A Comprehensive Survey on the Data-Driven Approaches used for Tackling the COVID-19 Pandemic

WALID SALAMEH[1,*], OLA M. SURAKHI[2], MOHAMMAD Y. KHANAFSEH[3]
[1]Computer Science Department,
Princess Sumaya University for Technology,
Amman 11941,
JORDAN

[2]Computer Science Department,
American University of Madaba,
Madaba 11821,
JORDAN

[3]Computer Science Department,
Birzeit University,
West Bank PO Box 14,
PALESTINE

Abstract: - The current evolution of Artificial Intelligence (AI) is fueled by the massive data sources generated by the Internet of Things (IoT), social media, and a diverse range of mobile and web applications. Machine learning (ML) and deep learning become the key to analyzing these data intelligently and developing complementary intelligent data-driven services in the healthcare sector. The world witnessed many AI-enabled tools that contributed to fighting against the COVID-19 pandemic and accelerated with unprecedented accuracy the development and the deployment of many countermeasures. The main objective of this study is to provide a comprehensive survey on the role of AI and ML methods in the healthcare sector. The study offers cases on how AI/ML can arm the world against future pandemics. Specifically, the study presents all available datasets, the main research problems related to COVID-19, and the solutions that AI and ML technologies offer. Finally, based on the analysis of the current literature, the limitations and open research challenges are highlighted. Our findings show that AI and ML technologies can play an essential role in COVID-19 forecasting, prediction, diagnosis, and analysis. In comparison, most of the previous works did not deploy a comprehensive framework that integrates the ML and DL with network security. This work emphasizes the mandate of including network security in all COVID-19 applications and providing complete and secure healthcare services.

Key-Words: - Artificial Intelligence, Artificial Neural Network, COVID-19, Data-driven, Diagnosis, Internet of Things, Machine Learning, Treatment.

Received: August 22, 2023. Revised: February 9, 2024. Accepted: March 6, 2024. Published: April 25, 2024.

1 Introduction

The first appearance of COVID-19 was in Wuhan City in China at the end of December 2019. COVID-19 was declared by the World Health Organization (WHO) to be widespread worldwide in March 2020, [1]. COVID-19 exponentially spread worldwide and exceedingly influenced the healthcare framework in many countries, increasing the confirmed positive cases and death rate, [2], [3]. Several countries started to force restrictions on the citizens to overcome the pandemic and stop the virus from spreading, such as lockdown and gathering restrictions, school and airport closures, and more, [4], [5].

Until now, a new generation of COVID-19 virus appears with no specific medications that can deal with it precisely. Some of these medications have been confirmed by the World Health Organization (WHO), [6], more efforts are still needed to provide a reliable solution to the COVID-19 Methods. An orderable solution is required to estimate future cases, analyze the effect of different features that help increase/decrease the spreading, and help develop a medical treatment to slow down the spreading.

Recently, researchers started to utilize Artificial Intelligence (AI) and Machine Learning (ML) algorithms to diagnose COVID-19 and investigate the effect of various health policies on COVID-19 spreading, [7], [8], [9]. The raw historical data for COVID-19 can be classified into three main groups: textual information, speech data, and image data. These data are used widely to diagnose COVID-19, predict its spreading, vaccine discovery, sentiment analysis of false news about COVID-19, etc. AI provides many methods that can help diagnose, trace, and forecast the spread of the COVID-19 virus based on these data types.

This paper focuses on analyzing how AI and ML technologies can be employed to provide solutions and assist in developing public health policies to mitigate the war on the COVID-19 pandemic from different perspectives. Studying the COVID-19 pandemic and utilizing AI and ML methods with the available datasets can bring forth numerous benefits: 1) Early Detection and Diagnosis. 2) Drug Discovery and Vaccine Development. 3) Optimizing Healthcare Resources. 4) Improving Public Health Policies. And 5) Continuous Learning and Improvement.

There are other survey and review papers on the same topic; we analyzed them and differentiated them from our article. This paper provides a taxonomy to identify four COVID-19 main research problems to which ML algorithms and methods are applied. Based on this taxonomy, we present a review of the related published papers. Then, we offer the limitations and challenges that impact this research area. Finally, we give some suggestions for improving the performance of using ML methods in the COVID-19 application management.

The main contributions of this paper can be summarized as follows:

- First, we overview the COVID-19 pandemic and the available datasets. The dataset is divided into three main categories: Images, Sound, and Textual dataset.
- Then, we listed the main research problems of the COVID-19 domain where AI and ML methods can be applied.
- A summary of the state-of-the-art works that utilize ML methods to solve the COVID-19 pandemic is given.
- Finally, we explore the services and solutions that AI and ML technologies offer. Along with the challenges and limitations in the same domain.

2 Related Survey Papers

A few research survey papers have been published that discussed the application of ML methods in COVID-19 applications. In this section, we summarized the recent works and differentiate them from our paper in three different aspects: 1) the COVID-19 applications that have been used on the ML methods, 2) the COVID-19 datasets used for a published reviewed paper that has been reviewed in the survey paper and 3) determine whether the authors highlight the type of COVID-19 features used for the task of modeling and forecasting in the published papers as shown in Table 1.

The authors in [10], surveyed the role of AI and ML in fighting against the COVID-19 pandemic. The authors discussed five main applications for the COVID-10 area and analyzed primary datasets in the same place. However, the effect of health policies and features for each work has not been studied. [11], reviewed the research papers that have been published in Science Direct, Springer, Hindawi, and MDPI in the area of COVID-19 and ML methods. The authors highlighted the findings of each work. In [12], the authors identified the role of different technologies in tackling COVID-19. Other authors reviewed Pubmed, Scopus, and Google databases for the research that applied AI technology to COVID-19. They suggested that AI techniques can predict the number of positive cases, [13]. In some COVID-19 applications, no related papers were surveyed on that domain. [14], studied some papers applying AI and ML to predict COVID-19 spread. [15], presented a review of the research papers that applied ML to one application of COVID-19, indicating the number of confirmed cases. [16], overviewed published articles that used ML and DL mechanisms for COVID-19 diagnosis. A similar overview is performed, [17].

The current study stands out from the available literature in several ways. Firstly, unlike existing papers, it introduces a taxonomy for systematically reviewing research papers concerning COVID-19 applications, datasets, and features utilized in machine learning (ML) mechanisms for modeling COVID-19 data. While prior literature may touch upon one or more of these aspects, none comprehensively categorize and review papers based on all three dimensions. Moreover, this study conducts a literature search specifically targeting relevant studies published in 2020 and 2021, using key terms such as Machine learning, Deep learning, Artificial Intelligence, COVID-19, dataset, and forecasting, and diagnosis. By focusing on recent research, it ensures relevance and timeliness in its review. Furthermore, the study goes beyond mere categorization by applying different ML and DL algorithms and methods to COVID-19 management across various domain applications. It meticulously classifies these works based on application types, providing detailed descriptions of the datasets used for modeling. Additionally, the study analyzes the performance results of these papers, emphasizing the importance of dataset features in influencing ML efficacy. By synthesizing insights from application types, datasets, and ML methods, the study aims to offer recommendations and suggestions for health sector officials to devise effective strategies in combatting the pandemic. This holistic approach distinguishes the current study as a valuable contribution to the field of COVID-19 research and ML applications.

