# Detection of COVID-19 infection from CT images using the medical photogrammetry technique

**Hatice Çatal Reis** *1 , **Veysel Türk** 2 , **Serhat Kaya** 3

*1Gümüşhane University, Department of Geomatics Engineering, Türkiye*
*2Harran University, Department of Computer Engineering, Türkiye*
*3Dicle University, Department of Mining Engineering, Türkiye*

**Abstract**

Medical data such as computed tomography (CT), magnetic resonance imaging (MRI), and Ultrasound images are used in medical photogrammetry. CT images have been used frequently in recent years for the diagnosis of COVID-19 disease, which has contagious and fatal symptoms. CT is an effective method for early detection of lung anomalies due to COVID-19 infection. Machine learning (ML) techniques can be used to detect and diagnose medical diseases. In particular, classification methods are applied for disease diagnosis and diagnosis. This study proposes traditional machine learning algorithms Random Forest, Logistic Regression, K-Nearest Neighbor and Naive Bayes, and an ensemble learning model to detect COVID-19 anomalies using CT images. According to the experimental findings, the proposed ensemble learning model produced an accuracy of 96.71%. This study can help identify the fastest and most accurate algorithm that predicts CT images with Covid-19 during the epidemic process. In addition, machine learning-based approaches can support healthcare professionals and radiologists in the diagnostic phase.

## 1. Introduction

COVID-19 has become the disease with the highest number of cases and deaths in the world in recent years [1]. COVID-19 is a highly contagious virus. The virus has undergone multiple mutations (Alpha, Beta, Delta, and Omicron) during its active periods, and the transmission rate has increased as it mutated [2]. Although not sure, the Omicron variant is transmitted 2/3 times faster than the Delta variant [2]. The COVID-19 pandemic has devastated social life and many areas of states, especially health, education, and the economy [3]. As of March 21, 2022, it has been determined that the number of infected people has reached 400 million, and there have been more than 6 million deaths [1]. By March 27, 2020, approximately sixty million educators and nearly two billion students in more than 180 countries were adversely affected by school closures [3]. With the closure of schools, distance education applications have become widespread; because of this, teachers and students had to adapt to the new teaching methodology [4].

Problems such as technological infrastructure inadequacies and adaptation problems of teachers and students in this process caused a decrease in the quality of education, and the increase in inequalities at the point of access to technology created inequality of opportunity in education [4, 5]. Such factors have caused significant damage to the education system. Due to the COVID-19 pandemic worldwide, approximately 1.6 billion students have received distance education [6]. The COVID-19 pandemic has also created adverse effects on economic balances. Situations such as job losses, cessation of trade/tourism, and closure of businesses during the pandemic have adversely affected their economic activities [7]. During the epidemic, 255 million people lost their jobs [8]. The COVID-19 virus affects humans' respiratory, renal, neuronal, gastrointestinal, and cardiovascular systems and causes severe pathologies in many organs, such as the heart and kidney, especially the

**\* Corresponding Author**

*(hatice.catal@yahoo.com.tr) ORCID ID 0000-0003-2696-2446
(veyselturk484@gmail.com) ORCID ID 0000-0003-1250-0590
(kserhat086@gmail.com) ORCID ID 0000-0002-8824-2340

lungs [9, 10]. The infection can cause serious complications in the elderly, in people with weakened immune systems, and in patients with chronic health problems such as cardiovascular diseases such as heart valve diseases/peripheral vascular diseases, diabetes, lung cancer, liver diseases, and hypertension. Therefore, it indicates that it is pathogenic [11]. Type I (hypoxemia, lack of oxygen in the blood) and type II (hypercapnia, increased carbon dioxide in the blood) respiratory failure, pulmonary edema, inflammation in the alveoli/bronchi, fibrosis (hardening of the lung tissue), thickening of the capillary wall in patients with COVID-19 infection Capillary occlusion, pericarditis (heart membrane inflammation), myocarditis (heart muscle inflammation), intravascular coagulation, heart attack, symptoms can be seen [12, 13, 14]. Puntmann et al. [15], conducted the study with patients with COVID-19 infection, myocarditis was observed in 60 of 100 patients, and cardiac involvement was observed in 78 patients. According to Douaud et al. [16], brain tissue damage, gray matter reduction, and brain volume reduction were detected in patients with COVID-19 infection. SARS-CoV-2 in the first stage of the disease stage; muscle pain, fatigue, and fever are common symptoms in the second stage; Ground-glass opacities (GGO) are seen in the majority of computed tomography (CT) images in the third stage; hypercoagulation in patients undergoing treatment, in the fourth and final stage; A picture of multi-organ failure occurs as a result of an excessive response of the immune system [17]. IgM-IgG antibody tests, reverse transcription polymerase chain reaction to diagnose and detect COVID-19 patients. (RT-PCR) test kits have been widely used. In addition, CT, and chest X-ray (CXR) techniques, which are among the medical photogrammetric techniques, are also among the essential methods used in the disease diagnosis process [18]. In some cases, IgM-IgG antibody tests and RT-PCR can give incorrect or inadequate information [19, 20]. Chest CT/CXR shows abnormalities in the lung [21, 22]. Therefore, a CT/CXR scan can play a crucial role in the early diagnosis of COVID-19 infection [23]. Chest CT has shown to be a useful complement to test kits, and it has been presented to have better accuracy in diagnosing COVID-19 [24-28].

Among the medical imaging techniques, CT and CXR are frequently used in disease diagnosis. Among the advantages and disadvantages of the CT imaging technique compared to the CXR imaging technique; (i) CT shows higher accuracy in diagnostic processes [29]. (ii) It is the gold standard in diagnosing pneumonia, especially in adult patients. (iii) CT imaging technique is more damaging than CXR because of the radiation it emits [29]. Moreover, (iv) the high cost of the devices is one of the obstacles to their widespread use in hospitals [29]. Among the advantages and disadvantages of the CXR imaging technique compared to the CT imaging technique; (i) CXR is non-invasive, and its radiation rate is lower than CT. (ii) It is widely used in emergency departments because of its cheapness [30]. (iii) CXR analysis process is more complex than CT [31]. (iv) It is sensitive to noisy areas. (v) It performs poorly detecting anomalies in relatively small areas [31, 32]. Here, medical images obtained with the CT imaging technique contain more detail than CXR; medical images obtained from the CT imaging technique were used in anomaly detection in this study due to factors such as its success in detecting anomalies.

The application of machine learning models is promising to increase the diagnostic accuracy in the disease detection process from radiological images [33]. ML methods include COVID-19, ex-ebola, cholera, H1N1 influenza, zika, oyster norovirus, etc., applied in pandemics [34-36]. Chen [37] used the Support Vector Machine (SVM) algorithm with 10-fold cross-validation (cv) approach for the detection of COVID-19 from 296 (148 COVID-19, 148 Non-COVID-19) chest CT images. In addition, the author used Histogram Equalization (HE) and Gray-Level Co-Occurrence Matrix (GLCM) techniques to increase the efficiency of the proposed method. In addition, for the detailed performance analysis of the proposed method with the SVM algorithm, an experimental comparison was made with the K-Nearest Neighbors (KNN) and Naive Bayes (NB) algorithms under the same conditions. As a result, while the HE-GLCM-SVM hybrid method produced 75.69% accuracy in the COVID-19 detection study, the HE-GLCM-KNN and HE-GLCM-NB methods produced 69.63% and 66.46% accuracy, respectively. However, the limitation of the study was that testing was carried out with limited samples. The methods proposed here can be tested with a large-scale dataset to validate their findings in the study. In addition, the data set used in the study is lossy data. For the disease detection process to be carried out with maximum efficiency, it is necessary to perform the classification process with raw data with a dicom extension. According to Hasoon et al. [38], KNN and SVM algorithms were used to detect COVID-19 from chest X-ray images. In the study using 5000 medical images, the proposed models were tested with the 5-fold CV approach. In the pre-processing steps in the study; image thresholding, image noise removal, morphological operation, and segmentation application; Region of Interest (ROI), Haralick texture features, Histogram of Gradient (HOG), Local binary pattern (LBP) methods were applied in feature extraction. In the proposed study, LBP-KNN was the most successful method, with an average accuracy of 98.66%. According to the results (i), the proposed method for classifying and early detecting COVID-19 disease presented successful data. (ii) CXR imaging technique is a successful method for detecting COVID-19 disease. The limitation of the study is that the proposed methods were tested with a relatively small dataset. A detailed analysis should be performed with a large-scale dataset to verify the model's performance. Barstugan et al. [39] used the SVM algorithm for the early-stage detection of COVID-19 disease, and the dataset used consists of 150 chest CT images. In classification, testing was carried out with the 2, 5, and 10-fold CV approach. The authors used the discrete wavelet transform (DWT), local directional Pattern (LDP), Gray Level Co-occurrence Matrix (GLCM), Gray-Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GRLLM) methods to improve the performance of classification. Extraction was performed. In the study performed with GLSZM feature extraction and SVM algorithm with 5-fold CV, an accuracy value of 98.71%

