



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.



## COVIDC: An expert system to diagnose COVID-19 and predict its severity using chest CT scans: Application in radiology

Wajid Arshad Abbasi<sup>a,\*</sup>, Syed Ali Abbas<sup>a</sup>, Saiqa Andleeb<sup>b</sup>, Ghafoor ul Islam<sup>b</sup>, Syeda Adin Ajaz<sup>a</sup>, Kinza Arshad<sup>a</sup>, Sadia Khalil<sup>a</sup>, Asma Anjam<sup>a</sup>, Kashif Ilyas<sup>a</sup>, Mohsib Saleem<sup>a</sup>, Jawad Chughtai<sup>a</sup>, Ayesha Abbas<sup>a</sup>

<sup>a</sup> Computational Biology and Data Analysis Lab., Department of Computer Science & Information Technology, King Abdullah Campus, University of Azad Jammu & Kashmir, Muzaffarabad, AJ&K, 13100, Pakistan

<sup>b</sup> Biotechnology Lab., Department of Zoology, King Abdullah Campus, University of Azad Jammu & Kashmir, Muzaffarabad, AJ&K, 13100, Pakistan

### ARTICLE INFO

#### Keywords:

COVID-19  
CT imaging  
Expert system  
Early diagnosis  
SARS-COV-2  
Pandemic  
Radiology

### ABSTRACT

Early diagnosis of Coronavirus disease 2019 (COVID-19) is significantly important, especially in the absence or inadequate provision of a specific vaccine, to stop the surge of this lethal infection by advising quarantine. This diagnosis is challenging as most of the patients having COVID-19 infection stay asymptomatic while others showing symptoms are hard to distinguish from patients having different respiratory infections such as severe flu and Pneumonia. Due to cost and time-consuming wet-lab diagnostic tests for COVID-19, there is an utmost requirement for some alternate, non-invasive, rapid, and discounted automatic screening system. A chest CT scan can effectively be used as an alternative modality to detect and diagnose the COVID-19 infection. In this study, we present an automatic COVID-19 diagnostic and severity prediction system called COVIDC (COVID-19 detection using CT scans) that uses deep feature maps from the chest CT scans for this purpose. Our newly proposed system not only detects COVID-19 but also predicts its severity by using a two-phase classification approach (COVID vs non-COVID, and COVID-19 severity) with deep feature maps and different shallow supervised classification algorithms such as SVMs and random forest to handle data scarcity. We performed a stringent COVIDC performance evaluation not only through 10-fold cross-validation and an external validation dataset but also in a real setting under the supervision of an experienced radiologist. In all the evaluation settings, COVIDC outperformed all the existing state-of-the-art methods designed to detect COVID-19 with an F1 score of 0.94 on the validation dataset and justified its use to diagnose COVID-19 effectively in the real setting by classifying correctly 9 out of 10 COVID-19 CT scans. We made COVIDC openly accessible through a cloud-based webserver and python code available at <https://sites.google.com/view/wajidarshad/software> and <https://github.com/wajidarshad/covidc>.

### 1. Introduction

Coronavirus disease 2019 (COVID-19) is a contagious infection caused by a family of viruses called coronaviridae. Since its identification at the end of 2019 in Wuhan, a city in the Hubei Province of China, this viral disease has spread rapidly around the globe [1]. The World Health Organization (WHO) has already declared this pandemic as a global health calamity and it has infected and killed millions of people worldwide [2]. Physicians say this is a viral infection that mainly infects the human respiratory system and antibiotics do not cure this disease [3, 4]. Meanwhile, currently, there is an inadequate provision of an

effective vaccine or medication to prevent or cure this lethal infection.

In the absence or insufficient supply of effective medication or vaccines to cure or prevent COVID-19 infection, early diagnosis of the disease is fundamental to avoid its further surge by advising or adopting quarantine or isolation. Most of the patients having COVID-19 infection stay asymptomatic while others show mild to moderate symptoms such as fever or cough, shortness of breath, and signs such as oxygen saturation or lung auscultation [5,6]. Signs and symptoms have the utmost importance to identify any infection and to perform further diagnostic tests. However, even symptomatic COVID-19 infection does not show any particular symptoms which distinguish this from other respiratory

\* Corresponding author.

E-mail addresses: [wajidarshad@gmail.com](mailto:wajidarshad@gmail.com), [wajidarshad@ajku.edu.pk](mailto:wajidarshad@ajku.edu.pk) (W.A. Abbasi).

<https://doi.org/10.1016/j.imu.2021.100540>

Received 26 December 2020; Received in revised form 17 February 2021; Accepted 19 February 2021

Available online 23 February 2021

2352-9148/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

infections such as severe flu and Pneumonia [6]. In this situation, all the suspected patients require screening and have to pass through the adopted diagnostic tests.

Currently, there are two types of tests to detect COVID-19: diagnostic tests and antibody tests [7]. A diagnostic test such as Reverse transcription-polymerase chain reaction (RT-PCR) can be used to detect an active coronavirus infection whereas, the antibody test looks for antibodies created in your body in the response to the disease. RT-PCR is currently the gold standard diagnostic practice to detect viral infection [8]. However, the standard confirmatory clinical RT-PCR test to detect COVID-19 is manual, complex, laborious, costly, and ineffective [9,10]. Moreover, the limited availability of test kits and domain experts further hamper the situation. Meanwhile, antibody tests cannot be used to diagnose COVID-19 as after infection, antibodies can take numerous days or weeks to develop and may stay in your blood a bit longer after recovery [7]. Keeping in view the limitation of prevailing diagnostic techniques and a rapid surge of infected patients, there is an utmost requirement for some alternate automatic screening systems that can be used by the physicians to swiftly identify and isolate the COVID-19 infected patients.

X-ray and Computed Tomography (CT) imaging modalities are the most common noninvasive diagnostic techniques that help medical practitioners to diagnose and treat many diseases. In the current framework, chest X-ray (CXR) and chest Computed Tomography (CCT) can also effectively be used to diagnose COVID-19 [11,12]. However, CCT due to 3D image and contrast dyes shows a significantly improved performance in COVID-19 diagnosis in comparison to simple 2D CXR [9]. In the meanwhile, CT scans of COVID-19 infected patients show diverse features and manual interpretation of these scans with subtle variations is quite challenging [13]. Moreover, the current enormous upsurge of infected patients makes it a challenging task for the domain experts to complete a timely diagnosis [14,15]. Therefore, some Computer-Aided Diagnostic (CAD) systems are required to better manipulate and understand the CCT images.