Table 1. Related Survey Papers

Survey Paper	COVID-19 Research Area	COVID-19 Datasets	COVID-19 Modelling Features
[10]	Diagnosis and Identification, Forecasting the spread, Association between COVID-19 infection and patient characteristics, Treatment development, Supporting applications	√	-
[11]	Diagnosis and Identification	-	-
[12]	Diagnosis and Identification, Forecasting the spread	-	-
[13]	Early detection and diagnosis of the infection, Monitoring the treatment, Contact tracing, Projection of cases and mortality, Development of drugs and vaccines, Reducing the workload of healthcare workers, Prevention of the disease	-	-
[14]	Outbreak prediction, Virus spreading, Diagnosis and treatment, vaccine discovery	-	-
[15]	Forecasting the spread	-	-
[16]	Diagnosis and Identification	√	-
[17]	Forecasting the spread	√	-
[18]	Diagnosis and Identification, virology and pathogenesis, drug and vaccine development, Forecasting the spread	√	-
[19]	Diagnosis and Identification, COVID-19 emotional and sentiment analysis from social media, knowledge-based discovery and semantic analysis from the collection of scholarly articles covering COVID-19, Forecasting the spread	-	-

3 COVID-19 Pandemic

3.1 Taxonomy of COVID-19 Research Problems of This Research Paper

The COVID-19 pandemic put the world and public health in a severe and critical challenge. As positive cases increased exponentially, the COVID-19 pandemic changed human life worldwide and influenced the economy and citizens' social life. A reliable estimation of the number of instances and feasible strategies to control the pandemic is urgently needed to provide a potential solution for the outbreak. AI and ML have been applied recently in many application areas of COVID-19 to help make informed decisions by policymakers to control the pandemic. The research investigation in this domain can be done under three primary methodologies as shown in Figure 1.

Tackling any research problem employing machine learning methodologies follows a systematic approach. Firstly, researchers delineate clear research questions pertinent to the domain, aligning with the objectives of understanding, diagnosing, treating, or preventing the spread of COVID-19. Next, they meticulously gather relevant datasets encompassing images, sound recordings, textual information, or other pertinent sources. Subsequently, employing sophisticated analytical techniques, researchers dissect the data, identifying patterns, trends, and correlations crucial for addressing the research questions at hand. Through feature extraction, key insights are distilled, laying the groundwork for the subsequent modeling phase. Finally, employing machine learning algorithms, these insights are encapsulated into robust models capable of providing predictive, diagnostic, or prescriptive solutions, thereby contributing to the ongoing efforts to combat the COVID-19 pandemic effectively.

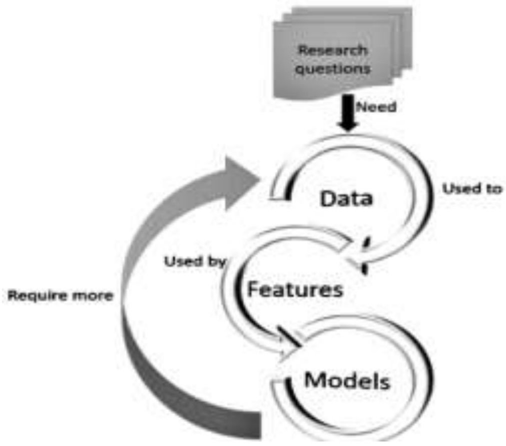

Fig. 1: Taxonomy of COVID-19 Research Problems of This Paper

3.2 COVID-19 Datasets

No country worldwide succeeded in providing reliable data about COVID-19 information. Most of the numerous deaths and people with COVID-19 are left unreported. Thus, the lack of accurate data has become a severe problem in the COVID-19 domain research. The research in this area integrates the data provided from different resources to create a valuable dataset that can be used by ML and DL methods to be applied in the COVID-19 domain system. In general, three types of COVID-19 datasets are collected from various resources: images, speech, and text.

Images Dataset

The image dataset comes in three forms: CT scan, X-ray, and Ultrasound, collected by the screening tools and devices for the patient with COVID-19. Most medical images need to be pre-processed before using them in the modeling process. This involves two main steps: segmentation and augmentation. The segmentation process highlights the infected area (the region of interest) to differentiate it from the other normal areas. The augmentation filters the image and transforms it into another format to increase the size of the dataset, [20].

The most common way to diagnose patients with respiratory diseases is by taking a medical image of the chest. In the case of COVID-19 infection, the chest area will appear to have a tissue or fluid that prevents gas exchange or the appearance of a shadow area in the X-ray image, [21]. Using this type of dataset can help perform the classification task on the patient's medical image to determine if he is infected with COVID-19 or not, or to evaluate the progress and changes during his infection time. Examples of the medical images dataset which is available online are a CT-scan dataset with 125 CT-chest images, [22], an X-ray dataset with 13,800 images for 13,000 patients collected from several online repositories, [23] and an Ultrasound dataset with 64 videos, [24].

Speech Dataset

This type of dataset is the least available online due to the lack of collected data, with a significant challenge in managing it. This dataset type comes in three forms: 1) the patient’s cough sound, 2) the patient's breathing sound, and 3) the patient’s voice. These data can be used to classify if the patient is infected with COVID-19 or not, determine his health status, and trace the patient’s medical progress, [25]. An online example of the sound dataset is Coswara, which contains a collection of recordings of breath and deep coughs of the patients with more information about the health status and some geographical information for each one, [26]. Another example of the online sound dataset for COVID-19 is SACRO, developed from South Africa through smartphones, [27]. Several studies showed that a sound dataset helps diagnose and identify the COVID-19 virus, [28], [29], [30]. A unique number of features can be extracted from the sound dataset. These features do not overlap with the other respiratory infection features. Thus, it can train complex ML models for accurate diagnosis and prediction. The WHO reported that coughing is a symptom for the COVID-19 patient and is considered the primary reason for spreading the virus, [31].

Textual Dataset

The textual datasets collected during the pandemic and found in the literature can be categorized into different groups. 1) Data about the patients and their symptoms (statistical data), 2) data about several reported cases which comes in a time-series format, 3) mobility data related to the transmission of the citizens, 4) policy data related to the government policies and rules that have been issued during the pandemic, and 5) data from the social media which reflect the semantic of the humans. The most popular textual dataset reported the number of people with COVID-19, the number of deaths, the number of recovered, the positive rate, etc. Almost every country worldwide has developed a statistical dataset to report these numbers. For example, John Hopkins University developed a real-time dataset to aggregate data about COVID-19 for the researcher for analyzing and modelling. The dataset is available online and can be accessed by anyone, [32]. Research in [33], investigated the effect of human mobility in China on the spatial distribution of COVID-19 using the mobility data. Another dataset developed by Oxford University contains data from various countries about the rules and restrictions that the government has applied to control the virus spreading, [34]. This data includes essential features such as travel limitation, social distancing, face masking, cancellation of public events, etc., which are necessary to show its influence on the virus transmission. Another dataset is developed from a tweet of 2500 participants for the semantic analysis to understand the emotions and worries toward the COVID-19 pandemic, [35]. A similar dataset was created from Arabic tweets, where the authors collected about 2,433,660 tweets from Arab countries to analyze the public response and behavior toward the pandemic, [36].

3.3 The Main Research Problems of COVID-19

Artificial Intelligence (AI) technology and machine learning science have been applied in many areas. The healthcare system is one of the most important fields where AI technology can provide many services to enhance its performance. The COVID-19 pandemic opened horizons for researchers to apply state-of-the-art technologies to improve health services and find a solution to stop the spreading of this pandemic. This section summarizes the various areas where AI technology can be applied in fighting against the COVID-19 pandemic.