was obtained. The proposed method in this study was trained with a small-scale dataset, and performance analysis was performed. The number of test data used to validate the model may need to be increased. Accordingly, a different and large-scale dataset should be used for detailed performance analysis of the proposed method. Yang et al. [40] had 180 (90 COVID-19 patients, 90 other types of pneumonia patients) chest CT images used the SVM algorithm to detect COVID-19. In the study, feature extraction was performed with A Gray Level Co-occurrence Matrix (GLCM), The Gray Run Length Matrix (GRLLM), The Neighborhood Gray Level Difference Matrix (NGLDM), The Gray Level Zone Length Matrix (GLZLM), and Histogram methods. The obtained features were classified using the SVM algorithm. In the application performed with GLCM-SVM, an accuracy value of 85.95% was obtained. Among the limitations of this study is the use of a small data set in the training and testing process of machine learning algorithms. In addition, machine learning algorithms were applied within the scope of the study in the disease diagnosis process, as well as convolution neural network-based deep learning architectures, which have been able to produce successful results in the field of image processing in recent years, could also be used in the disease diagnosis process. Although there is previous research on the COVID-19 disease detection process, more comprehensive analyzes are needed to test the success of machine learning algorithms in the disease detection process. In this study, we aimed to detect COVID-19 patients based on machine learning automatically. These models are Random Forest (RF), Logistic Regression (LR), k-Nearest Neighbors (KNN), Naive Bayes (NB), and Ensemble learning (EL) models. In the study, the features obtained by the deep learning model were used for the input data of machine learning algorithms. Deep learning is a machine learning method that consists of multi-layered neural networks. Deep learning can automatically discover complex patterns in data and essential features in representative data. Experimental stage;

(i) Deep features were extracted from the COVID-CT dataset with the DenseNet201 model previously trained with the ImageNet dataset.

(ii) feature vector obtained from deep learning algorithm was used as input data of RF, DT, KNN, and Gaussian NB machine learning algorithms. The classification performance of machine learning algorithms is directly proportional to the selection of hyper-parameters.

In this study, the Randomized Search CV algorithm is used in the process of determining the most suitable hyper-parameters for machine learning algorithms. (iii) Principal Component Analysis (PCA) method has been used to reduce the computational cost of the algorithms, eliminate noise and unnecessary information, and reduce the data size. The PCA method was applied to the feature vector obtained from the deep learning algorithm in the experimental process. PCA is a statistical method used to determine the relationships among the variables in a multivariate dataset and to determine the basic

structures of the variables. PCA calculates the correlation of variables in the data set and determines the principal components using eigenvalue analysis. Eigenvalues and eigenvectors are obtained from the correlation matrix. When the eigenvalues are ordered from largest to smallest, the principal components with the largest eigenvalues are selected. These principal components represent the most variable and essential features of the data. PCA compresses or reduces the size of data through selected principal components. During this process, essential properties of the data are preserved. Thus, the data set is expressed with fewer variables. (iv) Finally, another method applied within the scope of the study to increase the success of machine learning algorithms in the diagnosis process of COVID-19 patients is ensemble learning. The voting classifier classification method is used in the Ensemble learning method. How can we predict cost-effectively COVID-19 patients, and how can we choose the best algorithm? The accuracy of the method is a critical element in evaluating the classification performance. There are many possibilities to explore in medical imaging [41], and each discipline offers approaches with its workspace. The important thing is to present effective scientific development and study. Unlike the medical literature, we presented a machine learning-based diagnostic approach. Radiologists and clinic physicians must be the pioneer in the Pandemic in medical imaging and interpretation [42]; however, different experts can be required to contribute to the system and the process. In this challenging process, Machine learning-based approaches can offer a secondary perspective.

## 2. Material and Method

Machine learning is considered a branch of artificial intelligence. Machine-learning algorithms are proven as a valuable part of computer-aided diagnosis and decision [43]. Machine learning algorithms and ensemble learning methods were used to predict non-COVID-19/COVID-19 from Chest-CT data.

The study has two main steps in the classification process: training and testing. The suggested workflow diagram is given in Figure 1. The algorithms used for the study are Random Forest, Logistic Regression, KNN, and Naive Bayes, respectively. The Voting Classifier algorithm was used in the Ensemble learning application.

