A Review on the Use of Machine Learning Against the Covid-19 Pandemic

Sardar Asad Ali Biabani
Science and Technology Unit
Umm Al-Qura University
Makkah Al-Mukarramah, Saudi Arabia
sabiabani@uqu.edu.sa

Nahla A. Tayyib
Faculty of Nursing, Deanship of Scientific Research
Umm Al-Qura University
Makkah Al-Mukarramah, Saudi Arabia
natayyib@uqu.edu.sa

Abstract-Coronavirus-2019 disease (Covid-19) is a contagious respiratory disease that emerged in late 2019 and has been recognized by the World Health Organization (WHO) as a global pandemic in early 2020. Since then, researchers have been exploring various strategies and techniques to fight against this outbreak. The point when the pandemic appeared was also a period in which Machine Learning (ML) and Deep Learning (DL) algorithms were competing with traditional technologies, leading to significant findings in diverse domains. Consequently, many researchers employed ML/DL to speed up Covid-19 detection, prevention, and treatment. This paper reviews the state-of-the-art ML/DL tools used, thoroughly evaluating these techniques and their impact on the battle against Covid-19. This article aims to provide valuable insight to the researchers to assess the use of ML against the Covid-19 pandemic.

Keywords-Covid-19; machine learning; deep learning; coronavirus; evaluation; detection; prevention; treatment

I.   INTRODUCTION

Deep Learning techniques can be employed in many technical predicaments instead than conventional ones in order to save time and resources. This paper investigates and evaluates the use of Machine Learning (ML) and Deep Learning (DL) techniques on the fight against the Covid-19 pandemic. We will mainly focus on four states: At first, DL and ML models require a lot of data in their training process to provide accurate predictions. Therefore, we begin this study by reporting some Covid-19-related datasets used to train ML/DL tools. Then, we will investigate the way these tools are employed to prevent the disease, either by tracking the spread of the virus or by helping in the productions of a new vaccine. Thirdly, the paper will study the utilization of ML techniques in the discovery phase, presenting the models and algorithms adapted to diagnose and detect the virus. Finally, we will examine the same approaches, but this time applied to the drug discovery and treatment of the disease. This paper aims to present a review about the use of ML against Covid-19, to evaluate these techniques, and to analyze their effectiveness to counter the pandemic. However, before talking about implementing ML or DL techniques in any given field, essentially, when the models we are dealing with are on supervised learning methods, the fundamental problem to deal

with is data. Various types of Covid-19 datasets available for practitioners and researchers are examined below.

A broad range of data was collected and made open during this pandemic. It ranges from classic Computerized Tomography (CT) scans to X-Ray Chest (XRC) radiograph images, but one can also find datasets based on text data, video, or sound. Authors in [1] collected a dataset of 123 computerized CT scans. This dataset contains not only Covid-19 but also MERS, SARS, and ARDS (other coronaviruses) images. The authors in [2] provide a more extensive dataset of 275 CT scans of Covid-19 patients. A familiar and productive technique used to augment data in ML is segmentation. This technique was used in several studies [3] to develop more samples. In addition to segmentation, other studies collected negative Covid-19 CT images to combine them with positive cases for binary classification diagnosis models [4-6]. XRC images belong to a different type of data. One of the first efforts to collect an XRC Covid-19 dataset was [7]. The authors used a collection of more than 13000 images to train a neural network classification model that achieved 93.11% accuracy. Moreover, added preliminaries in the XRC datasets were developed in [1]. This resource framed the basis of many extensions [8-11], whether by appending more images of Covid patients, using data augmentation techniques such as segmentation or by adding non-Covid-19 images. In addition, an Italian hospital provided more than 4700 X-ray images of Covid-19 patients [12]. Another popular dataset type deals with tracking and visualizing Covid-19's new cases and deaths in real-time. The most known dataset of this kind is [13], while [14] was used to predict the pandemic's spread.

Besides these platforms that offer universal statistics, many country-focused datasets accumulate information about specific countries [15-17]. Although these datasets represent limited regions, they generally present richer resources since they consider demographic aspects, travel restrictions, and sometimes even personal data. In [18], the authors presented a collection of Covid-19, pneumonia, and healthy videos. A related dataset [19] consolidated data from some hospitals and web resources to present a collection of images and videos of lung ultrasounds. Additionally, the Coswara ongoing project [20], consists of collecting breathing, cough, and speech sounds from Covid-19 patients to build a diagnostic tool. Likewise,

another sound-based South African dataset was erected [21] on cough audio.