Diagnosis

The primary diagnosis of patients with COVID-19 methods is 1) molecular diagnostic, which includes RT-PCR test, and 2) medical images, [37]. The diagnosis can be performed either in the clinic or laboratory. These diagnoses result in different kinds of data such as text, speech, and image that can be utilized in the AI and ML methods to provide solutions and services for the healthcare system in the COVID-19 domain. For example, ML methods like Convolutional Neural Networks (CNN) can analyze medical images (CT scan and X-ray) to extract features and differentiate between COVID-19 and other diseases. On the other hand, the cough sound and the patient's breathing data can detect the COVID-19 symptoms, [38].

These speech datasets include a unique feature that confirms whether a patient is infected with COVID-19 or not. The ML methods with the deep architecture can model speech datasets and provide an accurate classification for COVID-19-positive cases. The use of such techniques is urgent to fast and automate the diagnosis of the COVID-19 virus.

Treatments

The exponential growth in the number of positive cases of COVID-19 and deaths put countries around the world in a critical situation. Clinical and scientists worldwide have been urged to work hard to search for a vaccine or drugs with efficient operations. The traditional way of drug development is a complex process that may need a lot of time. This contradicts the virus’s rapid spread, which requires a solution to curb it. AI technology and ML methods can speed up drug and vaccine discovery. Some of the operations where AI and ML technologies can be applied in drug development are: representing the relationship between entities such as pair of genes and the interactions between molecules, discovering new chemical compounds to identify COVID-19, predicting the best protein that could serve as an effective vaccine and more.

Forecasting

Since its appearance in 2019, the number of confirmed cases of COVID-19 has been increasing exponentially. It reached 290 million in Jan 2022. Determining the number of future positive cases is the primary key to planning against the pandemic. The statistical methods and ML techniques can analyze the pandemic status and forecast the growth of the virus spreading. Some of the exciting measurements indicated in the COVID-19 domain include the total number of people with COVID-19, the infection rate, and the number of confirmed deaths. This data can help provide a general insight about the quality of the healthcare system in a specific country and track the infection-spreading parameters that cause the increase/decreasing of the transmission.

Tracing

The AI tools and ML techniques could provide solutions to avoid virus spread by tracking and screening its transmission. The use of smart devices such as mobile phones and sensors can accelerate the development of a monitor system that tracks patients through AI applications, [39], [40]. A large amount of data can be aggregated from these devices. These data can be modeled by ML methods and algorithms and classified into different categories such as mild, urgent, etc. A decision-making process can be adopted to help decide whether the patient needs intensive care or respiratory support, etc. The continuous monitoring of the patients can help reduce the number of patients visiting the hospital (for mild cases), which prevents the healthcare system from failing, [41]. A recent improvement in the AI application for the COVID-19 domain is the development of medical chatbots based on ML and Deep Learning (DL) models, [42], [43], [44]. The chatbots can assist patients by continuous answering and guiding in dealing with the disease’s potential problems.

3.4 Types of Features Included in the Datasets

Many studies have been proposed to analyze the association and correlation between the spreading of COVID-19 disease and other features. The determination of the future severity of the pandemic depends on the spreading speed relating to a set of factors that may accelerate the spreading or not.

In this paper, we classify the features into three groups) The patient characteristics and health status, 2) the environment and meteorological data, and 3) the mobility and country policy. Analyzing the correlation between these features and the risks of the COVID-19 outbreak may help to point out the main reason for the virus spreading. In this way, an effective solution could be provided to the government authorities to control the outbreak and minimize the load on the healthcare sector.

Patient Characteristics and Health Status

Several studies investigate the correlation between human characteristics and the risk of being infected with COVID-19 or not. This includes many features such as blood type, age, gender, smoking or not, obesity, health status and historical diseases, and more, [45], [46], [47], [48], [49]. For example, the authors in [46], studied the relation between the blood type and the risk of COVID-19 infection and found that the blood type of group A has a higher risk of getting an infection than another blood type. A higher risk of disease with the patient age is found by many researchers based on the analysis of the association between patient age and the number of positive cases and mortality, [50], [51]. The results showed that patients above 75 years old are at a higher risk of fatality. Other studies analyzed the correlation between COVID-19 infection rate, death rate, and patient gender. In [52], a higher death rate in males than females are reported. Also, the immune response and the infection rate are different in males and females due to the biological differences in features, [53]. COVID-19 is a respiratory disease, and as smoking is a bad habit that destroys the lungs and weakens the immune system, the smoking patient becomes at a higher risk than the non-smoker patient, [54], [55]. The same results are conducted for the patient's health status. It was found that patients with other severe diseases such as diabetes are at a higher risk of having a respiratory illness like COVID-19, [56].

Environment and meteorological data

Previous studies have proved that meteorological information is vital in spreading the COVID-19 virus. This includes addressing the relationship between the weather conditions and the infection rate. Previous studies showed a correlation between temperature and the infection rate, [57], [58]. Other features that influence the spread of the disease are population size and the country's location. It has been shown by researchers in [59] that the population size is a crucial transmissibility feature that causes an increase or decrease in the spreading speed of COVID-19. The higher the population size, the faster spreading of the virus, [57]. Many other features related to the environmental country conditions can influence the spreading of the virus, such as sea level air pollution rate, [60], [61], [62]. In [61], the authors applied a study on 65 different countries to investigate the effect of many environmental features (wind speed, sea pressure, rainfall, etc.) on spreading the virus. The results showed a strong correlation between environmental features and virus spreading.

Country Policy and Mobility

In the pandemic policy, the governments play a vital role in preventing the virus from spreading quickly by managing the right policy decisions. It became a challenge for the authorities to balance the need to control the quarantine and other aspects like economic and social. Therefore, most countries force policies and strategies on several levels in response to the spreading of the COVID-19 pandemic and to tackle the emerging situations,[63]. For example, some countries force travel restrictions, school closures, lockdowns (complete or partial), and more. The effect of these features on controlling the virus spreading will lead the authorities in the governments to determine the most effective interventions in containing the outbreak, [64].

Many organizations have started to record these data and provide it publicly on the websites such as the World Health Organization (WHO), [65], the World Meter Corona Virus Statistics website, [66], the Centers for Disease Control, and Prevention (CDC), [67] and more.

These data can help model the number of infected cases (People with COVID-19) based on their correlation with these features. The analysis of this kind of work will provide a strategy for the country to be followed to give a practical solution that can help control the spreading of the virus. The authors in [64] applied a regression analysis to identify the main features that affect the cases (infected and recovered) of COVID-19 in Europe, the United States, and China. They found that government actions such as border closures, full lockdowns, and a high rate of COVID-19 testing were not associated with statistically significant reductions in the number of critical cases or overall mortality. Another study investigated the effect of social distancing, border restrictions, quarantine, and isolation on the local transmission of the COVID-19 virus in Hong Kong. They found that viral transmission is reduced when forcing such policies, [68].