### 2.1. Dataset

This section describes how the COVID-CT dataset is obtained and built (Table 1). We collected 757 CT images from https://github.com/UCSD-AI4H/COVID-CT [44]. GitHub is a public platform. These CT images have different sizes (e.g., 525x442, 450x319, 485x345 etc.) and different standards. All Chest-CT images were resized due to different scanning characteristics and image size differences. Table 1 represents the ML methods used and statistics of the dataset for COVID-19 prediction. Some examples of the dataset used in the study are given in Figure 2.
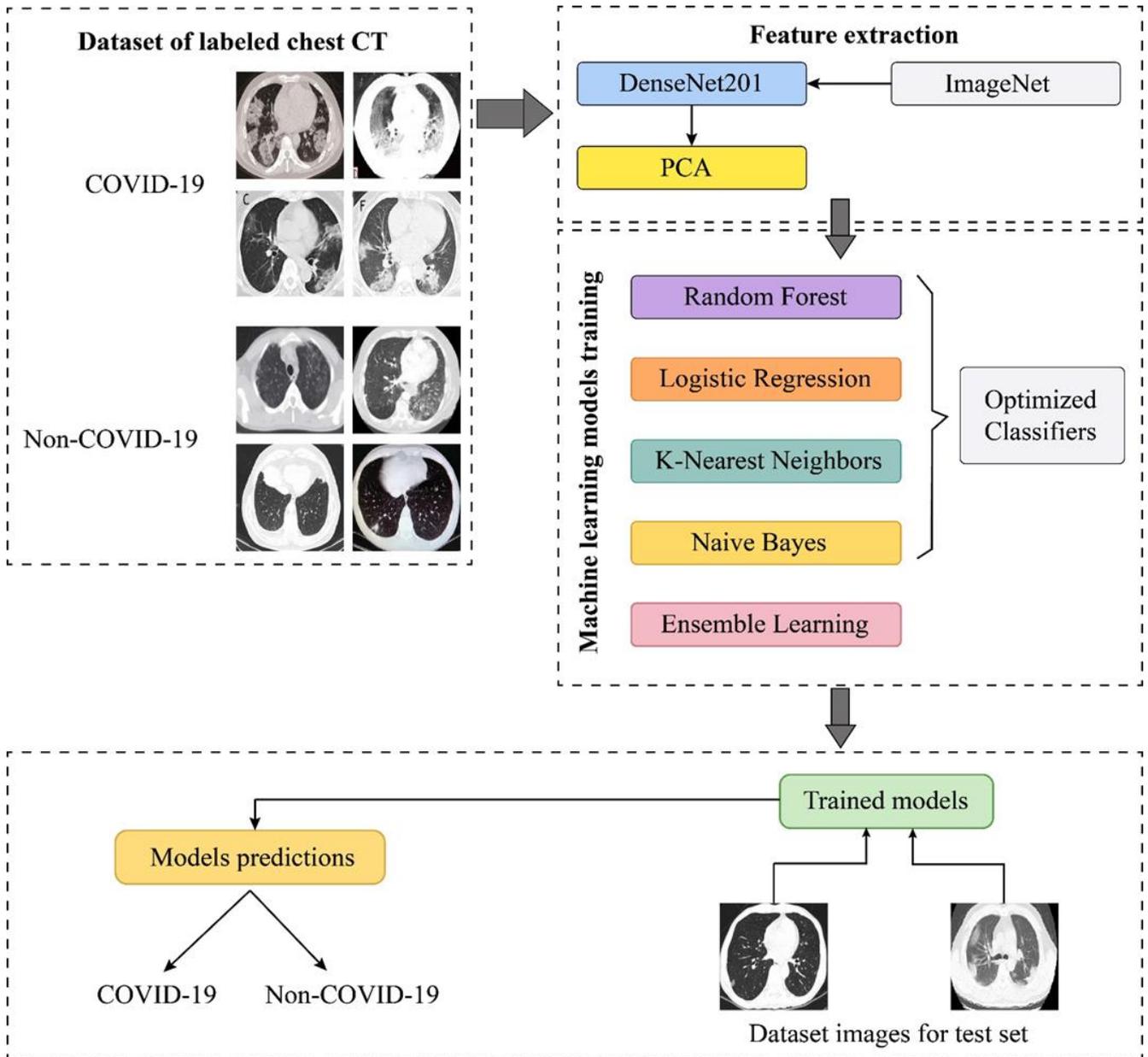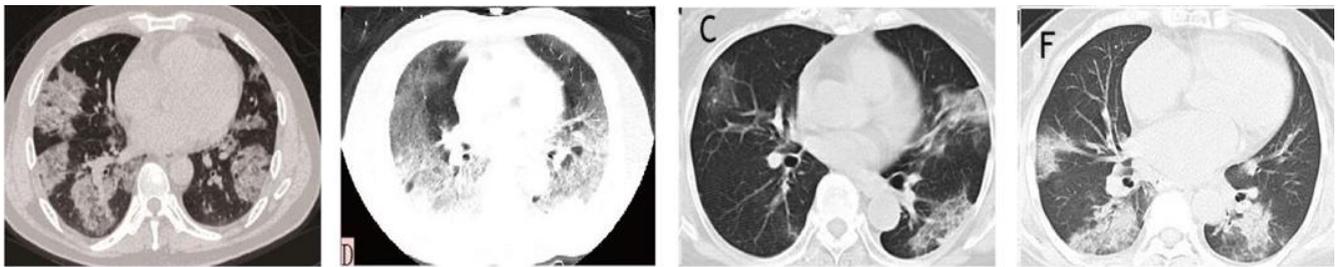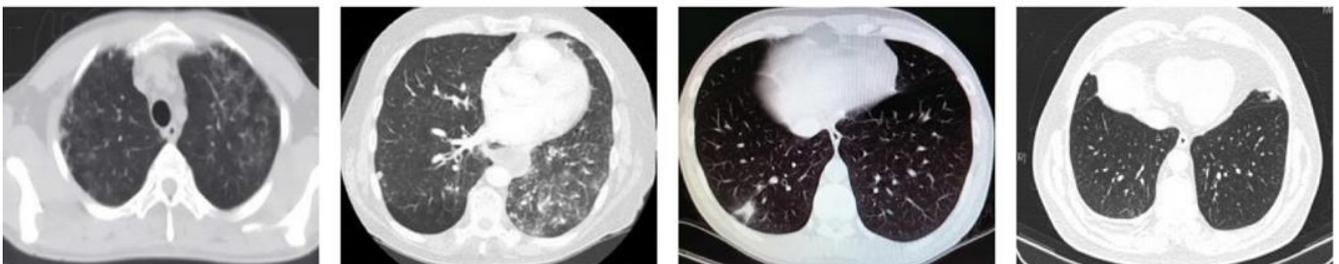
**Figure 1.** Flowchart of the study.



(a) COVID-19 infected patients



(b) Non-COVID-19 infected patients

**Figure 2.** Samples of CT images.

**Table 1.** Statistics of the dataset.

| Medical imaging technique | Label | Category | Number of images | Number of training | Number for testing |
|---|---|---|---|---|---|
| CT | 0 | Non-COVID-19 | 397 | 317 | 80 |
| | 1 | COVID-19 | 360 | 288 | 72 |

## 2.2. Implementation details of machine learning algorithms and pre-processing

This research study presents a methodology for classifying COVID-19 chest CT and normal chest CT images using machine learning architectures. We implement our models using the Python programming language and scikit-learn machine learning library. The study was done on a laptop with an Intel i5 processor, 6 GB of RAM, and a GTX 940MX NVidia GPU with 2GB of VRAM. In this study, feature extraction from the COVID-CT dataset was performed with the pre-trained DenseNet201 [45] deep learning model. Here are the preprocessing steps applied: (i) All images are resized from 224x224 pixels. The interpolation (INTER_CUBIC) technique was used in the resizing process. (ii) In the second step, min-max normalization was applied to the matrix obtained in the first step. Min-max normalization compresses the values in the dataset into a unit range. Thus, different variables/features share the same scale. This allows the model to learn the relevant features in a balanced way. (iii) Labels are assigned for data entries in the final stage. Accordingly, the "0" label for Non-COVID-19 images and the "1" label for COVID-19 images are defined.

Finally, feature vectors obtained from the deep learning model are used as input data for optimized machine learning algorithms. In addition, the PCA method has been applied to reduce the feature vector size obtained in the last step. In the experimental process, before and after the application of the PCA method, it was analyzed comparatively. In the application made with the PCA method, 200 features were used. Finally, within the scope of the study, the EL method was applied for a collective learning requirement with RF, DT, KNN, and Gaussian NB algorithms used in the diagnosis process of COVID-19 patients. Before and after applying the PCA method in the application process of the EL method, it was analyzed comparatively. The results are in the Results section. The trans-test approach was used in the studies performed with a machine learning algorithm and EL. Accordingly, the CT dataset consisting of Non-COVID-19/COVID-19 images is divided into 80% training and 20% test dataset.

The training process of machine learning algorithms was carried out with the training dataset. After the training process was completed, the performance of the models was performed with the test dataset. Accuracy, precision, recall, F1-score, MSE, RMSE, and confusion metrics were used to evaluate the experimental results.

## 2.3. Machine learning algorithm

In this section, the theoretical framework of traditional machine learning algorithms used for the detection of patients with COVID-19 infection is given. In this study, the Randomized Search CV algorithm in the sklearn-model_selection module was used to determine the hyperparameters of machine learning algorithms used to detect COVID-19 patients. The hyperparameters used in the algorithms are given in Table 2.

**Table 2.** Random Forest, Logistic Regression, K-Nearest Neighbors and Naive Bayes hyperparameters.

| ML algorithms | Hyperparameters | Defined Parameters |
|---|---|---|
| Random Forest | bootstrap | False |
| | max_features | "auto" |
| | n_estimator (number of trees) | 50 |
| | min_samples_split | 18 |
| | min_samples_leaf | 3 |
| | max_depth | 182 |
| | criterion | "gini" |
| | random_state | 0 |
| Logistic Regression | C | 1 |
| | penalty | "l2" |
| | max_iter | 79 |
| | solver | "saga" |
| | multi_class | "auto" |
| K-Nearest Neighbors | n_neighbors | 13 |
| | weights | "uniform" |
| | algorithm | "brute" |
| | leaf_size | 148 |
| | p | 2 |
| | metric | "euclidean" |
| Naive Bayes | priors | None |
| | var_smoothing | 1e-11 |

### 2.3.1. Random Forest

RF is a supervised machine learning method with decision networks based on the classification algorithm [46, 47], and it is often preferred in image classification [48]. The algorithm is trained by averaging the obtained samples according to the set of Decision Trees. It combines multiple classifiers to solve a complex problem and improves the model's performance [49]. A dataset contains large-size CT images, and the dataset is divided into subsets and sent to each decision tree. The Random Forest algorithm is noise-resistant and has better performance [50].

### 2.3.2. Logistic Regression

The logistic regression algorithm explains the connection between the two-dimensional response variable. The most significant feature of logistic regression is that variables do not require normal distribution [51]. Logistic Regression is divided into three types as Binomial, Multinomial, and Ordinal. We used binomial logistic regression, which has two possible dependent variables, such as COVID-19 or non-COVID-19.

### 2.3.3. K-Nearest Neighbors

The K-Nearest Neighbors is one of the traditional supervised machine learning algorithms. KNN is a non-parametric algorithm that estimates the data into categories and to which class the newest data will be included [52, 53]. Uses all of the data in the KNN dataset. The success of KNN depends on the Euclidean distance metric used to recognize neighbors close to the test data of the data to be classified [54]. Euclidean distance sends the newly added data to the nearest class with the help of Euclidean distance. In the algorithm, the number k has no optimal value. The trial-and-error method can determine the k metric by experimental processes.