Table I summarizes the cited datasets, the nature of the data they contain, and, more importantly, the number of samples in each one. Notably, the big problem in using these datasets to train DL models is that they do not have enough data to predict real-world situations accurately.

TABLE I. REVIEWED COVID-19 DATASETS

Ref.	Nature of the data	Number of samples
[1]	Frontal view X-rays	123
[2]	CT scans	275
[3]	CT scans	20 cases
[4]	CT scans	5900 images of 1200 patients
[5]	CT scans	888 CT scans of 888 subjects
[6]	CT scans	366,558 CT scan images
[7]	XRC	13000 images
[12]	X-ray images	4700
[19]	Videos	247

The following sections will provide an extensive literature review of the papers that used ML techniques to help the spreading prevention, detection, and treatment of Covid-19.

II. ML METHODS FOR COVID-19 PREVENTION

This section will examine the operation of ML techniques in the prevention of the spreading of Covid-19. To accomplish this, we will present two effective solutions that have been used in response to the outbreak: Vaccines (to reduce the spread of the infection) and spread estimation (to embrace the appropriate restrictions at the right time and place).

A. Vaccine Discovery

The invention of a vaccine is arguably the most prolific Covid-19 area of research. In fact, given its significance in putting an end to the limitations placed on people (lockdown, festivities, gatherings), the vaccine is the ultimate solution to the Covid-19 crisis. Unfortunately, the firms that use ML techniques in their vaccine development pipeline publish minimal data, if any [22]. A pioneering effort to use ML in the process of developing an effective Covid-19 vaccine was conducted in [23]. The authors trained an XGBoost model (a decision-tree-based ensemble ML algorithm) to predict the best candidate protein for the vaccine. Their work resulted in 6 potential candidates. The authors adopted an optimized ML-based technique to pick some suitable amino acid fibers to improve the composition of already discovered vaccines [24]. Separate studies [25-27] relied on Support Vector Machine (SVM) and other ML techniques in the same direction. SVM is a supervised learning ML algorithm generally used to classify data by figuring out the best hyper-plane that could separate them with the lowest error rate.

The most documented DL-based approach to explore the world of proteins is the AlphaFold project powered by Google DeepMind [28, 29]. It focuses on predicting the 3D structure of proteins based on their genetic sequence. During the first months of the pandemic, DeepMind published a prediction for many proteins related to SARS-CoV-2, the virus responsible for Covid-19. This work aimed to contribute to the understanding of the functioning of the virus.

B. Spread Estimation

The compartmental models are the most used models to forecast the spread of the pandemic [30]. They rely on stochastic frameworks to predict the right actions, decisions, and strategies to perform under any circumstances. In [31], the authors used this method to evaluate the impact of social distancing on the pandemic spread to predict the expected number of infections after the lockdown. A similar approach was used in [32] to predict the outbreak's growth in Wuhan, the Chinese city where the first Covid-19 cases were detected. The main recommendation of this work was that the lockdown and travel restrictions in China would help the control of the spread of the virus in the rest of the world. Compartmental models were further used in various studies [33-36] to predict the transmission of Covid-19 in Japan, Egypt, and other countries. Researchers used DL [37] and logistic growth [38] models to estimate the spread in China and South Korea. A complex approach was used in [39], where the authors used Markov Chain Monte Carlo and the number of confirmed deaths and cases to predict the reproduction rate. DL models were also used to predict a condition that refers to rapid breathing associated with Covid-19 infections. The authors in [40] relied on a bidirectional Gated Recurrent Unit (GRU) model and data coming from smartphones. Impressive results were provided by the authors in [41] by training a Deep Neural Network (DNN) on a sizeable real-time dataset collected from smartphones. A Long Short-Term Memory (LSTM) model was used in [42] to estimate the pandemic spread in Canada. The authors in [43] developed a technique that joined LSTM and GRU to design and train a model to measure the Covid-19 release and death cases. Table II sums up the primary discussed studies that used ML/DL tools for Covid-19 prevention.