4 Modelling COVID-19 Pandemic

4.1 Taxonomy of Machine Learning and Deep Learning in COVID-19 Pandemic Management

In this section, we summarized the role of the ML and DL technologies in facing the COVID-19 outbreak. The summary is tabular with a taxonomy for the recent works that applied ML and DL methods in COVID-19 management. The taxonomy divided these works based on the COVID-19 application domain that has been used. It includes four main sections: 1) Diagnosis and Identification, 2) Treatments Development (Vaccine and Drugs), 3) Forecasting the Spread, and 4) Patient Tracing. Table 2 (Appendix) listed the works that applied ML an DL methods in the Diagnosis and Identification of COVID-19.

In the Diagnosis section, the ML and DL methods have been applied to classify the dataset used (medical images, sound, or textual) for the disease diagnoses or identify the relationship between the dataset and COVID-19 symptoms. The performance of each method is then evaluated based on statistical measures such as precision, recall, F-measure, and accuracy that assess the results. These models and techniques could provide a solution concerning the decision-making process to control the outbreak done by government officials.

In the treatment development section, AI technology and ML algorithms have been applied widely to develop a drug or vaccine against COVID-19. Using such methods can reduce the time and the cost needed to design a sophisticated development drug pipeline, making them more effective in identifying a new antiviral drug.

In a recent publication, the authors proposed a deep-learning model to predict a drug that targets the SARS-CoV-2-related proteins, [79]. Pham et al., used the DeepCE algorithm to predict a treatment for COVID-19 by repurposing the drug compound, [80]. In another study, an ML model is used to predict a new indication about the possible herbal and drug combinations based on several positive drug-disease associations, [81]. In general, the AL technology can provide solutions for tracing the drug development and proving its effectiveness against the COVID-19 virus, [82], [83]. A deep learning-based pipeline model has been developed to screen the small molecules against the virus, [84], while other authors predicted antiviral peptides using ML algorithms, [85].

On the other hand, the vaccine becomes the best solution to impact the pandemic. While many companies developed different vaccine components, the AL technology can be utilized to analyze the issues related to the efficiency of these vaccine candidates. This will include the analysis of virus immunity, the side effects based on the historical health information, manufacturing and storage, and more. All of this development will help the improvement of the vaccine to be more effective and safer. Some researchers used ML methods to predict the best protein that could be the most suitable for an effective vaccine, [86], [87]. In [86], the authors used the XGBoost model, and in [87], the authors utilized generative deep learning models.

However, the development companies published little information about the methodologies pipeline of the vaccine process with minimal information about how they integrate the ML in the development process.

Forecasting is the most popular application where ML methods have been applied. Since its appearance in 2019, people with COVID-19 have increased exponentially worldwide. So, it becomes essential to determine the future severity of the outbreak. This includes the analysis of the pandemic status using ML and DL methods to extract features that may lead to the virus spreading. Table 3 (Appendix) summarized some recent works that applied ML or DL methods and algorithms in the COVID-19 forecasting spread modeling.

ML methods are applied on mobile devices, like mobiles, in the tracing application to track the patients, [39], [40]. These devices can provide many services, such as patient monitoring, diagnosing, and screening. Based on the data aggregated by the smart devices, the AI technologies and ML methods can be applied to analyze data and provide a helpful decision-making service such as deciding if the patient is in urgent need of an Intensive Care Unit (ICU) admission, [101]. Some of the AI applications where ML is applied to provide decision-making services include:

1) Classify the patient status into light, medium, or severe, [102], [103], [104].

2) Monitoring the patient's symptoms, [105], [106], [107].

3) Chatbots provide useful information and guidance for the patients, [42], [43], [44].

4) Social media sentiment analysis of population realization towards COVID-19, [108], [109], [110].

4.2 Challenges and Limitations

This section highlights the critical challenges that can be concluded from analyzing the literary works.

1) Most literary works evaluate their results based on statistical measurements like accuracy, F-score, etc. According to this evaluation, the proposed model is authenticated, and the results are accepted. Few works show the level of uncertainty.

2) Most of the works explicitly explained the proposed model's methodology. Thus, the model is transparent.

3) Few works used a pre-trained model. The use of previous models increases generalization and supports the idea of transfer learning that produces a robust forecasting model.

4) The type of dataset used in the proposed model affects the complexity level. As more features are included in the training model, the complexity increases. Most Gene, image, and sound datasets include complicated features that need more analysis and pre-processing steps.

5) The dataset size used in the model training influences the forecasting accuracy very much. A more extensive dataset will result in more accurate and satisfactory results, [111]. A collaboration between all medical sectors worldwide is highly recommended to integrate all data sources and expand the existing dataset.

6) Using the traditional ML methods results in a simpler model with satisfactory results. In contrast, the DL methods are more complicated, which results in a complex model with multi-hidden layers that increase training complexity, [112]

7) Different ML methods have different prediction performances and, therefore, can be used in other classes of COVID-19 applications.

8) The size of the dataset greatly influences the ML method's performance. As the dataset size increases, the model performance will be improved. Most of the COVID-19 existing datasets are limited, and the performance of the ML model cannot be generalized to include all the pandemic aspects.

9) The hybridization of the ML model outperforms single forecasting models, [113], [114]. The behavior of the trained data can be learned by many models in ensemble learning (hybrid approach), which results in a compatible forecasting value with the observed ones.

10) Few studies focused on modeling the role of the vaccine and drug in combating the pandemic.

11) While every region has its specific weather conditions, a few works include geographical location and weather conditions in the virus spreading.

12) Different features influenced the spreading of the virus. Few of these works analyzed the correlation between the COVID-19 spread and infection and the other features. Most of the correlation analysis included the confirmed, dead, and recovered cases.

13) The number of works conducted in the European countries was found to be the most than that in the other countries.

14) Security and privacy issues are critical when dealing with the healthcare system, [115], [116]. Integrating AL technologies and security applications such as Blockchain technology is recommended to maintain patient privacy while providing a high-level service using AI and ML methods.

15) DL methods and algorithms are complex and need a high computational resource to model, process, and work with big data. Integrating Fog computing and Edge technology is recommended to handle this challenge.

16) In general, ML and DL methods are black-box models, [117]. The interpretation of model behavior regarding the choice of features and generated accuracy is essential to explain it to medical experts and decision-makers.

17) Most ML and DL proposed in the literature show promising performance and good results in the forecasting, prediction, diagnosis, and analysis. In reality, most of these models do not deploy. A comprehensive framework that integrates the ML and DL models with network security is required to include all the COVID-19 applications and provide a complete healthcare service.

5 Conclusions and Future Works

In this paper, we have conducted a comprehensive survey about the application of AI technology and ML methods for intelligent data analysis and applications to tackle the COVID-19 pandemic. We have discussed and analyzed how ML methods can be used to provide a solution to mitigate the impact of the pandemic. It was found that AI solutions-based ML algorithms can be used for diagnosis, treatment, forecasting, and monitoring applications. While the ML algorithm needs to be trained through solid knowledge related to application data, this paper also surveyed the real-time datasets that are provided as open-source for the COVID-19 domain. ML algorithms require training on relevant data to learn patterns and make predictions. In the context of the COVID-19 pandemic, this means that algorithms need to be trained on datasets containing information about the virus, its transmission, symptoms, diagnosis, treatment, and other related factors. Solid knowledge about these aspects is crucial for effective algorithm training. In addition to discussing the training of ML algorithms, the paper also explores real-time datasets that are openly accessible in the COVID-19 domain. These datasets provide up-to-date information about various aspects of the pandemic, such as infection rates, hospitalization numbers, testing data, genomic sequences of the virus, and vaccination statistics. By surveying real-time datasets, the paper ensures that the ML algorithms discussed in the study are not only trained on relevant historical data but also have access to the latest information about the pandemic. This allows for more accurate and timely predictions, analyses, and decision-making processes. Moreover, open-access datasets promote transparency, collaboration, and reproducibility in research, enabling other researchers and stakeholders to verify findings and build upon them. Finally, we have discussed the limitations and challenges of the current AI solutions-based ML algorithms. These challenges create a research path for the future applications of ML methods in real COVID-19 environments.

