A Survey of the Application of Artifical Intellegence on COVID-19 Diagnosis and Prediction

Hana Alalawi
Computer Science Department
College of Computer and Information Systems
Umm Al-Qura University
Makkah, Saudi Arabia
s44280166@st.uqu.edu.sa

Manal Alsuwat
Computer Science Department
College of Computer and Information Systems
Umm Al-Qura University
Makkah, Saudi Arabia
s44280158@ st.uqu.edu.sa

Hosam Alhakami
Computer Science Department
College of Computer and Information Systems
Umm Al-Qura University
Makkah, Saudi Arabia
hhhakam@uqu.edu.sa

Abstract-The importance of classification algorithms has increased during the recent years. Classification is a branch of supervised learning with the goal of predicting class labels categorical of new cases. Additionally, with Coronavirus (COVID-19) propagation since 2019, the world still faces a great challenge in defeating COVID-19 even with modern methods and technologies. This paper gives an overview of classification algorithms to provide the readers with an understanding of the concept of the state-of-the-art classification algorithms and their applications used in the COVID-19 diagnosis and detection. It also describes some of the research published on classification algorithms, the existing gaps in the research, and future research directions. This article encourages both academics and machine learning learners to further strengthen the basis of classification methods.

Keywords-artificial intelligence; machine learning; deep learning; classification algorithms; COVID-19; medical image introduction

I. INTRODUCTION

Machine learning is one of the fastest growing fields in Artificial Intelligence (AI). Classification algorithms refer to the problem of predictive modeling, in which the class classification is predicted for a given example of input data. Many machine learning algorithms are used. This paper will consider k-Nearest Neighbor (KNN), Support Vector Machine (SVM), Neural Networks (NNs), Decision Tree (DT), Logistic Regression (LR), and Random Forest (RF). Machine learning is divided into 3 main fields based on the learning style: supervised learning, unsupervised learning, or reinforcement learning. This research will present an overview of the classification algorithms and their application in predicting COVID-19 cases, focusing on the strengths and weaknesses of each algorithm. This research surveys the machine learning classification algorithms used to predict and diagnose COVID-19.

II. COVID-19 CASES AND SYMPTOMS

Coronaviruses are a wide class of virus infections that produce illnesses that range from colds to more serious diseases, like the Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). In 2019, in Wuhan, China, a new coronavirus was discovered which since then has evolved to a global pandemic status. To diagnose COVID-19, quickly and efficiently, screening tools are needed. Machine learning algorithms with computer vision systems help diagnose COVID-19. This section highlights the research focused on machine learning and classification algorithms for COVID-19 diagnosis.

A. Symptoms of COVID-19

According to the World Health Organization [1], COVID-19 influences people in various ways. Most people who contract it have mild to moderate symptoms and improve without hospitalization. The common characteristic symptoms include high temperature, dry cough, and exhaustion. The less common symptoms are aches, sore throat, diarrhea, conjunctivitis, headaches, loss of sense of taste or smell, rashes, and discoloration of the fingers or toes. The most dangerous symptoms are breathing difficulties or shortness of breath, the losing the ability to speak or move and, chest pain or pressure.

B. The Way COVID-19 Spreads

Individuals can catch COVID-19 from other people who have the virus. The disease spreads principally by droplets, which are spread by a person with COVID-19 through coughing, sneezing, or talking. If people breathe in these droplets from an infected person, they, too, can become

Corresponding author: Hana Alalawi

infected with COVID-19. Thus, people should keep at least 1m distance from one another.

C. The Precautions to Stop the Spread of COVID-19

Since December 2019, COVID-19 has influenced all countries in the world. As of May 2020, the number of confirmed cases and deaths worldwide has sharply increased [2]. The pandemic has negatively impacted different industries, such as transportation, health services, and tourism. Many countries took quick action to stop the spread of the virus, which included awareness campaigns, social distancing, lockdowns, and quarantines, and transferring all work and schooling online. Virus researchers and doctors began searching for a cure for the virus. COVID-19 is now the most significant cause of mortality in many countries, specifically the United States, Spain, Italy, China, and Iran. Figure 1 presents the most recent number of individuals diagnosed with COVID-19 around the world, while Figure 2 illustrates daily COVID-19 deaths per million people. It can be clearly seen that the number of confirmed cases and the number of deaths fluctuate from time to time.
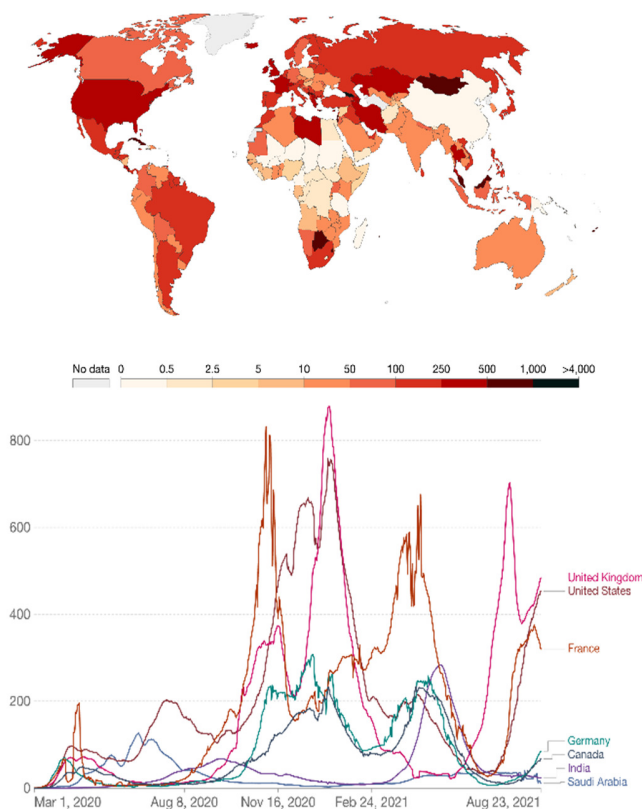
algorithms in terms of real-time processing and prediction. Authors in [2] used a novel dataset with different algorithms. After comparing the results, the boosted RF classifier was identified as a good model for the problem. It achieved an accuracy of 94% and an F1 score of 0.86. Authors in [5] provided a comprehensive survey of various classifications by introducing basic classification algorithms and discussing some of the main classification methods, including Bayesian networks, KNN classification, RF extrapolation, and SVM. They also discussed the problems and potential solutions associated with each. Authors in [6] focused on implementing AI and machine learning in healthcare by providing better solutions with a high rate of accuracy to overcome the challenges presented by COVID-19. They presented a review of the previous studies that applied AI algorithms and techniques to tackle COVID-19, such as support vector regression algorithms, supervised multi-layered classifiers (XGBoost), deep learning LSTM networks, and regression trees. They presented the results associated with the pros and cons for each suggested model.
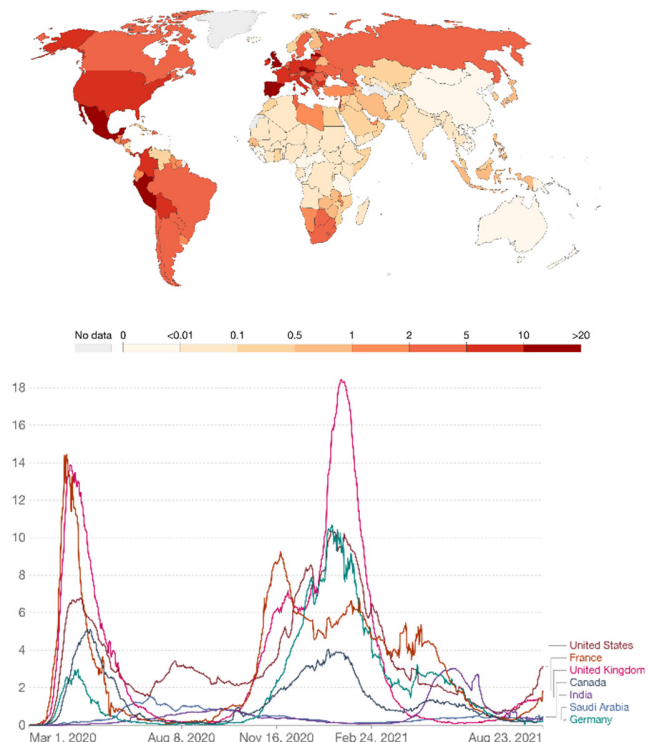

Fig. 1.   Confirmed COVID-19 cases per million. Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/coronavirus-testing [3].

