COVID-19 Diagnosis with Artificial Intelligence

Rassa Ghavami Modegh, Ahmad Salimi, Sepehr Ilami, Amir H. Dehqan, Hamed Dashti, Shaghayegh Haghjooy Javanmard, Hossein Ghanaati, and Hamid R. Rabiee

Abstract Pandemics are epidemics that spread rapidly over countries or continents. They are large-scale outbreaks of contagious diseases that can significantly increase morbidity and mortality and cause substantial economic and social harm. If not properly controlled, the hospitals, physicians, and care staff become overloaded, driving the exponential growth of the infected population. Therefore, the early diagnosis of the infected would help to break the transmission chain. Artificial intelligence (AI) methods and deep networks have proven their outstanding performance in many fields. In this chapter, we discuss the potential use-cases of AI in controlling the COVID-19 pandemic. We introduce general concepts of AI in the diagnosis and screening of COVID-19. We also highlight common problems and pitfalls in developing AI methods with experiments on actual data. We ascertain the certainty of interpretable AI models for a trustful diagnosis. We study the state-of-the-art techniques in the diagnosis of COVID-19. We examine some of the experiences in controlling the COVID-19 pandemic and discuss why AI was not used to its full potential in the COVID-19 pandemic. Finally, we propose some future works to prevent the discussed problems in future pandemics.

R. G. Modegh · A. Salimi · S. Ilami · H. R. Rabiee (✉)
Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
e-mail: rabiee@sharif.edu

R. G. Modegh · A. Salimi · S. Ilami · A. H. Dehqan · H. Dashti · H. R. Rabiee
AI-Med, Medric Group, AI & Data Science Innovation Center, Sharif University of Technology, Tehran, Iran

A. H. Dehqan
Iran University of Medical Sciences, Tehran, Iran

S. H. Javanmard
Applied Physiology Research Center, Isfahan Cardiovascular Research Institute, Isfahan University of Medical Science, Isfahan, Iran

H. Ghanaati
Department of Radiology, Tehran University of Medical Sciences, Tehran, Iran

Acronyms

AI	artificial intelligence
ML	machine learning
NLP	natural language processing
KNN	K-nearest neighbors
SVM	support vector machine
AUC	area under the curve
CT	computed tomography
MRI	magnetic resonance imaging
CXR	chest X-ray
API	application programming interface
PCR	polymerase chain reaction
SOTA	state of the art
PET	positron emission tomography
NIH	National Institutes of Health
WHO	World Health Organization
HMS	health monitoring system
PHO	Ontario Health Services
RSNA	Radiological Society of North America
GPS	global positioning system

1 Introduction

Pandemics are epidemics that spread rapidly over the globe. If not properly controlled, the hospitals, physicians, and healthcare staff become overloaded, which results in substantial morbidity and mortality and causes significant economic and social harm. During the pandemics such as COVID-19, patients flock to hospitals at any time of the day, and many others check in with the mildest symptoms. Healthcare workers need to be present at all times, and severe fatigue can affect their performance. As a gold standard, the Polymerase Chain Reaction (PCR) test is used to confirm the presence of COVID-19. However, PCR is time-consuming and has high false-negative rates. Therefore, in some medical centers, they have been replaced by Computerized Tomography (CT)-scan imaging (A non-invasive imaging technique that gives detailed three-dimensional images from the body). CT-scan diagnosis by a specialized radiologist is faster, contains more details about pneumonia, and can

provide a quantitative measurement of the severity of infection [1]. However, experts are not always available, and pandemics can exacerbate these conditions.

In these situations, accurate and rapid diagnosis of the disease effectively prevents the outbreak of a pandemic. However, factors such as the increasing workload of physicians and lack of expert radiologists in a pandemic make screening and identifying suspicious patients difficult and slow [2]. However, due to the high transmission rate of the COVID-19 pandemic, rapid diagnosis of the disease is necessary. Artificial intelligence (AI) systems can pressure off the medical centers by assisting physicians in speeding up the diagnosis and treatment procedures. They can also help the physicians to make early predictions and efficiently manage the available resources to control the spread of disease.

Artificial Intelligence is the ability of a computer to perform tasks commonly associated with intelligent beings [3]. The core of an AI system is its knowledge processing unit. The system uses this unit to acquire knowledge and perform specific tasks [4]. In Machine learning (ML), an important domain in AI, rules are learned from data and used to make accurate decisions. Computers have advantages over humans that make them more suitable to perform specific tasks. They can perform calculations much faster, have more memory to remember what is essential, and are perpetually available. They never get tired, never change modes, and no internal state affects their decision-making process. For example, the COVID-19 diagnosis of CT-scan takes up to 15 minutes for a radiologist, while it takes only a few seconds for an AI-based method [5]. Another challenge in the COVID-19 pandemic is distinguishing between COVID-19 pneumonia and other bacterial and non-COVID-19 viral pneumonia types. AI can help inexperienced physicians improve their ability to correctly diagnose different types of infections [6]. In addition, compared with an experienced thoracic radiologist, AI can achieve higher sensitivity than radiologists at the early stages of the disease when the lack of human-visible abnormalities in CT-scan exist [7].

Resource planning is utilized in hospitals to determine whether the hospital can care for a patient's needs. However, planning becomes a challenge during a pandemic with overburdened hospitals and limited resources. The goal is to save as many lives as possible. In this case, they have no choice but to admit patients based on the severity of their condition and the possibility of their survival. Some patients do not need to be treated in the hospital. However, for others, hospitalization is essential. Therefore, determining the condition of each patient and predicting the required facilities becomes vital. AI can help with resource planning by learning complex patterns from symptoms, historical and clinical data. Currently, many AI models are being used to predict overall mortality for the next few months or even a year [8]. These methods use various information such as blood biomarkers, age, gender, and disease background of the patient to achieve accurate predictions [9].

During a pandemic, patients with varying degrees of severity are admitted to hospitals. However, not all of them need to be hospitalized. Moreover, hospitals can be a source of further infection, especially during a pandemic. For example, a study reported that 41% of 138 hospitalized patients with COVID-19 were infected after being admitted to a hospital [10]. The Health Monitoring System (HMS) allows

patients to be monitored remotely. HMS is an advanced technology alternative to traditional patient and health management. It consists of a wearable wireless device, such as a wristband with a sensor, equipped with software for the specialist to access important medical information [11]. AI can help to monitor the patients on two levels. First, it can assess the condition more frequently than a physician to warn for irregularities. Second, it can also predict the patient's health condition to adopt the required precautions in the coming days.

Infectious diseases such as COVID-19 can spread from person to person by physical contact, droplets, saliva, or airborne [12]. Because of the rapid spread of the COVID-19 virus, it has a high rate of infection [13]. To prevent the outbreak of the disease, one approach is to implement a patient tracking system that immediately alerts individuals who have had recent contact with known cases of the disease during their viral period and prompt them to become isolated [14]. There are several ways to track down individuals who have had contact or interaction with patients affected by the COVID-19 virus. For example, by collecting data from Bluetooth-based tracking applications, GPS and social graphs, video surveillance, and CCTV cameras. Further information can be gathered from card transaction data, internet search and social media monitoring, text data, and network-based Application Programming Interfaces (APIs), which are intermediary programs that connect two other apps together. Automated systems powered by AI can be designed to analyze different data types and model a tracking system to control the pandemic [14–16]. However, challenges such as technical limitations, socioeconomic disparities, data privacy and security risks, and ethical issues still lie ahead [17, 18].

The following section describes the general concepts and guidelines for using an appropriate AI system to manage the COVID-19 pandemic, focusing on diagnostic and screening tasks. Then, we explain the potential problems and pitfalls of AI-based methods using experiments with actual data. In section three, we elaborate on the use of AI to facilitate various operations in a pandemic. In section four, we describe the State-Of-The-Art (SOTA) AI systems used to diagnose COVID-19. The last section discusses the obstacles to exploiting AI's full potential in the current COVID-19 pandemic and proposes future works to make AI more effective in a pandemic.

2 A Guideline to Develop AI Models for Diagnosis and Screening

Continuous screening and rapid diagnosis are two of the most frequent tasks in controlling a pandemic. The function of the standard COVID-19 screening test is to identify people with a higher risk of spreading the disease. Since asymptomatic transmission plays a significant role in a pandemic, using screening tests to classify patients becomes very important to reduce the outbreak of the disease. The purpose of these tests is not to confidently detect the virus but to identify the suspicious patients and isolate them to minimize the rate of infection. For example, chest X-ray (CXR)

images are used for screening COVID-19 as they are widely available, inexpensive, fast, and contain helpful information for detecting COVID-19 infections [19–23]. In addition, some studies have shown that magnetic resonance imaging (MRI), PET imaging, and ultrasound can be used for diagnosis and screening. However, they are not usually used in clinical practice [24–30].