In envisioning future research directions, the integration of emerging technologies such as blockchain holds promise in devising comprehensive frameworks to address the COVID-19 pandemic while safeguarding privacy and security concerns. Blockchain, renowned for its decentralized and immutable nature, offers a potential solution to the challenge of securely managing and sharing sensitive health data amidst the pandemic. By leveraging blockchain technology, researchers can design frameworks that ensure the integrity and confidentiality of COVID-19-related data, while facilitating seamless data exchange among healthcare providers, researchers, and public health authorities. This can enable more effective contact tracing, monitoring of infection spread, and vaccine distribution efforts, all while preserving individual privacy rights.

Furthermore, to enhance the performance of prediction models, future research endeavors could focus on the inclusion of additional features from diverse domains, such as meteorological data and medical records. Incorporating meteorological data, including temperature, humidity, and air quality indices, into predictive models can provide valuable insights into the environmental factors influencing virus transmission dynamics. Similarly, integrating comprehensive medical data, including patient demographics, comorbidities, and treatment histories, can enrich predictive models by capturing a more nuanced understanding of individual susceptibility to COVID-19 and disease progression patterns. By harnessing these multidimensional datasets, researchers can refine prediction models to yield more precise and accurate results, ultimately empowering healthcare professionals and policymakers with actionable insights for mitigating the impact of the pandemic.

Acknowledgement:
This work has been conducted during the Sabbatical year of Prof. Walid A. Salameh, thus, we would like to thank and acknowledge PSUT for their continuous support.

References:

[1] Arden, Madelynne A., and Joseph Chilcot. “Health Psychology and the Coronavirus (COVID-19) Global Pandemic: A Call for Research.” *British Journal of Health Psychology*, vol. 25, no. 2, Mar. 2020, pp. 231–32.

[2] Manderson, Lenore, and Susan Levine. “COVID-19, Risk, Fear, and Fall-Out.” *Medical Anthropology*, vol. 39, no. 5, Mar. 2020, pp. 367–70.

[3] Sornette, Didier, et al. “Interpreting, Analysing and Modelling COVID-19 Mortality Data.” *Nonlinear Dynamics*, vol. 101, no. 3, Aug. 2020, pp. 1751–76.

[4] A. Sandford, “Coronavirus: Half of humanity now on lockdown as 90 countries call for confinement,” *Euronews*, vol. 3, no. 04, p. 2020, 2020.

[5] Shakil, M. H., Munim, Z. H., Tasnia, M., & Sarowar, S. (2020). COVID-19 and the environment: A critical review and research agenda. Science of the Total Environment, vol. 745, pp.141022.

[6] Ndwandwe, Duduzile, and Charles S. Wiysonge. “COVID-19 Vaccines.” *Current Opinion in Immunology*, vol. 71, Aug. 2021, pp. 111–16

[7] Alimadadi, Ahmad, et al. “Artificial Intelligence and Machine Learning to Fight COVID-19.” *Physiological Genomics*, vol. 52, no. 4, Apr. 2020, pp. 200–02.

[8] 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, p.

110059, 2020

[9] M. van der Schaar, A. M. Alaa, A. Floto, A. Gimson, S. Scholtes, Wood, E. McKinney, D. Jarrett, P. Lio, and A. Ercole, “How artificial intelligence and machine learning can help healthcare systems respond to covid-19,” *Machine Learning*, vol. 110, no. 1, pp. 1–14, 2021

[10] N. El-Rashidy, S. Abdelrazik, T. Abuhmed, E. Amer, F. Ali, J.-W. Hu, and S. El-Sappagh, “Comprehensive survey of using machine learning in the covid-19 pandemic,” *Diagnostics*, vol. 11, no. 7, p. 1155, 2021.

[11] Kwekha-Rashid, Ameer Sardar, et al. “Coronavirus Disease (COVID-19) Cases Analysis Using Machine-Learning Applications.” *Applied Nanoscience*, vol. 13, no. 3, May 2021, pp. 2013–25

[12] O. E. Mbunge, B. Akinnuwesi, S. G. Fashoto, A. S. Metfula, and P. Mash-wama, “A critical review of emerging technologies for tackling covid-19 pandemic,” *Human behavior and emerging technologies*, vol. 3, no. 1, pp. 25–39, 2021

[13] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, “Artificial intelli-gence (ai) applications for covid-19 pandemic,” *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 4, pp. 337–339, 2020

[14] Q.-V. Pham, D. C. Nguyen, T. Huynh-The, W.-J. Hwang, and P. N. Pathirana, “Artificial intelligence (ai) and big data for coronavirus (covid-19) pandemic: A survey on the state-of-the-arts,” *IEEE access*, vol. 8, p. 130820, 2020

[15] A. Ahmad, S. Garhwal, S. K. Ray, G. Kumar, S. J. Malebary, and M. Barukab, “The number of confirmed cases of covid-19 by using machine learning: Methods and challenges,” *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2645–2653, 2021

[16] Z. A. A. Alyasseri, M. A. Al-Betar, I. A. Doush, M. A. Awadallah, K. Abasi, S. N. Makhadmeh, O. A. Alomari, K. H. Abdulkareem, Adam, R. Damasevicius et al., “Review on covid-19 diagnosis models based on machine learning and deep learning approaches,” *Expert systems*, vol. 39, no. 3, p. e12759, 2022

[17] M. R. H. Mondal, S. Bharati, and P. Podder, “Diagnosis of covid-19 using machine learning and deep learning: A review,” *Current Medical Imaging*, vol. 17, no. 12, pp. 1403–1418, 2021.

[18] J. Chen, K. Li, Z. Zhang, K. Li, and P. S. Yu, “A survey on applications of artificial intelligence in fighting against covid-19,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 8, pp. 1–32, 2021

[19] 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, 2021.

[20] J. Zhao, Y. Zhang, X. He, and P. Xie, *“Covid-ct-dataset: a ct scan dataset about covid-19,”* arXiv preprint arXiv:2003.13865, vol. 490, 2020, [Online]. https://github.com/UCSD-AI4H/COVID-CT (Accessed Date: June 17, 2020).

[21] X. Li, W. Zeng, X. Li, H. Chen, L. Shi, X. Li, H. Xiang, Y. Cao, Chen, C. Liu et al., “Ct imaging changes of corona virus disease 2019 (covid-19): a multi-center study in southwest china,” *Journal of translational medicine*, vol. 18, no. 1, pp. 1–8, 2020.

[22] J. P. Cohen, P. Morrison, and L. Dao, *“Covid-19 image data collection,”* arXiv preprint arXiv:2003.11597, 2020, [Online]. https://github.com/ieee8023/covid-chestxray-dataset (Accessed Date: March 25, 2020).

[23] 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, pp. 1–12, 2020.