In Computer-Aided Diagnostic (CAD) system design using CT scans, machine learning in the form of deep learning has successfully been applied in diagnosing lung diseases [16–19]. Also to diagnose COVID-19 using CCT images, several machine learning-based methods with varying sources and amount of training data have been proposed in the literature (for details please see “Related Works” in the next section). Most of these previously proposed methods in the literature are based on Convolution Neural Network (CNN) deep learning approach [20]. However, for the deep learning approaches to perform effectively, a healthy amount of data is usually required for parameter tuning which is not readily available so far. Moreover, all the published studies focused only to diagnose COVID-19 whereas, its severity prediction still requires an efficient method. Therefore, to address the issues of data scarcity in deep learning and the absence of an efficient method for COVID-19 severity prediction, in this article we have proposed a method called COVIDC (COVID-19 detection using CT scans) to diagnose and predict the severity of the COVID-19 infection with CCT images. Our proposed method is different from previously proposed methods for the diagnosis of the COVID-19 and its severity prediction in that it uses a combination of transfer and shallow learning. Using this method, we can extract efficient feature representation of CT scans by using transfer learning with off-the-shelf pre-trained models on ImageNET and still be able to avoid the curse of dimensionality by engaging shallow learning algorithms such as Support Vector Machines (SVMs) [21]. This has led to improved accuracy in comparison to methods that rely only on the deep learning paradigm.

## 2. Related Works

Since COVID-19 identification at the end of 2019, several studies have been proposed in the literature to diagnose COVID-19 with chest computed tomography (CT) by using artificial intelligence (AI)

techniques. We have searched existing techniques in the online literature using different keywords such as “Machine learning and COVID-19”, “COVID-19 diagnosis with CT scans”, and “COVID-19 diagnosis with CT scans and machine learning”. While going through the online existing literature (peer-reviewed) published in reputed journals, we found a plethora of machine learning techniques to diagnose COVID-19 using chest CT scans with varying sources and amount of training data [22–37]. All these previously published techniques can be categorized into three main classes as follows: Deep learning-based, transfer learning with fine-tuning a customized fully connected layer, shallow learning with handcrafted textured features.

Previously published studies employing deep learning-based machine learning techniques to detect COVID-19 mostly used Convolution Neural Network (CNN) [20] based architectures in their proposed design [23–25,29,31–37]. However, deep learning approaches to generalize well normally require an enormous amount of data which is not readily available right now [22].

Due to data scarcity issues in CNN-based deep learning, some studies have also been proposed in the literature using transfer learning [22,28,30]. Ahuja et al. used data augmentation and transfer learning with ResNet18, ResNet50, ResNet101, and SqueezeNet pretrained networks to classify COVID-19 images. Ahuja et al. concluded that ResNet18 outperformed other models by obtaining a 0.97 AUC score [22]. Ko et al. proposed a fast-track COVID-19 classification model network based on transfer learning with VGG16, ResNet-50, InceptionV3, and Xception as backbone pre-trained models to diagnose COVID-19. They considered 1194 COVID-19, 264 low-quality COVID-19 (only for testing), and 2239 pneumonia, normal, and other disease CT scans in their study. They also used rotation and zoom data augmentation procedures to maximize the number of training samples. It was concluded that FCNet based on ResNet-50 outperformed other pre-trained models and achieved 96.97% of accuracy in the external validation data set of COVID-19 pneumonia images [28,38]. Pathak et al. proposed a transfer learning-based system to detect COVID-19 in CT scans with the ResNet50 to extract the features and a 2D convolutional neural network for the classification [30]. The proposed system achieved a 93.01% of accuracy with 10-fold cross-validation on 413 COVID and 439 non-COVID scans. These previously published methods to detect COVID-19 using transfer learning still have room for improvement because in transfer learning, fine-tuning the customized fully connected layer requires a healthy amount of data due to high dimensional features space generated by the off-the-shelf pre-trained models.

To solve data scarcity issues in deep learning, Kang et al. proposed a shallow learning technique for COVID-19 classification using different types of handcrafted features extracted from CT images [27]. They used 2522 CT images (1495 are from COVID-19 patients, and 1027 are from community-acquired pneumonia) for the classification purpose. The proposed method achieved an overall accuracy of 95.5%. However, even after learning multiple representations of CT scans, the method proposed by Kang et al. still has room for performance improvement because it is often hard to represent subtle variations in CT scans using handcrafted texture descriptors [39]. Moreover, almost all the above studies have been proposed only to diagnose COVID-19 but its efficient severity detection is still an open research question. Furthermore, to the best of our knowledge, all of the above-discussed studies only discussed the technical details of their proposed methodology without providing any easily accessible interface to end-users. To overcome these shortcomings and to deal with the issues of data scarcity in deep/transfer learning and handcrafted textured features, in this study we have proposed a novel COVID-19 diagnosis and its severity predictor called COVIDC (COVID-19 detection using CT scans) by employing a combination of transfer and shallow learning.

## 3. Methods

In the following sections, we give the detail of our experimental

procedure to design and test the generalization performance of COVIDC (COVID-19 diagnosis using CT images).

### 3.1. Dataset and preprocessing

In this study, we have collected three different datasets of Chest Computed Tomography (CCT) images from publicly open online repositories [40–42]. For COVID-19 diagnosis, we have used two datasets: i) training set (1433 COVID and 1229 non-COVID CT images) [41], ii) external validation set (349 COVID and 397 non-COVID) [42]. For COVID-19 severity prediction, we have used a dataset of 141 CT images of severe COVID-19 patients and 135 CT images of less-severe [40]. All the images in these datasets have gone through the essential preprocessing required for feature extraction including image resizing (313 × 313 pixels), de-noising, and contrast stretching [43].

### 3.2. Proposed methodology

The methodology we have adopted for this study has been shown in Fig. 1. We give the detail of our proposed methodology to design and develop CT image-based COVID-19 diagnosis and its severity predictor using machine learning as follows.