### 2.3.4. Naive Bayes

Naive Bayes is the oldest machine learning method [43]. This method does not include the same iterative training process as many other machine learning methods. The Bernoulli classifier works the independent Booleans variables, and it is a fast and easy ML algorithm to predict a class of data. The Bernoulli Naive Bayes classifier assumes that features take only two values (COVID-19 or non-COVID-19).

### 2.3.5. Ensemble Learning

Ensemble learning is a machine learning algorithm. It strategically combines multiple heterogeneous/homogeneous classifiers to create a high-performance model in classification/regression applications [55]. Many studies show that the ensemble learning method is widely used in different problems [55]. In this study, the Voting Classifier classifier was used. In this research, the Voting Classifier function in the ensemble module of the Scikit-learn library was used to classify the COVID-19 disease proposed by the voting ensemble learning method. Other hyperparameters include the estimators (['RF', rf), ('GNB', gnb), ('LR', lr), ('KNN', knn)]), voting=(hard)).

### 2.4. Quantitative Analysis

Receiver operating characteristics (ROC) curves, $F_1$-score, Precision, Recall, Accuracy, Mean Square Error (MSE), and Root Mean Square Error (RMSE) metrics were used to perform quantitative analysis in the classification performed with machine learning algorithms. In this section, mathematical expressions of evaluation metrics are given. The Confusion Matrix (CM) is summarized in Table 3.

**Table 3.** Confusion matrix

|  | Predicted Positive | Predicted Negative |
|---|---|---|
| Actual Positive | True Positive (TP) | False Negative (FN) |
| Actual Negative | False Positive (FP) | True Negative (TN) |

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$F_1 - \text{Score} = \frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (4)$$

$$\text{MSE} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \quad (5)$$

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (6)$$

In Equation (5-6); "$n$" total number of data, "y" actual values "ŷ" predictive values.

## 3. Results

In our experiment, we used different machine-learning methods for evaluation. This study evaluated the applicability of four machine learning models (Random Forest, Logistic Regression, K-Nearest Neighbors, and Naive Bayes) and an ensemble learning model for diagnosing COVID-19 patients. According to the accuracy metric value before the PCA application, Random Forest, K-Nearest Neighbors algorithms, and the proposed ensemble learning method produced the most successful results with a value of 96.71%. On the other hand, in the COVID-19 disease diagnosis process, Naive Bayes has the lowest success rate of 94.74% according to the accuracy metric in this study (Table 4). Our other findings, according to the F1-score metric value, the Random Forest produced the most successful results with a rate of 96.60%. In comparison, Naive Bayes produced the lowest result, with a rate of 94.59%.

The Precision, Recall, F1-score, MSE, and RMSE metrics of the Ensemble Learning method were 95.89%, 97.22%, 96.55%, 0.0329, and 0.1814, respectively.

The confusion matrix of machine learning algorithms is given in Figure 3.

The statistical results obtained according to the dataset classes before the PCA application are given in Table 5.
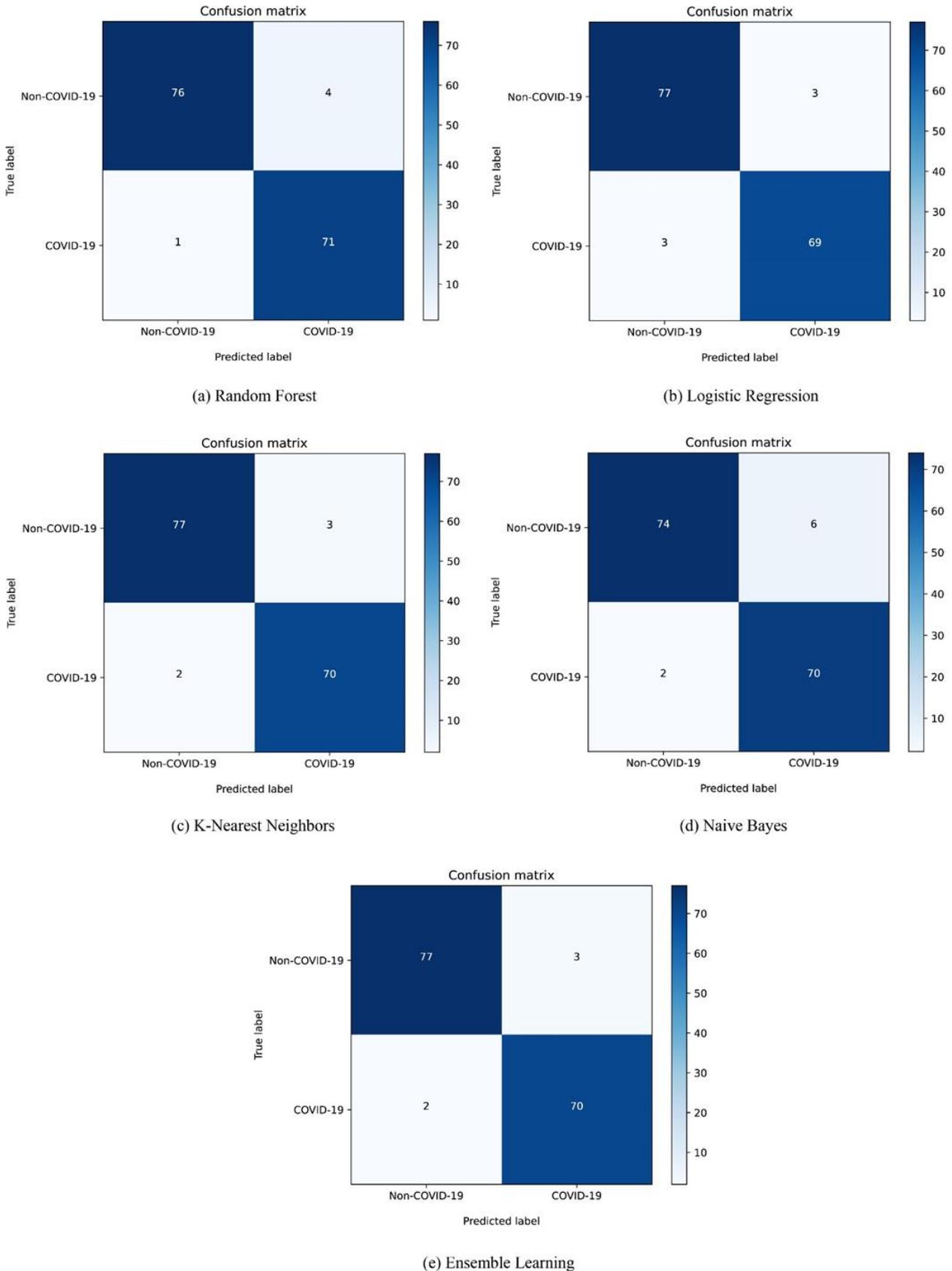


(a) Random Forest



(b) Logistic Regression



(c) K-Nearest Neighbors



(d) Naive Bayes



(e) Ensemble Learning

**Figure 3.** Confusion matrix visualization of machine learning algorithms (without PCA).

**Table 4.** Comparing the performance of algorithms (without PCA).

| Models | Accuracy | Precision | Recall | $F_1$-score | MSE | RMSE |
|---|---|---|---|---|---|---|
| Random Forest | 0.9671 | 0.9467 | 0.9861 | 0.9660 | 0.0329 | 0.1814 |
| Logistic Regression | 0.9605 | 0.9583 | 0.9583 | 0.9583 | 0.0395 | 0.1987 |
| K-Nearest Neighbors | 0.9671 | 0.9589 | 0.9722 | 0.9655 | 0.0329 | 0.1814 |
| Naive Bayes | 0.9474 | 0.9211 | 0.9722 | 0.9459 | 0.0526 | 0.2294 |
| Ensemble Learning | 0.9671 | 0.9589 | 0.9722 | 0.9655 | 0.0329 | 0.1814 |

**Table 5.** Results of before/after pre-training precision, recall, F1-score metrics (without PCA).

| Model | Class | Precision | Recall | $F_1$-score | Number for testing |
|---|---|---|---|---|---|
| Random Forest | Non-COVID-19 | 0.9870 | 0.9500 | 0.9682 | 80 |
| | COVID-19 | 0.9467 | 0.9861 | 0.9660 | 72 |
| Logistic Regression | Non-COVID-19 | 0.9625 | 0.9625 | 0.9625 | 80 |
| | COVID-19 | 0.9583 | 0.9583 | 0.9583 | 72 |
| K-Nearest Neighbors | Non-COVID-19 | 0.9747 | 0.9625 | 0.9686 | 80 |
| | COVID-19 | 0.9589 | 0.9722 | 0.9655 | 72 |
| Naive Bayes | Non-COVID-19 | 0.9737 | 0.9250 | 0.9487 | 80 |
| | COVID-19 | 0.9211 | 0.9722 | 0.9459 | 72 |
| Ensemble Learning | Non-COVID-19 | 0.9747 | 0.9625 | 0.9686 | 80 |
| | COVID-19 | 0.9589 | 0.9722 | 0.9655 | 72 |

In Table 5, the testing process was carried out with machine learning models and 152 CT images.