Fig. 2.   Confirmed COVID-19 deaths per million. Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/coronavirus-testing [3].

During the COVID-19 pandemic, in May 2020, authors in [4] provided a pre-tuned boosted RF classification algorithm to help the medical industry overcome the gaps of the traditional healthcare system through applying machine learning

Early detection of COVID-19 is essential. Previous studies have shown that in initial COVID-19 screenings and in all patient-care environments, blood tests are relatively rapid, cheap, and simple to conduct. Authors in [7] analyzed the state-of-the-art COVID-19 detection techniques, and clinical data were provided in order to encourage researchers to develop better disease prediction modeling. Authors in [8] provided a

comprehensive overview of the methods used for COVID-19 detection as well as a compendium of available open-source datasets related to COVID-19. Sixty newspaper articles, reports, factsheets, and websites dealing with COVID-19 were studied and reviewed. Most mathematical models were based on Susceptible-Infected-Recovered (SIR) models and Susceptible-Exposed-Infected-Removed (SEIR) models, while Convolutional Neural Networks (CNNs) was the main AI implementations used with X-ray and CT images. Mathematical modeling and AI have proved to be reliable tools to mitigate this pandemic.

III. RESEARCH METHODOLOGY

For searching and selecting previous studies, the Saudi Digital Library (SDL), IEEE Xplore, ScienceDirect, Google Scholar databases were consulted. These databases are considered appropriate for COVID-19 prediction and cover the newest literature. The research studies retrieved from these databases are relevant and guaranteed to comprehend the purpose of AI systems in COVID-19. The databases were searched in a 2-year period between 2019 and 2021. The following keywords were utilized to direct the process of searching: classification algorithms, classification applications, classification review, machine learning, coronavirus, and COVID-19. Only modern studies and articles containing relevant information about the application of classification techniques concerning COVID-19 were selected and extracted. All publication types that were published in conferences, journals, and review papers were selected. The primary result of the search process identified 47 articles that met the review's eligibility rules, which are the novelty of the paper and the use of AI. After examining the full texts, 40 were selected and 4 articles were excluded because other methods were used in diagnosing COVID-19 instead of AI. It is difficult to present a comprehensive review of all machine learning classification methods in one article; consequently, this survey only focused on commonly used classification techniques.

IV. CLASSIFICATION ALGORITHMS

Classification algorithms are the most widely used techniques in the machine learning field. In general, the models in machine learning are divided into 3 general approaches based on their learning style: supervised learning, unsupervised learning, and reinforcement learning. The classification is a way of learning as the machine learns how to assign the labels to each class of the data. There are several models of classification algorithms, such as LR, DT classifier, SVM, NNs, KNN, and RF. Table I shows a quick comparison of all the mentioned classification algorithms in terms of their advantages and disadvantages.

A. Support Vector Machine

SVM classification has become popular due to its impressive performance, which is comparable to the performance of advanced NNs trained in complex tasks, using high computing cost algorithms. SVMs were originally designed as an effective approach for pattern recognition and classification [9]. Immediately after their introduction, these algorithms were used for some classification problems and applications, such as speech recognition [10], computer vision,

and image processing [11]. The SVM algorithm is primarily used for classification techniques. SVM generates a hyperplane that separates data into various classes. SVM can resolve either linear or non-linear issues. The main objective of the SVM is to discover a hyperplane in N-dimensional space (N matches the number of features), which clearly classifies the datasets or points of data. The accuracy of the outcome is directly related to the selected hyperplane. It discovers a plane with the maximum distance among the data points of two classes. The hyperplane is illustrated graphically as a line that divides one class from another. The data points are on different sides of the hyperplane and are assigned to different classes. The hyperplane dimension is dependent on the feature quantity. If the input is two features, then the hyperplane is a line and a 2D plane if the true of input is three. The SVM is also utilized in medical diagnosis to detect anomalies and in air quality management systems and financial analyses.

B. K-Nearest Neighbor

The KNN algorithm supposes that similar objects are close to one another. This algorithm is one of the most powerful supervised classification algorithms. KNN classification models are often used for searching the nearest neighbor class to predict the target name or label. The KNN is a simple algorithm used to save all available cases and classify new cases according to a measure of similarity, which is measured by a distance function (Manhattan, Euclidean, Minkowski, or Hamming). Nearest neighbor techniques are categorized as structureless KNN techniques and structure-based KNN techniques. Structure-based KNN techniques are based on structures of data, such as the Orthogonal Structure Tree (OST), axis tree, k-d tree, and ball tree. Additionally, data and site patterns suggesting suspicious behavior are analyzed by KNN algorithms.

C. Neural Networks

More accurately Artificial Neural Networks (ANNs). An ANN is a series of neurons associated with synapses that mimic the human brain structure and has several components of processing. The human brain, however, is much more complicated. There are neurons or nodes in a neural network. Each of these neurons accepts data, processes them, and transfers them to a different neuron. The various units are communicating by transmitting signals to each other. An ANN is constructed as a directed graph with nodes and edges that link the nodes. The edges of each node are the interconnections. NNs can essentially be used for any task, from spam filtering to computer vision. They are typically utilized for machine translation, identification of anomalies and risk management, as well as language and face recognition.

D. Logistic Regression

LR is a machine-learning algorithm that predicts the likelihood of the response variables by given a set of explanatory independent variables. It is a supervised binary classification algorithm [12]. The response variables are coded as binary values, 0 and 1, and based on the values of the binary target variables, the data will be classified. The LR types are binary or binomial, multinomial, and ordinal. The values of the target variable in the binary LR have two possible types, 0 and

1, to represent gender (male and female), success or failure etc. Multinomial LR is used when the response variable has an unordered 3 or more possible values to represent the data, such as class types A, B, C, and D. The ordinal LR is the classification in which the response variable has 3 or more ordered values, such as student's scores (high, middle, and low). The LR algorithm has a solid statistical and mathematical background, but it is sensitive to outliers and prone to multicollinearity.

E. Decision Tree

The DT classifier is a supervised machine learning algorithm that builds a tree structure to classify the labeled data based on a condition. The root node is the topmost node in the tree, and the leaf nodes represent the outcomes, where the edges connect the nodes to the leaf nodes.

F. Random Forest

RF is a classification and regression model, however, it is usually used in classification problems. It is a combination of random DTs selected to find the average of the best predictions [13]. To measure the best separates of the data, the RF uses the most voted class approach to select the best solutions [14]. This algorithm is used in different fields in real life applications, such as gene selection, finance analysis, stock market, e-commerce, and medical diagnoses.

TABLE I.   ADVANTAGES AND DISADVANTAGES OF DIFFERENT CLASSIFICATION ALGORITHMS

Models	Advantages	Disadvantages
SVM	- Efficient, better accuracy. - Not biased by outliers. - Non-overfitting sensitive.	- Problems in selecting the right kernel function. - Datasets have different results based on the kernel function. - Not optimal for non-linear situations or problems, not the best option for a huge number of features.
KNN	- Simple to understand. - Fast and efficient.	Sensitive to noise; the k number of neighbors should be picked manually.
NN	- Identifies complex associations between independent and reliable variables efficiently. - Capable of managing noisy data.	- Prone to local minima and overfitting. - ANN processing is complicated to interpret and demands a long processing time.
LR	- Reasonable accuracy for datasets and works perfectly if the data set is separated linearly.	- Cannot solve non-linear problems, since it has a linear decision surface. - In real-world situations, linearly separable data are rarely found. - Sensitive to outliers.
DT	- Easily processing large dimension data Interpretability works for both linear and nonlinear tasks, with no need for scaling or normalizing features.	Poor performance, and overfitting can easily occur with very small datasets.
RF	- Powerful and accurate. - Good performance on a variety of problems.	- No interpretability, overfitting can easily occur. - Need to choose the number of trees manually.