TABLE II. ML STUDIES ON THE PREVENTION AGAINST COVID-19

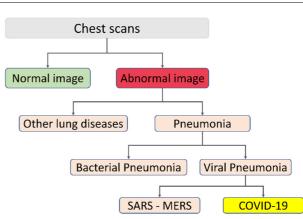| Role	Ref	Model
Vaccine discovery	[23]	XGBoost
	[24]	Combinational ML system
	[25, 26]	SVM
	[28, 29]	DNN
Spread estimation	[31-36]	Compartmental models
	[38]	Logistic growth models
	[39]	Markov Chain Monte Carlo
	[40]	Bidirectional GRU
	[42]	LSTM
	[43]	GRU + LSTM

III. ML METHODS FOR COVID-19 DETECTION

A prominent area where Artificial Intelligence (AI) methods shined during the Coronavirus pandemic is disease diagnosing and detecting. Several published papers highlight the use of DL models on Covid-19 identification. It is also worth noting that many of the studies we found are on CT chest scans [44]. The authors in [45] used a DL model to perform visual feature extraction from CT chest images. This project aimed to implement a classifier discriminating between Covid-19 and other pneumonia diseases. The researchers in [46] achieved promising results (86.30% accuracy) when developing a classification model to tackle the same problem. Authors in [47] proposed a similar work by implementing a DL
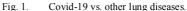
algorithm to detect Covid-19 infections through CT segmentation. A robust system was developed in [48], where the authors created and trained a Convolutional Neural Network (CNN) on lung images to detect Covid-19 cases. The authors in [8] proposed another CNN model trained on CT chest scans and aimed to distinguish people affected by Covid-19 from others. Authors in [47] took a different approach than we have seen hitherto. They presented an approach based on Deep Transfer Learning to build and train a model to classify chest CT images. Moreover, the researchers in [49] opted for Transfer Learning (TL) to design a classification model to detect patients infected by Covid-19 based on their CT chest scans. However, instead of applying the DenseNet201 pre-trained model as the basis of their architecture, the authors proposed the ResNet50 model. Also, based on CT chest images, the authors of [50] proposed a model based on ResNet-50 and DPN-92 for classifying infected and non-infected patients. TL was also proposed in [51], where the Inception pre-trained model was used to create a diagnosis model. The authors of [52] intended the same goal. However, their approach was slightly different since they trained their model on X-ray images. Similarly, the authors in [9] proposed a TL model based on pre-trained models such as VGG19, VGG16, ResNet, DenseNet, and InceptionV3. It should be mentioned that these DNN models are well-known for reaching state-of-the-art results for many applications, especially in computer vision classification. Authors in [53] proposed a new classification model based on the Random Forest algorithm and X-ray scans from people affected by Covid-19 and other lung diseases. An aggregated model was developed in [54], where pre-trained models such as ResNet, VGG, and AlexNet were used, along with SVM. The pre-trained models were used for feature extraction, while the SVM model handled the classification. In another direction, some studies were exploring a completely different path to achieve the same goal of detecting Covid-19, namely sound-based and video-based techniques. For example, the authors of [18] used videos of healthy people and patients affected by Covid-19 and pneumonia. These data then fit a pre-trained model (VGG-16) used in the prediction phase. Moreover, authors in [55, 56] shown that respiratory sounds (mainly cough and breathing sounds) could be employed to detect people affected by Covid-19. The authors of [55] built binary classifiers for distinguishing Covid-19 positive patients from other respiratory diseases. Figure-1 illustrates the intersections between Covid-19 and other lung diseases and why it is difficult to distinguish one from the others.

IV. ML METHODS FOR COVID-19 TREATMENT

So far, this article explored the use of ML methods in the detection and prevention against Covid-19. However, the most straightforward and essential action to fight the outbreak is to assist the affected people to recover. It can eventually be solved by finding drugs that have a practical impact in fighting the new virus. This section will discover how ML models contributed to alleviating this problem and how efficient they were. Drug discovery is one of the trickiest and most challenging areas in medicine. Furthermore, it is known for taking many years to give valuable results.

Fig. 1.   Covid-19 vs. other lung diseases.

Many researchers have turned their attention towards DL techniques instead of traditional technologies to speed up this process. For instance, the authors in [57, 58] used LSTM, CNN, and other DL models in their quest to find acting antivirals among the existing known drugs. Several other studies focused on Biomedical Knowledge Graphs (BKGs), a technique used to detect similarities and relationships between drugs, viral proteins, and genes. Authors in [59] utilized BKGs to identify a promising treatment for Covid-19. Authors in [60] used the same technique to describe the relations between gene-disease pairs. The ML model showed relations between the viral protein and some previously known drugs, which led to the prediction of many potential drug candidates that could be effective against Covid-19 [61]. Similar work has been done in [62], in which the researchers used a DL model to find candidate drugs. A multitask ML model and more than 2000 human proteins were used in [63] to find relations between Covid-19 and Known Drug Targets (KDTs). The drug representation was extracted in [64] using a deep graph network and known biological interactions.