### 3.3. Features extraction and preprocessing

Deep learning, where we learn automatically an efficient feature representation of an image, is very famous and has successfully been used in different image classification and analysis tasks. However, to apply deep learning efficiently, we normally require an extensive amount of data to better learn the actual distribution of the image samples. This data hunger is a challenge so far to apply the deep learning approach efficiently in our case to diagnose and predict the severity of COVID-19 where we face data scarcity. Therefore, to overcome this challenge, we have used the transfer learning strategy for feature extraction. Using transfer learning, feature extraction from the available CT images in our datasets has been performed using different off-the-shelf CNN based pre-trained models on ImageNet. For this purpose, we have used different off-the-shelf CNN based pre-trained models such as Resnet50 [44], InceptionV3 [45], Xception [46], VGG16 [47], NAS-NetLarge [48], DenseNet121 [49]. We have used these models for feature extraction by simply loading a pre-trained model without its classifier part by specifying the “include\_top” argument as “False”. The selection of these pre-trained CNN based models was based on their reported accuracy. Preprocessing and resizing required by these pre-trained models have been applied before extracting the features map.

### 3.4. Classification algorithms

Most of the machine learning studies involved in COVID-19 diagnosis have used pure deep learning strategies or transfer learning with a customized fully connected layer to be fine-tuned. However, to train deep neural networks from scratch or to fine-tune the parameters of the customized fully connected layers on the top of a pre-trained model, we need an extensive amount of data and high-performance computational resources, which is mostly impractical [50]. Therefore, we have proposed a different machine learning-based approach for the preliminary diagnosis of COVID-19 and its severity prediction using digital chest CT images. The novelty of the proposed approach is that it uses a blend of pre-trained CNN based models to extract features and shallow learning algorithms such as SVMs for the classification purpose. The proposed scheme is somehow based on the paradigm of transfer learning.

In this work, our dataset consists of examples of the form  $(I_i, y_i)$  where  $I_i$  is a chest CT image and  $y_i \in \{+1, -1\}$  is its associated label. For COVID-19 identification,  $y_i$  indicate whether  $I_i$  belongs to a COVID-19 patient (+1) or not (−1) and for COVID-19 severity prediction  $y_i$  indicate severe (+1) or mild (−1). For a given chest CT image  $I_i$ , we extract deep feature maps denoted by  $x_i$ . Our objective is to learn two separate functions for the diagnosis and severity prediction of COVID-19. For this purpose, we have used three different machine learning-based classification algorithms: classical Support Vector Machine (SVM), Random Forest (RF), and Gradient Boosting Machine (XGBoost) [51–53]. The detail of these shallow machine learning algorithms is as follows.

#### 3.4.1. Support Vector classification (SVC)

We used SVM for the diagnosis of COVID-19 and its severity prediction by learning a function  $f(x) = \langle w, x \rangle$  with  $w$  as parameters to be learned from the training data  $\{(x_i, y_i) | i = 1, 2, \dots, N\}$ . The optimal value of the  $w$  is obtained in SVM by solving the following optimization problem [52].

$$\min_{w, \xi} \frac{1}{2} \lambda w^2 + \sum_{i=1}^N \xi_i \quad (1)$$

Such that for all  $i : y_i w \cdot x_i \geq 1 - \xi_i, \xi_i \geq 0$

The objective function in Eq. (1) maximizes the margin while minimizing margin violations (or slacks  $\xi$ ) [52]. The hyperparameter  $\lambda = \frac{1}{c}$  controls the tradeoff between margin maximization and margin violation. We used both linear and radial basis function (RBF) kernels and coarsely optimized the values of  $\lambda$  and  $\gamma$  using grid search with scikit-learn (version:0.23) [54].

#### 3.4.2. Random Forest Classification (RFC)

Random Forest Classification (RFC) uses an ensemble learning technique based on bagging. A random forest operates by constructing several decision trees in parallel during training and outputs the mean of the classes as the prediction of all trees [51]. It usually performs better

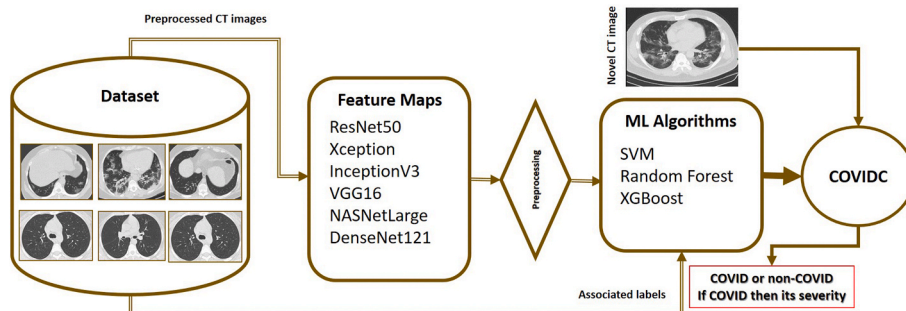


Fig. 1. A proposed methodology for the development of computer-aided COVID-19 diagnosis and its severity prediction system (COVIDC) using Machine learning and chest CT images. This system has been trained using shallow learning algorithms such as SVMs with chest CT scans by extracting feature maps involving pre-trained off-the-shelf models. COVIDC can be used to predict whether a novel test CT image has COVID-19 or not.

on problems having features with non-linear relationships. Each classification tree in the RF is constructed on randomly sampled subsets of input features. In this study, we have optimized RF for the number of decision trees in the forest, the maximum number of features considered for splitting a node, the maximum number of levels in each decision tree, and a minimum number of samples required to split. We have also seen this classification technique effectively in action in many other studies [55–58].

### 3.4.3. XGBoost classification (XGBC)

XGBoost is also an ensemble learning technique based on the boosting that combines weak learners into a strong learner in an iterative fashion [53,59]. We have used trees as default base learners in XGBoost ensembles. In this study, we have optimized XGBoost in terms of the learning rate, maximum depth, the number of boosting iterations, booster, and subsample ratio with a grid search and using a python-based package xgboost 0.7 [59].