V.   APPLICATIONS OF CLASSIFICATION ALGORITHMS IN COVID-19 DETECTION AND DIAGNOSIS

In coronavirus research, diagnosis is an essential preventive phase because it has a similar appearance to other forms of pneumonia. Consequently, the discovery of COVID-19 is critically important and vital in its initial stages. Different diagnostic methods for COVID-19, including a range of techniques of medical imagery, blood tests (CBCs), and PCR, have been suggested. The World Health Organization states diagnoses of COVID-19 disease have to be checked with reverse transcriptase-polymerase chain reaction (RT-PCR) tests [15]. According to the U.S. Food and Drug Administration [39], there are 2 main types of tests that have been used to diagnose the virus, RT-PCR and imaging methods such as X-ray and CT-scanning [16, 17]. RT-PCR testing, however, takes time and this can be dangerous. For this reason, the initial detection of COVID-19 is often first conducted with medical imaging and the RT-PCR test is then carried out. The latest results from radiological imagery indicate that these images provide valuable details on the COVID-19 virus. The use of advanced AI technology combined with radiological imaging can be useful to accurately diagnose this disease. X-ray imaging has the benefits of being cheap and of low risk for human health [25, 26]. In an X-ray, it is a relatively complex process to detect COVID-19. In these photos, the radiologist should carefully note the very long and problematic spots carrying water and pus. Therefore, illnesses such as pulmonary tuberculosis may be erroneously identified by a radiologist or a doctor as COVID-19. So, the X-ray method has a big error rate. Compute Tomography (CT) images are often utilized to detect the virus more precisely. However, CT images are much more costly for patients than X-rays. Detecting the effect of different respiratory system diseases such as ARDS, streptococcus pneumonia, chlamydia pneumonia, cavitating pneumonia, pneumococcal pneumonia, aspiration pneumonia, and COVID-19 on human lungs by using X-ray imaging techniques can be challenging as these diseases have the same impact on human lungs.

In raw Chest X-Ray (CXR) samples using deep NNs, authors in [18] implemented a personalized network (DarkCovidNet) for the automatic detection of COVID-19. The proposed model provides accurate classifications concerning Binary (COVID vs. No-Findings) and Multi-class (COVID vs. No-Findings vs. Pneumonia) classification with outstanding performance precision, 98.08% for binary- and 87.02% for multi- classification. Multi-classification demonstrated that the expert system can be applied to assist the radiologists in validating the examination process quickly and accurately.

Coronavirus causes different symptoms such as cough, fever, dyspnea, musculoskeletal, and anosmia/dysgeusia compared to the common cold, flu, and influenzas [19]. Authors in [20] proposed a method for binary classification and multi-classification problems by using deep learning CNN networks for COVID-19 detection. The model attains 99.6% accuracy on X-ray images and 71.81% on CT-scan images.

CNNs have been used to classify the images of the patients (COVID, normal, and pneumonia classes). The accuracy of the CNN multiclass model with Adam, sProp, and Sgdm optimizers was 96.4%, 95.4%, and 95.5% respectively. On the other hand, the individual classification accuracy with Adam, RmsProp, and Sgdm optimizers was 97.7%, 94.8%, and 97.8.

Authors in [21] proposed a method that used the Decompose, Transfer, and Compose (DeTraC) technique to predict and classify the chest images of COVID-19 patients. DeTraC deals with image recognition and classification successfully. The authors applied the CNN DeTraC in an X-ray images dataset gathered from many hospitals around the world. The model was trained in 3 steps: first was the extraction of the deep features of the images. Second, the optimization step is done by using Stochastic Gradient Descent (SGD). Finally, the images were classified by using the last layer of the DeTraC network. The model achieved a good performance with 95.12% accuracy, 97.91% sensitivity, and 91.87% specificity. On the other hand, authors in [22] proposed a new deep CNN framework to detect and classify the patient's X-Ray images as positive or negative cases. The proposed framework, COVIDX-Net, includes 7 different CNN layers built by 7 models such as VGG16, InceptionV3, V3, DenseNet, and Google MobileNet to analyze and detect viruses. The COVIDX-Net achieved a good performance in VGG19 and DenseNet with F1-scores of 0.89 and 0.9, respectively. Furthermore, authors in [23] introduced another CNN model to detect the confirmed cases of the coronavirus. Authors in [17] used machine-learning for the classification of COVID-19 and non-COVID lung scan slices with 89.8% classification accuracy. Authors in [24] used chest X-ray images for COVID-19 diagnosis, and their classification accuracy was 99.56% for the disease class. Authors in [27] presented an early COVID-19 detection using SVM. The algorithm has applied to CT of abdominal images. From 150 CT images, 4 various dimensional datasets were generated and 4 different methods were employed for extracting chest CT picture features. Then, SVM was utilized to distinguish COVID-19 patients. During the classification process, 10 times cross-validation was used. The achieved classification accuracy was 99.68%.

In coronaviruses detection and prediction, deep learning with a CNN is one of the most accurate solutions. Diagnosing the infected patients currently is done by either CT-scan or X-ray chest imaging. Authors in [28] used a dataset provided by the Kaggle to develop a COVID-19 detector, the dataset contained Chest X-Ray (CXR) images of COVID-19 infected patients and of uninfected people. The proposed CNN-based COVID-19 detector was trained on augmented data and achieved high accuracy in the prediction of the patients' needs in the next 7 days. In addition, the prediction of the number of deaths, new confirmed cases and recovered cases was 94.18%, 99.94%, and 90.29% respectively. Implementing the VGG16 CNN in the detector enhanced the model's accuracy. Authors in [29] introduced a new open-source network called COVID-Net. COVID-Net is a deep CNN to detect the COVID-19 positive cases by analyzing CXR images. COVIDx is the largest open-source dataset of COVID-19 patient's CSR images that have been applied to train and assess the suggested COVID-Net. The model achieved good test accuracy equal to

93.3%. Authors in [30] introduced a technique of learning pipeline and multi-view representation of COVID-19 classification utilizing different types of features extracted from CT images. 2522 CT images were used (1495 for COVID-19 patients, and 1027 for Community-Acquired Pneumonia (CAP)). The classification was carried out with different machine learning models, i.e. LR, SVM, KNN, Gaussian-Naive-Bayes classifier, and NNs. The proposed method outperformed the examined machine learning models with accuracy of 95.5%, sensitivity of 96.6%, and specificity of 93.2%.

For the selection of features and classification of COVID-19, authors in [31] suggested two optimization algorithms. There are 3 cascade phases in the proposed structure. Initially, features are being extracted using a CNN called AlexNet from the CT scans. Then, the proposed feature selection algorithm is implemented, a guided Whale Optimization Algorithm (WOA) based on Stochastic Fractal Search (SFS). The selected features were then balanced. Finally, using voting classification and guided WOA based on Particle Swarm Optimization, different classifiers' predictions were collected to determine the most voted class. For evaluating their proposed model, two data sets were used with positive and negative COVID-19 clinical CT-images. The proposed voting classification (PSO-guided-WOA) was 99.5%, higher than the other voting classifiers. Authors in [32] proposed a new architecture for COVID-19 prediction consisting of 3 principal manners: random selection of the region of interest, training of the CNN network to obtain features, and model training on a fully connected network and prediction of various classifiers. This method achieved a classification accuracy of 82.9% utilizing the modified Inception (M-Inception) deep model developed using CT images. Authors in [33] created a large dataset of CT scan and X-ray images from multiple sources and presented two CNN approaches to classify them as having COVID-19 or not, which are transfer learning AlexNet and a simple CNN architecture. They achieved 98% accuracy via the pre-trained model and 94.1% by using the modified CNN. Authors in [34] reviewed the possibility of ML models for diagnosis corona virus using X-Ray images. They used LR and CNN and Generative Adversarial Network (GAN) for data augmentation in order to overcome the overfitting problem. The model gives 95.2–97.6% accuracy without PCA and 97.6–100% with PCA.