ML and DL techniques were also applied to narrow the possibilities in the molecular docking process, one of the most computationally expensive, yet frequently used techniques in drug design. The authors used a neural network to narrow 3 million candidates to 1000 [65]. A Random Forest model to identify 187 molecules in the coronavirus S-protein was proposed in [66]. Moreover, the ML framework to predict viral protein activities was developed in [67]. The ensemble model used in this framework studied 19 drugs and ranked them based on their capacity to prevent the coronavirus proteases. The authors also used docking to assess critical aptitude.

AI was also used to find new chemical compounds. The authors in [65] used the protein homology model along with 28 DL models and Reinforcement Learning (RL) techniques to evaluate drugs based on some predefined metrics. Furthermore, the authors in [68] used RL in their quest to discover new SARS drugs. More precisely, they used Deep Q Learning and evaluated the discovered molecules according to 3 predetermined factors.

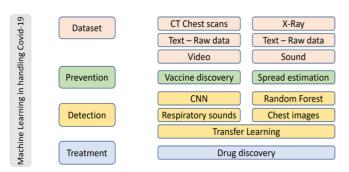Figure-2 summarizes the ML-based initiatives in handling the Covid-19 crisis.

Fig. 2.   ML solutions towards the Covid-19 pandemic.

V.   EVALUATION AND DISCUSSION

Many researchers have used AI as a forefront tool in the fight against Covid-19. Choosing this technology instead of traditional approaches was driven by the impact that DL techniques have gained during the last decade. This section addresses these contributions evaluates their impact, and presents their limitations. ML and DL approaches have contributed to the findings and improvements in the fight against the pandemic. We presented many examples in this paper to back up this claim. However, our study also shows that AI did not play a game-changing role in this fight. We dedicate this section to present our observations and explain why this is the case.

The companies that used ML tools in their research and development process generally publish little relevant information about their findings, particularly in drug discovery and treatment of the disease, although the biggest challenge the development of effective DL solutions faces is the lack of data. DL is known to be more effective when an abundant amount of (useful) data is available. However, as we have seen above, Covid-19 datasets are generally composed of a few samples. For instance, we have seen models built upon data collected from only 28 cases. The number of samples is not the only issue with Covid-19 datasets. The quality of these samples also raises another challenge. Many datasets are being collected from web portals, affecting the quality of images and the information credibility. Another issue regarding Covid-19 data consists of their non-representativeness. There is a substantial disparity in the number of tests between countries, classes, and cities in the same country, affecting the scaling of the trained models and resulting in inaccurate predictions. The challenge of extracting these samples could explain the lack of quality and quantity of data related to Covid-19. Besides, one should not forget that we are dealing with personal data, so privacy constraints come into play. One solution that has been used in lots of studies to handle this problem is data augmentation. Simultaneously with data augmentation, TL could present a great opportunity using the already available models and data developed for similar medical scenarios.

It is essential to note that the dilemma with data issues when using DL models is that we can have a well-trained model that gives high performance in the limited data but fails to scale when new situations arrive. Therefore, it is unlikely to get a solution applied in the real world. Our study shows that most of the contributions are generally about predicting the spread of the pandemic or detecting the virus. There are relatively few studies in the drug and vaccine discovery area. This is in line with the previous remark about real-world implementations since these areas present a more significant risk and therefore do not allow a margin of error to play with non-scalable and non-precise models. Our study also revealed that the primary source of Covid-19 data is in the form of images. Therefore, CNNs constitute the basis for many propositions. There is also the issue regarding privacy protection and the issue of the interpretability of DL models and how much we could trust their predictions in critical situations.

Table III exhibits the problems of existing DL/ML approaches when applied to the Covid-19 pandemic.

