (%20 of all dataset) and 70 are Covid-19, 80 are non-Covid-19. Number of TP is 60 and TN is 76 for RF classifier.

In the Table 8 the feature vector is set as combination of GLSZM, NGTDM, GLDM features. When the results of Table 8 evaluated it is seen that there are not any significant difference of accuracy for several classifier when the NGTDM features used instead of GLRLM features compared to Table 7. The mean accuracy of all classifiers is 79.35% where the best accuracy is 83.51 for RF classifier.

The results of the last combination for texture features are made with GLCM, NGTDM, GLDM given in Table 9. This combination decreases the accuracy and other scores for RF, KNN and XGBoost and only there are increase in SVM results. The mean accuracy of all classifiers is 78.57% where the best accuracy is 80.82 for RF classifier.

It is clear from Table 3-9 that the classification accuracy of texture features are reaches a maximum of 85.52% for RF classifier given in Table 6. This accuracy is acceptable for classification without radiologist comments. However, to improve the results of COVID-19 classification we also used First order statistics features and shape features of CT scan images. The results are given in Table 10 and Table 11.

The results of Shape features given in the Table 10 are very high where the mean accuracy and F1 scores in the Table 10 are about 90%. Shape features are better than texture feature for Covid-19 classification of CT images. Because the lower accuracy for shape feature is obtained by KNN as 88.75% and this is better than all previous results of Table 3-9.

Table 5
Results of GLSZM features.

GLSZM	ACC	PRE	SENS	F1
SVM	73.99%	74.35%	80.13%	78.21%
RF	81.25%	82%	84.14%	82.85%
KNN	78%	79.3%	81.64%	80.4%
XGBoost	80.54%	82.11%	82.8%	82.18%

Table 6
Results of GLCM, GLRLM, GLSZM combination features.

GLCM + GLRLM + GLSZM	ACC	PRE	SENS	F1
SVM	82.85%	84.04%	83.79%	83.41%
RF	85.52%	86.22%	85.77%	86%
KNN	80.9%	84%	81.1%	80.2%
XGBoost	84.32%	85.14%	85.78%	85.45%

Table 7
Results of GLDM, GLRLM, GLSZM combination features.

GLDM + GLRLM + GLSZM	ACC	PRE	SENS	F1
SVM	78.95%	75.88%	89.66%	81.54%
RF	83.24%	82.82%	86.57%	85%
KNN	82.98%	84.48%	81.8%	83.54%
XGBoost	82.44%	84.72%	83.48%	84.47%

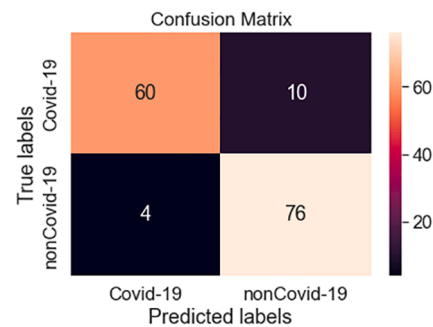


Fig. 9. Confusion matrix of RF classifier results for GLDM, GLRLM, GLSZM feature combination.

Table 8
Results of GLSZM + NGTDM + GLDM combination features.

GLSZM + NGTDM + GLDM	ACC	PRE	SENS	F1
SVM	70.11%	72.72%	74.34%	72.15%
RF	83.51%	84.25%	85.29%	84.64%
KNN	80.96%	81.33%	82.87%	82.37%
XGBoost	82.83%	82.88%	85.41%	83.51%

Table 9
Results of GLCM + NGTDM + GLDM combination features.

GLCM + NGTDM + GLDM	ACC	PRE	SENS	F1
SVM	74.78%	76.87%	79.98%	79.97%
RF	80.82%	81%	83.6%	83%
KNN	78.28%	78.78%	79.4%	78.7%
XGBoost	80.41%	82.57%	78.42%	81.73%

Table 10
Results of the Shape features.

Shape	ACC	PRE	SENS	F1
SVM	91.3%	91.3%	90%	90.64%
RF	93.82%	92.34%	93.37%	92%
KNN	88.75%	87.95%	91%	89.38%
XGBoost	91.42%	91.83%	93.22%	92.37%

Table 11
Results of the First Order Statistics features.