During the coronavirus period, the collected data were often heterogeneous. Their features were categorical and numerical, or specific categorical such as the occurrence of cough, while the numerical features are quantitative, for example the temperature of the body. It is difficult to process these two forms. Moreover, this information may be incomplete as certain features have missing values. Incompleteness and heterogeneity are major challenges in identifying COVID-19 cases for data collected from many people. Heterogeneity and missing values are not directly handled by the common classification algorithms. To deal with these two issues, authors in [34] proposed the KNN Variant (KNNV) classification algorithm for incomplete and heterogeneous datasets as tools for detecting COVID19 cases accurately and efficiently. The two key concepts of the proposed algorithm are that the parameter K is chosen adaptively and that it calculates the

distances from other instances in a new way for each case being classified. The datasets were manually generated from a dataset of 68 COVID-19 and 62 non-COVID-19 cases from the SIRM database. Flu cases from the Influenza Research Database (IRD) were used. Then, the 2 datasets were merged and shuffled randomly to obtain an IHC dataset with two subsets: COVID-19 and flu. The results indicate that the suggested algorithm outperforms other known algorithms in terms of accuracy, recall, and F score. Authors in [35] published the initial research of applying an RF algorithm to determine the COVID-19 in clinically available blood test data. The researchers obtained 253 blood samples, including 105 confirmed COVID-19 cases, from different hospitals in Lanzhou, China. The method uses 11 key features out of the 49 clinical blood indexes that have been derived from the RF algorithm. The trained model, with 11 main features, exhibited 95.12% sensitivity, 96.97% specificity, and an overall 96.95% accuracy. In external blood samples, the algorithm's efficiency was slightly lower, with 95.12% sensitivity, 96.97% specificity, and 95.95% accuracy. The used dataset has a very high positive ratio of 41.5% which can affect performance. Moreover, the model in [2] revealed a high positive correlation between the patient's age and gender with the possibility of getting infected by COVID-19. Authors in [36] introduced a study of analyzing X-ray images of patient's and healthy person's lungs to compare and detect the level of infection on the lungs, and then classify it into level-based effect. Because of the complexity and the limited access to X-ray images, complexity-based theory was used to analyze the images.

During the COVID-19 epidemic, classifying the patients into infected or healthy is critical to prevent the spread of the virus. Authors in [37] suggested a technique based on deep learning by utilizing an SVM classifier to detect and classify the X-ray images of the patients' lungs as infected or healthy. Due to the limitation of diagnostic kits, the diagnosis of the disease has been used by X-ray imaging machines, as they are widely available in clinics and hospitals. The proposed system, based on deep CNNs, was used to analyze the X-ray image datasets. The system was trained on two datasets, namely one dataset of positive and one of negative cases. The datasets have been used to extract the images' features based on deep architectures such as GoogleNet, ResNet50, and Inception V3. The combination of the SVM classifier, the open-source dataset, and deep feature extraction techniques achieved accuracy of 95.38%. Authors in [38] proposed a COVID-19 classification by using CT chest images to classify the patients into two groups, +Ve (infected) and -Ve (not-infected). The model, called Multi-Objective Differential Evolution (MODE), with CNN has been implemented in a CT-chest images dataset, but the deep CNN showed some hyperparameter-tuning issues. However, the image features have been extracted by implementing many convolutions and the max-pooling layer was used to minimize the spatial size. The suggested model showed a good rate of accuracy, F-measure, and Kappa with values of 97.89%, 2.0928%, 1.9276% respectively.

Authors in [40] introduced a deep learning-based transfer model to diagnose COVID-19 cases by using CT-scan and CXR images. The algorithm was trained on 3 datasets. The model showed the detection results faster than the RT-PCR method. In addition, they discovered a pattern between the infected patients with COVID-19 and pneumonia. Performing an X-ray or CT-scan in an infected lung a shadowy area called Ground Glass Opacity (GGO) will appear. Applying deep learning led to fast and efficient detection results. The aim of the model is the binary classification of the chest images in a fast and accurate way. Image classification was conducted and 3 experiments were proposed and applied on the dataset. The model achieved accuracy of 95.61%. Authors in [7] established a community learning model known as ERLX in routine blood tests diagnostics for COVID-19. The proposed model used structural diversity by 2 classificatory levels. To improve predictive capabilities, the prediction from the first classification (extra trees, RF, LR) was supplied to the second classifier (XGBoost). A series of steps in data preparation were performed using the KNN algorithm to control the null values of the data collection. Isolation Forest (iForest) was used to eliminate outlier data, and SMOTE to equalize data distribution. By comparing the findings with the state-of-the-art studies available for a publicly available dataset from the Albert Einstein Hospital in Brazil, the efficacy and reliability of the model for diagnosing COVID-19 was demonstrated. Ensemble models are more efficient, robust, and flexible since diversity is the fundamental guiding principle for capture underlying training data structure. The ensemble model achieved outstanding performance with an overall accuracy of 99.88%, sensitivity of 98.72%, and specificity of 99.99%.

Due to the small number of X-rays and CT images with COVID-19, training the machine learning and deep network models is extremely difficult. Therefore, transfer-learning could be an applicable solution that has been extensively adopted in several latterly submitted COVID-19 detection methods. A deep transfer learning-based approach has been suggested in [41] using COVID-19, normal, and viral pneumonia CXR images to detect COVID-19 pneumonia automatically, using deep CNNs. The results show that transfer learning has proven to be successful, robust, and easily deployable for detecting COVID-19 with 98% accuracy. However, the conventional transfer-learning system which uses a developed, pre-trained deep network in the images database to transmit the original information might not be an excellent option, since the features of X-rays and CT-image of COVID-19 are completely different from pictures for other apps. But after this research, researchers gave hope that COVID-19 can be recognized from other infections and normal lung conditions by utilizing CXR imaging. Authors in [42] report a COVID-19 prediction analysis in CXR imagery by applying transfer learning. They compared 4 common deep CNN predictors (ResNet18, ResNet50, SqueezeNet, and DenseNet-161). After training, the models presented an average specificity rate of ~90% and 97.5% sensitivity. Authors in [43] suggested a weakly supervised deep-learning technique to detect and identify CT image infection with COVID-19 using 3D CT volumes for classification of COVID-19 and lesion location. The proposed method can reduce the manual labeling requirements of CT images, but it can still reliably detect infections and draw a distinction between COVID-19 and non-COVID-19 images. The model achieved 96.2% accuracy,

94.5% sensitivity, 95.3% specificity, and an AUC value of 0.970.

Authors in [44, 45] stated that for computerized COVID-19 pneumonia detection, CT scans are utilized in deep learning systems. Although CT scans contribute more comprehensive information, X-rays are faster, easier to take, less unhealthy, and cheaper. Nevertheless, it is extremely hard to effectively train a very deep network because of the scarcity of COVID-19 X-rays at that time. The accuracy of their research in COVID-19 classification exhibited 96% AUC, 90% sensitivity, and 96% specificity and 99.6% AUC, 98.2% sensitivity, and 92.2% specificity respectively. In terms of analyzing and detecting the radiological images automatically, CNN deep learning techniques are widely used [46]. Usually in deep learning techniques, model training is conducted with large datasets, but due to the limitation of the infected X-ray or CT scan images, the network weights have been fine-tuned from pre-trained networks. Authors in [47] proposed a classification model to classify the chest CT scan of COVID-19 patients as positive (infected) or negative (uninfected). The model was based on Deep Transfer Learning (DTL) with DenseNet201 and CNN. It had 99.82% training accuracy and the achieved testing and validation accuracies were 96.25% and 97.4% respectively.

Authors in [48] proposed a deep DT classifier consisting of 3 stages. Each DT was trained using a deep learning model with a NN based on a PyTorch frame for the detection of COVID-19 in chest X-ray images. The average acquired accuracy was 95%. Authors in [49] utilized a pretrained DenseNet121 model for the classification of segmented 3D lung regions. The dataset consisted of 2724 CT scans from 2617 sufferers. Lung regions were separated via utilizing the 3d Anisotropic Hybrid Network (AH-Net) architecture, achieving 90.8% accuracy, 93% specificity, and 94.9% AUC score. The significance of using AI in the medical domain to diagnose and predict diseases is in making several decisions based on big data analysis. Authors in [50] suggested a deep learning-based model to predict the coronavirus severity level by using CT scan images of infected patients. The model was trained in 303 chest images and was tested in 105 images, achieving 97.4% training and 81.9% testing accuracy.