Before the PCA application, according to Table 5, the Random Forest algorithm produced 96.82% F1-scores in detecting 80 Non-COVID-19 images and 96.60% F1-scores in detecting 72 COVID-19 images. The K-Nearest Neighbors algorithm produced 96.86% F1-scores in detecting 80 Non-COVID-19 images and 96.55% F1-scores in detecting 72 COVID-19 images. The naive Bayes algorithm produced a 94.87% F1-score in detecting 80 Non-COVID-19 images and a 94.59% F1-score in detecting 72 COVID-19 images. While the Ensemble Learning method produces 97.47%, 96.25%, and 96.86% values, respectively, according to the Precision, Recall, and F1-score metrics in detecting 80 Non-COVID-19 images, it has Precision, Recall, F1-score metrics in detecting 72 COVID-19 images. It produced 95.89%, 97.22%, and 96.55% values, respectively.

Before the PCA application, in Figure 3a, the Random Forest model produced 5 incorrect (FN=1, FP=4) prediction results on 152 test images. However, at the same time, it performed successfully by making correct predictions for 147 (TN=76 and TP=71) test data exhibited. In Figure 3b, the Logistic Regression model produced 6 incorrect (FN=3, FP=3) prediction results on 152 test images, while correct predictions for 146 (TN=77 and TP=69) test data. In Figure 3c, the K-Nearest Neighbors model produced 5 incorrect (FN=2, FP=3) prediction results on 152 test images. At the same time, it predicted correctly for 147 (TN=77 and TP=70) test data and produced the most successful result with the Random Forest algorithm. In Figure 3d, the Naive Bayes model, which produced the lowest result before PCA, produced 8 incorrect (FN=2, FP=6) prediction results on

152 test images, while it predicted correctly for 144 (TN=74 and TP=70) test data. Finally, in Figure 3e, the Ensemble Learning method produced 5 incorrect (FN=2, FP=3) prediction results on 152 test images. In contrast, correct predictions for 147 (TN=77 and TP=70) test data, K-Nearest Neighbors, and Random Forest algorithms produced the most successful result.

After applying the PCA method, the Logistic Regression algorithm produced the most successful result with a value of 96.05%, according to the accuracy metric value. In the COVID-19 disease diagnosis process, Random Forest has the lowest success rate of 92.76% according to the accuracy metric value in this study (Table 6). Our other findings; according to the F1-score metric value, the Logistic Regression algorithm produced the most successful result with a rate of 95.83%, while Random Forest produced the lowest result at 92.52%. The Accuracy, Precision, Recall, F1-score, MSE, and RMSE metrics of the Ensemble Learning method were 93.42%, 91.89%, 94.44%, 93.15%, and 0.0658, 0.2565, respectively. The confusion matrix of machine learning algorithms is given in Figure 4.

After applying the PCA method, the statistical results obtained according to the dataset classes are given in Table 7. In Table 7, the testing process was carried out with machine learning models and 152 CT images.
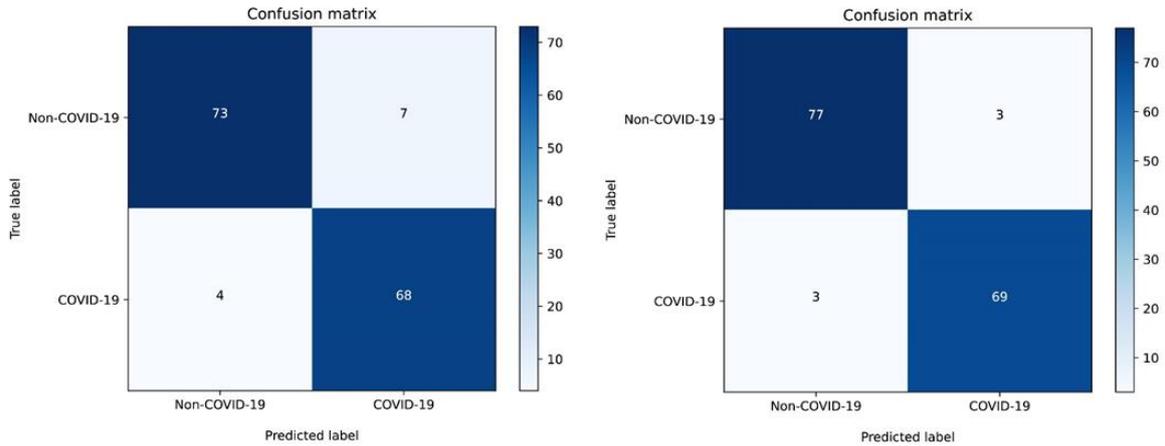
After applying the PCA method, according to Table 7, the Logistic Regression algorithm produced 96.25% F1-scores in detecting 80 Non-COVID-19 images and 95.83% F1-scores in detecting 72 COVID-19 images. The Random Forest algorithm produced 92.99% F1-scores in detecting 80 Non-COVID-19 images and 92.52% F1-scores in detecting 72 COVID-19 images.

**Table 6.** Comparing the performance of algorithms (with PCA).

| Models | Accuracy | Precision | Recall | $F_1$-score | MSE | RMSE |
|---|---|---|---|---|---|---|
| Random Forest | 0.9276 | 0.9067 | 0.9444 | 0.9252 | 0.0724 | 0.2690 |
| Logistic Regression | 0.9605 | 0.9583 | 0.9583 | 0.9583 | 0.0395 | 0.1987 |
| K-Nearest Neighbors | 0.9408 | 0.9200 | 0.9583 | 0.9388 | 0.0592 | 0.2433 |
| Naive Bayes | 0.9342 | 0.9189 | 0.9444 | 0.9315 | 0.0658 | 0.2565 |
| Ensemble Learning | 0.9342 | 0.9189 | 0.9444 | 0.9315 | 0.0658 | 0.2565 |

**Table 7.** Results of before/after pre-training precision, recall, F1-score metrics (with PCA).

| Model | Class | Precision | Recall | $F_1$-score | Number for testing |
|---|---|---|---|---|---|
| Random Forest | Non-COVID-19 | 0.9481 | 0.9125 | 0.9299 | 80 |
| | COVID-19 | 0.9067 | 0.9444 | 0.9252 | 72 |
| Logistic Regression | Non-COVID-19 | 0.9625 | 0.9625 | 0.9625 | 80 |
| | COVID-19 | 0.9583 | 0.9583 | 0.9583 | 72 |
| K-Nearest Neighbors | Non-COVID-19 | 0.9610 | 0.9250 | 0.9427 | 80 |
| | COVID-19 | 0.9200 | 0.9583 | 0.9388 | 72 |
| Naive Bayes | Non-COVID-19 | 0.9487 | 0.9250 | 0.9367 | 80 |
| | COVID-19 | 0.9189 | 0.9444 | 0.9315 | 72 |
| Ensemble Learning | Non-COVID-19 | 0.9487 | 0.9250 | 0.9367 | 80 |
| | COVID-19 | 0.9189 | 0.9444 | 0.9315 | 72 |