First Order	ACC	PRE	SENS	F1
SVM	94.36%	100%	88.9%	94.39%
RF	94.1%	96.1%	93.59%	94.25%
KNN	83.9%	90.86%	81.17%	84.98%
XGBoost	98.65%	99.73%	96.98%	98.35%

The confusion matrix for shape features and RF classifier is given in Fig. 10. Number of TP is achieved 64 and TN is 77 for RF classifier. This matrix obtained from a fold of the 10-fold cross validation process.

In this paper the last classification is made with First order statistics features and it is obtained superior performances. According to Table 11, the accuracy of classification achieves 98.65% for XGBoost, 94.36% for SVM, 94.1% for RF. Only accuracy of KNN is 83.9%. The precision, sensitivity and F1 scores are also have superior performances.

Fig. 11 brief the confusion matrix for First order statistics features and XGBoost classifier. Number of TP is achieved 70 and TN is 77 for XGBoost classifier. From the 150 test CT images almost all patients detect with true class.

The classification performance of the proposed method is compared with the literature studies in the Table 12. It is understood from the table that our proposed feature extraction method has significant accuracy rate when compared the literature. The results of our paper are also better than most of literature study. Only [6] and [10] are little better than our result, but their dataset are very clean and collected from one source.

In the literature, some methods, given in [28]-[35], adopted same dataset for their experimental results, similar to proposed method. The dataset given in [17] is publicly available. For a fair comparison, the results of proposed method compared with the method proposed in [28]-[35]. These studies use different variants of convolutional neural network for classification of Covid-19. Performance comparison between proposed method and compared methods is given in Table 13. As can be seen in Table 13, the proposed method outperforms competitors in terms of accuracy, precision and F1 metrics. Sensitivity results of our study is better than most of methods.

3.1. Additional experiments

The 10-fold cross validation is secure and robust method to test performance of studies. Although the cross-validation approach is a reliable method, some authors of medical studies use the external validation which is also known as independent test approach to avoid potential bias of cross-validation. In the independent test, a part of dataset, namely 80% of the whole dataset, is used for the training process and validation. The rest of the dataset which corresponds to the unseen data, is used for the independent test in order to evaluate the performance of the model. So, in the additional tests we re-set our model for

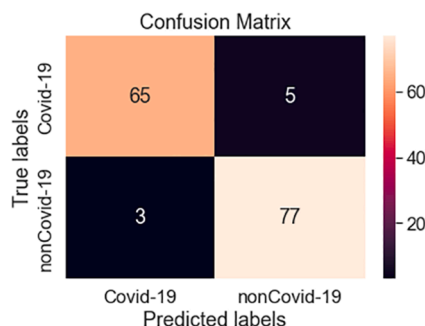


Fig. 10. Confusion matrix of RF classifier results for shape features.

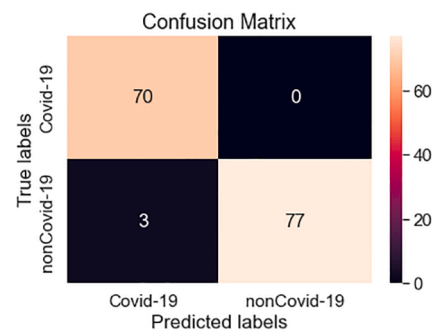


Fig. 11. Confusion matrix of XGBoost classifier results for First order statistics features.

Table 12
The performance comparison of the proposed with the literature studies.

Method	Dataset Size	Accuracy(%)
Ardakani et al. [4]	612	91.94
Al-Karawi et al. [5]	470	95.37
Barstugan et al. [6]	150	99.64
Dey et al. [7]	400	87.00
Liu et al. [8]	88	94.16
Özkaya et al. [9]	150	98.27
Kassani et al. [10]	274	99.00
Shi et al. [11]	2,685	89.40
Zheng et al. [12]	540	90.80
Xu et al. [13]	618	86.70
Song et al. [14]	274	93.00
Wang et al. [15]	1065	89.50
Alsharman et al. [16]	812	82.14
Our Proposed	746	98.65

Table 13
The performance comparison between proposed method and the literature studies under same dataset.