> On the other side, the goal of diagnostic tests is to detect the presence or absence of the virus in suspected patients of COVID-19. These tests may also identify the cause, location, and severity of the disease. If the screening result indicates the presence of the disease, diagnostic tests are required to confirm the infection. For instance, the PCR or CT-scan tests can confirm the presence of COVID-19 [1, 31–33].

There is a wide range of data sources for the detection of a virus. Usually, the easier, faster, and more available the data collection, the less accurate it becomes. Generally speaking, during the COVID-19 pandemic, the data range starts with mobile applications widely available and fast. It continues to PCR testing and CT-scan images, which are more accurate and less available. As mentioned above, screening tests require simple and rapid data collection, while diagnostic tests require more accurate data.

As described above, screening is conducted over a larger population. However, an expert is not available to assess the test in many cases. Although diagnosis is performed over a selected population, their number is still high in a pandemic, resulting in the extreme tiredness of the experts. In addition, there is a shortage of experts in some regions. AI can help in both tasks and assist the medical staff in making more accurate and rapid diagnoses. We provide a guideline for how an AI system can be trained for such tasks. We also describe the general concepts and the common problems and pitfalls in screening and diagnosis with several real-world examples [34–37].

2.1 Training, Testing, and Further Validation of AI Models

In the supervised learning approach for an ML model, everything starts from data. We do not embed the knowledge of how to decide or how to calculate in the model. Instead, we let the model learn its optimal configuration from the data. ML problems fall into three general categories of classification, regression, and clustering. There are a limited number of groups in classification problems, and the goal is to identify the group for each sample in the dataset. For example, in classification, the goal is to determine whether a person has a specific disease. In regression, each sample is associated with a continuous number, and the goal is to calculate that number. Calculating the percentage of infected lung regions in pulmonary diseases with pneumonia is an example of this category. In clustering, the samples do not have any assigned groups or numbers, and the goal is to group samples with similar patterns in the collected data. For example, clustering patients with different symptoms in an unknown disease can lead experts to the disease's possible stroke paths.

The procedure of deploying an AI model for a classification problem is depicted in Fig. 1. In a classification problem, the dataset consists of multiple samples from each
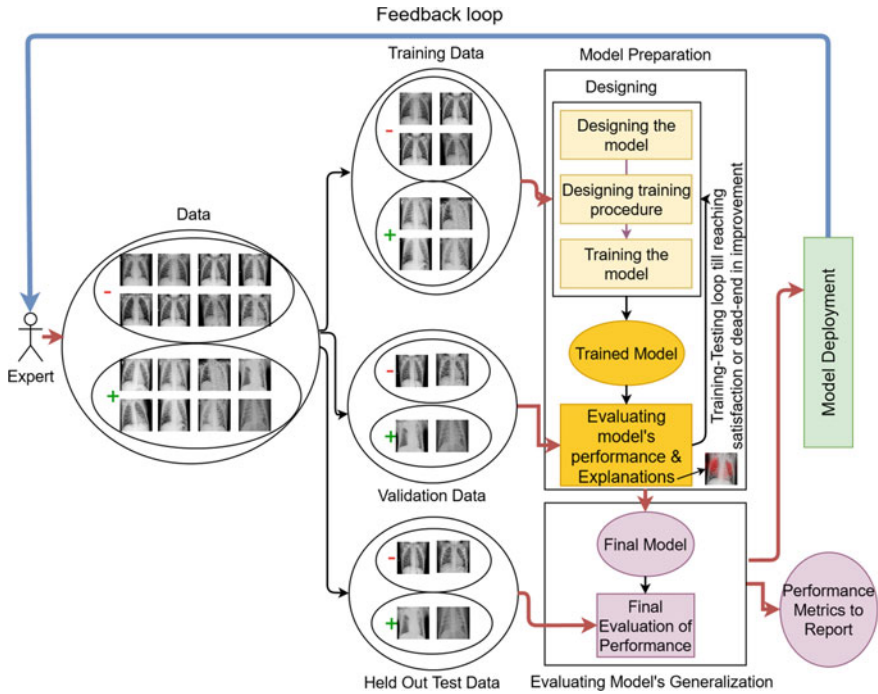
Fig. 1 The procedure of deploying an AI model for a classification problem

group or class. At first, the dataset is divided into three sets; training, validation, and test. Next, the training data is used to obtain the parameters of the model. A model may perform well on the training data but only learns the appearance of the samples as they are repeated in the training process. This phenomenon is called overfitting. To avoid over-fitting, multiple augmentations are applied to the input sample. For image data, rotation, geometric transformations, and changing brightness are examples of common augmentations. Finally, the trained model is evaluated on the validation data to make sure it has learned the general properties of the data. In addition to evaluation metrics, interpretations of the network may also be used as an extra check to validate the reasons behind the model's decisions.

Different architectures and training procedures lead to various performances of the model on training and validation sets. The one with the best performance on validation data is selected as the final model among the trained models. In evaluating the performance on validation data, redesigning, and retraining the model, the validation set is observed multiple times. Therefore, the model's performance on the validation set cannot be a fair metric for the model's generalizability. Therefore, a test set is held out to evaluate the model's performance on completely unseen data.

In some cases, models are designed to fit specific data appearances. As a result, they may fail on other samples, e.g., samples of another imaging device, harming the model's generalizability. Therefore, collecting the test data from sources other than

the training and validation data can improve the model's generalizability. At the final step, the model is deployed and tested by an expert. The feedbacks are collected and returned to the model to make it optimal.

2.2 Data Gathering and Soundness

As mentioned above, an ML algorithm learns the model from data. Therefore, it is essential to have a diverse dataset to enforce the model learn different observations. Otherwise, it is limited to what it has seen and may perform poorly for other datasets. For example, it is crucial to consider patients with different stages of the disease for diagnosis, including healthy people in the dataset. If the training dataset only consists of the severe stage of the disease, it may fail to detect the early stages of the disease.

Another important point is to ensure the labels are correctly assigned to the dataset samples because the model learns from those labeled samples. If an expert labels the samples, the model will become biased towards the expert's opinion. Having confirmation from experts, using laboratory tests, and post-data-gathering samples' observations ensure correctness and make the model more robust. Another problem induced by the inexactness of the labels is the trainability of the model. When there are many mistakes in labeling, the model fails to train well. Therefore, we designed an experiment to investigate the effect of inappropriate data labeling on AI models. We selected two groups of data, one from NIH's chest-X-ray dataset [38] and the other from RSNA's chest-X-ray dataset [39]. The NIH's dataset contains 112120 images that are labeled using natural language analysis on the radiological reports. Therefore, they may be inexact. On the other hand, RSNA's dataset is a subset of NIH's dataset that at least two experts annotate to make it more trusted. Therefore, we gathered "No Finding" and "Pneumonia" samples of NIH's dataset for the first group, including 45449 negative and 696 positive samples. We collected "Healthy" and "Pneumonia" samples of RSNA's dataset for the second group, including 8851 negative and 9555 positive samples.

To eliminate the effect of unequal group sizes, some of the negative samples of NIH and some of the positive samples of RSNA were randomly discarded. As a result, 8851 negative samples and 696 positive samples remained for both groups. Next, 80% of the datasets Were selected as the training set and the rest as the validation set. Finally, two well-known models, Inception V3 [40] and ResNet18 [41], were trained with the Adam optimizer [42]. We utilized transfer learning and initialized the weights using the trained models available in the Torchvision package [43]. For both models, we first tuned the last fully connected layer and then one block before that. To avoid the effect of imbalanced classes, batches of size 32 containing 16 positives and 16 negative samples were used for training. Data augmentation techniques such as random brightness, cropping and resizing, rotating, and shearing was used to expand the dataset for generalization.

The epoch acquiring the best average sensitivity and specificity on the validation data was selected for all the models. The fine-tuning was done in the second stage

Table 1 The effects of inappropriate data labeling. Sens, sensitivity; spec, specificity

Train set	NIH (inappropriate labeling)		RSNA (appropriate labeling)	
Stage	Validation	Test	Validation	Test
Metric	Sens Spec Avg	Sens Spec Avg	Sens Spec Avg	Sens Spec Avg
Inception v3	56.67 86.45 71.56	59.12 77.79 **68.46**	88.33 95.23 91.78	78.55 80.41 **79.48**
ResNet 18	66.67 77.92 72.29	49.45 64.62 **57.04**	87.22 96.24 91.73	54.38 94.82 **74.6**

and continued until the model achieved more than 98% sensitivity and specificity for the training data. The epoch with the best performance on validation data was selected as the final model. We tested the models with the dataset of COVID-19 radiography database [44], Chest X-Ray Pneumonia [45], and Kaggle VinBigData [46] to evaluate their generalizability. The results for validation and test datasets are presented in Table 1.