(a) Random Forest

(b) Logistic Regression
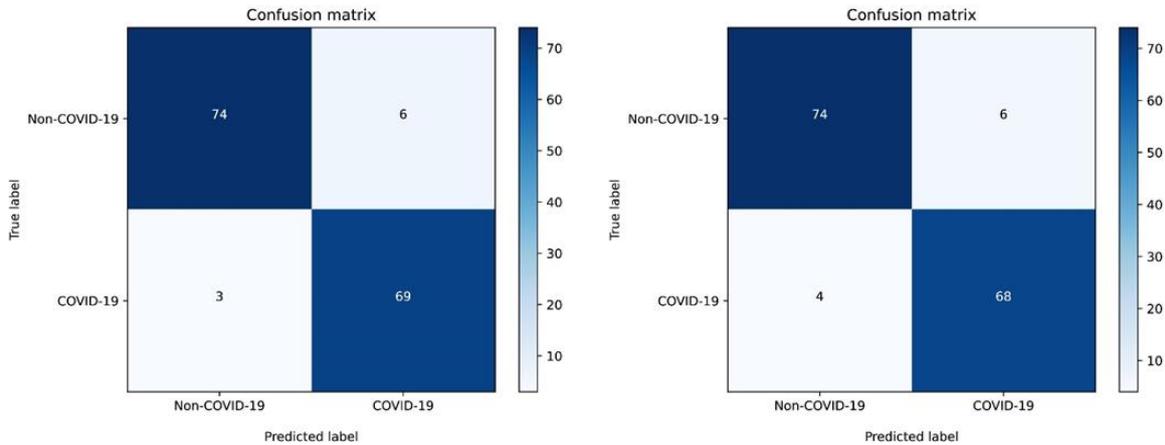
(c) K-Nearest Neighbors

(d) Naive Bayes
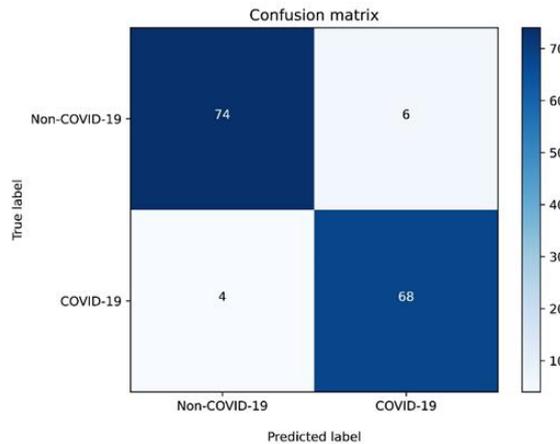
(e) Ensemble Learning

**Figure 4.** Confusion matrix visualization of machine learning algorithms (with PCA).

While the Ensemble Learning method produces 94.87%, 92.50%, and 93.67% values, respectively, according to the Precision, Recall, and F$_1$-score metrics in detecting 80 Non-COVID-19 images, it has Precision, Recall, F$_1$-score metrics in detecting 72 COVID-19 images. It produced 91.89%, 94.44%, and 93.15% values, respectively.

In Figure 4a, after applying the PCA method, the Random Forest model that produced the lowest result made 11 incorrect (FN=4, FP=7) prediction results on 152 test images while 141 (TN=73 and TP=68) correct for test data guessed.

In Figure 4b, the Logistic Regression model produced 6 incorrect (FN=3, FP=3) prediction results on 152 test images, while it correctly predicted 146 (TN=77 and TP=69) test data, achieving the most successful performance.

In Figure 4c, the K-Nearest Neighbors model produced 9 incorrect (FN=3, FP=6) prediction results on

152 test images, while it correctly predicted 143 (TN=74 and TP=69) test data.

In Figure 4d, the Naive Bayes model produced 10 incorrect (FN=4, FP=6) prediction results on 152 test images, while it predicted correctly for 142 (TN=74 and TP=68) test data.

Finally, in Figure 4e, the Ensemble Learning method produced 10 incorrect (FN=4, FP=6) prediction results on 152 test images, while it correctly predicted 142 (TN=74 and TP=68) test data.

According to Figure 5, before the PCA method was applied, the machine learning algorithms Random Forest, K-Nearest Neighbors, and Ensemble Learning produced the most successful results. In contrast, Naive Bayes had the most unsuccessful results.

After applying the PCA method, Logistic Regression produced the most successful result, while Random Forest produced the most unsuccessful result.
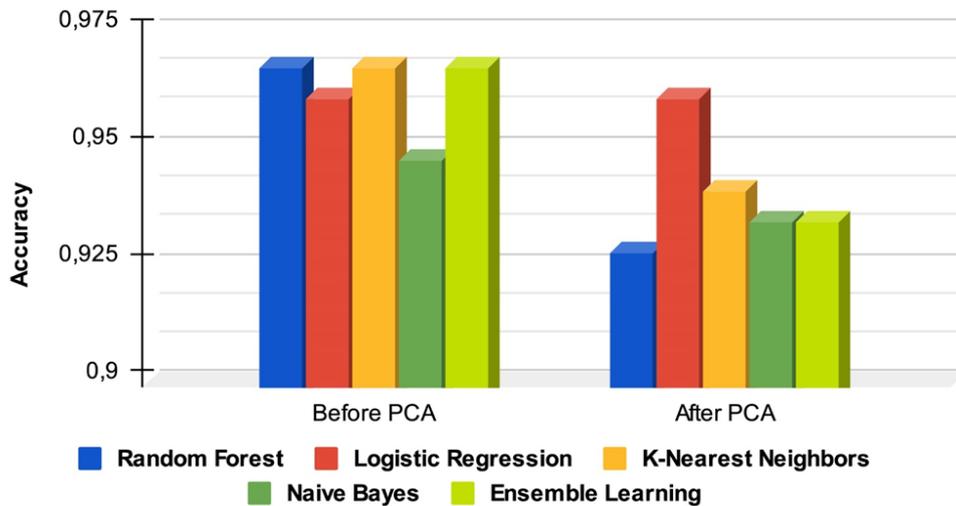


**Figure 5.** Chest CT dataset pre and after-PCA method accuracy values.

## 4. Discussion

In this paper, the definition of the most well-known machine learning-based methods were presented and explained to predict Covid-19-CT images. This study presents a comparative analysis for Random Forest, k-Nearest Neighbors, Logistic Regression, Naive Bayes algorithms of ML models to predict/diagnose the COVID-19 outbreak and suggests machine learning as a useful tool to predict COVID-19. The paper further suggests that the Ensemble learning method can realize better prediction. According to the statistical results, the ensemble learning model produced successful results.

The study can offer radiologists and clinical doctors a

second perspective. COVID-19 diagnosis performed using machine learning-based algorithms can help radiologists and physicians report, interpret, and gain time. The experimental results of the proposed study with the Covid-19 detection studies in the literature are compared and given in Table 8.

According to Table 6, our proposed Ensemble Learning method achieved 96.71% accuracy with 757 images. The LBP-KNN method achieved an accuracy rate of 98.66% with 5000 images. In the study carried out with the MobileNet model, one of the deep neural networks, an accuracy rate of 94.74% was achieved with 757 medical images.

**Table 8.** Experimental comparison results of the proposed method and the studies in the literature.

| Reference | Method | Number of images | Accuracy (%) |
|---|---|---|---|
| Chen [37] | HE-GLCM-SVM | 296 | 75.69 |
| Hasoon et al. [38] | LBP-KNN | 5000 | 98.66 |
| Barstugan et al. [39] | SVM | 150 | 98.71 |
| Yang et al. [40] | GLCM-SVM | 180 | 85.95 |
| Turk et al. [56] | MobileNet | 757 | 94.74 |
| Proposed | Ensemble Learning | 757 | 96.71 |

## 5. Conclusion

In this study, machine learning methods are recommended for diagnosing COVID-19 patients. In detecting COVID-19, the Ensemble Learning method was used with Random Forest, Logistic Regression, K-Nearest Neighbors, and Naive Bayes algorithms. In the study, machine learning algorithms are optimized with Randomized Search CV method. In preparing the input dataset of machine learning models, the DenseNet201 deep learning model was pre-trained with ImageNet. In addition, the PCA method was also used in this process.