[24] C. McDermott, M. Łacki, 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. p. 2, 2021

[25] J. Han, K. Qian, M. Song, Z. Yang, Z. Ren, S. Liu, J. Liu, H. Zheng, Ji, T. Koike et al., “An early study on intelligent analysis of speech under covid-19: Severity, sleep quality, fatigue, and anxiety,” arXiv preprint arXiv:2005.00096, 2020, [Online]. https://arxiv.org/abs/2005.00096 (Accessed Date: May 14, 2020).

[26] N. Sharma, P. Krishnan, R. Kumar, S. Ramoji, S. R. Chetupalli, K. Ghosh, S. Ganapathy et al., “Coswara–a database of breathing cough, and voice sounds for covid-19 diagnosis,” arXiv preprint arXiv:2005.10548, 2020, [Online].

https://arxiv.org/abs/2005.10548 (Accessed Date: August 11, 2020).

[27] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "Covid-19 cough classification using machine learning and global smartphone record-ings," *Computers in Biology and Medicine*, vol. 135, p. 104572, 2021.

[28] A. A. Farid, G. I. Selim, and H. A. A. Khater, "A novel approach of ct images feature analysis and prediction to screen for corona virus disease (covid-19)", *International Journal of Scientific & Engineering Research*, 11(3), 1141–49 2020, 2020.

[29] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z. A. Fayad, N. Zhang, Diao, B. Lin, X. Zhu, K. Li et al., "Chest ct findings in coronavirus disease-19 (covid-19): relationship to duration of infection," *Radiology*, vol. 295, no. 3, June 2020, p. 200463

[30] X. Mei, H.-C. Lee, K.-y. Diao, M. Huang, B. Lin, C. Liu, Z. Xie, Ma, P. M. Robson, M. Chung et al., "Artificial intelligence–enabled rapid diagnosis of patients with covid-19," Nature medicine, vol. 26, no. 8, pp. 1224–1228, 2020.

[31] World Health Organization, "Coronavirus disease 2019 (covid-19): situation report, 73," World Health Organization, 2020, [Online]. https://www.who.int/publications/m/item/situation-report---73 (Accessed Date: April 24, 2024).

[32] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track covid-19 in real time," *The Lancet infectious diseases*, vol. 20, no. 5, pp. 533–534, 2020.

[33] M. U. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, O. C.-. D. W. Group†, L. du Plessis, N. R. Faria, R. Li et al., "The effect of human mobility and control measures on the covid-19 epidemic in china," *Science*, vol. 368, no. 6490, pp. 493–497, 2020.

[34] T. Hale, N. Angrist, B. Kira, A. Petherick, T. Phillips, and S. Webster, *"Variation in government responses to covid-19,"* 2020, medRxiv 2020.07.04.20145334, https://doi.org/10.1101/2020.07.04.20145334.

[35] B. Kleinberg, I. van der Vegt, and M. Mozes, *"Measuring emo-tions in the covid-19 real world worry dataset,"* arXiv preprint arXiv:2004.04225, 2020, [Online]. https://arxiv.org/abs/2004.04225 (Accessed Date: May 14, 2020).

[36] S. Alqurashi, A. Alhindi, and E. Alanazi, "Large arabic twitter dataset on covid-19," arXiv preprint arXiv:2004.04315, 2020, [Online]. https://arxiv.org/abs/2004.04315 (Accessed Date: April 22, 2020).

[37] S. Alpdagtas, E. Ilhan, E. Uysal, M. Sengor, C. B. Ustundag, and Gunduz, "Evaluation of current diagnostic methods for covid-19," *APL bioengineering*, vol. 4, no. 4, p. 041506, 2020

[38] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, Wang, K. Wang et al., "Clinically applicable ai system for accu-rate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography," *Cell*, vol. 181, no. 6, pp. 1423–1433, 2020.

[39] N. El-Rashidy, S. El-Sappagh, S. Islam, H. M El-Bakry, and S. Abdel-razek, "Mobile health in remote patient monitoring for chronic diseases: Principles, trends, and challenges," *Diagnostics*, vol. 11, no. 4, p. 607, 2021

[40] A. Rahman, M. S. Hossain, N. A. Alrajeh, and F. Alsolami, "Adver-sarial examples—security threats to covid-19 deep learning systems in medical iot devices," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9603–9610, 2020.

[41] G. Yang, H. Zhang, D. Firmin, and S. Li, "Recent advances in artificial intelligence for cardiac imaging," *Computerized Medical Imaging and Graphics* vol. 90, June 2021, p. 101928

[42] A. Gagneux-Brunon, M. Detoc, S. Bruel, B. Tardy, O. Rozaire, Frappe, and E. Botelho-Nevers, "Intention to get vaccinations against covid-19 in french healthcare workers during the first pandemic wave: a cross-sectional survey," *Journal of Hospital Infection*, vol. 108, pp. 168–173, 2021.

[43] M. Alhasan and M. Hasaneen, "Digital imaging, technologies and artifi-cial intelligence applications during covid-19 pandemic," *Computerized Medical Imaging and Graphics*, vol. 91, p. 101933, 2021.

[44] Z. Zeng, P.-J. Chen, and A. A. Lew, "From high-touch to high-tech: Covid-19 drives robotics adoption," *Tourism geographies*, vol. 22, no. 3, pp. 724–734, 2020.

[45] C. A. Latz, C. DeCarlo, L. Boitano, C. Png,

R. Patell, M. F. Conrad, Eagleton, and A. Dua, "Blood type and outcomes in patients with covid-19," *Annals of hematology*, vol. 99, no. 9, pp. 2113–2118, 2020

[46] J. Zhao, Y. Yang, H. Huang, D. Li, D. Gu, X. Lu, Z. Zhang, L. Liu, Liu, Y. Liu et al., "Relationship between the abo blood group and the coronavirus disease 2019 (covid-19) susceptibility," *Clinical Infectious Diseases*, vol. 73, no. 2, pp. 328–331, 2021

[47] A. Algunmeeyn, F. El-Dahiyat, M. M. Altakhineh, M. Azab, and Z.-U.- Babar, "Understanding the factors influencing healthcare providers' burnout during the outbreak of covid-19 in jordanian hospitals," *Journal of pharmaceutical policy and practice*, vol. 13, no. 1, pp. 1–8, 2020.

[48] A. S. Abdelhafiz, A. Ali, H. H. Ziady, A. M. Maaly, M. Alorabi, and E. A. Sultan, "Prevalence, associated factors, and consequences of burnout among egyptian physicians during covid-19 pandemic," *Frontiers in public health*, vol. 8:590190. Dec. 2020

[49] F. Bekele, T. Sheleme, G. Fekadu, and K. Bekele, "Patterns and associated factors of covid-19 knowledge, attitude, and practice among general population and health care workers: A systematic review," *SAGE open medicine*, vol. 8, p. 2050312120970721, 2020

[50] C. Hoffmann and E. Wolf, "Older age groups and country-specific case fatality rates of covid-19 in europe, usa and canada," Infection, vol. 49, no. 1, pp. 111–116, 2021.

[51] N. Sudharsanan, O. Didzun, T. Barnighausen,¨ and P. Geldsetzer, "The contribution of the age distribution of cases to covid-19 case fatality across countries: a nine-country demographic study," *Annals of internal medicine*, vol. 173, no. 9, pp. 714–720, 2020.