### 3.5. Online access to COVIDC

We have developed and deployed a webserver of COVIDC which uses the optimal machine learning model for COVID-19 diagnosis and its severity prediction. This webserver takes a chest CT image and performs COVID-19 diagnosis and its severity prediction for it. After the successful submission of a CT image, users get COVIDC predictions in the form of COVID-19 identification and its severity prediction. This process has broken into two steps: 1) whether the uploaded image represents a COVID-19 patient or not (COVID vs non-COVID), 2) if it is a COVID-19 infection then how severe it is. The webserver is available at <https://sites.google.com/view/wajidarshad/software>.

## 4. Results

### 4.1. Experimental setup

We have used two different datasets for COVID-19 diagnosis: train-test set (1433 COVID and 1229 non-COVID CT images) [41], ii) external validation set (349 COVID and 397 non-COVID) [42], and reported performance metrics on both the datasets. For the train-test set, we have used 10-fold cross-validation (CV). In a 10-fold CV, we have shuffled images in our datasets and then split them into 10 groups. Ten models have been trained and evaluated with each group given a chance to be held out as the test set [60]. Average performance metrics across folds have been reported. For the external validation set, we trained the classification models using the whole train-test set and tested them on the validation set. We have used the area under the ROC curve (ROC), the area under the precision-recall curve (PR), and F-measure (F1: Harmonic mean of the model's precision and recall) as performance measures for model evaluation and performance assessment [60–62]. We used grid search over training data to find the optimal values of hyper-parameters of different classification models. The detailed discussion of results showing the predictive performance of our proposed

method over cross-validation and also on an external validation dataset is given as follows.

### 4.2. Classification results of COVID versus non-COVID CT images

We have trained various shallow machine learning models for the classification of COVID versus non-COVID CT scan images with a range of deep learning-based feature maps and evaluated using both 10-fold cross-validation (CV) and on an external validation dataset. The results of our evaluation in both settings are shown in Tables 1 and 2 and Fig. 2. Using 10-fold CV, we observed a maximum F1-score of 0.98 along with 0.99, and 0.99 as the area under the ROC curve, and the area under the PR curve, respectively with Support Vector Classifier and DenseNet121 feature map (Table 1). The F1-score of 0.98 and area under the PR curve of 0.99 show the perfect sensitivity and precision of our proposed method while classifying COVID and non-COVID CT scans efficiently. Higher values of sensitivity show that our trained model produces fewer false negatives (classifying COVID as non-COVID) which is a genuine requirement in this case.

Meanwhile, using an external validation dataset for the evaluation of our machine learning models trained on the training set, we observed a similar trend of performance with an F1-score of 0.94 along with 0.98, and 0.98 as the area under the ROC curve, and the area under the PR curve, respectively with Support Vector Classifier and DenseNet121 features map (Table 2, Fig. 2(A) and (B)). As our external validation dataset is completely different (collected in a different country with different settings) and images have huge variations with our training set. Therefore, achieving an F1-score of 0.94 shows a significant consistent generalization performance of our proposed method in diverse CT imaging environments. Moreover, the significantly better performance of SVM in comparison to Random Forest (RF) and Extreme Boosting Machine (XGB) is attributed to its ability to deal well with high dimensional features with fewer examples as in our case. Moreover, features extracted through pre-trained deep learning models perform better than handcrafted ones as in the case of the study performed by Chandra et al., [15].

### 4.3. Classification results of COVID-19 severity prediction with CT images

Along with COVID-19 diagnosis using CT images, we have also tried to predict the severity of COVID-19 infection using the same image modality. For this purpose, we have trained several shallow machine learning models for the classification of severe versus mild COVID-19 infection using chest CT scans of COVID-19 infected patients with a range of deep learning-based feature maps and evaluated using 10-fold cross-validation (CV). To the best of our knowledge, this is the first study on COVID-19 severity prediction posed as a classification problem using machine learning with CT image modality. The results of our evaluation are shown in Table 3. Using 10-fold CV, we observed a maximum F1-score of 1.0 along with 1.00, and 0.99 as the area under the ROC curve, and the area under the PR curve, respectively with Support Vector Classifier and DenseNet121 feature map (Table 3). The F score of 1.0

**Table 1**

Predictive performance for COVID-19 diagnosis across different classification models and feature maps using 10-fold CV (COVID vs non-COVID).

Feature Map	SVC			RFC			XGBC		
	ROC	PR	F1	ROC	PR	F1	ROC	PR	F1
<b>Resnet50</b>	0.98 ± 0.03	0.98 ± 0.03	0.95	0.96 ± 0.03	0.96 ± 0.03	0.90	0.97 ± 0.01	0.97 ± 0.01	0.92
<b>Xception</b>	0.98 ± 0.04	0.98 ± 0.04	0.96	0.96 ± 0.03	0.96 ± 0.03	0.89	0.97 ± 0.01	0.97 ± 0.01	0.92
<b>InceptionV3</b>	0.98 ± 0.03	0.98 ± 0.03	0.96	0.96 ± 0.03	0.96 ± 0.03	0.89	0.97 ± 0.01	0.97 ± 0.01	0.92
<b>VGG16</b>	0.98 ± 0.03	0.98 ± 0.03	0.95	0.96 ± 0.03	0.96 ± 0.03	0.90	0.96 ± 0.01	0.96 ± 0.01	0.88
<b>NASNetLarge</b>	0.97 ± 0.04	0.97 ± 0.04	0.94	0.95 ± 0.03	0.95 ± 0.03	0.88	0.97 ± 0.01	0.97 ± 0.01	0.90
<b>DenseNet121</b>	<b>0.99 ± 0.03</b>	<b>0.99 ± 0.03</b>	<b>0.98</b>	<b>0.96 ± 0.01</b>	<b>0.96 ± 0.01</b>	<b>0.90</b>	<b>0.98 ± 0.03</b>	<b>0.98 ± 0.03</b>	<b>0.94</b>