In the study's first phase, DenseNet201 architecture was used to obtain the COVID-CT dataset features. In the second step, the PCA method was applied to reduce the data size of the obtained feature vector. In the last stage, the Ensemble Learning method was applied for collective learning with optimized machine learning (RF, DT, KNN and Gaussian NB) algorithms used in the diagnosis process of COVID-19 patients. Experimental results are presented comparatively before and after PCA. Accordingly, in the pre-PCA diagnostic study, Random Forest, K-Nearest Neighbors, and Ensemble Learning methods produced the most successful results of 96.71%, while the Naive Bayes algorithm was the most unsuccessful model with an accuracy rate of 94.74%. In the post-PCA diagnostic study, the Logistic Regression algorithm produced the most successful result with an accuracy rate of 96.05%. In contrast, the Random Forest algorithm was the most unsuccessful model, with an accuracy rate of 92.76%. Our study produced very satisfactory results in this state.

However, different models are needed to reach a definite conclusion, including large datasets and deep learning algorithms. Machine learning techniques can be used effectively in disease detection as a secondary view.

## Author contributions

**Hatice Çatal Reis:** Conceptualization, Methodology, Writing, and Editing-Reviewing. **Veysel Türk:** Software, Visualization, Investigation, Writing and Editing. **Serhat Kaya:** Data obtained, Pre-processing.

## Conflicts of interest

The authors declare no conflicts of interest.

## References

1. Shiri, I., Salimi, Y., Pakbin, M., Hajianfar, G., Avval, A. H., Sanaat, A., ... & Zaidi, H. (2022). COVID-19 prognostic modeling using CT radiomic features and machine learning algorithms: Analysis of a multi-institutional dataset of 14,339 patients. Computers in Biology and Medicine, 145, 105467. https://doi.org/10.1016/j.compbiomed.2022.105467

2. Del Rio, C., Omer, S. B., & Malani, P. N. (2022). Winter of Omicron—the evolving COVID-19 pandemic. Jama, 327(4), 319-320. https://doi.org/10.1001/jama.2021.24315.

3. Başeğmez, M., & Aydin, C. C. (2022). Türkiye'de COVID-19 sürecinde alınan önlemler çerçevesinde okul bahçe ve sınıflarının CBS ile değerlendirilmesi. Geomatik, 7(3), 209-219. https://doi.org/10.29128/geomatik.971403

4. Crawford, J., & Cifuentes-Faura, J. (2022). Sustainability in higher education during the COVID-19 pandemic: A systematic review. Sustainability, 14(3), 1879. https://doi.org/10.3390/su14031879

5. Maatuk, A. M., Elberkawi, E. K., Aljawarneh, S., Rashaideh, H., & Alharbi, H. (2022). The COVID-19 pandemic and E-learning: challenges and opportunities from the perspective of students and instructors. Journal of Computing in Higher Education, 34(1), 21-38. https://doi.org/10.1007/s12528-021-09274-2

6. Yig, K. G. (2023). Experiences of mathematics education teacher candidates in the emergency remote education: Reflections on the new normal. Mehmet Akif Ersoy University Journal of Education Faculty, (65), 549-577. https://doi.org/10.21764/maeuefd.1162499.

7. Panneer, S., Kantamaneni, K., Akkayasamy, V. S., Susairaj, A. X., Panda, P. K., Acharya, S. S., ... & Pushparaj, R. R. B. (2022). The great lockdown in the wake of COVID-19 and its implications: lessons for low and middle-income countries. International Journal of Environmental Research and Public Health, 19(1), 610. https://doi.org/10.3390/ijerph19010610

8. De Miquel, C., Domènech-Abella, J., Felez-Nobrega, M., Cristóbal-Narváez, P., Mortier, P., Vilagut, G., ... & Haro, J. M. (2022). The mental health of employees with job loss and income loss during the COVID-19 pandemic: the mediating role of perceived financial stress. International Journal of Environmental Research and Public Health, 19(6), 3158. https://doi.org/10.3390/ijerph19063158

9. Aldhafiri, F. K. (2022). COVID-19 and gut dysbiosis, understanding the role of probiotic supplements in reversing gut dysbiosis and immunity. Nutrition Clinique et Métabolisme, 36(3), 153-161. https://doi.org/10.1016/j.nupar.2022.01.003

10. Alves, M. H. M. E., Mahnke, L. C., Macedo, T. C., dos Santos Silva, T. K., & Junior, L. B. C. (2022). The enzymes in COVID-19: A review. Biochimie, 197, 38-48. https://doi.org/10.1016/j.biochi.2022.01.015

11. Chowdhury, M. E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., ... & Islam, M. T. (2020). Can AI help in screening viral and COVID-19 pneumonia?. IEEE Access, 8, 132665-132676. https://doi.org/10.1109/ACCESS.2020.3010287.

12. Siripanthong, B., Asatryan, B., Hanff, T. C., Chatha, S. R., Khanji, M. Y., Ricci, F., ... & Chahal, C. A. A. (2022). The pathogenesis and long-term consequences of COVID-19 cardiac injury. Basic to Translational Science, 7(3_Part_1), 294-308. https://doi.org/10.1016/j.jacbts.2021.10.011.

13. Raghav, A., Khan, Z. A., Upadhayay, V. K., Tripathi, P., Gautam, K. A., Mishra, B. K., ... & Jeong, G. B. (2021). Mesenchymal stem cell-derived exosomes exhibit

promising potential for treating SARS-CoV-2-infected patients. Cells, 10(3), 587. https://doi.org/10.3390/cells10030587

14. Cui, X., Chen, W., Zhou, H., Gong, Y., Zhu, B., Lv, X., ... & Ma, H. (2021). Pulmonary edema in COVID-19 patients: mechanisms and treatment potential. Frontiers in Pharmacology, 12, 664349. https://doi.org/10.3389/fphar.2021.664349

15. Puntmann, V. O., Carerj, M. L., Wieters, I., Fahim, M., Arendt, C., Hoffmann, J., ... & Nagel, E. (2020). Outcomes of cardiovascular magnetic resonance imaging in patients recently recovered from coronavirus disease 2019 (COVID-19). JAMA cardiology, 5(11), 1265-1273. https://doi.org/10.1001/jamacardio.2020.3557.

16. Douaud, G., Lee, S., Alfaro-Almagro, F., Arthofer, C., Wang, C., McCarthy, P., ... & Smith, S. M. (2022). SARS-CoV-2 is associated with changes in brain structure in UK Biobank. Nature, 604(7907), 697-707. https://doi.org/10.1038/s41586-022-04569-5

17. Sagris, M., Theofilis, P., Antonopoulos, A. S., Oikonomou, E., Tsioufis, K., & Tousoulis, D. (2022). Genetic predisposition and inflammatory inhibitors in COVID-19: where do we Stand?. Biomedicines, 10(2), 242. https://doi.org/10.3390/biomedicines10020242

18. Çatal Reis, H. (2018). Bone anomaly of the foot detection using medical photogrammetry. International Journal of Engineering and Geosciences, 3(1), 1-5. https://doi.org/10.26833/ijeg.333686.

19. Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., ... & Xia, L. (2020). Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology, 296(2), E32-E40. https://doi.org/10.1148/radiol.2020200642

20. La Salvia, M., Secco, G., Torti, E., Florimbi, G., Guido, L., Lago, P., ... & Leporati, F. (2021). Deep learning and lung ultrasound for Covid-19 pneumonia detection and severity classification. Computers in Biology and Medicine, 136, 104742. https://doi.org/10.1016/j.compbiomed.2021.104742

21. Fang, Y., Zhang, H., Xie, J., Lin, M., Ying, L., Pang, P., & Ji, W. (2020). Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology, 296(2), E115-E117. https://doi.org/10.1148/radiol.2020200432

22. Gupta, A., Gupta, S., & Katarya, R. (2021). InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using Chest X-ray. Applied Soft Computing, 99, 106859. https://doi.org/10.1016/j.asoc.2020.106859

23. Zu, Z. Y., Jiang, M. D., Xu, P. P., Chen, W., Ni, Q. Q., Lu, G. M., & Zhang, L. J. (2020). Coronavirus disease 2019 (COVID-19): a perspective from China. Radiology, 296(2), E15-E25. https://doi.org/10.1148/radiol.2020200490

24. Ojha, V., Mani, A., Pandey, N. N., Sharma, S., & Kumar, S. (2020). CT in coronavirus disease 2019 (COVID-19): a systematic review of chest CT findings in 4410 adult patients. European radiology, 30, 6129-6138. https://doi.org/10.1007/s00330-020-06975-7