TABLE III.   PROBLEMS WITH THE USE OF AI REGARDING COVID-19

Problem	Details	Proposed solutions
Competition	Closed models and datasets	Open-source initiatives
Data	Number of samples	Open-source initiatives
	Quality of samples	Data augmentation
	Non-representativeness	TL
Model	Scalability	Data quantity and quality
	Privacy	Differential privacy
	Interpretability	More research

VI.   CONCLUSION

According to the official public service announcement on Coronavirus from the WHO, Covid-19 caused almost 5 million deaths while there are more than 238 million confirmed cases. This enormously heavy toll has generated the commitment of the research and development communities to alleviate this crisis. It turned out that ML and DL tools have been widely used in many applications during the last decade, including healthcare applications [69, 70]. To render valuable guidance to researchers in evaluating the use of ML/DL models against Covid-19, this paper studied the available Covid-19 datasets, the indispensable fuel for any DL-based approach. Then, the article reviewed how ML/DL tools were used in preventing, diagnosing and treating/curing the Covid-19 virus. We presented the models that advanced to detect the virus and then find drugs to treat the disease.

Finally, the last section of the paper is devoted to the thorough evaluation of these techniques, along with a discussion about their effectiveness in fighting Covid-19. The results that came from this study reveal that ML/DL techniques are not mature enough to be used in critical situations when people's lives are at stake, mainly due to the scarcity of the required medical data that are needed to train the models. The quality of the available data samples often raises some concerns. We proposed several potential solutions to remedy these circumstances, such as encouraging open-source initiatives, data augmentation, TL, and differential privacy.

REFERENCES

[1] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," *arXiv:2003.11597 [cs, eess, q-bio]*, Mar. 2020, Accessed: Dec. 13, 2021. [Online]. Available: http://arxiv.org/abs/2003.11597.

[2] 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: Dec. 13, 2021. [Online]. Available: http://arxiv.org/abs/2003.13865.

[3] M. Jun *et al.*, "COVID-19 CT Lung and Infection Segmentation Dataset." Zenodo, Apr. 20, 2020, https://doi.org/10.5281/zenodo.3757476.

[4] The National Library of Medicine, "MedPix." https://medpix.nlm.nih.gov/home (accessed Dec. 13, 2021).

[5] "LUNA16 - Grand Challenge," *grand-challenge.org*. https://luna16.grand-challenge.org/ (accessed Dec. 13, 2021).

[6] D. J. Bell, "COVID-19 | Radiology Reference Article," *Radiopaedia*. https://radiopaedia.org/articles/covid-19-4 (accessed Dec. 13, 2021).

[7] 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.

[8] 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.

[9] 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.

[10] A. Sharma, S. Rani, and D. Gupta, "Artificial Intelligence-Based Classification of Chest X-Ray Images into COVID-19 and Other Infectious Diseases," *International Journal of Biomedical Imaging*, vol. 2020, Oct. 2020, Art. no. e8889023, https://doi.org/10.1155/2020/8889023.

[11] J. Shuja, E. Alanazi, W. Alasmary, and A. Alashaikh, "COVID-19 open source data sets: a comprehensive survey," *Applied Intelligence*, vol. 51, no. 3, pp. 1296–1325, Mar. 2021, https://doi.org/10.1007/s10489-020-01862-6.

[12] A. Signoroni *et al.*, "BS-Net: Learning COVID-19 pneumonia severity on a large chest X-ray dataset," *Medical Image Analysis*, vol. 71, Jul. 2021, Art. no. 102046, https://doi.org/10.1016/j.media.2021.102046.

[13] "COVID Live Update: 270,678,636 Cases and 5,325,414 Deaths from the Coronavirus - Worldometer." https://www.worldometers.info/coronavirus/ (accessed Dec. 13, 2021).

[14] Datopian, "Novel Coronavirus 2019," *DataHub*. https://datahub.io/core/covid-19 (accessed Dec. 13, 2021).

[15] B. Xu *et al.*, "Epidemiological data from the COVID-19 outbreak, real-time case information," *Scientific Data*, vol. 7, no. 1, Mar. 2020, Art. no. 106, https://doi.org/10.1038/s41597-020-0448-0.

[16] B. Xu and M. U. G. Kraemer, "Open access epidemiological data from the COVID-19 outbreak," *The Lancet. Infectious Diseases*, vol. 20, no. 5, May 2020, Art. no. 534, https://doi.org/10.1016/S1473-3099(20)30119-5.

[17] A. J. Kucharski *et al.*, "Early dynamics of transmission and control of COVID-19: a mathematical modelling study," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 553–558, May 2020, https://doi.org/10.1016/S1473-3099(20)30144-4.

[18] C. McDermott, M. Lacki, B. Sainsbury, J. Henry, M. Filippov, and C. Rossa, "Sonographic Diagnosis of COVID-19: A Review of Image Processing for Lung Ultrasound," *Frontiers in Big Data*, vol. 4, Mar. 2021, Art. no. 612561, https://doi.org/10.3389/fdata.2021.612561.