As results show, both models trained on the NIH dataset fail to achieve high-performance metrics as the models trained on RSNA with the validation set. This proves that in datasets of equal size, certainty on labels helps exceedingly in training the model. Apart from the low sensitivity achieved on the test data due to the shortage of positive samples in the selected groups, the models trained with RSNA samples have achieved higher performance metrics for the test data.

2.3 Data Diversity, the Problem of Batch Effect and Generalization

In the COVID-19 pandemic, many published papers have used small datasets and reported high evaluation results. However, in most cases, the samples related to different classes were collected from other sources. Therefore, these models are susceptible to bias toward the appearance of samples for each dataset instead of learning to solve a general problem. As a result, the models lack generalization and may fail when applied to unseen data. This problem is not specific to small datasets, and large datasets may also have the same issue. For example, suppose the test set was gathered from the same training and validation datasets. In that case, the problem would not be identified in the final evaluation stage, and a problematic model will be deployed.

We designed an experiment to show the destructive effect of bias and batch effect in an intense investigation. We selected two groups of data from Kaggle's chest Xray Pneumonia and Kaggle's RSNA challenges. Both challenges were intended for detecting pneumonia from chest X-ray images. We selected pneumonia samples from the RSNA dataset and healthy samples from the chest-X-ray Pneumonia dataset for the first group, including 9555 positive and 1583 negative samples. We selected the same amount of positive and negative samples from both datasets for the second group. We selected 80% of the datasets as the training set and 10% as the validation

Table 2 The problem of batch effect and generalization caused by a biased dataset

Dataset	Group 1—biased dataset		Group 2—unbiased dataset	
Stages	Validation	Test	Validation	Test
Metrics	Sens Spec Avg	Sens Spec Avg	Sens Spec Avg	Sens Spec Avg
Inception v3	99.19 100 99.59	24.62 5.08 **14.85**	93.14 94.81 93.98	97.95 89.28 **93.61**
ResNet 18	99.65 100 99.82	16.15 2.60 **9.37**	94.88 94.07 94.48	98.21 92.33 **95.27**

set. 10% of each primary dataset was put aside as the shared test dataset. We eliminated the positive samples related to the RSNA dataset and the negative samples of the chest X-ray Pneumonia from the test set to bold the differences. The training and model selection was performed similarly to the scheme mentioned in Sect. 2.2. The evaluation results of the models are presented in Table 2. As the table shows, the models trained on the first group have reached high sensitivities and specificities on the validation set, which has the same bias as the training data. One may think the models perform even better than the second group considering the performance metrics of the validation sets. However, as the evaluation of test data with inverse bias shows, the models of the first group have become entirely biased toward the appearances of the samples for each dataset. In contrast, the evaluation metrics of the second group on the test set are similar to the validation set results.

2.4 Interpreting the Black-Box Deep AI Models

Deep neural networks have shown excellent performances, achieving high accuracies in many domains, even better than human experts. However, most of these models are black-boxes, meaning the internal decision-making mechanism of the network at the intermediate layers is not known. Therefore, their high accuracy is not sufficient to build trust toward them. Consequently, they may have performed well due to wrong reasons that are irrelevant to the domain-specific concerns [47]. In recent years, researchers have focused on interpreting the black-box models. Interpreting means explaining the reasons behind the decision of a model in a human-understandable way [48]. These interpretations help identify the bias in the model's decisions, make sure the model has been fair, monitor the model performance based on the reasons behind the decisions, and learn unknown domains from the model [47].

To show the trace of bias in the models, we interpreted the decisions of the above ResNet18 model using Guided Grad-CAM [49]. The interpretations will show which parts of the images were considered by each model to make a decision. Figure 2 presents the interpretation results of the models on seven random pneumonia samples from the test dataset. As the figure shows, the model trained on the biased dataset pays more attention to the outside regions of the lung, while the other model focuses on lung infections. It shows that the model trained on the biased dataset has learned
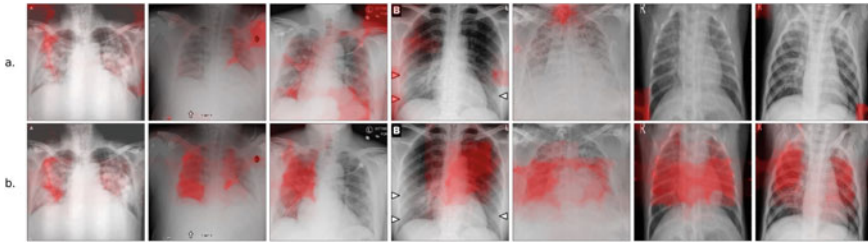
Fig. 2 Comparing interpretation results on ResNet18 using Guided Grad-CAM, trained on **a** the biased, and **b** the unbiased datasets

the apparent differences between the two datasets as the distinguishing feature of the patient and healthy images, instead of learning the patterns of pneumonia.

3 State-Of-The-Art AI Technologies for COVID-19 Diagnosis

Physicians in many countries have adopted imaging techniques and test kits to obtain more accurate diagnosis results of COVID-19 and determine the severity of the disease for each patient. CT-scan imaging is more accurate and sensitive among other imaging techniques, and many researchers have worked on the automatic diagnosis of COVID-19 using CT-scan images. Some have also focused on estimating the severity metrics for the patients. In addition to detecting infected from healthy individuals, some have also aimed to solve the more difficult problem of distinguishing COVID-19 from other lung diseases such as community-acquired pneumonia. Unfortunately, a fair comparison between published results is impossible due to the lack of benchmark datasets. Therefore, most studies applied similar steps with a limited number of choices on different datasets. Consequently, we discuss the existing challenges and the general methods used by other studies to overcome these challenges in the case of COVID-19 diagnosis. The details and performance measures for the selected studies are given in Table 3.

Diagnosis with CT-scan images is not a straightforward problem. It includes many challenges, and researchers have used different methods to overcome these challenges. For example, CT-Scan samples may have a different number of images, from 30 to 800, depending on the thickness of the cuts. This poses a significant challenge to the learning process. Some researchers have focused on high-resolution samples to keep the sensitivity and accuracy high [50]. In contrast, the others have trained their model on large cohorts with different thicknesses to have a more generalized model [6, 51, 52].

The COVID-19 CT-scans datasets usually have labeled samples. This is because carefully annotating images associated with infection and images with disease features is time-consuming, while radiologists focus on their primary tasks during the

Table 3 Details and performance metrics of the SOTA studies in diagnosing COVID-19

References	Classification groups	Training data	Test data	Evaluation metrics on test data
[50]	COVID-19 vs CAP vs Non-pneumonia	1165 COVID-19 1560 CAP 1193 Non-pneumonia [6 centers]	127 COVID-19 175 CAP 132 Non-pneumonia	Sensitivities for each group: COVID-19: 90% CAP: 87% Non-pneumonia: 94% AUC for COVID-19: 0.96
[6]	COVID-19 vs. nonviral CAP vs Influenza A/B vs Non-pneumonia	1294 COVID-19 666 nonviral CAP 70 Influenza A/B 1233 Non-pneumonia [3 centers + 2 public datasets]	1235 COVID-19 668 nonviral CAP 62 Influenza A/B 1234 Non-pneumonia [3 centers + 2 public datasets] 2113 COVID-19 1528 CAP 1333 Non-pneumonia [2 public holdout datasets]	Sensitivities for each group: COVID-19: 87.04% CAP: 96.88% Influenza A/B: 83.08% Non-pneumonia: 93.44% 2 Other public datasets: COVID-19: 84.66% CAP: 88.24% Non-pneumonia: 81.91%
[53]	COVID-19 vs others	703 COVID-19 684 non-COVID-19 [more than 7 centers]	326 COVID-19 1011 non-COVID-19 [more than 7 centers]	Sensitivity: 84.0% Specificity: 93.0% AUC: 94.9%
[54]	COVID-19 vs Influenza A vs Healthy	189 COVID-19 194 Influenza A 145 Healthy (3 centers)	30 COVID-19 30 Influenza A 30 Healthy	Sensitivities of each group: COVID-19: 86.67% Influenza: 83.33% Healthy: 90.0%
[51, 57]	COVID-19 vs CAP vs Normal	[Inexact numbers] 3000 COVID-19 750 CAP 500 Normal [from 15 countries]	[Inexact numbers] 250 COVID-19 200 CAP 100 Normal [from 15 countries]	Sensitivities for each group: [slice-based evaluation] COVID-19: 96.2% CAP: 98.2% Normal: 99.0%

(continued)

Table 3 (continued)

References	Classification groups	Training data	Test data	Evaluation metrics on test data
[52]	COVID-19 and non-COVID patients vs. Healthy COVID-19 vs non-COVID patients	1995 COVID-19 513 non-COVID-19 patients 517 Healthy [from 6 centers]	222 COVID-19 57 non-COVID-19 patients 519 Healthy [from 6 centers] 1435 COVID-19 [from 6 other hospitals]	Sensitivities: Diseased: 97.25% Healthy: 87.50% COVID vs Other diseases: COVID-19: 98.15% Other: 81.03% [for unseen centers] Diseased: 95.80% COVID-19 vs others: 95.73%

pandemic. Having many images per sample and being trained only on the samples' labels is like looking for a needle in a haystack for a model that does not know about the disease. When we train a model to distinguish between groups of samples, the model tries to learn the groups' differences. Searching in a large sample makes the training process harder and slower. Moreover, a sample of almost 200 images can fill the memory of a common GPU with 12 GB of RAM. Therefore the training becomes even more difficult.