25. Tabatabaei, S. M. H., Talari, H., Moghaddas, F., & Rajebi, H. (2020). CT features and short-term prognosis of COVID-19 pneumonia: a single-center study from Kashan, Iran. Radiology: Cardiothoracic Imaging, 2(2), e200130. https://doi.org/10.1148/ryct.2020200130

26. Wang, Y., Dong, C., Hu, Y., Li, C., Ren, Q., Zhang, X., ... & Zhou, M. (2020). Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study. Radiology, 296(2), E55-E64. https://doi.org/10.1148/radiol.2020200843

27. Hani, C., Trieu, N. H., Saab, I., Dangeard, S., Bennani, S., Chassagnon, G., & Revel, M. P. (2020). COVID-19 pneumonia: a review of typical CT findings and differential diagnosis. Diagnostic and Interventional Imaging, 101(5), 263-268. https://doi.org/10.1016/j.diii.2020.03.014

28. Wang, K., Kang, S., Tian, R., Zhang, X., & Wang, Y. (2020). Imaging manifestations and diagnostic value of chest CT of coronavirus disease 2019 (COVID-19) in the Xiaogan area. Clinical Radiology, 75(5), 341-347. https://doi.org/10.1016/j.crad.2020.03.004

29. Liu, G., Wang, G., Yang, Z., Liu, G., Ma, H., Lv, Y., ... & Zhu, W. (2022). A lung ultrasound-based nomogram for the prediction of refractory Mycoplasma pneumoniae pneumonia in hospitalized children. Infection and Drug Resistance, 15, 6343-6355. https://doi.org/10.2147/IDR.S387890

30. Kaur, N., & Mittal, A. (2022). CADxReport: Chest x-ray report generation using co-attention mechanism and reinforcement learning. Computers in Biology and Medicine, 145, 105498. https://doi.org/10.1016/j.compbiomed.2022.105498

31. Gakhar, M., & Aggarwal, A. (2022). ThoraciNet: thoracic abnormality detection and disease classification using fusion DCNNs. Physical and Engineering Sciences in Medicine, 45, 961-970. https://doi.org/10.1007/s13246-022-01137-z

32. Packhäuser, K., Gündel, S., Münster, N., Syben, C., Christlein, V., & Maier, A. (2022). Deep learning-based patient re-identification is able to exploit the biometric nature of medical chest X-ray data. Scientific Reports, 12(1), 14851. https://doi.org/10.1038/s41598-022-19045-3

33. Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., ... & Schönlieb, C. B. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nature Machine Intelligence, 3(3), 199-217. https://doi.org/10.1038/s42256-021-00307-0

34. Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R., Reuter, U., ... & Atkinson, P. M. (2020). Covid-19 outbreak prediction with machine learning. Algorithms, 13, 249. https://doi.org/10.3390/a13100249.

35. Canayaz, M., Şehribanoğlu, S., Özdağ, R., & Demir, M. (2022). COVID-19 diagnosis on CT images with Bayes optimization-based deep neural networks and machine learning algorithms. Neural Computing and Applications, 34(7), 5349-5365. https://doi.org/10.1007/s00521-022-07052-4

36. Muurlink, O. T., Stephenson, P., Islam, M. Z., & Taylor-Robinson, A. W. (2018). Long-term predictors of dengue outbreaks in Bangladesh: A data mining approach. Infectious Disease Modelling, 3, 322-330. https://doi.org/10.1016/j.idm.2018.11.004

37. Chen, Y. (2021). Covid-19 classification based on gray-level co-occurrence matrix and support vector machine. COVID-19: Prediction, Decision-making, and its Impacts, 47-55. https://doi.org/10.1007/978-981-15-9682-7_6

38. Hasoon, J. N., Fadel, A. H., Hameed, R. S., Mostafa, S. A., Khalaf, B. A., Mohammed, M. A., & Nedoma, J. (2021). COVID-19 anomaly detection and classification method based on supervised machine learning of chest X-ray images. Results in Physics, 31, 105045. https://doi.org/10.1016/j.rinp.2021.105045

39. Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. arXiv preprint https://doi.org/10.48550/arXiv.2003.09424.

40. Yang, N., Liu, F., Li, C., Xiao, W., Xie, S., Yuan, S., ... & Jiang, G. (2021). Diagnostic classification of coronavirus disease 2019 (COVID-19) and other pneumonias using radiomics features in CT chest images. Scientific Reports, 11, 17885. https://doi.org/10.1038/s41598-021-97497-9

41. Currie, G., Hawk, K. E., Rohren, E., Vial, A., & Klein, R. (2019). Machine learning and deep learning in medical imaging: intelligent imaging. Journal of Medical Imaging and Radiation Sciences, 50(4), 477-487. https://doi.org/10.1016/j.jmir.2019.09.005

42. Tang, A., Tam, R., Cadrin-Chênevert, A., Guest, W., Chong, J., Barfett, J., ... & Canadian Association of Radiologists (CAR) Artificial Intelligence Working Group. (2018). Canadian Association of Radiologists white paper on artificial intelligence in radiology. Canadian Association of Radiologists Journal, 69(2), 120-135. https://doi.org/10.1016/j.carj.2018.02.0

43. Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. Radiographics, 37(2), 505-515. https://doi.org/10.1148/rg.2017160130

44. 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 https://doi.org/10.48550/arXiv.2003.13865.

45. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, 4700-4708. https://doi.org/10.1109/CVPR.2017.243.

46. Avci, C., Budak, M., Yagmur, N. & Balcik, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. International Journal of Engineering and Geosciences, 8(1), 1-10. https://doi.org/10.26833/ijeg.987605.

47. Erdem, F., Derinpinar, M.A., Nasirzadehdizaji, R., Oy, S., Seker, D. Z. & Bayram, B. (2018). Rastgele orman yöntemi kullanılarak kıyı çizgisi çıkarımı İstanbul Örneği. Geomatik, 3 (2), 100-107. https://doi.org/10.29128/geomatik.362179.

48. Akar, O. & Tunc Gormus, E. (2019). Göktürk-2 ve Hyperion EO-1 uydu görüntülerinden rastgele orman sınıflandırıcısı ve destek vektör makineleri ile arazi kullanım haritalarının üretilmesi. Geomatik, 4 (1), 68-81. https://doi.org/10.29128/geomatik.476668.

49. Duman, H. S., & Başaraner, M. (2022). Şekil göstergeleri ve topluluk öğrenmesi sınıflandırma algoritmaları ile bina detaylarının şekil karmaşıklık analizi. Geomatik, 7(3), 197-208. https://doi.org/10.29128/geomatik.947334

50. Gong, M., Bai, Y., Qin, J., Wang, J., Yang, P., & Wang, S. (2020). Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin. Journal of Building Engineering, 27, 100950. https://doi.org/10.1016/j.jobe.2019.100950

51. Nguyen, P. T., Ha, D. H., Avand, M., Jaafari, A., Nguyen, H. D., Al-Ansari, N., ... & Pham, B. T. (2020). Soft computing ensemble models based on logistic regression for groundwater potential mapping. Applied Sciences, 10(7), 2469. https://doi.org/10.3390/app10072469

52. Siddique, M. A. B., Sakib, S., & Rahman, M. A. (2019, December). Performance analysis of deep autoencoder and NCA dimensionality reduction techniques with KNN, ENN and SVM classifiers. 2nd International Conference on Innovation in Engineering and Technology (ICIET), 1-6. https://doi.org/10.1109/ICIET48527.2019.9290722

53. Apaydin, C., & Abdikan, S. (2021). Fındık bahçelerinin Sentinel-2 verileri kullanılarak piksel tabanlı sınıflandırma yöntemleriyle belirlenmesi. Geomatik, 6(2),107-114 https://doi.org/10.29128/geomatik.705988

54. Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J. J., ... & Ahmad, A. (2020). Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: Hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. Remote Sensing, 12(2), 266. https://doi.org/10.3390/rs12020266

55. Hou, S., Liu, Y., & Yang, Q. (2022). Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning. Journal of Rock Mechanics and Geotechnical Engineering, 14(1), 123-143. https://doi.org/10.1016/j.jrmge.2021.05.004

56. Turk, V., Catal Reis, H. & Kaya, S. (2022). Automatic prediction of covid-19 from chest-computed tomography (CT) images using deep learning architectures. Gumushane University Journal of Science. https://doi.org/10.17714/gumusfenbil.1002738.