Based on coronavirus confirmed cases, countries have been classified into categories based on the risk rate. Authors in [51] proposed an AI-guided method to predict the long-term country-specific risk of COVID-19 by the Bayesian optimization guided shallow LSTM. On average the model showed an accuracy of 77.4%. Applying the model during the pandemic posed many challenges such as the limitation of the dataset, data uncertainty, and data confusion. Authors in [52] provided a system based on deep learning to COVID-19 detection using CNN and convolutional (ConvLSTM) with 100% accuracy and F1 score, while authors in [53] obtained 89.5% accuracy, 88% specificity and 87% sensitivity in CT images. Authors in [54] got AUC of 99% and recall of 93% in COVID-19 diagnosis. Authors in [55] used the ResNet-50 deep learning model to classify COVID-19 in CT scans with 96% AUC. Authors in [56] identified coronavirus cases rapidly using the RestNet-100 CNN along with LR achieving accuracy of 99.15%. Authors in [57] proposed COVID-19 pneumonia

detection using a small number of COVID-19 CXR images. They optimized the features using the CNN model CONVNet with 98.1% accuracy. Authors in [58] investigated the performance of KNN, DT, NB, SVM, and LR to discover COVID-19 cases. The results were 93% accuracy for NB and DT, 93.60% for SVM, 93.50% for KNN, and 92.80% for LR and they concluded that ML methods have an important use in identifying COVID-19. Authors in [59] proposed a detection-classification approach to diagnose COVID-19 by examining the patients' lungs using CT-scan and X-ray images. Their approach was compared with the SVM classifier. The comparison was made on the classification accuracy as a performance indicator. The classification–detection model achieved 84% and 75% with CXRs and CT-scans, respectively.

Table II provides a summary of all the reviewed studies.

VI. COVID-19 IMAGE DATASETS

CT scanning and X-ray imaging are the most common used techniques in respiratory disease diagnosis. For decades, radiologists and other healthcare professionals have used them. Discovering COVID-19 early is crucial to prevent its spread by using diagnostic imagery methods. To show the damage of COVID-19 on the patient's lungs, many health and technology centers publish data about COVID-19, pneumonia, and acute respiratory distress syndrome. Authors in [61] produced the first public dataset of the frontal view of X-ray images. The dataset was collected from different open sources by the University of Montreal. It consists of more than 400 CT and X-ray images of COVID-19, SARS, MERS-CoV, varicella, influenza, and herpes patients. The dataset is continuously updated by the newest COVID-19 cases through the GitHub link [62]. The dataset in [63] is a public dataset of 5,863 X-Ray images categorized into normal/healthy lungs and pneumonia-infected lungs. Another dataset of chest X-ray images of COVID-19 cases was published in 2020 [64]. The dataset contains X-ray images of COVID-19-infected lungs along with normal lungs in addition to viral pneumonia images. It consists of 3,487 images of normal, COVID-19, and viral patients. There is also a public dataset available on the Kaggle website [65], which contains a mixture of X-ray and CT scan images of patients diagnosed with COVID-19. Patients' images were extracted from public articles in RSNA, Radiopaedia, and SIRM [66]. ChestX-ray8 is a large public database provided in [32]. It consists of frontal-view X-ray images collected from 1992 to 2015. The dataset contains 108,948 images of more than 32,717 patients. Images are labeled and divided among 8 common thoracic pathology diseases. Authors in [67] produced a publicly available dataset of CT scan images of COVID-19 cases. The COVID-CT dataset contains 349 CT scans of confirmed cases, which were collected from relative sources, such as medRxiv, NEJM, JAMA, Lancet, and bioRxiv. The COVID-19 CT lung and infection segmentation dataset [68] contains 100 CT images collected from more than 40 patients with COVID-19. The COVID-19 imaging database is another open-source dataset used by the British Society of Thoracic Imaging [69]. This dataset is used for learning and education purposes and it consists of 59 images of SARS-CoV2 patients.

VII. DISCUSSION

This study focused on the latest studies that used AI applications to confront COVID-19. In imaging-based COVID-19 classification, AI has been successfully applied. However, much work remains to be done. A growing number of different classification and prediction models use AI along with more publicly available datasets. In this paper, many recent research studies were reviewed. These studies involved deep learning and machine learning for COVID-19 classification. Deep learning approaches, such as CNNs, automatically perform the function extraction process. Little research, however, has been done on traditional machine learning models, like KNN, SVM, RF, and DT (see Table II). Most of this research used medical images to diagnose COVID-19 because they provide more details and depict the diseased areas of the body. The AI-based classification of CXRs has the untapped potential to meet this need.

Latest research has shown that compared with CT and Magnetic Resonance Imaging (MRI), CXR-related diagnosis is the most widely used because of its low cost, short therapy time, and low radiation exposure. However, the reviewed studies that used these images have some weaknesses. Some did not involve all age groups or adequately considered the patient's gender. Improving upon this can help facilitate the classification process of machine learning models. Given the small number of CXR and CT pictures available for COVID-19 at that time, the AI models were trained on fewer pictures, which may lead to overfitting in classification models and low accuracy. Deep learning models must be based on large datasets to work efficiently. For this reason, some techniques, such as augmentation, have been used to generalize models and to increase the size of the datasets. Researchers constantly work to update classification models and to increase their reliability and usefulness in real-life circumstances using the available datasets. In addition, the availability of more and more diversified training images would make the creation of strong and scalable classification models easier. AI-based disease classification can be merged with confirmed lab tests. It would be an excellent suggestion to help diagnose and evaluate recovering/infected patients utilizing fast AI-based tools.

Deep learning has become the dominant approach for detecting and diagnosing COVID-19. The image data in COVID-19 applications, however, could be inconsistent and inaccurate, posing a challenge for the creation of an exact segmentation and diagnostic network. For this situation, weakly supervised deep learning approaches may help. Some researchers diagnose COVID-19 cases employing machine learning methods by utilizing deep learning strategies via selecting the features of the images and obtaining high quality results. Finally, all the medical images used in the previous studies were not standardized and from different sources, in addition to the different medical devices that took these pictures. All these factors make a comprehensive comparison difficult.

VIII. CONCLUSION

The negative impact of COVID-19 has increased since December 2019. AI methods and applications in the medical fields showed success in fighting the virus. In this paper, an extensive review of the latest studies in COVID-19 classification and prediction was presented with a description of the frequently used methods and techniques in disease diagnosis, detection, and prediction. The number of contributions in this field is growing exponentially due to their success in preventing the spread of COVID-19. This paper summarized the research papers in classification algorithms and their applications in COVID-19 prediction models with a review of the method's aim in addition to the models' performance. Among the published studies, the use of deep learning in radiology imaging machines improved the models' performance.

IX. FUTURE DIRECTIONS AND CHALLENGES

The researchers in machine learning –related COVID-19 detection may face several challenges, especially when dealing with data provided from different sources and different formats.

- Large-scale training data are scarce and difficult to get.

Many deep learning techniques rely on a large dataset to train models, such as medical images to build an automatic system for prediction [70]. Due to the rapid explosion of COVID-19, the interpretation and labeling of training samples are time-consuming and require expert physicians, which results in insufficient data [71].

- Inappropriate data

Diverse online publications have published incorrect reports concerning COVID-19. Because of this problem, the use and reliability of AI-based methods have been reduced.

- Data protection and privacy.

To prevent the spread of COVID-19 and conduct proper contact tracing, patient's personal information is often collected, such as their identity number, contact information, and medical data. Therefore, protecting and maintaining the patient's privacy during treatment is incredibly important.

- Incorrect structured and unstructured data.

Incorrect information in text descriptions and medical images poses a challenge to researchers and machine learning models. A huge amount of data from various sources may be impossible to be used. AI also faces problems, such as dealing with unbalanced datasets, difficulty in screening, triaging patients due to restrictions [72], and social distancing, as well as dealing with poor quality data. Regarding future research directions, AI can be used in remote video investigation and consultations, biological research and knowledge of important protein structures and virus sequences, impact assessment of COVID-19, drug development, patient contact tracking, as well as diagnosis and treatment of COVID-19.