Integrating human knowledge about the disease to the method, architecture of the model, and training process can significantly reduce the training challenges. It can also help the model learn relevant differences, which helps distinguish the groups of samples. For example, since we know COVID-19 affects the lungs, it would be logical to search for the marks of the disease in the lung areas of CT-scan images. Many researchers have adopted this knowledge in a pre-processing step. Some have used image processing techniques such as automatic thresholding to separate the lung areas from images. Some have trained a network over their own annotated private dataset to detect the lungs [6, 50–54]. Therefore, the rest of the images are eliminated except for the lungs, and the model skips the irrelevant areas during the training process.

Researchers have tried to ease the challenge of training with different methods. Some have adopted a fully supervised approach by providing slice-based labels and training their model to classify the slices rather than the whole image [51]. While being very helpful in training the model, this solution suffers from being biased to the error of the annotator. Even if the error were reduced by aggregating the annotations from multiple radiologists, there would still be a problem with the lack of human-detectable signs. This means the model will at most be able to distinguish what radiologists can distinguish. In practice, we have many unlabeled samples and only a limited number of labeled samples. Therefore, Some researchers have adopted semi-fully supervised methods for the training process. This is possible by adopting

a two termed loss function; one for the sample-level prediction and the other for the slice-level prediction. The first term is calculated over all the samples of the training batch, but the second term only focuses on samples with specific slice labels. Here the sample-level loss considers the slices with a lack of human observable marks. Thus, the slices with a negative label belonging to a positive sample should not be included as healthy slices in the training process for the annotated samples. Finally, many researchers have trained their models merely based on the sample-level labels and achieved notable results [6, 50, 52, 53].

COVID-19 is detected mainly by ground-glass opacity infections in the peripheral of the lung. Some researchers [52, 55] have developed their method based on this knowledge. First, they utilize the existing models and tools for pulmonary lesion detection that specify the regions of the lesions [56]. In the next step, they use deep learning models to classify the type of lesions and aggregate the results of the lesions to decide for the whole sample. Unfortunately, these methods may also miss the samples with the lack of human-detectable marks.

Diagnosing COVID-19 from CT-scan images is inherently a 3D problem. This is because the images of the CT-scan are cuts through another dimension, named depth. Some researchers have treated CT-scan images as 3D data, extended the 2D models to work with 3D data, and used the sample-level labels to train the model [6, 53, 54]. These models utilize 3D convolutional kernels and demand much more GPU memory than the 2D models during the training, and have more extensive computations in the evaluation phase. In addition, they are more sensitive to the thicknesses of the slices as convolutional networks are not invariant to different scales. Others have worked merely on 2D models and classified slices rather than the whole samples [51]. Other researchers have just aggregated the features extracted from the slices before making the final decision using a pooling layer [50] or an LSTM network [6]. However, they miss the information that the peripheral slices can add to each slice. Other researchers have adopted a hybrid 2D-3D model to have advantages of both schemes. They have used 2D models to extract information from individual slices and added the peripheral slices as extra channels to the input slices. Another problem with the 2D and the hybrid methods is their requirement for slice-level labels in the training phase. Some have overcome this problem by aggregating the features extracted from slices before the decision-making section of the network. To solve the problem with GPU's memory, they have subsampled uniformly from different parts of the sequence of images for each CT-scan. It is more probable to capture the disease-related marks in at least one of the chosen slices in subsampling. Some others have used a weakly supervised approach to dynamically guess the slices related to the disease in each round of training [52]. Moreover, for the 2D and hybrid models, an additional post-processing step must be considered to reduce the false positives. Many have used Markov models to extract more reasonable probabilities for slices based on the probabilities of the peripheral slices.

To train the models for detecting infections and diagnosing COVID-19, pixel-level labels are required. This takes even more effort from radiologists for annotations. Researchers have used fully supervised, semi-supervised, and weakly supervised schemes for training the model based on slices of CT-scan images. There have also

been other weakly supervised schemes that use the interpretations of the model to perform sample-level or slice-level predictions for finding the input areas that affect the output. They use these areas as ground truth and train the infection detection model by utilizing them [below]. In detecting infection, fully supervised and semi-supervised methods can lead to more accurate results. In contrast, the weakly supervised methods would only detect approximate zones due to many pixels and the inexact labels used in training.

4   AI and Pandemic

This section addresses other areas where AI can help speed up processes in managing the pandemic.

4.1   AI for Status Prediction

Two different approaches have been taken to study COVID-19 mortality; predicting large-scale mortality and predicting the death of each individual according to its condition. The first approach is commonly used to predict mortality in a city or country. It uses the distribution of mortality rates in recent days and weeks to predict the near future. To achieve higher accuracy, parameters such as hospital facilities, human mobility, non-pharmaceutical interventions, demographics, historical air quality, and econometrics in the area can also be considered [58].

The second approach predicts the probability of death for each individual. It uses different models like Neural Networks, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, and Decision Tree to obtain the highest accuracy. For example, in our study, the neural network model achieved 89.98% accuracy in predicting the mortality of COVID-19 patients.

In addition, in [59] they trained an AI model on a large dataset of hospitals in Ontario, with nearly 70 thousand patients. Data were collected from the Ontario Health Services (PHO) and Canada Health Services (PHAC) datasets. The model achieved an Area Under the Curve (AUC) of 90% for PHO and 93.5% for PHAC datasets.

4.2   Utilization of AI in Vaccine Discovery

To combat the COVID-19 pandemic, vaccination is the most effective strategy. Most SARS–CoV-2 vaccines in advanced clinical trials are based on modern vaccine approaches that rely on introducing the specific parts of the virus or their genes into the body to generate a targeted immune response. Thus, current methods

have shifted away from live-attenuated and inactivated whole-pathogen vaccines to purified antigens and epitopes.

Vaccine development is complex and often requires extensive, time-consuming, and resource-intensive studies to determine efficacy and potential side effects. AI can help speed up the lengthy and costly process of vaccine development.

The vaccine design process has been revolutionized by reverse vaccinology, which focuses on finding potential vaccine candidates by analyzing pathogens' protein-coding genome (proteome). The SARS–CoV-2 consists of four structural proteins, E (envelope), M (membrane), N (nucleocapsid), and S (spike), as well as several non-structural proteins. The AI approaches can facilitate antigen selection, epitope prediction, immune response modeling, and affinity with human leukocyte antigen alleles in the vicinity of COVID-19 to select the best possible ones. Numerous epitope prediction studies focused on this protein since the S protein is involved in viral entry and provokes an immune response. BNT162b2, mRNA-1273, and AZD1222 are three recently approved vaccines against SARS–CoV-2, all of which use the S protein. Thus, AI approaches can identify specific epitopes from a significant number of potential SARS SARS–CoV-2 peptides capable of inducing a robust and protective immune response.

Another critical problem that could be solved by an AI-based method in predicting the immunogenicity of the developed vaccine. In search for SARS–CoV-2 proteins associated with optimal immune response, computational biology can identify gene coding proteins associated with COVID-19 severity. In addition, a cellular immune response network can be constructed using host-virus and virus-host interaction data.

One of the significant challenges in vaccine development lies in the mutations of SARS–CoV-2 strains. Therefore, there is a need to test whether a selected epitope is conserved across mutations and multiple populations for vaccine development. Despite the development of several machine learning-based classifiers for allergenicity and toxicology, there is currently no method for predicting the toxicity of all vaccine components in combination, which computational network analysis could achieve [60]. Furthermore, breeding SARS–CoV-2 variants resistant to the approved vaccines is not impossible. Therefore, more robust and precise AI-in-silico approaches should be developed to a better vaccine against the new variants of the virus.