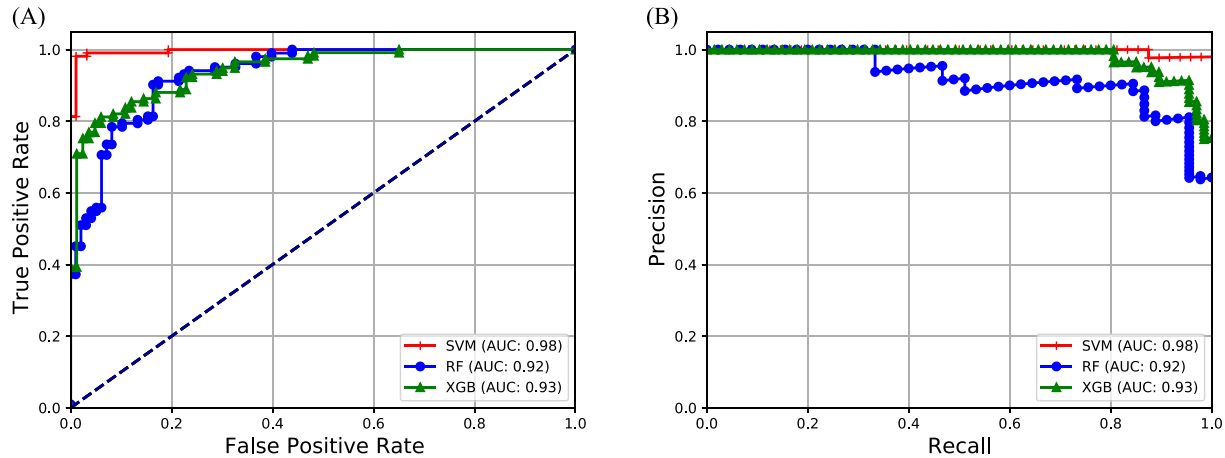
ROC (Area under the ROC curve), PR (Area under the precision-recall curve), F1 (F1 Score), SVC (Support Vector classifier), RF (Random Forest classifier), XGBC (XGBoost classifier). Bold-faced values indicate the best performance for each model.

**Table 2**

Predictive performance for COVID-19 diagnosis across different classification models and feature maps on external validation dataset (COVID vs non-COVID).

Feature Map	SVC			RFC			XGBC		
	ROC	PR	F1	ROC	PR	F1	ROC	PR	F1
Resnet50	0.96	0.96	0.90	0.92	0.92	0.83	0.90	0.91	0.82
Xception	0.96	0.96	0.90	0.90	0.90	0.82	0.90	0.91	0.82
InceptionV3	0.96	0.96	0.90	0.90	0.90	0.83	0.90	0.91	0.82
VGG16	0.96	0.96	0.89	0.90	0.90	0.82	0.90	0.91	0.82
NASNetLarge	0.94	0.94	0.88	0.88	0.88	0.80	0.87	0.88	0.80
DenseNet121	<b>0.98</b>	<b>0.98</b>	<b>0.94</b>	<b>0.92</b>	<b>0.92</b>	<b>0.83</b>	<b>0.93</b>	<b>0.93</b>	<b>0.84</b>

ROC (Area under the ROC curve), PR (Area under the precision-recall curve), F1 (F1 Score), SVC (Support Vector classifier), RF (Random Forest classifier), XGBC (XGBoost classifier). Bold-faced values indicate the best performance for each model.



**Fig. 2.** Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves showing predictive performance of COVIDC for the classification of chest CT images across different classifiers (SVM, RF, XGB) and DenseNet feature map on an external validation dataset. COVID vs non-COVID: ROC(A), PR(B).

**Table 3**

Predictive performance for COVID-19 severity prediction across different classification models and feature maps using 10-fold CV.

Feature Map	SVC			RFC			XGBC		
	ROC	PR	F1	ROC	PR	F1	ROC	PR	F1
Resnet50	0.99 ± 0.01	0.99 ± 0.01	0.98	0.97 ± 0.03	0.97 ± 0.03	0.96	0.98 ± 0.01	0.98 ± 0.01	0.95
Xception	0.99 ± 0.01	0.99 ± 0.01	0.98	0.96 ± 0.03	0.96 ± 0.03	0.96	0.98 ± 0.01	0.98 ± 0.01	0.96
InceptionV3	0.99 ± 0.01	0.99 ± 0.01	0.99	<b>0.99 ± 0.01</b>	<b>0.99 ± 0.01</b>	<b>0.98</b>	<b>0.98 ± 0.01</b>	<b>0.98 ± 0.01</b>	<b>0.98</b>
VGG16	0.99 ± 0.01	0.99 ± 0.01	0.98	0.95 ± 0.03	0.94 ± 0.03	0.94	0.98 ± 0.01	0.98 ± 0.01	0.97
NASNetLarge	0.98 ± 0.01	0.98 ± 0.01	0.97	0.96 ± 0.03	0.96 ± 0.03	0.92	0.97 ± 0.01	0.97 ± 0.01	0.95
DenseNet121	<b>1.00 ± 0.01</b>	<b>0.99 ± 0.01</b>	<b>1.00</b>	0.98 ± 0.01	0.98 ± 0.01	0.97	0.98 ± 0.05	0.98 ± 0.05	0.97

ROC (Area under the ROC curve), PR (Area under the precision-recall curve), F1 (F1 Score), SVC (Support Vector classifier), RF (Random Forest classifier), XGBC (XGBoost classifier). Bold-faced values indicate the best performance for each model.

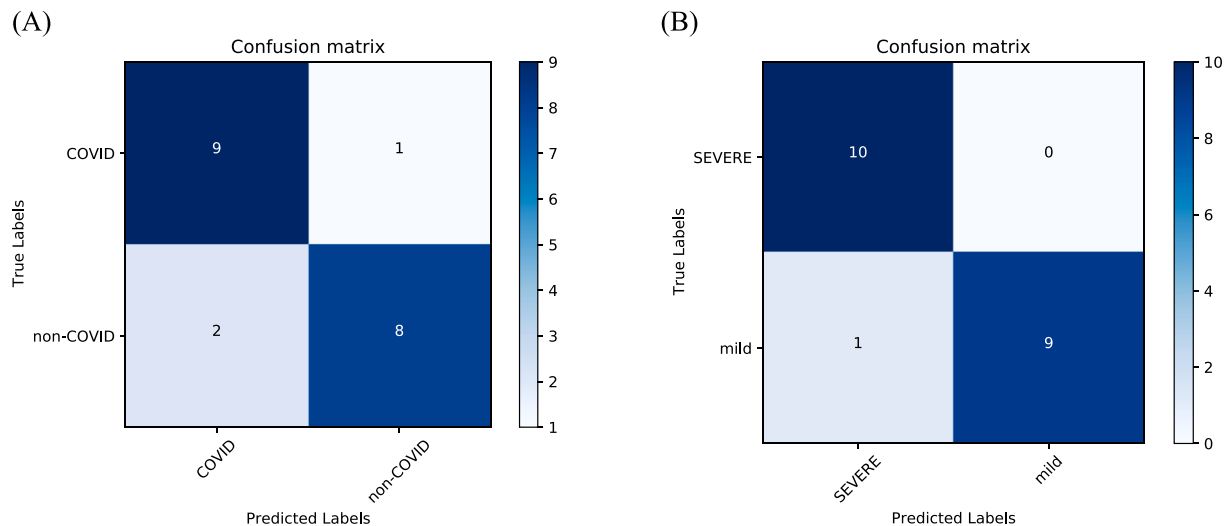
with an area under the PR of 0.99 shows that our proposed classification model produces almost zero false positives (classifying mild COVID-19 as severe) and false negatives (classifying severe COVID-19 as mild). These results show that we have been able to classify the severe and mild COVID-19 patients with perfect accuracy using our proposed method called COVIDC.