[19] "covid19_ultrasound/data at master jannisborn/covid19_ultrasound," *GitHub*. https://github.com/jannisborn/covid19_ultrasound (accessed Dec. 13, 2021).

[20] "Project Coswara | IISc." https://coswara.iisc.ac.in (accessed Dec. 13, 2021).

[21] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "COVID-19 cough classification using machine learning and global smartphone recordings," *Computers in Biology and Medicine*, vol. 135, Aug. 2021, Art. no. 104572, https://doi.org/10.1016/j.compbiomed.2021.104572.

[22] N. El-Rashidy *et al.*, "Comprehensive Survey of Using Machine Learning in the COVID-19 Pandemic," *Diagnostics*, vol. 11, no. 7, Jul. 2021, Art. no. 1155, https://doi.org/10.3390/diagnostics11071155.

[23] E. Ong, M. U. Wong, A. Huffman, and Y. He, "COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning," *Frontiers in Immunology*, vol. 11, 2020, Art. no. 1581, https://doi.org/10.3389/fimmu.2020.01581.

[24] S. Deoras, "How ML Is Assisting In Development Of Covid-19 Vaccines," *Analytics India Magazine*, Jul. 28, 2020. https://analyticsindiamag.com/how-ml-is-assisting-in-development-of-covid-19-vaccines/ (accessed Dec. 13, 2021).

[25] M. S. Rahman *et al.*, "Epitope-based chimeric peptide vaccine design against S, M and E proteins of SARS-CoV-2, the etiologic agent of COVID-19 pandemic: an in silico approach," *PeerJ*, vol. 8, Jul. 2020, Art. no. e9572, https://doi.org/10.7717/peerj.9572.

[26] M. Prachar *et al.*, "COVID-19 Vaccine Candidates: Prediction and Validation of 174 SARS-CoV-2 Epitopes," Mar. 2020, https://doi.org/10.1101/2020.03.20.000794.

[27] 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.

[28] "AlphaFold: Using AI for scientific discovery," *Deepmind*. https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery (accessed Dec. 13, 2021).

[29] "Scientists Use Cloud-Based Supercomputing and AI to Develop COVID-19 Treatments and Vaccine Models," *Hospimedica.com*, Sep. 16, 2020. https://www.hospimedica.com/covid-19/articles/294784537/scientists-use-cloud-based-supercomputing-and-ai-to-develop-covid-19-treatments-and-vaccine-models.html (accessed Dec. 13, 2021).

[30] N. Wang, Y. Fu, H. Zhang, and H. Shi, "An evaluation of mathematical models for the outbreak of COVID-19," *Precision Clinical Medicine*, vol. 3, no. 2, pp. 85–93, Jun. 2020, https://doi.org/10.1093/pcmedi/pbaa016.

[31] M. Liu, R. Thomadsen, and S. Yao, "Forecasting the spread of COVID-19 under different reopening strategies," *Scientific Reports*, vol. 10, no. 1, Nov. 2020, Art. no. 20367, https://doi.org/10.1038/s41598-020-77292-8.

[32] M.-T. Li *et al.*, "Analysis of COVID-19 transmission in Shanxi Province with discrete time imported cases," *Mathematical biosciences and engineering: MBE*, vol. 17, no. 4, pp. 3710–3720, May 2020, https://doi.org/10.3934/mbe.2020208.

[33] A. R. Ives and C. Bozzuto, "Estimating and explaining the spread of COVID-19 at the county level in the USA," *Communications Biology*, vol. 4, no. 1, pp. 1–9, Jan. 2021, https://doi.org/10.1038/s42003-020-01609-6.

[34] M. A. Ibrahim and A. Al-Najafi, "Modeling, Control, and Prediction of the Spread of COVID-19 Using Compartmental, Logistic, and Gauss Models: A Case Study in Iraq and Egypt," *Processes*, vol. 8, no. 11, Nov. 2020, Art. no. 1400, https://doi.org/10.3390/pr8111400.

[35] T. Kuniya, "Prediction of the Epidemic Peak of Coronavirus Disease in Japan, 2020," *Journal of Clinical Medicine*, vol. 9, no. 3, Mar. 2020, Art. no. 789, https://doi.org/10.3390/jcm9030789.