4.3 AI in Controlling the Pandemic

COVID-19, as a global health crisis, forced the health care providers to seek new technologies to monitor and control the spread of the pandemic. The extraordinary amount of data derived from public health surveillance, real-time epidemic outbreak monitoring, trend forecasting, regular situation briefings, and medical records must be managed to control and anticipate new diseases.

AI-based methods can track the spread of the virus in real-time, plan public health interventions and monitor their effectiveness. Indeed, the flexibility, rapid analysis

Table 4 Possible applications of artificial intelligence and big data for the management of the COVID-19 outbreak

Time scale	Possible application	Example
Short-term (weeks)	Rapid identification of an ongoing outbreak	AI can facilitate real-time epidemiological data collection, risk-assessment, decision-making processes, and design/implementation of public health interventions
	Diagnosis and prognosis of COVID-19 cases	Recognition of specific diagnostic and prognostic features
Medium-term (months)	Identification of a potential therapeutic option	Identify already existing drugs/discovering new molecules
Long-term (decades)	Enhancing cities and favoring the development of healthy, intelligent, resilient cities	Design newly standardized protocols for sharing data and information during emergencies

and identification of patterns, ability to adapt based on a new understanding of the disease process, self-improvement as new data become available, and lack of human bias in the analysis make AI a promising new tool for pandemic management. Table 4 indicates some possible applications of AI in the control and management of the COVID-19 pandemic.

Information obtained from patient tracking plays a vital role for general public health governance in designing, planning, and organizing to cope with the pandemic [15]. Researchers in [18] have listed 36 countries that have successfully implemented mobile app patient tracking systems. There are several ways to achieve this goal as follows.

4.3.1   Patient Tracking Using Mobile Apps

A successful example is the QR-code-based screening app used in Hubei, China, to monitor people's movement. A similar approach has been used in Taiwan to track high-risk individuals based on their travel history to affected areas [61, 62]. Tracking information from Internet-based searches can also help to predict future outbreaks [61]. For example, using WeChat text data in the context of COVID-19 was another successful approach to predicting disease outbreaks in China [63]. In addition, by searching for "fever" and "cough" in Google Trends, the researchers discovered that these words had a significant association with the COVID-19 outbreak and subsequent hospitalizations or deaths [22].

4.3.2 Patient Tracking Using Video Surveillance

Using video surveillance to detect proximity and social distance has advantages over other approaches such as Bluetooth or GPS technologies. However, the methods mentioned above lead to a high rate of false positives due to their low spatial resolution. Other studies show deep-learning approaches for analyzing CCTV cameras in the workplace to monitor workers' activities and detect violations [64, 65]. In another study, a facial touch detection system (that detects when someone unconsciously touches their face) was also developed to ensure security and collect more data that can be used in related applications [66].

4.3.3 Patient Tracking Using Natural Language Processing (NLP)

An important application of machine learning in patient tracking is determining public opinion and society's perception of social distancing [67]. For example, the results of a text classification study showed that public opinion plays a vital role in guiding relevant decisions (sentiment analysis to ensure that they have been well taught) [14].

There are also challenges with current mobile apps [14]. For example, technical limitations include a lack of highly skilled developers and companies developing and deploying the tracking system. Different countries like the US, UK, and Australia have come to other solutions to overcome this problem [18].

Unfortunately, there is not enough quantitative information about the contact tracing apps used in the above countries to compare their performance and other features comprehensively.

However, in [68], the authors mentioned a COVID-19 Tracing App Scale (COVIDTAS) framework to compare these apps. COVIDTAS was adopted from a framework in [69] and developed based on features such as usability, technology, privacy, tracking effectiveness, and factors reflecting user experience and sentiment.

4.4 Wearable Sensors Application in COVID-19 Pandemic

Other trending approaches that are progressively developing are COVID-19 wearable devices. Despite the classic and conventional techniques used in clinical settings like PCR tests or imaging modalities, these approaches have continuous access to health records of potential COVID-19 subjects all day long. At present, devices like smartwatches or wristbands are recording health data used for screening purposes. Moreover, wearable devices like smart lenses, smart on-teeth sensors, smart masks, and smart biosensors are gaining more attention. However, possible precautions of these technologies, such as data privacy, must be considered [70, 71].

More recently, observational studies on wearable biosensors for remote monitoring of COVID-19 subjects by AI algorithms have revealed promising performance

in detecting COVID-19 patients. Un and his colleagues designed their observational study to show the potentials of wearable biosensors and AI in clinical monitoring. They illustrated that wearable biosensors with AI reached a high correlation with manual procedures in predicting clinical worsening events, as well as prolonged hospitalization [72].

Additionally, there are other related works in literature, including newly-developed wearable devices and state-of-the-art AI algorithms that can predict the potential outcomes of disease. As another example, a wireless skin-interfaced device attaching to the suprasternal notch designed by FitBit can sense multiple features like body movement, heart rate, respiratory-related signals, and other signs and symptoms like body temperature and cough. Therefore, by an ongoing fusion of these features by utilizing AI techniques, one can classify or predict diseases (e.g., being COVID-19 or not) more accurately [73].

5   Future Directions

5.1   The COVID-19 Pandemic Experience

Since the COVID-19 outbreak in late 2019, the disease has become a potential threat to global health. Facing a pandemic is cross-sectoral. These sectors, including economy, social, cultural, environmental, and political, are collectively called social determinants of health.

In a 2000 report, the World Health Organization (WHO) outlined the primary role of health systems in achieving their goals through a set of six fundamental building blocks, including service delivery, health workforce, health information systems, medical products, vaccines and technology, health financing, and governance. Information and AI technologies that serve these six principles can increase the resilience of health systems.

The strategy of countries on health systems plays a crucial role in controlling pandemics. COVID-19 pandemic showed a fundamental weakness in international organizations and governments to face a pandemic. For example, WHO made basic mistakes in managing the COVID-19 pandemic. Without sufficient information, it declared the virus is non-communicable and considered the use of the mask unnecessary for a long time. A year and a half after the pandemic, COVAX vaccine distribution policies cannot provide comprehensive vaccination plans. In the COVID-19 pandemic, medical advances in specialized fields did not help significantly. Unfortunately, at the beginning of this crisis, due to incorrect policy-making, the influx of hospitals and emergency rooms caused the spread of the virus and increased mortality.

Since it is time-consuming to find effective vaccines and treatments at the beginning of a pandemic, the most important global action is identifying and tracking patients and implementing smart distancing between citizens. The experience and

evidence of successful countries in the field of COVID-19 pandemic management in the world have shown that most of the successful management activities are based on the correct and timely application of AI solutions. The first step in managing the information of any epidemic is to identify and record the information of the target population. Accurate identification and registration of patients in different disease states are possible through the following:

1. Identifying the exact number of cases and the actual geographical prevalence,
2. Preparation of contact tracking map of infected people and carriers and the possibility of predicting the future pattern of outbreaks,
3. Ability to control the movement of the population for the best type of social/physical distancing policy, and
4. Anticipating the needs and resources of care and treatment, focusing on efficient distribution of resources.

The following steps are suggested for implementing a control and management system:

1. Implementing a national pandemic web-based system,
2. Identifying the status of each individual; without symptoms, suspicious symptoms, definitive infected, and convalescence categories,
3. Launching a status inquiry system
4. Requiring citizens to carry pandemic IDs,
5. Electronic screening process through registering geographical position and an AI algorithm to compute the risk of being infected by being near the patients and suspected people,
6. Tracking patients through their phones and GPS,
7. Establishing an AI-based face tracking system for patients to prevent them from entering crowded areas,
8. Developing AI-based screening systems that recognized patients through their temperature or other characteristics,
9. Developing AI-based chatbots, and
10. Preparing a telemedicine system to checkup patients in remote locations.

We should adopt new data collection and analysis strategies using emerging technologies. For example, the Internet of Things (IoT) refers to the interconnected network of physical objects such as sensors, health measuring devices, intelligent sensors, home appliances, automotive devices, etc. IoT enables objects to sense, process, and communicate with each other and automatically interact with people and provide intelligent service to users. The IoT platforms can also be used over cloud computing platforms to provide systematic and intelligent prevention and control of COVID-19, which includes five steps: symptom detection, quarantine monitoring, disease contact detection, and social distancing, disease prognosis, and disease mutation tracking. If IoT, cloud, and AI are appropriately utilized, they can provide rapid and efficient healthcare services, especially in the perspective of COVID-19.