Method	ACC	PRE	SENS	F1
Saeedi et al. [28]	90.61	89.76	90.80	90.28
Pham [29]	96.20	92.22	95.78	96.00
Sakagianni et al. [30]	88.10	88.57	86.11	87.32
Elaziz et al. [31]	78.30	78.50	78.30	78.40
Madhavi et al. [32]	97.50	96.73	98.34	97.48
Shaik et al. [33]	97.79	97.77	97.84	97.78
Cruz [34]	86.70	88.17	83.67	85.86
Polsinelli et al. [35]	85.03	85.01	87.55	86.26
Proposed Method	98.65	99.73	96.98	98.35

independent test experiments.

In our new experiment, the existing dataset [17] is divided into two parts. 80% dataset is used for training process and the rest of the dataset, namely 20%, is used for the testing process. The performance of the trained model is tested for these test sequences. All feature vectors are tested with four different classifiers. According to the experimental results it is seen that, the independent test results supported the results of proposed method with 10-fold cross validation. All these results were compatible with the previous results. In fact, a little improvement has been achieved for gray-level based feature results. Table 14 shows the additional test results for important part of feature vectors.

4. Discussion and conclusion

This paper firstly aims to classify Covid-19 disease from CT images. Another important aim is to evaluate the effect of several feature extraction methods on classification accuracy. In the previous methods built on sophisticated methods and features, however the medical images have gray level intensity. So, intensity-based features allow to

Table 14
Independent test set scores of proposed method with COVID-CT Dataset.

Feature vector	Classifier	ACC	PRE	SENS	F1
GLSZM	SVM	70.67	69.15	81.25	74.71
	RF	83.33	82.35	87.50	84.85
	KNN	74.00	78.87	70.00	74.17
	XGBoost	82.00	80.46	87.50	83.83
GLCM + GLRLM + GLSZM	SVM	82.67	80.00	90.00	84.71
	RF	88.00	85.23	93.75	89.29
	KNN	80.67	80.00	85.00	82.42
	XGBoost	86.00	85.54	88.75	87.12
GLSZM + NGTDM + GLDM	SVM	81.33	77.66	91.25	83.91
	RF	85.33	84.52	88.75	86.59
	KNN	81.33	80.23	86.25	83.13
	XGBoost	85.33	84.52	88.75	86.59
Shape	SVM	90.00	90.12	91.25	90.68
	RF	94.00	94.00	94.00	94.00
	KNN	92.00	92.23	92.00	91.97
	XGBoost	94.67	94.67	94.67	94.67
First Order	SVM	96.67	96.89	96.67	96.67
	RF	96.67	96.67	96.67	96.67
	KNN	87.33	87.35	87.33	87.34
	XGBoost	98.67	98.67	98.67	98.67

investigate important properties of images. the Covid-19 process need to detect the diagnosis quickly. In this paper, the approximate time for the training part of [17] is 1.5 s, 17 ms, 6 ms and 0.25 s for SVM, RF, KNN and XGBoost respectively, and the test times for each image in test dataset is lower than 1 ms. The time results of referenced studies that mentioned the time are as follows: [7] requires a mean time of 173 ± 11 s, [12] needs an average of 1.93 s, [31] needs 3,123 s. Proposed method has superior performance by using correct features for medical images.

The dataset used for the proposed method was collected from several papers and it was captured by different CT modalities. The CT images has different dimensions, aspect ratio and gray levels, although proposed method achieves 98.65% accuracy. Many existing literatures used clear dataset that collected from one source.

In the proposed method we focused on both classification and watch feature effects on classification accuracy. As a result of experimental test, we achieved very good success. In the individual feature vectors of gray level, the best results given by GLSZM with 78.45% mean accuracy for all classifiers where RF classifier has 81.25% and XGBoost has 80.54% accuracies. The GLCM features gives low results when used individually but gives higher results in the combined feature vectors. The best accuracy of the combined feature vector is given by combination of GLDM, GLRLM and GLSZM features as 85.52%. Here we must note that again: each feature of the GLCM or GLRLM is calculated for each 0, 45, 90, 135 degrees angles separately. So they do not need angle based trial. Also, we add trial and error by using different distance values for all feature extraction method, but the results did not change significantly.