#### 4.4. Predictive performance of COVIDC in its real use

The real generalization performance of our optimal trained models (COVIDC) has been evaluated under the supervision of an experienced radiologist at Abbas Institute of Medical Sciences (AIMS) hospital located in Muzaffarabad, Azad Jammu & Kashmir, Pakistan. For this purpose, 20 CT scans for COVID-19 diagnosis (10 COVID and 10 non-COVID) and 20 CT scans for severity (10 severe and 10 milds) have been used. These scans were used as novel input CT scans to our proposed method called COVIDC in raw digital form as these were not included in the training set. Results obtained through this evaluation are

shown as confusion matrices in Fig. 3. For COVID vs non-COVID, our proposed system (COVIDC) has been able to classify correctly 9 out of 10 provided COVID CT images as COVID and 1 as non-COVID (Fig. 3(A)). Similarly, for the provided non-COVID CT images, our system classified correctly 8 out of 10 images as non-COVID, and 2 as COVID (Fig. 3(A)). This performance is reasonably good as we have less misclassification rate for COVID and this is ideally required to isolate the infected patients. These results help us to interpret that this system can reliably be used to advise isolation or quarantine while suspecting COVID-19 infection.

Over severe vs mild COVID-19 classification task, our proposed system (COVIDC) has been able to classify correctly 10 out of 10 provided COVID CT images as severe (Fig. 3(B)). Similarly, for the provided mild COVID CT images, our system classified correctly 9 out of 10 images as mild, and 1 as severe (Fig. 3(B)). This performance of our proposed system in severity prediction shows that this system can reliably be used to advise intensive care during COVID-19 infection. Overall these results justify the use of our proposed system in real settings.



**Fig. 3.** Confusion matrices: Showing COVIDC performance in real setting. A) COVID vs non-COVID; B) COVID severity prediction.

## 5. Discussion

In the absence or inadequate provision of specific drugs or vaccines for the treatment or prevention of COVID-19 its quick and automatic diagnosis and severity prediction is crucial. After the successful application of CT scans with machine learning in many clinical and radiological studies, it has become an important avenue to be explored to detect and severity prediction of COVID-19. Several machine learning methods have been proposed which use digital CT scans along with deep learning or transfer learning by fine-tuning the customized fully connected layer. These deep and transfer learning-based approaches face data scarcity issues. Meanwhile, studies based on shallow learning with handcrafted features do not achieve an ideal accuracy. Furthermore, all the existing methods are proposed only to detect COVID-19 and its severity prediction was still an open research question. To knob these issues and to get the best of both worlds, we have proposed a novel approach that uses a combination of transfer and shallow learning to detect COVID-19 and predict its severity. In this approach, we have used various off-the-shelf CNN based pre-trained models on ImageNet to get feature maps and shallow learning algorithms such as SVMs to get the final trained model. We have used various feature maps and shallow machine learning algorithms along with different valuation techniques to come up with a model having improved generalization performance with an F1 score of 0.94 along with are under the PR curve of 0.98. Direct performance comparison of our proposed method with existing methods is not possible due to variations in the amount of training data and performance metrics. However, we give a brief comparison based on learning strategies. In this regard, the performance of our proposed method is comparable to a transfer learning-based method proposed by Ahuja et al. with an F1 score even without using image augmentation to increase the dataset [22]. If we compare our method with existing shallow learning-based methods, our method gives improved performance in comparison to a method proposed by Kang et al. trained using handcrafted features with an overall accuracy of 86%. Improved performance of our COVIDC over the existing state-of-the-art methods [22, 27] and in a real setting suggests that the proposed method can effectively be used to diagnose COVID-19 and to predict its severity.

## 6. Conclusions

In this study, we have proposed a system called COVIDC for the preliminary diagnosis and severity prediction of COVID-19 infection using chest CT scans. The stringent performance evaluation through 10-fold CV, on an external validation dataset, and in a use under real setting

shows that our proposed system can effectively be used not only to diagnose COVID-19 infection but also its severity prediction. The use of our proposed system can help to reduce the surge of COVID-19 by advising timely isolation and further diagnostic tests such as RT-PCR to the suspected patients. This system can also aid in avoiding massive casualties by deciding on intensive care in case of severe COVID-19 infection. The key findings of the study are listed as follows:

- Our proposed system performed significantly better in comparison to the state-of-the-art existing systems during the rigorous adopted evaluation criteria even in the presence of substantial variations in the used CT images.
- Our proposed system not only diagnose COVID-19 but also predict its severity.
- We have made our proposed system accessible through a publicly open cloud-based webserver and open-source code.

## Availability of data and materials

All data generated or analyzed during this study are included in this paper or available at online repositories. A Python implementation of the proposed method together with a webserver is available at <https://sites.google.com/view/wajidarshad/software> and <https://github.com/wajidarshad/covidc>.

## Authors' contributions

WAA conceived the idea, developed the scientific workflow, performed the experiments, analyzed and interpreted the results, and was a major contributor in manuscript writing. SAA contributed to the analysis of the results and writing of the manuscript. SA, GI, SAA, KA, SK, AA, KI, MS, JC, and AA helped in results interpretation and validation, formal analysis, and manuscript writing. All authors have read and approved the final manuscript.