TABLE II.   SUMMARY OF THE RESEARCH ON COVID-19 DIAGNOSIS

Ref.	Aim	Methods	Datasets	Performance
[2]	Proposed a fine-tuned RF model supported with AdaBoost algorithm for covid-19 prediction.	Boosted RF algorithm	Novel Coronavirus 2019 dataset.	Revealed a high positive correlation between the patient's age and sex with the possibility of getting infected by Covid-19.
[3]	ERLX model developed for the detection and classification of COVID-19.	Extra trees, RF, LR, XGBoost	Routine blood examinations of COVID-19 cases at Albert Einstein Hospital.	99.88% accuracy, 98.72% sensitivity, 99.99% specificity.
[4]	Classifying COVID-19 and non-COVID classes based on lung CT scans	NB, KNN, DT, RF, SVM	CT scans from available datasets, including 500 photos (250 normal and 250 COVID-19).	Accuracy of NB, KNN, DT at 80%, of RF at 79%, and of SVM at 81%
[5]	Implemented a personalized network (DarkCovidNet) for automatic COVID-19 detection	DarkNet classifier	127 infected patients, 500 normal cases, and 500 pneumonia frontal X-ray images.	98.08% for binary and 87.02% for multi-class classification accuracy.
[6]	Proposed a method for binary classification and multi-classification problems by using a deep learning CNN network for the identification of COVID-19.	Deep CNN	910 X-rays and CT scans of COVID-19.	99% accuracy on X-ray images, and 71% on CT-scan images.
[7]	Used DeTraC to predict and classify the chest images of COVID-19 sufferers	DeTraC with CNN	CXR photos from many hospitals worldwide.	95.12% accuracy, 97.91% sensitivity, 91.87% specificity.
[8]	Prediction of patient's images as positive or negative cases.	COVIDX-Net deep CNN framework	50 X-ray photos with 25 positive COVID-19 confirmed cases.	The COVIDX-Net achieved a good performance in VGG19 and DenseNet including f1-scores of 0.89 and 0.9, respectively.
[9]	Detect the confirmed cases of coronavirus	Deep CNN with ResNet50, ResNet101, ResNet 152, Inception V3, and Inception-ResNet V2	Dr. Joseph Cohen open-source dataset and CXR images (pneumonia).	Accuracy of 96.1%, 99.5, and 99.7 for dataset-1, dataset-2, and dataset-3 respectively
[10]	Utilizing CNN to diagnose cases of COVID-19 with chest X-rays	Xception and ResNet50V2-based CNN	X-rays datasets of 180 images of COVID-19.	99.56% accuracy and 80.53% recall.
[11]	Early COVID-19 detection	SVM	150 CT images for COVID-19.	99.68% accuracy.
[12]	A detector trained on augmented data with high accuracy in the prediction of the patients' needs in the next 7 days.	Deep learning CNN	COVID-19 X-ray images	High accuracy of 94.18%, 99.94%, and 90.29% respectively in predicting the number of deaths, newly confirmed cases, and recovered cases.
[13]	A new open-source network called COVID-Net to distinguish the Covid-19 positive cases by analyzing CXR images.	Deep CNN	COVIDx is the largest open-source dataset of Covid-19 patient's CXR images.	93.3% accuracy.
[14]	COVID-19 classification based on pipeline and multiview representation learning method.	Pipeline and multiview representation learning	CT images of 1495 COVID-19 1027 CAP.	95.5% accuracy, 96.6% sensitivity, 93.2% specificity.
[15]	Two optimization algorithms were introduced in order to select and classify features of the COVID-19.	CNN, Guided WOA	334 CT images for COVID-19 and 794 CT images for non-COVID-19.	99.5% accuracy.
[16]	COVID-19 prediction consisting of 3 main processes.	CNN	CT images of 453 COVID-19 cases	82.9% accuracy.
[17]	KNNV classification algorithms for incomplete and heterogeneous datasets as tools for detecting COVID19 cases	KNNV	IHC dataset with 68 COVID patients and 62 flu patients.	All metrics indicate that KNNV is much stronger than the relevant algorithms.
[18]	Discover COVID-19 from clinically available blood test data	RF	Hospitals, Lanzhou, China.	95.12% sensitivity, 96.97% specificity, 96.95% accuracy.
[19]	Suggested a system based on deep CNNs to analyze X-ray image datasets.	Deep CNN	One dataset with 25 cases of COVID-19+ and 25 of COVID-19- X-ray images and another with 133 CRX images as COVID-19- and 133 as COVID-19+.	Achieved accuracy in terms of testing, F1 score, MCC, and Kappa, 95.38%, 95.52%, 91.41%, 90.76 % respectively.
[20]	Classify the patients into +Ve (infected) and -Ve (not-infected) groups.	MODE with CNN	X-ray and CT scan images	Accuracy, F-measure, and Kappa statistics were 97.89%, 2.0928%, 1.9276% respectively
[21]	A deep learning-based transfer model to diagnose COVID-19 cases. Binary classification in a fast and accurate way.	Deep learning-based transfer model	Three datasets SARS-COV2 CT-scan, CRX, and COVID-CRX	95.61% accuracy.
[22]	Automatic detection of COVID-19 pneumonia utilizing DCNNs.	DCNN	864 COVID-19, 1345 pneumonia, and 1341 CXR images.	98% accuracy.
[58]	Building automatic CT picture analysis tools based on AI for Coronavirus identification, quantification, and tracking.	Deep Learning	CT image datasets, the testing set had 157 international patients	99.6% AUC, 98.2% sensitivity, 92.2% specificity.

[54]	A weakly-supervised DL for identifying and classifying COVID-19 from CT scans and reduce the requirements of labeling the CT images.	Weakly-supervised deep learning system using CNN	150 3D chest CT scans of COVID-19, CAP, and NP patients respectively	96.2% accuracy, 94.5% sensitivity, 95.3% specificity, 97.0% AUC.
[23]	COVID-19 prediction in CXR imaging using transfer learning.	Deep Transfer Learning (CNN)	5,000 CXRs from publicly available datasets	Specificity rate of ~90% with a sensitivity range of 97.5%
[54]	A weakly-supervised DL for identifying and classifying COVID-19 from CT scans and reduce the requirements of labeling the CT images.	Weakly-supervised deep learning system using CNN	150 3D chest CT scans of COVID-19, CAP, and NP patients respectively	96.2% accuracy, 94.5% sensitivity, 95.3% specificity, 97.0% AUC.
[16]	A fully automated chest CT framing for detecting COVID-19	Deep Learning	3D chest CT exams acquired from 6 medical centers, consisting of 1296 images of COVID-19 patients	96% AUC, 90% sensitivity, 96% specificity.
[59]	A deep learning-based transfer model to diagnose COVID-19 cases by utilizing CT-scan and X-ray images of the human's chest. The aim is the binary classification of the chest images in a fast and accurate way.	Deep learning-based transfer model	Three datasets SARS-COV2 CT-scan, Chest X-Ray images, and COVID-CXRs	95.61% accuracy.
[60]	A COVID-19 assessment DT classifier from CXR images.	Deep learning-based DT classifier	CXR	95.0% accuracy.
[28]	Utilized pre-trained model DenseNet121 to classify the segmented 3D lung regions.	AH-Net DenseNet121	CT images of 1029 COVID-19 1695 non-COVID-19 patients.	0.90% accuracy, 93% specificity, and 94% AUC.
[62]	Predict the severity level of Coronavirus by using CT scan images of infected patients.	Deep learning	The dataset contains 303 CT-images as a training set and 105 images as testing set.	97.4% training and 81.9% testing accuracy.
[43]	Classify the countries based on the risk rate.	AI-guided method with Bayesian optimizer	COVID-19 open-source dataset and weather dataset, the number of cases confirmed, number of cases recovered, and the total number of deaths are included.	77.4% accuracy.
[44]	Create a large dataset and classify COVID-19	CNN model and pre-trained AlexNet	170 X-ray images and 361 CT images of COVID-19	AlexNet =98%, CNN =94.1%.
[62]	Diagnose corona virus with PCA.	LR, CNN	X-ray images.	95.2–97.6% accuracy without PCA and 97.6–100% with PCA.
[63]	COVID-19 detection	LSTM and CNN	CT and X-ray images	100% for accuracy and F1 score.
[64]	COVID-19 detection	Deep learning	CT images of 259 patients.	89.5% accuracy, 88% specificity, 87% sensitivity.
[65]	COVID-19 diagnosis	Deep learning-based CT diagnosis	CT scans of 88 patients of COVID-19, 101 of bacterial pneumonia, 86 healthy persons.	AUC of 0.99 and recall of 0.93.
[66]	Classify COVID-19	ResNet-50 deep learning	CT scan	96% AUC
[67]	Identify corona virus rapidly	ResNet-100	CT scan	99.15% accuracy.
[37]	Feature optimization using CNN	CNN model (COVXNet)	X-rays	98.1% accuracy.
[38]	Investigate the performance of ML to discover the COVID-19.	KNN, DT, NB, SVM, LR	Digital data of 1,048,576 patients.	93% accuracy for NB and DT, 93.60% for SVM, 93.50% for KNN, 92.80% for LR.
[39]	Classify COVID-19 using CT-scan and X-ray imaging techniques.	ResNet50, InceptionResNetV2, Xception, VGGNet16, SVM	CXR images and CT scans.	Accuracy of 84% and 75% for X-Ray and CT-scans, respectively.