[52] Capuano, Annalisa, et al. "Covid-19 Kills More Men Than Women: An Overview of Possible Reasons." *Frontiers in Cardiovascular Medicine*, vol. 7, July 2020. Crossref, https://doi.org/10.3389/fcvm.2020.00131.

[53] C. P. Tadiri, T. Gisinger, A. Kautzky-Willer, K. Kublickiene, M. T. Herrero, V. Raparelli, L. Pilote, and C. M. Norris, "The influence of sex and gender domains on covid-19 cases and mortality," *Cmaj*, vol. 192, no. 36, pp. E1041–E1045, 2020.

[54] R. N. van Zyl-Smit, G. Richards, and F. T. Leone, "Tobacco smoking and covid-19 infection," *The Lancet Respiratory Medicine*, vol. 8, no. 7, 664–665, 2020

[55] A. Gulsen,¨ B. A. Yigitbas, B. Uslu, D. Dromann,¨ and O. Kilinc, "The effect of smoking on covid-19 symptom severity: systematic review and meta-analysis," *Pulmonary medicine*, vol. 2020, Sept. 2020, pp. 1–11.

[56] P. Wang, J. Sha, M. Meng, C. Wang, Q. Yao, Z. Zhang, W. Sun, X. Wang, G. Qie, X. Bai et al., "Risk factors for severe covid-19 in middle-aged patients without comorbidities: a multicentre retrospective study," *Journal of Translational Medicine*, vol. 18, no. 1, pp. 1–12, 2020.

[57] K. Aabed and M. M. Lashin, "An analytical study of the factors that influence covid-19 spread," *Saudi Journal of Biological Sciences*, vol. 28, no. 2, pp. 1177–1195, 2021

[58] L. Y. K. Nakada and R. C. Urban, "Covid-19 pandemic: environmental and social factors influencing the spread of sars-cov-2 in sao˜ paulo, brazil," *Environmental Science and Pollution Research*, vol. 28, no. 30, 40 322–40 328, 2021.

[59] X. Zhang, H. Liu, H. Tang, M. Zhang, X. Yuan, and X. Shen, "The effect of population size for pathogen transmission on prediction of covid-19 spread," *Scientific reports*, vol. 11, no. 1, pp. 1–12, 2021

[60] X. Xing, Y. Xiong, R. Yang, R. Wang, W. Wang, H. Kan, T. Lu, D. Li, J. Cao, J. Penuelas˜ et al., "Predicting the effect of confinement on the covid-19 spread using machine learning enriched with satellite air pollution observations," *Proceedings of the National Academy of Sciences*, vol. 118, no. 33, 2021.

[61] S. Lin, Y. Fu, X. Jia, S. Ding, Y. Wu, and Z. Huang, "Discovering correlations between the covid-19 epidemic spread and climate," *Inter-national Journal of Environmental Research and Public Health*, vol. 17, no. 21, p. 7958, 2020.

[62] S. Wang, Y. Zhang, J. Ma, S. Zhu, J. Shen, P. Wang, and H. Zhang, "Responses of decline in air pollution and recovery associated with covid-19 lockdown in the pearl river delta," *Science of the Total Environment*, vol. 756, p. 143868, 2021

[63] A. Dewi, A. Nurmandi, E. Rochmawati, E. P. Purnomo, M. D. Rizqi, A. Azzahra, S.

Benedictos, W. Suardi, and D. T. K. Dewi, "Global pol-icy responses to the covid-19 pandemic: proportionate adaptation and policy experimentation: a study of country policy response variation to the covid-19 pandemic," *Health Promotion Perspectives*, vol. 10, no. 4,359, 2020

[64] R. Chaudhry, G. Dranitsaris, T. Mubashir, J. Bartoszko, and S. Riazi, "A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on covid-19 mortality and related health outcomes," *EClinicalMedicine*, vol. 25, p. 100464, 2020

[65] WHO, "Coronavirus disease (COVID-19) pandemic ," Geneva: WHO, 2020, [Online]. https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (Accessed Date: April 22, 2024).

[66] D. Worldometer, "Covid-19 coronavirus pandemic," World Health Organization, 2020, [Online]. https://www.worldometers.info/coronavirus (Accessed Date: April 13, 2024).

[67] A. W. CDC, "Centers for disease control and prevention," 2020, [Online]. https://www.cdc.gov/coronavirus/2019-ncov/ (Accessed Date: April 22, 2024).

[68] B. J. Cowling, S. T. Ali, T. W. Ng, T. K. Tsang, J. C. Li, M. W. Fong, Q. Liao, M. Y. Kwan, S. L. Lee, S. S. Chiu et al., "Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in hong kong: an observational study," *The Lancet Public Health*, vol. 5, no. 5, pp. e279–e288, 2020

[69] F. Gabbay, S. Bar-Lev, O. Montano, and N. Hadad, "A lime-based explainable machine learning model for predicting the severity level of covid-19 diagnosed patients," *Applied Sciences*, vol. 11, no. 21, p. 10417, 2021

[70] S. S. Aljameel, I. U. Khan, N. Aslam, M. Aljabri, and E. S. Alsulmi, "Machine learning-based model to predict the disease severity and outcome in covid-19 patients," *Scientific programming*, vol. 2021, Apr. 2021, pp. 1–10 .

[71] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Song et al., "Using artificial intelligence to detect covid-19 and community-acquired pneumonia based on pulmonary ct: evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020

[72] J. Pu, J. Leader, A. Bandos, J. Shi, P. Du, J. Yu, B. Yang, S. Ke,Guo, J. B. Field et al., "Any unique image biomarkers associated with covid-19?" *European radiology*, vol. 30, no. 11, pp. 6221–6227, 2020.

[73] 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, 2020

[74] M. Aly, K. H. Rahouma, and S. M. Ramzy, "Pay attention to the speech: Covid-19 diagnosis using machine learning and crowdsourced respiratory and speech recordings," *Alexandria Engineering Journal*, vol. 61, no. 5, pp. 3487–3500, 2022

[75] K. K. Lella and A. Pja, "Automatic diagnosis of covid-19 disease using deep convolutional neural network with multi-feature channel from respiratory sound data: cough, voice, and breath," *Alexandria Engineering Journal*, vol. 61, no. 2, pp. 1319–1334, 2022

[76] D. Arias-Garzon,´ J. A. Alzate-Grisales, S. Orozco-Arias, H. B. Arteaga-Arteaga, M. A. Bravo-Ortiz, A. Mora-Rubio, J. M. Saborit-Torres, J. A. M. Serrano, M. de la Iglesia Vaya,´ O. Cardona-Morales et al., "Covid-19 detection in x-ray images using convolutional neural net-works," *Machine Learning with Applications*, vol. 6, p. 100138, 2021

[77] L. A. Amar, A. A. Taha, and M. Y. Mohamed, "Prediction of the final size for covid-19 epidemic using machine learning: A case study of egypt," *Infectious Disease Modelling*, vol. 5, pp. 622–634, 2020.

[78] H. Burdick, C. Lam, S. Mataraso, A. Siefkas, G. Braden, R. P. Dellinger, A. McCoy, J.-L. Vincent, A. Green-Saxena, G. Barnes et al., "Prediction of respiratory decompensation in covid-19 patients using machine learning: The ready trial," *Computers in biology and medicine*, vol. 124, p. 103949, 2020

[79] B. R. Beck, B. Shin, Y. Choi, S. Park, and K. Kang, "Predicting commercially available antiviral drugs that may act on the novel coronavirus (sars-cov-2) through a drug-target interaction deep learning model,"

Computational and Structural Biotechnology Journal, vol. 18, pp. 784–790, 2020.