## 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.

## Acknowledgments

The authors would like to acknowledge the services of all those who

have provided CT scans data as open access repositories. We also thank the reviewers and the editor for their valuable feedback and suggestions to improve the presentation of this work.

## References

- [1] WHO. Director-General's remarks at the media briefing on COVID-19. <http://www.who.int/director-general/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020>. [Accessed 8 December 2020].
- [2] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;20:533–4. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1).
- [3] Cheng S-C, Chang Y-C, Fan Chiang Y-L, Chien Y-C, Cheng M, Yang C-H, et al. First case of coronavirus disease 2019 (COVID-19) pneumonia in taiwan. *J Formos Med Assoc* 2020;119:747–51. <https://doi.org/10.1016/j.jfma.2020.02.007>.
- [4] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [5] Niazzkar M, Niazzkar HR. COVID-19 outbreak: application of multi-gene genetic programming to country-based prediction models. *ELECTRON J GEN MED* 2020;17:e247. <https://doi.org/10.29333/ejgm/8232>.
- [6] Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease. *Cochrane Database Syst Rev* 2020;2020. <https://doi.org/10.1002/14651858.CD013665>.
- [7] Commissioner O. Coronavirus disease 2019 testing basics. FDA; 2020.
- [8] Sheikhzadeh E, Eissa S, Ismail A, Zourab M. Diagnostic techniques for COVID-19 and new developments. *Talanta* 2020;220:121392. <https://doi.org/10.1016/j.talanta.2020.121392>.
- [9] Borakati A, Perera A, Johnson J, Sood T. Diagnostic accuracy of X-ray versus CT in COVID-19: a propensity-matched database study. *BMJ Open* 2020;10:e042946. <https://doi.org/10.1136/bmjopen-2020-042946>.
- [10] Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahub ZB, et al. Can AI help in screening Viral and COVID-19 pneumonia? *IEEE Access* 2020;8:132665–76. <https://doi.org/10.1109/ACCESS.2020.3010287>.
- [11] Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* 2020;296:E32–40. <https://doi.org/10.1148/radiol.2020200642>.
- [12] Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. *Radiology* 2020;296:E41–5. <https://doi.org/10.1148/radiol.2020200343>.
- [13] Kwee TC, Kwee RM. Chest CT in COVID-19: what the radiologist needs to know. *Radiographics* 2020;40:1848–65. <https://doi.org/10.1148/rg.2020200159>.
- [14] Asnaoui KE, Chawki Y, Idri A. Automated methods for detection and classification pneumonia based on X-ray images using deep learning. *ArXiv:2003.14363* 2020. <https://arxiv.org/abs/2003.14363>.
- [15] Chandra TB, Verma K, Singh BK, Jain D, Netam SS. Coronavirus disease (COVID-19) detection in Chest X-Ray images using majority voting based classifier ensemble. *Expert Syst Appl* 2021;165:113909. <https://doi.org/10.1016/j.eswa.2020.113909>.
- [16] Angelini E, Dahan S, Shah A. Unravelling machine learning – insights in respiratory medicine. *Eur Respir J* 2019;54(6):1901216. <https://doi.org/10.1183/13993003.01216-2019>.
- [17] Walsh SLF, Humphries SM, Wells AU, Brown KK. Imaging research in fibrotic lung disease; applying deep learning to unsolved problems. *The Lancet Respiratory Medicine* 2020;8:1144–53. [https://doi.org/10.1016/S2213-2600\(20\)30003-5](https://doi.org/10.1016/S2213-2600(20)30003-5).
- [18] Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *The Lancet Respiratory Medicine* 2018;6:837–45. [https://doi.org/10.1016/S2213-2600\(18\)30286-8](https://doi.org/10.1016/S2213-2600(18)30286-8).
- [19] Wang S, Zhou M, Liu Z, Liu Z, Gu D, Zang Y, et al. Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. *Med Image Anal* 2017;40:172–83. <https://doi.org/10.1016/j.media.2017.06.014>.
- [20] LeCun Y, Haffner P, Bottou L, Bengio Y. Object recognition with gradient-based learning. In: Forsyth DA, Mundy JL, di Gesù V, Cipolla R, editors. *Shape, contour and grouping in computer vision*. Berlin, Heidelberg: Springer; 1999. p. 319–45. [https://doi.org/10.1007/3-540-46805-6\\_19](https://doi.org/10.1007/3-540-46805-6_19).
- [21] Debie E, Shafi K. Implications of the curse of dimensionality for supervised learning classifier systems: theoretical and empirical analyses. *Pattern Anal Appl* 2019;22:519–36. <https://doi.org/10.1007/s10044-017-0649-0>.
- [22] Ahuja S, Panigrahi BK, Dey N, Rajinikanth V, Gandhi TK. Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices. *Appl Intell* 2020;1–15. <https://doi.org/10.1007/s10489-020-01826-w>.
- [23] Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. Artificial intelligence augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other origin at chest CT. *Radiology* 2020;296:E156–65. <https://doi.org/10.1148/radiol.2020201491>.
- [24] Han Z, Wei B, Hong Y, Li T, Cong J, Zhu X, et al. Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning. *IEEE Trans Med Imag* 2020;39:2584–94. <https://doi.org/10.1109/TMI.2020.2996256>.
- [25] Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun* 2020;11:4080. <https://doi.org/10.1038/s41467-020-17971-2>.
- [26] Javor D, Kaplan H, Kaplan A, Puchner SB, Krestan C, Baltzer P. Deep learning analysis provides accurate COVID-19 diagnosis on chest computed tomography. *Eur J Radiol* 2020;133. <https://doi.org/10.1016/j.ejrad.2020.109402>.
- [27] Kang H, Xia L, Yan F, Wan Z, Shi F, Yuan H, et al. Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning. *IEEE Trans Med Imag* 2020;39:2606–14. <https://doi.org/10.1109/TMI.2020.2992546>.
- [28] Ko H, Chung H, Kang WS, Kim KW, Shin Y, Kang SJ, et al. COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation. *J Med Internet Res* 2020;22:e19569. <https://doi.org/10.2196/19569>.
- [29] Mishra AK, Das SK, Roy P, Bandyopadhyay S. Identifying COVID19 from chest CT images: a deep convolutional neural networks based approach. *Journal of Healthcare Engineering* 2020;2020:e8843664. <https://doi.org/10.1155/2020/8843664>.
- [30] Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S, Shukla PK. Deep transfer learning based classification model for COVID-19 disease. *IRBM* 2020. <https://doi.org/10.1016/j.irbm.2020.05.003>. In press.
- [31] Qian X, Fu H, Shi W, Chen T, Fu Y, Shan F, et al. Lung-sys: a deep learning system for multi-class lung pneumonia screening from CT imaging. *IEEE Journal of Biomedical and Health Informatics* 2020;24:3539–50. <https://doi.org/10.1109/JBHI.2020.3030853>.
- [32] Shah V, Keniya R, Shridharani A, Punjabi M, Shah J, Mehendale N. Diagnosis of COVID-19 using CT scan images and deep learning techniques. *Emergency Radiology* 2021. <https://doi.org/10.1007/s10140-020-01886-y>. In press.
- [33] Silva P, Luz E, Silva G, Moreira G, Silva R, Lucio D, et al. COVID-19 detection in CT images with deep learning: a voting-based scheme and cross-datasets analysis. *Inform Med Unlocked* 2020;20:100427. <https://doi.org/10.1016/j.imu.2020.100427>.
- [34] Singh D, Kumar V, Vaishali, Kaur M. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks. *Eur J Clin Microbiol Infect Dis* 2020;39:1379–89. <https://doi.org/10.1007/s10096-020-03901-z>.
- [35] Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J* 2020;56(2):2000775. <https://doi.org/10.1183/13993003.00775-2020>.
- [36] X W, X D, Q F, Q Z, J F, H M, et al. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. *IEEE Trans Med Imag* 2020;39:2615–25. <https://doi.org/10.1109/tmi.2020.2995965>.
- [37] Xu X, Jiang X, Ma C, Du P, Li X, Lv S, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* 2020;6:1122–9. <https://doi.org/10.1016/j.eng.2020.04.010>.
- [38] Ozsahin I, Sekeroglu B, Musa MS, Mustapha MT, Uzun Ozsahin D. Review on diagnosis of COVID-19 from chest CT images using artificial intelligence. *Computational and Mathematical Methods in Medicine* 2020;2020:e9756518. <https://doi.org/10.1155/2020/9756518>.
- [39] Zare MR, Alebiosu DO, Lee SL. Comparison of Handcrafted Features and Deep Learning in Classification of Medical X-ray Images. *Fourth international conference on information retrieval and knowledge management (CAMP)*, vol. 2018; 2018. p. 1–5. <https://doi.org/10.1109/INFRKM.2018.8464688>.
- [40] Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. COVID-19 image data collection: prospective predictions are the future. *arXiv:2006.11988* 2020. <https://arxiv.org/abs/2006.11988>.
- [41] Soares E, Angelov P, Biaso S, Froes MH, Abe DK. SARS-CoV-2 CT-scan dataset: a large dataset of real patients CT scans for SARS-CoV-2 identification. *MedRxiv* 2020. <https://doi.org/10.1101/2020.04.24.20078584>.
- [42] Yang X, He X, Zhao J, Zhang Y, Zhang S, Xie P. COVID-CT-dataset: a CT scan dataset about COVID-19. *arXiv:2003.13865* 2020. <https://arxiv.org/abs/2003.13865>.
- [43] Chandra TB, Verma K. Analysis of quantum noise-reducing filters on chest X-ray images: a review. *Measurement* 2020;153:107426. <https://doi.org/10.1016/j.measurement.2019.107426>.
- [44] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *arXiv:1512.03385* 2015. <https://arxiv.org/abs/1512.03385>.
- [45] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. 2015. *ArXiv:151200567* [Cs].
- [46] Chollet F. Xception: deep learning with depthwise separable convolutions. 2017. *ArXiv:161002357* [Cs].
- [47] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *ArXiv*; 2015, *arXiv:1409.1556*.
- [48] Zoph B, Vasudevan V, Shlens J, Le QV. Learning transferable architectures for scalable image recognition. *ArXiv*; 2018, 1707.07012.
- [49] Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Arxiv*; 2018, 1608.06993.
- [50] Wang D, Mo J, Zhou G, Xu L, Liu Y. An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. *PLoS One* 2020;15:e0242535. <https://doi.org/10.1371/journal.pone.0242535>.
- [51] Breiman L. Random forests. *Mach Learn* 2001;45:5–32. <https://doi.org/10.1023/A:1010933404324>.
- [52] Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995;20:273–97. <https://doi.org/10.1023/A:1022627411411>.
- [53] Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001;29:1189–232. <https://doi.org/10.1214/aos/1013203451>.
- [54] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;12:2825–30.