5.2   Toward a Universal Crowd-Sourcing and Validating Framework for AI Models

The AI methods have declared they can come in handy during the pandemics, although they have not shown a significant impact on the case of COVID-19 [74]. It was expected to have the AI methods emerge before the first COVID-19 peak in diagnosis and screening, which means scientists had few months to train sophisticated models. As mentioned before, having a large dataset with a wide variety is crucial to have generalizable models, while there was no public dataset in the said period for the COVID-19 pandemic. This led scientists to search the literature, hoping to find sufficient data, and another group searched among the care centers looking for data. However, sufficient positive cases cannot be located in one medical center in the emergence of a pandemic. Besides, there is numerous paperwork to satisfy the privacy of data, which takes even more time. In other words, much time would be spent on gathering the required data while it can be spent on designing and training a suitable model. These problems can be solved if there is a universal crowd-sourcing public dataset. There might be a lack of samples in a small area. However, there is undoubtedly enough data all over the world. Many people wish to donate their samples and contribute to making a universal public dataset to help scientists battling against the pandemic. This framework should have strict confirmation policies to ensure data validity and trustability. Other scientists like physicians and experts can also use this worldwide public dataset to study the disease comprehensively.

Unfortunately, the hotness of a pandemic causes a paper storm. In this situation, it is difficult and time-consuming to find state-of-the-art methods. The review papers ease this process by summarizing many studies. Nevertheless, there is a trade-off between the coverage of ideas and faster release. For example, in screening and diagnosing COVID-19 using medical images, there are more than 2000 papers, howbeit no review paper has covered more than a few methods. Another problem is that not many experts trust the papers in the arXiv, and it takes a long time to be published in a peer-reviewed journal or conference. These problems can be solved by having a universal public framework for sharing the data, papers, and results in a structured way. For example, in the case of diagnosing COVID-19, it could be a tabled data containing the date, the number of training samples, the number of test samples, classification groups, evaluation metrics on the private test set and also on public datasets, and some extra tags for the general methodologies like fully supervised, semi-supervised, weakly supervised, or unsupervised. Having tabled data like this could make searching the literature much easier and faster by utilizing AI-based filters and sorting tools. There are websites like https://paperswithcode.com/ that have aimed at a similar goal by introducing datasets and grouping the studies.

The experience of the COVID-19 pandemic showed that the potentials of AI methods did not widely explore to combat the pandemics. One reason is that people cannot easily trust the reported results. The results cannot be trusted unless one can trust the data and run the proposed trained model on the same data. Some organizations like Kaggle and Dream Challenge aim to solve global problems by challenging

and validating the trained models with private test datasets. During a pandemic, the challenges can help to a great extend. Unfortunately, in the case of diagnosing COVID-19 using CT-scan images, there were no such challenges.

In conclusion, in the case of the COVID-19 pandemic, the absence of the proposed framework is evident. As a result, there was a delay in the emergence of sophisticated and trustable AI models, duplicate ideas on different private datasets, and no fair comparison. This happened while the studies could have been complementary and helped in the evolution of working solutions. The recent experience proved that the world had not been prepared for a pandemic. Efficient adoption of emerging technologies such as IoT, cloud, and AI can significantly help control and manage pandemics.

6   Conclusion

This chapter briefly introduced AI, its strong potentials, and its capability of making manual procedures faster and more accurate during pandemics. Considering its benefits, AI has a high potential to help in a pandemic where hospitals are overloaded, and health experts cannot respond on time. In recent years, AI models have shown outstanding performance in many health applications, especially during pandemics. Therefore, it is essential for the health staff who work with AI systems to know these models and how they operate. So, they neither overestimate nor underestimate them. We demonstrated how AI models are trained. We set up an experiment on actual data to show that models cannot be trained perfectly on any data. A more precisely labeled dataset can lead to much better training. We also showed that not every model that has reported high performance could be confidently trusted. Moreover, we set up an extreme experiment to show how a model can become biased to dataset-specific features, resulting in increased performance in the dataset and low performance in other general datasets. In this regard, we introduced the concept of explainability of AI models. We showed how explaining decisions could help understand and trust the model's findings, which can lead to understanding whether the model's decision on a single case is reliable and if the model is performing rationally in general or has become biased. When a model is biased to unrelated features, it will not behave as expected. In conclusion, the models need to be evaluated on a diverse set of test data representing the actual population's distribution so that the experts can trust the reported performance metrics.

We described the general applications of AI in pandemics and the corresponding SOTA studies. We focused more deeply on the role of AI models in diagnosis and screening as two essential requirements of breaking out a pandemic. Several methods of diagnosis and screening were used during the COVID-19 pandemic. As stated before, evaluating the models on large test sets can indicate their similar performance on the general population and, therefore, their applicability on large-scale situations. Consequently, we selected SOTA methods acquiring the desired quality. Due to the high overlap between the methods, we described the general methodologies and mentioned differences rather than the extensive description of those methods.

Finally, we demonstrated experiences during the COVID-19 pandemic that avoided exploiting the AI's full potentials. We stated the lack of valuable data and the tremendous required effort to gather datasets are the major delaying factors in delivering effective AI systems for the pandemic's peak time. We also described the effect of the so-called paper storm as another delaying factor. The literature review shows that most studies have reported similar performance on different private datasets. However, their performance on general datasets is unknown. Therefore, creating an open dataset could facilitate the delivery of effective AI solutions in a pandemic. We reported the lack of trust towards the reported performance metrics as a prohibitory reason for using AI models by experts. To solve the problems mentioned above, we proposed a unified framework for gathering data, effectively managing the paper storm, and trustfully evaluating the models that can be more effective in the next pandemic.

References

1. Zhang, K., Liu, X., Shen, J., Li, Z., Sang, Y., Wu, X., Zha, Y., Liang, W., Wang, C., Wang, K., Ye, L., Gao, M., Zhou, Z., Li, L., Wang, J., Yang, Z., Cai, H., Xu, J., Yang, L., Cai, W., Xu, W., Wu, S., Zhang, W., Jiang, S., Zheng, L., Zhang, X., Wang, L., Lu, L., Li, J., Yin, H., Wang, W., Li, O., Zhang, C., Liang, L., Wu, T., Deng, R., Wei, K., Zhou, Y., Chen, T., Lau, J.Y.-N., Fok, M., He, J., Lin, T., Li, W., Wang, G.: Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. Cell **181**(6), 1423–1433.e11 (2020). https://doi.org/10.1016/j.cell.2020.04.045
2. Huang, S., Yang, J., Fong, S., Zhao, Q.: Artificial intelligence in the diagnosis of COVID-19: challenges and perspectives. Int. J. Biol. Sci. **17**(6) 1581–1587 (2021). https://doi.org/10.7150/ijbs.58855
3. Copeland, B.J.: Artificial intelligence. In: Encyclopedia Britannica. Available via DIALOG., 2020. https://www.britannica.com/technology/artificial-intelligence. Accessed 21 April 2021
4. Kurfess, F.J.: Artificial intelligence. In: Encyclopedia of Physical Science and Technology, pp. 609–629. Elsevier (2003). https://doi.org/10.1016/b0-12-2274105/00027-2
5. McCall, B.: COVID-19 and artificial intelligence: protecting health-care workers and curbing the spread. Lancet Digit. Health **2**(4), e166–e167 (2020). https://doi.org/10.1016/s2589-7500(20)30054-6
6. Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., Deng, L., Zheng, C., Zhou, J., Shi, H., Feng, J.: Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. Nat. Commun. **11**(1) (2020). https://doi.org/10.1038/s41467-020-18685-1
7. Mei, X., Lee, H.-C., yue Diao, K., Huang, M., Lin, B., Liu, C., Xie, Z., Ma, Y., Robson, P.M., Chung, L., Bernheim, A., Mani, V., Calcagno, C., Li, K., Li, S., Shan, H., Lv, J., Zhao, T., Xia, J., Long, Q., Steinberger, S., Jacobi, A., Deyer, T., Luksza, M., Liu, F., Little, B.P., Fayad, Z.A., Yang, Y.: Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat. Med. **26**(8), 1224–1228 (2020). https://doi.org/10.1038/s41591-020-0931-3
8. Cerna, A.E.U., Jing, L., Good, C.W., vanMaanen, D.P., Raghunath, S., Suever, J.D., Nevius, C.D., Wehner, G.J., Hartzel, D.N., Leader, J.B., Alsaid, A., Patel, A.A., Kirchner, H.L., Pfeifer,

J.M., Carry, B.J., Pattichis, M.S., Haggerty, C.M., Fornwalt, B.K.: Deep-learning-assisted analysis of echocardiographic videos improves predictions of all-cause mortality. Nat. Biomed. Eng. **5**(6), 546–554 (2021). https://doi.org/10.1038/s41551-020-00667-9