An important finding of this paper is that the shape features and first order statistics features are more valuable than gray level texture features. Covid-19 disease effect the lung surfaces and change its shape. Therefore, the shape feature help to classifier about better learning for classification. The same conditions are also goes on for first order statistics features. The accuracy results for shape features are about 90% where they varied between 93.82% and 88.61%, for several classifiers. The mean results of first order statistics features are achieve 92.75% where the lower is 83.9% and the best accuracy is 98.65% which is also the best for all our results. For shape features, we achieve over 90% with RF and XGBoost classifier as 93.82% and 91.42% respectively. First order features achieve 98.65%, 94.36% and 94.1% for XGBoost, SVM and RF respectively. These results obtained with the reliable and robust 10-fold cross validation. The results of proposed method are also verified with independent test. Both results have the significant superiority when compared with the similar literature studies.

Another important finding of this paper is about classifiers. The SVM

classifier shows lack of confidence for individual set of matrices however it has superior performance for first order features. RF classifier gives best result for most of feature vectors. The results of XGBoost are also very considerable. So, we recommend the reader of this paper to using RF and XGBoost classifier for Covid-19 classification. In the future, we aim to test our method on the several datasets and improve our model for more quick and more accurate diagnosis of the Covid-19 disease from CT images.

CRedit authorship contribution statement

Farid Fuad Al-Areqi: Software, Validation, Writing – original draft.
Mehmet Zeki Konyar: Supervision, Conceptualization, Methodology, Visualization, Investigation, Writing – review & editing, Software, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] A. Tahamtan, A. Ardebili, Real-time RT-PCR in COVID-19 detection: issues affecting the results, *Expert Rev. Mol. Diagn.* 20 (5) (2020) 453–454.
- [2] M. Xu, D. Wang, H. Wang, X. Zhang, T. Liang, J. Dai, X. Yu, COVID-19 diagnostic testing: Technology perspective, *Clin. Transl. Med.* 10 (4) (2020).
- [3] H. Mohammad-Rahimi, M. Nadimi, A. Ghalyanchi-Langeroudi, M. Taheri, S. Ghafouri-Fard, Application of machine learning in diagnosis of COVID-19 through X-ray and CT images: a scoping review, *Front. Cardiovas. Med.* 8 (2021) 185.
- [4] A.A. Ardakani, U.R. Acharya, S. Habibollahi, A. Mohammadi, COVIDiag: a clinical CAD system to diagnose COVID-19 pneumonia based on CT findings, *European radiology* 31 (1) (2021) 121–130.
- [5] D. Al-Karawi, S. Al-Zaidi, N. Polus, S. Jassim, Machine learning analysis of chest CT scan images as a complementary digital test of coronavirus (COVID-19) patients, *MedRxiv.* (2020).
- [6] Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. *arXiv preprint arXiv:2003.09424.*
- [7] N. Dey, V. Rajnikanth, S.J. Fong, M.S. Kaiser, M. Mahmud, Social group optimization-assisted Kapur's entropy and morphological segmentation for automated detection of COVID-19 infection from computed tomography images, *Cogn. Comput.* 12 (5) (2020) 1011–1023.
- [8] C. Liu, X. Wang, C. Liu, Q. Sun, W. Peng, Differentiating novel coronavirus pneumonia from general pneumonia based on machine learning, *Biomed. Eng. Online* 19 (1) (2020) 1–14.
- [9] U. Ozkaya, S. Ozturk, M. Barstugan, Coronavirus (COVID-19) classification using deep features fusion and ranking technique, in: *Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach*, Springer, Cham, 2020, pp. 281–295.
- [10] S.H. Kassania, P.H. Kassanib, M.J. Wesolowski, K.A. Schneider, R. Deters, Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach, *Biocybernet. Biomed. Eng.* 41 (3) (2021) 867–879.
- [11] F. Shi, L. Xia, F. Shan, B. Song, D. Wu, Y. Wei, D. Shen, Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification, *Phys. Med. Biol.* 66 (6) (2021).
- [12] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, X. Wang, Deep learning-based detection for COVID-19 from chest CT using weak label, *MedRxiv.* (2020).
- [13] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su, G. Lang, Y. Li, H. Zhao, J. Liu, K. Xu, L. Ruan, J. Sheng, Y. Qiu, W. Wu, T. Liang, L. Li, A deep learning system to screen novel coronavirus disease 2019 pneumonia, *Engineering* 6 (10) (2020) 1122–1129.
- [14] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Chong, J. Shen, Y. Zha, Y. Yang, Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, *IEEE/ACM Trans. Comput. Biol. Bioinform.* 18 (6) (2021) 2775–2780.
- [15] S. Wang, B.o. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, B.o. Xu, A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19), *European radiology* 31 (8) (2021) 6096–6104.
- [16] N. Alsharman, I. Jawarneh, GoogleNet CNN neural network towards chest CT-coronavirus medical image classification, *J. Comput. Sci.* 16 (5) (2020) 620–625.
- [17] "GitHub - UCSF-AI4H/COVID-CT: COVID-CT-Dataset: A CT Scan Dataset about COVID-19." <https://github.com/UCSF-AI4H/COVID-CT> (accessed Oct. 09, 2021).
- [18] Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., & Xie, P. (2020). COVID-CT-dataset: a CT scan dataset about COVID-19. *arXiv preprint arXiv:2003.13865.*