[36] T.-M. Chen, J. Rui, Q.-P. Wang, Z.-Y. Zhao, J.-A. Cui, and L. Yin, "A mathematical model for simulating the phase-based transmissibility of a novel coronavirus," *Infectious Diseases of Poverty*, vol. 9, no. 1, Feb. Art. no. 24, 2020, https://doi.org/10.1186/s40249-020-00640-3.

[37] P. Boldog, T. Tekeli, Z. Vizi, A. Denes, F. A. Bartha, and G. Rost, "Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China," *Journal of Clinical Medicine*, vol. 9, no. 2, Feb. 2020, Art. no. 571, https://doi.org/10.3390/jcm9020571.

[38] J. Hilton and M. J. Keeling, "Estimation of country-level basic reproductive ratios for novel Coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices," *PLOS Computational Biology*, vol. 16, no. 7, 2020, Art. no. e1008031, https://doi.org/10.1371/journal.pcbi.1008031.

[39] S. Y. Lee, B. Lei, and B. Mallick, "Estimation of COVID-19 spread curves integrating global data and borrowing information," *PLOS ONE*, vol. 15, no. 7, 2020, Art. no. e0236860, https://doi.org/10.1371/journal.pone.0236860.

[40] Y. Wang *et al.*, "Unobtrusive and Automatic Classification of Multiple People's Abnormal Respiratory Patterns in Real Time Using Deep Neural Network and Depth Camera," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8559–8571, Sep. 2020, https://doi.org/10.1109/JIOT.2020.2991456.

[41] C. T. Rentsch *et al.*, "Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: A nationwide cohort study," *PLOS Medicine*, vol. 17, no. 9, 2020, Art. no. e1003379, https://doi.org/10.1371/journal.pmed.1003379.

[42] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons & Fractals*, vol. 135, Jun. 2020, Art. no. 109864, https://doi.org/10.1016/j.chaos.2020.109864.

[43] S. Dutta and S. K. Bandyopadhyay, "Machine Learning Approach for Confirmation of COVID-19 Cases: Positive, Negative, Death and Releas." 2020, https://doi.org/10.1101/2020.03.25.20043505.

[44] V. Kumar Singh, M. Abdel-Nasser, N. Pandey, and D. Puig, "LungINFseg: Segmenting COVID-19 Infected Regions in Lung CT Images Based on a Receptive-Field-Aware Deep Learning Framework," *Diagnostics*, vol. 11, no. 2, Feb. 2021, Art. no. 158, https://doi.org/10.3390/diagnostics11020158.

[45] M. Zietz, J. Zucker, and N. P. Tatonetti, "Associations between blood type and COVID-19 infection, intubation, and death," *Nature Communications*, vol. 11, no. 1, Nov. 2020, Art. no. 5761, https://doi.org/10.1038/s41467-020-19623-x.

[46] V. Lopez and M. Cukic, "A dynamical model of SARS-CoV-2 based on people flow networks," *Safety Science*, vol. 134, Feb. 2021, Art. no. 105034, https://doi.org/10.1016/j.ssci.2020.105034.

[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. Ahuja, B. K. Panigrahi, N. Dey, V. Rajinikanth, and T. K. Gandhi, "Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices," *Applied Intelligence*, vol. 51, no. 1, pp. 571–585, Jan. 2021, https://doi.org/10.1007/s10489-020-01826-w.

[49] A. Sedik *et al.*, "Deploying Machine and Deep Learning Models for Efficient Data-Augmented Detection of COVID-19 Infections," *Viruses*, vol. 12, no. 7, Jul. 2020, Art. no. 769, https://doi.org/10.3390/v12070769.

[50] B. Wang *et al.*, "AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system," *Applied Soft Computing*, vol. 98, Jan. 2021, Art. no 106897, https://doi.org/10.1016/j.asoc.2020.106897.

[51] 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.

[52] D. Dansana *et al.*, "Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm," *Soft Computing*, Aug. 2020, https://doi.org/10.1007/s00500-020-05275-y.

[53] F. Shi *et al.*, "Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification," *Physics in Medicine & Biology*, vol. 66, no. 6, Mar. 2021, Art. no. 065031, https://doi.org/10.1088/1361-6560/abe838.

[54] 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*, vol. 5, no. 4, pp. 643–651, 2020.

[55] C. Brown *et al.*, "Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data," in *26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, San Diego, CA, USA, Aug. 2020, pp. 3474–3484, https://doi.org/10.1145/3394486.3412865.