9. Ko, H., Chung, H., Kang, W.S., Park, C., Kim, D.W., Kim, S.E., Chung, C.R., Ko, R.E., Lee, H., Seo, J.H., Choi, T.-Y., Jaimes, R., Kim, K.W., Lee, J.: An artificial intelligence model to predict the mortality of COVID-19 patients at hospital admission time using routine blood samples: Development and validation of an ensemble model. J. Med. Internet Res. **22**(12), e25442 (2020). https://doi.org/10.2196/25442

10. Wang, Y., Wang, Y., Chen, Y., Qin, Q.: Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control 21 measures. J. Med. Virol. **92**(6), 568–576 (2020). https://doi.org/10.1002/jmv.25748

11. Gudadhe, A., Genani, G.: Wearable health monitoring systems. The Industrial Design Engineering Wiki (2021). http://wikid.io.tudelft.nl/WikID/index.php/Wearable_Health_Monitoring_Systems. Accessed 15 June 2021

12. W.W.H. Organization: Health topic, coronavirus disease overview (2020). https://www.who.int/health-topics/coronavirus#tab=tab_1. Accessed 29 May 2020

13. Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., Fraser, C.: Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science **368**(6491), eabb6936 (2020). https://doi.org/10.1126/science.abb6936

14. Lalmuanawma, S., Hussain, J., Chhakchhuak, L.: Applications of machine learning and artificial intelligence for covid-19 (SARS-CoV-2) pandemic: a review. Chaos, Solitons Fractals **139**, 110059 (2020). https://doi.org/10.1016/j.chaos.2020.110059

15. Golinelli, D., Boetto, E., Carullo, G., Nuzzolese, A.G., Landini, M.P., Fantini, M.P.: Adoption of digital technologies in health care during the COVID-19 pandemic: Systematic review of early scientific literature. J. Med. Internet Res. **22**(11), e22280 (2020). https://doi.org/10.2196/22280

16. Saeedvand, S., Jafari, M., Aghdasi, H.S., Baltes, J., Rahmani, A.M.: Deep Learning: A Taxonomy of Modern Weapon to Combat Covid-19 Pandemic (2020)

17. Mbunge, E.: Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls. Diabetes & Metab.Ic Syndr.: Clin. Res. Rev. **14**(6), 1631–1636 (2020). https://doi.org/10.1016/j.dsx.2020.08.029

18. Health, T.L.D.: Pandemic versus pandemonium: fighting on two fronts. Lancet Digit. Health **2**(6), e268 (2020). https://doi.org/10.1016/s2589-7500(20)30113-8

19. Das, D., Santosh, K.C., Pal, U.: Truncated inception net: COVID-19 outbreak screening using chest x-rays. Phys. Eng. Sci. Med. **43**(3), 915–925 (2020). https://doi.org/10.1007/s13246-020-00888-x

20. Basu, S., Mitra, S., Saha, N.: Deep learning for screening covid-19 using chest x-ray images. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 2521–2527. IEEE (2020)

21. Li, X., Li, C., Zhu, D.: Covid-mobilexpert: on-device covid-19 screening using snapshots of chest x-ray (2020). arXiv preprint arXiv:2004.03042

22. Punn, N.S., Sonbhadra, S.K., Agarwal, S., Rai, G.: Monitoring covid-19 social distancing with person detection and tracking via fine-tuned Yolo v3 and deepsort techniques (2020). arXiv preprint arXiv:2005.01385

23. Nayak, S.R., Nayak, D.R., Sinha, U., Arora, V., Pachori, R.B.: Application of deep learning techniques for detection of covid-19 cases using chest x-ray images: a comprehensive study. Biomed. Signal Process. Control **64**, 102365 (2021)

24. Fields, B.K., Demirjian, N.L., Dadgar, H., Gholamrezanezhad, A.: Imaging of covid-19: Ct, mri, and pet. In: Seminars in Nuclear Medicine. Elsevier (2020)

25. Yang, S., Zhang, Y., Shen, J., Dai, Y., Ling, Y., Lu, H., Zhang, R., Ding, X., Qi, H., Shi, Y., et al.: Clinical potential of ute-mri for assessing covid-19: patient-and lesion-based comparative analysis. J. Magn. Reson. Imaging **52**(2), 397–406 (2020)

26. Langenbach, M.C., Hokamp, N.G., Persigehl, T., Bratke, G.: Mri appearance of covid-19 infection. Diagn. Interv. Radiol. **26**(4), 377 (2020)
27. Abdelmohsen, M.A., Alkandari, B.M., Gupta, V.K., ElBeheiry, A.A.: Diagnostic value of abdominal sonography in confirmed covid-19 intensive care patients. Egypt. J. Radiol. Nucl. Med. **51**(1), 1–7 (2020)
28. Evans, K.D., Yang, Q., Liu, Y., Ye, R., Peng, C.: Sonography of the lungs: diagnosis and surveillance of patients with covid-19. J. Diagn. Med. Sonogr. **36**(4), 370–376 (2020)
29. Kaufman, A.E., Naidu, S., Ramachandran, S., Kaufman, D.S., Fayad, Z.A., Mani, V.: Review of radiographic findings in covid-19. World J. Radiol. **12**(8), 142 (2020)
30. Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., Xue, Z., Liao, H., Chen, F., Yang, F., et al.: The role of imaging in the detection and management of covid-19: a review. IEEE Rev. Biomed. Eng. **14**, 16–29 (2020)
31. Al-Tawfiq, J.A., Memish, Z.A.: Diagnosis of sars-cov-2 infection based on ct scan vs rt-pcr: reflecting on experience from mers-cov. J. Hosp. Infect. **105**(2), 154–155 (2020)
32. Poortahmasebi, V., Zandi, M., Soltani, S., Jazayeri, S.M.: Clinical performance of rt-pcr and chest ct scan for covid-19 diagnosis; a systematic review. Adv. J. Emerg. Med. **4**(2s), e57–e57 (2020)
33. Gietema, H.A., Zelis, N., Nobel, J.M., Lambriks, L.J., Van Alphen, L.B., Oude Lashof, A.M., Wildberger, J.E., Nelissen, I.C., Stassen, P.M.: Ct in relation to Rt-PCR in diagnosing covid-19 in the Netherlands: a prospective study. PloS one **15**(7), e0235844 (2020)
34. Chowdhury, E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z.B., Islam, K.R., Khan, M.S., Iqbal, A., Al Emadi, N., et al., Can AI help in screening viral and covid-19 pneumonia? IEEE Access **8**, 132665–132676 (2020)
35. Castiglioni, I., Ippolito, D., Interlenghi, M., Monti, C.B., Salvatore, C., Schiaffino, S., Polidori, A., Gandola, D., Messa, C., Sardanelli, F.: Artificial intelligence applied on chest x-ray can aid in the diagnosis of covid-19 infection: a first experience from Lombardy, Italy. MedRxiv (2020)
36. Mangal, A., Kalia, S., Rajgopal, H., Rangarajan, K., Namboodiri, V., Banerjee, S., Arora, C.: Covidaid: Covid-19 detection using chest X-ray (2020). arXiv preprint arXiv:2004.09803
37. Borkowski, A.A., Viswanadhan, N.A., Thomas, L.B., Guzman, R.D., Deland, L.A., Mastorides, S.M.: Using artificial intelligence for covid-19 chest x-ray diagnosis. Fed. Pract. **37**(9), 398 (2020)
38. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: ChestX-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2017). https://doi.org/10.1109/cvpr.2017.369
39. Pneumonia detection challenge database. RSNA: Radiological Society of North America (2018). https://www.kaggle.com/c/rsnapneumonia-detection-challenge/. Accessed 10 April 2021
40. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2016). https://doi.org/10.1109/cvpr.2016.308
41. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2016). https://doi.org/10.1109/cvpr.2016.90
42. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization (2014). arXiv preprint arXiv:1412.6980
43. Marcel, S., Rodriguez, Y.: Torchvision the machine-vision package of torch. In: Proceedings of the international conference on Multimedia—MM '10. ACM Press (2010). https://doi.org/10.1145/1873951.1874254
44. Chowdhury, M.E.H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z.B., Islam, K.R., Khan, M.S., Iqbal, A., Emadi, N.A., Reaz, M.B.I., Islam, M.T.: Can AI help in screening viral and COVID-19 pneumonia? IEEE Access **8**, 132665–132676 (2020). https://doi.org/10.1109/access.2020.3010287

45. Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M.K., Pei, J., Ting, M.Y., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., Shi, A., Zhang, R., Zheng, L., Hou, R., Shi, W., Fu, X., Duan, Y., Huu, V.A., Wen, C., Zhang, E.D., Zhang, C.L., Li, O., Wang, X., Singer, M.A., Sun, X., Xu, J., Tafreshi, A., Lewis, M. A., Xia, H., Zhang, K.: Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell **172**(5), 1122–1131.e9 (2018). https://doi.org/10.1016/j.cell.2018.02.010