REFERENCES

[1] World Health Organization, "Q&As on COVID-19 and related health topics." https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub (accessed Oct. 28, 2021).

[2] S. Parasa, N. Reddy, D. O. Faigel, A. Repici, F. Emura, and P. Sharma, "Global Impact of the COVID-19 Pandemic on Endoscopy: An International Survey of 252 Centers From 55 Countries," *Gastroenterology*, vol. 159, no. 4, pp. 1579-1581.e5, Oct. 2020, https://doi.org/10.1053/j.gastro.2020.06.009.

[3] "Coronavirus (COVID-19) Testing," *Our World in Data*, Mar. 05, 2020. https://ourworldindata.org/coronavirus-testing (accessed Oct. 28, 2021).

[4] C. Iwendi *et al.*, "COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm," *Frontiers in Public Health*, vol. 8, 2020, Art. no. 357, https://doi.org/10.3389/fpubh.2020.00357.

[5] A. A. Soofi and A. Awan, "Classification Techniques in Machine Learning: Applications and Issues," *Journal of Basic & Applied Sciences*, vol. 13, pp. 459–465, 2017.

[6] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, Oct. 2020, Art. no. 110059, https://doi.org/10.1016/j.chaos.2020.110059.

[7] M. AlJame, I. Ahmad, A. Imtiaz, and A. Mohammed, "Ensemble learning model for diagnosing COVID-19 from routine blood tests," *Informatics in Medicine Unlocked*, vol. 21, Jan. 2020, Art. no. 100449, https://doi.org/10.1016/j.imu.2020.100449.

[8] Y. Mohamadou, A. Halidou, and P. T. Kapen, "A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19," *Applied Intelligence*, vol. 50, no. 11, pp. 3913–3925, Nov. 2020, https://doi.org/10.1007/s10489-020-01770-9.

[9] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, https://doi.org/10.1007/BF00994018.

[10] K. Aida-zade, A. Xocayev, and S. Rustamov, "Speech recognition using Support Vector Machines," in *10th International Conference on Application of Information and Communication Technologies*, Baku, Azerbaijan, Oct. 2016, pp. 1–4, https://doi.org/10.1109/ICAICT.2016.7991664.

[11] Y. Tian, E. Li, L. Yang, and Z. Liang, "An image processing method for green apple lesion detection in natural environment based on GA-BPNN and SVM," in *International Conference on Mechatronics and Automation*, Changchun, China, Aug. 2018, pp. 1210–1215, https://doi.org/10.1109/ICMA.2018.8484624.

[12] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *Journal of Biomedical Informatics*, vol. 35, no. 5, pp. 352–359, Oct. 2002, https://doi.org/10.1016/S1532-0464(03)00034-0.

[13] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, no. 2, pp. 197–227, Jun. 2016, https://doi.org/10.1007/s11749-016-0481-7.

[14] M. Zakariah, "Classification of large datasets using Random Forest Algorithm in various applications: Survey," *International Journal of Engineering and Innovative Technology*, vol. 4, no. 3, pp. 189–198, 2014.

[15] C. Sohrabi *et al.*, "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)," *International Journal of Surgery*, vol. 76, pp. 71–76, Apr. 2020, https://doi.org/10.1016/j.ijsu.2020.02.034.

[16] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, Jun. 2020, https://doi.org/10.1007/s13246-020-00865-4.

[17] S. Kadry, V. Rajinikanth, S. Rho, N. S. M. Raja, V. S. Rao, and K. P. Thanaraj, "Development of a Machine-Learning System to Classify Lung CT Scan Images into Normal/COVID-19 Class," *arXiv:2004.13122 [cs, eess, stat]*, Apr. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://arxiv.org/abs/2004.13122.

[18] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, Jun. 2020, Art. no. 103792, https://doi.org/10.1016/j.compbiomed.2020.103792.

[19] A. Carfì, R. Bernabei, and F. Landi, "Persistent Symptoms in Patients After Acute COVID-19," *JAMA*, vol. 324, no. 6, pp. 603–605, Aug. 2020, https://doi.org/10.1001/jama.2020.12603.

[20] M. K. Nath, A. Kanhe, and M. Mishra, "A Novel Deep Learning Approach for Classification of COVID-19 Images," in *5th International Conference on Computing Communication and Automation*, Greater Noida, India, Oct. 2020, pp. 752–757, https://doi.org/10.1109/ICCCA49541.2020.9250907.

[21] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, Feb. 2021, https://doi.org/10.1007/s10489-020-01829-7.

[22] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," *arXiv:2003.11055 [cs, eess]*, Mar. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://arxiv.org/abs/2003.11055.

[23] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 1207–1220, Aug. 2021, https://doi.org/10.1007/s10044-021-00984-y.

[24] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images," *Informatics in Medicine Unlocked*, vol. 19, 2020, Art. no. 100360, https://doi.org/10.1016/j.imu.2020.100360.

[25] W. H. Self, D. M. Courtney, C. D. McNaughton, R. G. Wunderink, and J. A. Kline, "High Discordance of Chest X-ray and CT for Detection of Pulmonary Opacities in ED Patients: Implications for Diagnosing Pneumonia," *The American journal of emergency medicine*, vol. 31, no. 2, pp. 401–405, Feb. 2013, https://doi.org/10.1016/j.ajem.2012.08.041.

[26] G. D. Rubin *et al.*, "The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational Consensus Statement from the Fleischner Society," *Radiology*, vol. 296, no. 1, pp. 172–180, Jul. 2020, https://doi.org/10.1148/radiol.2020201365.

[27] M. Barstugan, U. Ozkaya, and S. Ozturk, "Coronavirus (COVID-19) Classification using CT Images by Machine Learning Methods," *arXiv:2003.09424 [cs, eess, stat]*, Mar. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://arxiv.org/abs/2003.09424.

[28] M. Alazab, A. Awajan, A. Mesleh, A. Abraham, V. Jatana, and S. Alhyari, "COVID-19 Prediction and Detection Using Deep Learning," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 12, pp. 168–181, 2020.

[29] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Scientific Reports*, vol. 10, no. 1, Nov. 2020, Art. no. 19549, https://doi.org/10.1038/s41598-020-76550-z.

[30] H. Kang *et al.*, "Diagnosis of Coronavirus Disease 2019 (COVID-19) With Structured Latent Multi-View Representation Learning," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2606–2614, Aug. 2020, https://doi.org/10.1109/TMI.2020.2992546.

[31] E.-S. M. El-Kenawy, A. Ibrahim, S. Mirjalili, M. M. Eid, and S. E. Hussein, "Novel Feature Selection and Voting Classifier Algorithms for COVID-19 Classification in CT Images," *IEEE Access*, vol. 8, pp. 179317–179335, 2020, https://doi.org/10.1109/ACCESS.2020.3028012.

[32] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," in *Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, Jul. 2017, pp. 3462–3471, https://doi.org/10.1109/CVPR.2017.369.

[33] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms," Mar. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://128.84.4.34/abs/2004.00038.