[56] M. Faezipour and A. Abuzneid, "Smartphone-Based Self-Testing of COVID-19 Using Breathing Sounds," *Telemedicine journal and e-health*, vol. 26, no. 10, pp. 1202–1205, Oct. 2020, https://doi.org/10.1089/tmj.2020.0114.

[57] N. Bung, S. R. Krishnan, G. Bulusu, and A. Roy, "De novo design of new chemical entities for SARS-CoV-2 using artificial intelligence," *Future Medicinal Chemistry*, vol. 13, no. 6, pp. 575–585, Mar. 2021, https://doi.org/10.4155/fmc-2020-0262.

[58] K. Avchaciov, O. Burmistrova, and P. Fedichev, *AI for the repurposing of approved or investigational drugs against COVID-19*. 2020.

[59] P. Richardson *et al.*, "Baricitinib as potential treatment for 2019-nCoV acute respiratory disease," *Lancet*, vol. 395, no. 10223, pp. e30–e31, 2020, https://doi.org/10.1016/S0140-6736(20)30304-4.

[60] J. Fauqueur, A. Thillaisundaram, and T. Togia, "Constructing large scale biomedical knowledge bases from scratch with rapid annotation of interpretable patterns," *arXiv:1907.01417 [cs]*, Jul. 2019, Accessed: Dec. 13, 2021. [Online]. Available: http://arxiv.org/abs/1907.01417.

[61] Y. Ge *et al.*, "A data-driven drug repositioning framework discovered a potential therapeutic agent targeting COVID-19," Mar. 2020, https://doi.org/10.1101/2020.03.11.986836.

[62] X. Zeng *et al.*, "Repurpose Open Data to Discover Therapeutics for COVID-19 Using Deep Learning," *Journal of Proteome Research*, vol. 19, no. 11, pp. 4624–4636, Nov. 2020, https://doi.org/10.1021/acs.jproteome.0c00316.

[63] C. Loucera, M. Esteban-Medina, K. Rian, M. M. Falco, J. Dopazo, and M. Pena-Chilet, "Drug repurposing for COVID-19 using machine learning and mechanistic models of signal transduction circuits related to SARS-CoV-2 infection," *Signal Transduction and Targeted Therapy*, vol. 5, Dec. 2020, Art. no. 290, https://doi.org/10.1038/s41392-020-00417-y.

[64] K. Hsieh *et al.*, "Drug Repurposing for COVID-19 using Graph Neural Network with Genetic, Mechanistic, and Epidemiological Validation," Sep. 2020. Accessed: Dec. 13, 2021. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2020arXiv200910931H.

[65] A. Zhavoronkov, "Artificial Intelligence for Drug Discovery, Biomarker Development, and Generation of Novel Chemistry," *Molecular Pharmaceutics*, vol. 15, no. 10, pp. 4311–4313, Oct. 2018, https://doi.org/10.1021/acs.molpharmaceut.8b00930.

[66] R. Batra, H. Chan, G. Kamath, R. Ramprasad, M. J. Cherukara, and S. K. R. S. Sankaranarayanan, "Screening of Therapeutic Agents for COVID-19 Using Machine Learning and Ensemble Docking Studies," *The Journal of Physical Chemistry Letters*, vol. 11, no. 17, pp. 7058–7065, Sep. 2020, https://doi.org/10.1021/acs.jpclett.0c02278.

[67] R. Mall, A. Elbasir, H. A. Meer, S. Chawla, and E. Ullah, "Data-Driven Drug Repurposing for COVID-19," Jul. 2020, https://doi.org/10.26434/chemrxiv.12661103.v1.

[68] B. Tang, F. He, D. Liu, M. Fang, Z. Wu, and D. Xu, "AI-aided design of novel targeted covalent inhibitors against SARS-CoV-2," Mar. 2020, https://doi.org/10.1101/2020.03.03.972133.

[69] B. Trstenjak, D. Donko, and Z. Avdagic, "Adaptable Web Prediction Framework for Disease Prediction Based on the Hybrid Case Based Reasoning Model," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1212–1216, Dec. 2016, https://doi.org/10.48084/etasr.753.

[70] K. Koklonis, M. Sarafidis, M. Vastardi, and D. Koutsouris, "Utilization of Machine Learning in Supporting Occupational Safety and Health Decisions in Hospital Workplace," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7262–7272, Jun. 2021, https://doi.org/10.48084/etasr.4205.