46. Nguyen, H.Q., Lam, K., Le, L.T., Pham, H.H., Tran, D.Q., Nguyen, D.B., Le, D.D., Pham, C.M., Tong, H.T., Dinh, D.H., et al.: Vindr-cxr: An open dataset of chest x-rays with radiologist's annotations (2020). arXiv preprint arXiv:2012.15029. https://www.kaggle.com/c/vinbigdata-chest-xray-abnormalities-detection/. Accessed 20 March 2021

47. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A survey of methods for explaining black box models. ACM Comput. Surv. **51**(5), 1–42 (2019). https://doi.org/10.1145/3236009

48. Gilpin, L.H., Bau, D., Yuan, B.Z., Bajwa, A., Specter, M., Kagal, L.: Explaining explanations: An overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE (2018). https://doi.org/10.1109/dsaa.2018.00018

49. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE (2017). https://doi.org/10.1109/iccv.2017.74

50. Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., Cao, K., Liu, D., Wang, G., Xu, Q., Fang, X., Zhang, S., Xia, J., Xia, J.: Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. Radiology **296**(2), E65–E71 (2020). https://doi.org/10.1148/radiol.2020200905

51. Gunraj, H., Sabri, A., Koff, D., Wong, A.: Covid-net ct-2: Enhanced deep neural networks for detection of covid-19 from chest ct images through bigger, more diverse learning. arXiv preprint arXiv:2101.07433 (2021)

52. Modegh, R.G., Hamidi, M., Masoudian, S., Mohseni, A., Lotfalinezhad, H., Kazemi, M.A., Moradi, B., Ghafoori, M., Motamedi, O., Pournik, O., et al.: Accurate and rapid diagnosis of covid-19 pneumonia with batch effect removal of chest ct-scans and interpretable artificial intelligence (2020). arXiv preprint arXiv:2011.11736

53. Harmon, S.A., Sanford, T.H., Xu, S., Turkbey, E.B., Roth, H., Xu, Z., Yang, D., Myronenko, A., Anderson, V., Amalou, A., Blain, M., Kassin, M., Long, D., Varble, N., Walker, S.M., Bagci, U., Ierardi, A.M., Stellato, E., Plensich, G.G., Franceschelli, G., Girlando, C., Irmici, G., Labella, D., Hammoud, D., Malayeri, A., Jones, E., Summers, R.M., Choyke, P.L., Xu, D., Flores, M., Tamura, K., Obinata, H., Mori, H., Patella, F., Cariati, M., Carrafiello, G., An, P., Wood, B.J., Turkbey, B.: Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat. Commun. **11**(1) (2020). https://doi.org/10.1038/s41467-020-17971-2

54. Wang, X., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., Liu, W., Zheng, C.: A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. IEEE Trans. Med. Imaging **39**(8), 2615–2625 (2020). https://doi.org/10.1109/tmi.2020.2995965

55. Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Ni, Q., Chen, Y., Su, J., Lang, G., Li, Y., Zhao, H., Liu. J., Xu, K., Ruan, L., Sheng, J., Qiu, Y., Wu, W., Liang, T., Li, L.: A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering **6**(10), 1122–1129 (2020). https://doi.org/10.1016/j.eng.2020.04.010

56. X. Li, Y. Zhou, P. Du, G. Lang, M. Xu, and W. Wu, "A deep learning system that generates quantitative CT reports for diagnosing pulmonary tuberculosis," *Applied Intelligence*, vol. 51, no. 6, pp. 4082–4093, 2021. [Online]. Available: https://doi.org/10.1007/s10489-020-02051-1

57. Gunraj, H., Wang, L., Wong, A.: COVIDNet-CT: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest CT images. Frontiers in Medicine **7** (2020). https://doi.org/10.3389/fmed.2020.608525

58. Arik, S.O., Li, C.-L., Yoon, J., Sinha, R., Epshteyn, A., Le, L.T., Menon, V., Singh, S., Zhang, L., Yoder, N., et al.: Interpretable sequence learning for covid-19 forecasting (2020). arXiv preprint arXiv:2008.00646

59. Snider, B., Phillips, P., MacLean, A., McBean, E., Gadsden, S.A., Yawney, J.: Artificial intelligence to predict the risk of mortality from covid-19: Insights from a Canadian application (2020). https://doi.org/10.1101/2020.09.29.20201632

60. Hwang, W., Lei, W., Katritsis, N.M., MacMahon, M., Chapman, K., Han, N.: Current and prospective computational approaches and challenges for developing COVID-19 vaccines. Adv. Drug Deliv. Rev. **172**, 249–274 (2021). https://doi.org/10.1016/j.addr.2021.02.004

61. Wang, C.J., Ng, C.Y., Brook, R.H., Response to COVID-19 in Taiwan. JAMA **323**(14) 1341 (2020). https://doi.org/10.1001/jama.2020.3151

62. Wang, W., Wang, Y., Zhang, X., Li, Y., Jia, X., Dang, S.: WeChat, a Chinese social media, may early detect the SARS-CoV-2 outbreak in 2019 (2020). https://doi.org/10.1101/2020.02.24.20026682

63. Ciaffi, J., Meliconi, R., Landini, M.P., Ursini, F.: Google trends and COVID-19 in Italy: could we brace for impact? Intern. Emerg. Med. **15**(8), 1555–1559 (2020). https://doi.org/10.1007/s11739-020-02371-7

64. Beyan, C., Bustreo, M., Shahid, M., Bailo, G.L., Carissimi, N., Bue, A.D.: Analysis of face-touching behavior in large scale social interaction dataset. In: Proceedings of the 2020 International Conference on Multimodal Interaction. ACM (2020). https://doi.org/10.1145/3382507.3418876

65. Raamkumar, A.S., Tan, S.G., Wee, H.L.: Use of health belief model-based deep learning classifiers for COVID-19 social media content to examine public perceptions of physical distancing: model development and case study. JMIR Public Health Surveill. **6**(3), e20493 (2020). https://doi.org/10.2196/20493

66. Jelodar, H., Wang, Y., Orji, R., Huang, S.: Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. IEEE J. Biomed. Health Inform. **24**(10), 2733–2742 (2020)

67. Bhattacharya, S., Maddikunta, P.K.R., Pham, Q.-V., Gadekallu, T.R., S.R.K.S., Chowdhary, C.L., Alazab. M., Piran, M.J.: Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey. Sustain. Cities Soc. **65**, 102589 (2021). https://doi.org/10.1016/j.scs.2020.102589

68. Raman, R., Achuthan, K., Vinuesa, R., Nedungadi, P.: COVIDTAS COVID-19 tracing app scale—an evaluation framework. Sustainability **13**(5), 2912 (2021). https://doi.org/10.3390/su13052912

69. Vinuesa, R., Theodorou, A., Battaglini, M., Dignum, V.: A socio-technical framework for digital contact tracing. Results Eng. **8**, 100163 (2020). https://doi.org/10.1016/j.rineng.2020.100163

70. Ates, H.C., Yetisen, A.K., Güder, F., Dincer, C.: Wearable devices for the detection of COVID-19. Nat. Electron. **4**(1), 13–14 (2021). https://doi.org/10.1038/s41928-020-00533-1

71. Best, J.: Wearable technology: covid-19 and the rise of remote clinical monitoring. BMJ, p. n413 (2021). https://doi.org/10.1136/bmj.n413

72. Un, K.-C., Wong, C.-K., Lau, Y.-M., Lee, J.C.-Y., Tam, F.C.-C., Lai, W.-H., Lau, Y.-M., Chen, H., Wibowo, S., Zhang, X., Yan, M., Wu, E., Chan, S.-C., Lee, S.-M., Chow, A., Tong, R.C.-F., Majmudar, M.D., Rajput, K.S., Hung, I.F.-N., Siu, C.-W.: Observational study on wearable biosensors and machine learning-based remote monitoring of COVID-19 patients. Scientific Reports **11**(1) (2021). https://doi.org/10.1038/s41598-021-82771-7

73. Jeong, H., Rogers, J.A., Xu, S.: Continuous on-body sensing for the COVID-19 pandemic: gaps and opportunities. Sci. Adv. **6**(36), eabd4794 (2020). https://doi.org/10.1126/sciadv.abd4794

74. Hu, Y., Jacob, J., Parker, G.J., Hawkes, D.J., Hurst, J.R., Stoyanov, D.: The challenges of deploying artificial intelligence models in a rapidly evolving pandemic. Nat. Mach. Intell. **2**(6), 298–300 (2020)