[34] J. Rasheed, A. A. Hameed, C. Djeddi, A. Jamil, and F. Al-Turjman, "A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 13, no. 1, pp. 103–117, Mar. 2021, https://doi.org/10.1007/s12539-020-00403-6.

[35] J. Wu *et al.*, "Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results," Apr. 2020. https://doi.org/10.1101/2020.04.02.20051136.

[36] H. Namazi and V. V. Kulish, "Complexity-based classification of the coronavirus disease (covid-19)," *Fractals*, vol. 28, no. 05, Aug. 2020, Art. no. 2050114, https://doi.org/10.1142/S0218348X20501145.

[37] P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, "Detection of Coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine," *International Journal of Mathematical Engineering and Management Sciences*, pp. 643–651, 2020.

[38] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 39, no. 7, pp. 1379–1389, Jul. 2020, https://doi.org/10.1007/s10096-020-03901-z.

[39] U.S. Food and Drug Administration, "FDA Approves First COVID-19 Vaccine," *FDA*, Aug. 23, 2021. https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine (accessed Oct. 28, 2021).

[40] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, and V. Singh, "A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images," *Chaos, Solitons & Fractals*, vol. 140, Nov. 2020, Art. no. 110190, https://doi.org/10.1016/j.chaos.2020.110190.

[41] S. Asif, Y. Wenhui, H. Jin, and S. Jinhai, "Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Network," in *6th International Conference on Computer and Communications*, Chengdu, China, Dec. 2020, pp. 426–433, https://doi.org/10.1109/ICCC51575.2020.9344870.

[42] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Medical Image Analysis*, vol. 65, Oct. 2020, Art. no. 101794, https://doi.org/10.1016/j.media.2020.101794.

[43] S. Hu *et al.*, "Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification From CT Images," *IEEE Access*, vol. 8, pp. 118869–118883, 2020, https://doi.org/10.1109/ACCESS.2020.3005510.

[44] L. Li *et al.*, "Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT," *Radiology*, Mar. 2020, Art. no. 200905, https://doi.org/10.1148/radiol.2020200905.

[45] O. Gozes *et al.*, "Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis," *arXiv:2003.05037 [cs, eess]*, Mar. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://arxiv.org/abs/2003.05037.

[46] M. Kaur, H. K. Gianey, D. Singh, and M. Sabharwal, "Multi-objective differential evolution based random forest for e-health applications," *Modern Physics Letters B*, vol. 33, no. 05, Feb. 2019, Art. no. 1950022, https://doi.org/10.1142/S0217984919500222.

[47] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 15, pp. 5682–5689, Oct. 2021, https://doi.org/10.1080/07391102.2020.1788642.

[48] S. H. Yoo *et al.*, "Deep Learning-Based Decision-Tree Classifier for COVID-19 Diagnosis From Chest X-ray Imaging," *Frontiers in Medicine*, vol. 7, 2020, Art. no. 427, https://doi.org/10.3389/fmed.2020.00427.

[49] S. A. Harmon *et al.*, "Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets," *Nature Communications*, vol. 11, no. 1, Aug. 2020, Art. no. 4080, https://doi.org/10.1038/s41467-020-17971-2.

[50] L. Xiao *et al.*, "Development and Validation of a Deep Learning-Based Model Using Computed Tomography Imaging for Predicting Disease Severity of Coronavirus Disease 2019," *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020, Art. no. 898, https://doi.org/10.3389/fbioe.2020.00898.

[51] R. Pal, A. A. Sekh, S. Kar, and D. K. Prasad, "Neural Network Based Country Wise Risk Prediction of COVID-19," *Applied Sciences*, vol. 10, no. 18, Jan. 2020, Art. no. 6448, https://doi.org/10.3390/app10186448.

[52] A. Sedik, M. Hammad, F. E. Abd El-Samie, B. B. Gupta, and A. A. Abd El-Latif, "Efficient deep learning approach for augmented detection of Coronavirus disease," *Neural computing & applications*, pp. 1–18, Jan. 2021, https://doi.org/10.1007/s00521-020-05410-8.

[53] S. Wang *et al.*, "A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19)," *European Radiology*, vol. 31, no. 8, pp. 6096–6104, Aug. 2021, https://doi.org/10.1007/s00330-021-07715-1.

[54] Y. Song *et al.*, "Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021, https://doi.org/10.1109/TCBB.2021.3065361.

[55] S. Serte and H. Demirel, "Deep learning for diagnosis of COVID-19 using 3D CT scans," *Computers in Biology and Medicine*, vol. 132, p. 104306, May 2021, https://doi.org/10.1016/j.compbiomed.2021.104306.

[56] M. Kavitha, T. Jayasankar, P. M. Venkatesh, G. Mani, C. Bharatiraja, and B. Twala, "COVID-19 Disease Diagnosis using Smart Deep Learning Techniques," *Journal of Applied Science and Engineering*, vol. 24, no. 3, pp. 271–277, 2021, https://doi.org/10.6180/jase.202106_24(3).0001.

[57] T. Mahmud, M. A. Rahman, and S. A. Fattah, "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and

[58] M. Malik *et al.*, "Determination of COVID-19 Patients Using Machine Learning Algorithms," *Intelligent Automation and Soft Computing*, vol. 31, no. 1, pp. 207–222, 2022.

[59] X. Deng, H. Shao, L. Shi, X. Wang, and T. Xie, "A Classification–Detection Approach of COVID-19 Based on Chest X-ray and CT by Using Keras Pre-Trained Deep Learning Models," *Computer Modeling in Engineering & Sciences*, vol. 125, no. 2, pp. 579–596, Nov. 2020, https://doi.org/10.32604/cmes.2020.011920.

[60] A. Hamed, A. Sobhy, and H. Nassar, "Accurate Classification of COVID-19 Based on Incomplete Heterogeneous Data using a KNN Variant Algorithm," *Arabian Journal for Science and Engineering*, vol. 46, no. 9, pp. 8261–8272, Sep. 2021, https://doi.org/10.1007/s13369-020-05212-z.

[61] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 Image Data Collection: Prospective Predictions Are the Future," *arXiv:2006.11988 [cs, eess, q-bio]*, Dec. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://arxiv.org/abs/2006.11988.

[62] J. P. Cohen, *ieee8023/covid-chestxray-dataset*. 2021. Accessed: Oct. 28, 2021. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset.

[63] "Chest X-Ray Images (Pneumonia)." https://kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed Oct. 28, 2021).

[64] M. E. H. Chowdhury *et al.*, "Can AI Help in Screening Viral and COVID-19 Pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020, https://doi.org/10.1109/ACCESS.2020.3010287.

[65] "COVID-19 Radiography Database." https://kaggle.com/tawsifurrahman/covid19-radiography-database (accessed Oct. 28, 2021).

[66] "COVID-19 X rays." https://kaggle.com/andrewmvd/convid19-x-rays (accessed Oct. 28, 2021).

[67] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-Dataset: A CT Scan Dataset about COVID-19," *arXiv:2003.13865 [cs, eess, stat]*, Jun. 2020, Accessed: Oct. 28, 2021. [Online]. Available: http://arxiv.org/abs/2003.13865.

[68] "COVID-19," *Medical segmentation*. http://medicalsegmentation.com/covid19/ (accessed Oct. 28, 2021).

[69] "COVID-19 British Society of Thoracic Imaging Database | The British Society of Thoracic Imaging." https://www.bsti.org.uk/training-and-education/covid-19-bsti-imaging-database/ (accessed Oct. 28, 2021).

[70] P. Chakraborty and C. Tharini, "Pneumonia and Eye Disease Detection using Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5769–5774, Jun. 2020, https://doi.org/10.48084/etasr.3503.

[71] H. H. A. Owida, A. Al-Ghraibah, and M. Altayeb, "Classification of Chest X-Ray Images using Wavelet and MFCC Features and Support Vector Machine Classifier," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7296–7301, Aug. 2021, https://doi.org/10.48084/etasr.4123.

[72] N. K. Al-Shammari, H. B. Almansour, and M. B. Syed, "Development of an Automatic Contactless Thermometer Alert System Based on GPS and Population Density," *Engineering, Technology & Applied Science Research*, vol. 11, no. 2, pp. 7006–7010, Apr. 2021, https://doi.org/10.48084/etasr.4103.