- [19] J.J.M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H.J.W.L. Aerts, Computational radiomics system to decode the radiographic phenotype, *Cancer research* 77 (21) (2017) e104–e107.
- [20] N.A. Saleh, H.M. Ertunç, R.A. Saleh, M.A. Rassam, A simple mask detection model based on a multi-layer perception neural network, in: 2021 International Conference of Technology, Science and Administration (ICTSA), IEEE, 2021, pp. 1–5.
- [21] V. Parekh, M.A. Jacobs, Radiomics: a new application from established techniques, *Expert Rev. Prec. Med. Drug Dev.* 1 (2) (2016) 207–226.
- [22] K. Cheng, A. Lin, J. Yuvaraj, S.J. Nicholls, D.T. Wong, Cardiac computed tomography radiomics for the non-invasive assessment of coronary inflammation, *Cells* 10 (4) (2021) 879.
- [23] S.B. Imandoust, M. Bolandraftar, Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background, *Internat. J. Eng. Res. Appl.* 3 (5) (2013) 605–610.
- [24] "ML Course." <https://prasanth-ntu.github.io/html/ML-Course-12.html> (accessed Oct. 09, 2021).
- [25] L. Huang, Y. Li, S. Chen, Q. Zhang, Y. Song, J. Zhang, M. Wang, Building safety monitoring based on extreme gradient boosting in distributed optical fiber sensing, *Optical Fiber Technol.* 55 (2020) 102149.
- [26] Gray Level Co-occurrence Matrix (GLCM) Features, <https://pyradiomics.readthedocs.io/en/latest/features.html#module-radiomics.glcm> (accessed Jan. 09, 2022).
- [27] Gray Level Run Length Matrix (GLRLM) Features, <https://pyradiomics.readthedocs.io/en/latest/features.html#module-radiomics.glrlm> (accessed Jan. 09, 2022).
- [28] Saeedi, A., Saeedi, M., & Maghsoudi, A. (2020). A novel and reliable deep learning web-based tool to detect COVID-19 infection from chest CT-scan.
- [29] T.D. Pham, A comprehensive study on classification of COVID-19 on computed tomography with pretrained convolutional neural networks, *Sci. Rep.* 10 (1) (2020) 1–8.
- [30] A. Sakagianni, Setting up an easy-to-use machine learning pipeline for medical decision support: A case study for COVID-19 diagnosis based on deep learning with CT scans, *Import. Health Inform. Public Health Dur. Pandemic* 272 (2020) 13.
- [31] M. Abd Elaziz, A. Dahou, N.A. Alsaleh, A.H. Elsheikh, A.I. Saba, M. Ahmadein, Boosting COVID-19 image classification using MobileNetV3 and aquila optimizer algorithm, *Entropy* 23 (11) (2021) 1383.
- [32] M. Madhavi, P. Supraja, Efficient explainable deep learning technique for COVID-19 diagnosis based on computed Tomography scan images of lungs. *AIP Conference Proceedings*, AIP Publishing LLC, 2022 (Vol. 2385, No. 1, p. 050001).
- [33] N.S. Shaik, T.K. Cherukuri, Transfer learning based novel ensemble classifier for COVID-19 detection from chest CT-scans, *Comput. Biol. Med.* 141 (2022) 105127.
- [34] J.F.H. Santa Cruz, An ensemble approach for multi-stage transfer learning models for COVID-19 detection from chest CT scans, *Intellig. Based Med.* 5 (2021).
- [35] M. Polsinelli, L. Cinque, G. Placidi, A light CNN for detecting COVID-19 from CT scans of the chest, *Pattern Recogn. Lett.* 140 (2020) 95–100.