- [55] Abbasi WA, Yaseen A, Hassan FU, Andleeb S, Minhas FUA. ISLAND: in-silico proteins binding affinity prediction using sequence information. *BioData Min* 2020;13:20. <https://doi.org/10.1186/s13040-020-00231-w>.
- [56] Ballester PJ, Mitchell JBO. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics* 2010;26:1169–75. <https://doi.org/10.1093/bioinformatics/btq112>.
- [57] Li H, Leung K-S, Wong M-H, Ballester PJ. Substituting random forest for multiple linear regression improves binding affinity prediction of scoring functions: cyscore as a case study. *BMC Bioinf* 2014;15:291. <https://doi.org/10.1186/1471-2105-15-291>.
- [58] Moal IH, Agius R, Bates PA. Protein-protein binding affinity prediction on a diverse set of structures. *Bioinformatics* 2011;27(21):3002–9. <https://doi.org/10.1093/bioinformatics/btr513>.
- [59] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22Nd ACM SIGKDD international conference on knowledge discovery and data mining*. New York, NY, USA: ACM; 2016. p. 785–94. <https://doi.org/10.1145/2939672.2939785>.
- [60] Abbasi WA, Minhas FUA. Issues in performance evaluation for host–pathogen protein interaction prediction. *J Bioinform Comput Biol* 2016;14:1650011.
- [61] Davis J, Goadrich M. The Relationship Between Precision-Recall and ROC Curves. *Proceedings of the 23rd International Conference on Machine Learning*, New York, NY, USA 2006:233–40. <https://doi.org/10.1145/1143844.1143874>.
- [62] Tharwat A. *Classification assessment methods*. Emerald Publishing Limited; 2020. 10.1016/j.aci.2018.08.003.