[80] T.-H. Pham, Y. Qiu, J. Zeng, L. Xie, and P. Zhang, "A deep learning framework for high-throughput mechanism-driven phenotype com-pound screening and its application to covid-19 drug repurposing," *Nature machine intelligence*, vol. 3, no. 3, pp. 247–257, 2021

[81] E. Kim, A.-s. Choi, and H. Nam, "Drug repositioning of herbal compounds via a machine-learning approach," BMC *Bioinformatics*, vol. 20, no. 10, pp. 33–43, 2019

[82] R. Chakraborty and S. Parvez, "Covid-19: An overview of the current pharmacological interventions, vaccines, and clinical trials," *Biochemical Pharmacology*, vol. 180, p. 114184, 2020

[83] Y. Zhu, J. Li, and Z. Pang, "Recent insights for the emerging covid-19: Drug discovery, therapeutic options and vaccine development," *Asian Journal of Pharmaceutical Sciences*, vol. 16, no. 1, pp. 4–23, 2021.

[84] H. Zhang, K. M. Saravanan, Y. Yang, M. Hossain, J. Li, X. Ren, Y. Pan, Y. Wei et al., "Deep learning based drug screening for novel coronavirus 2019-ncov," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 12, no. 3, pp. 368–376, 2020.

[85] Z. Xu, L. Yang, X. Zhang, Q. Zhang, Z. Yang, Y. Liu, S. Wei, and W. Liu, "Discovery of potential flavonoid inhibitors against covid-19 3cl proteinase based on virtual screening strategy," *Frontiers in molecular biosciences*, p. 247, vol. 7, Sept. 2020.

[86] A. Kumar, P. K. Gupta, and A. Srivastava, "A review of modern technologies for tackling covid-19 pandemic," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 4, pp. 569–573, 2020.

[87] Amilpur, S., & Bhukya, R. (2022). Predicting novel drug candidates against Covid-19 using generative deep neural networks. Journal of Molecular Graphics and Modelling, vol. 110, pp. 108045.

[88] F. Rustam, A. A. Reshi, A. Mehmood, S. Ullah, B.-W. On, W. Aslam, and G. S. Choi, "Covid-19 future forecasting using supervised machine learning models," *IEEE access*, vol. 8, pp. 101 489–101 499, 2020

[89] J. CSSE, "Covid-19 data repository by the center for systems science and engineering (csse) at johns hopkins university," Github repository, 2020, [Online]. https://covidbehaviors.org/ (Accessed Date: April 22, 2024).

[90] E. A. Rashed and A. Hirata, "One-year lesson: Machine learning prediction of covid-19 positive cases with meteorological data and mobility estimate in Japan," *International Journal of Environmental Research and Public Health*, vol. 18, no. 11, p. 5736, 2021.

[91] A. I. Shahin and S. Almotairi, "A deep learning bilstm encoding-decoding model for covid-19 pandemic spread forecasting," *Fractal and Fractional*, vol. 5, no. 4, p. 175, 2021.

[92] R. G. da Silva, M. H. D. M. Ribeiro, V. C. Mariani, and L. dos San-tos Coelho, "Forecasting brazilian and american covid-19 cases based on artificial intelligence coupled with climatic exogenous variables," *Chaos, Solitons & Fractals*, vol. 139, p. 110027, 2020

[93] S. A. Alanazi, M. Kamruzzaman, M. Alruwaili, N. Alshammari, S. A. Alqahtani, and A. Karime, "Measuring and preventing covid-19 using the sir model and machine learning in smart health care," *Journal of* healthcare engineering, vol. 2020, Nov. 2020, pp. 1–12.

[94] A. Roy and B. Kar, "Characterizing the spread of covid-19 from human mobility patterns and sociodemographic indicators," in *Proceedings of the 3rd ACM SIGSPATIAL international workshop on advances in resilient and intelligent cities*, 2020, pp. 39–48.

[95] G. Pinter, I. Felde, A. Mosavi, P. Ghamisi, and R. Gloaguen, "Covid-19 pandemic prediction for hungary; a hybrid machine learning approach," *Mathematics*, vol. 8, no. 6, p. 890, 2020

[96] G. Pandey, P. Chaudhary, R. Gupta, and S. Pal, "Seir and regression model based covid-19 outbreak predictions in india," arXiv preprint arXiv:2004.00958, 2020, [Online]. https://arxiv.org/abs/2004.00958. (Accessed Date: April 1, 2020).

[97] R. Sujath, J. M. Chatterjee, and A. E. Hassanien, "A machine learning forecasting model for covid-19 pandemic in india," *Stochastic Envi-ronmental Research and Risk Assessment*, vol. 34, no. 7, pp. 959–972,

2020.

[98] P. Wang, X. Zheng, J. Li, and B. Zhu, "Prediction of epidemic trends in covid-19 with logistic model and machine learning technics," *Chaos, Solitons & Fractals*, vol. 139, p. 110058, 2020

[99] S. Dutta and S. K. Bandyopadhyay, "Machine learning approach for confirmation of covid-19 cases: Positive, negative, death and release," *Iberoamerican Journal of Medicine*, vol. 2, no. 3, pp. 172–177, 2020

[100] A. Tob´ıas, "Evaluation of the lockdowns for the sars-cov-2 epidemic in italy and spain after one month follow up," *Science of the Total Environment*, vol. 725, p. 138539, 2020

[101] N. El-Rashidy, S. El-Sappagh, S. Islam, H. M. El-Bakry, and S. Abdel-razek, "End-to-end deep learning framework for coronavirus (covid-19) detection and monitoring," *Electronics*, vol. 9, no. 9, p. 1439, 2020.

[102] A. Shelke, M. Inamdar, V. Shah, A. Tiwari, A. Hussain, T. Chafekar, and N. Mehendale, "Chest x-ray classification using deep learning for automated covid-19 screening," *SN computer science*, vol. 2, no. 4, pp. 1–9, 2021

[103] T. A. Soomro, L. Zheng, A. J. Afifi, A. Ali, M. Yin, and J. Gao, "Artificial intelligence (ai) for medical imaging to combat coronavirus disease (covid-19): A detailed review with direction for future re-search," *Artificial Intelligence Review*, pp. 1–31, 2021

[104] S. Rahmatizadeh, S. Valizadeh-Haghi, and A. Dabbagh, "The role of artificial intelligence in management of critical covid-19 patients," *Journal of Cellular and Molecular Anesthesia*, vol. 5, no. 1, pp. 16–22, 2020

[105] M. Z. Alom, M. Rahman, M. S. Nasrin, T. M. Taha, and V. K. Asari, "Covid mtnet: Covid-19 detection with multi-task deep learning approaches," arXiv preprint arXiv:2004.03747, 2020, [Online]. https://arxiv.org/abs/2004.03747 (Accessed Date: April 18, 2020).

[106] 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, p. 1400, 2020.

[107] M. Tavakoli, J. Carriere, and A. Torabi, "Robotics, smart wearable technologies, and autonomous intelligent systems for healthcare during the covid-19 pandemic: An analysis of the state of the art and future vision," *Advanced Intelligent Systems*, vol. 2, no. 7, p. 2000071, 2020

[108] J. Torales, M. O'Higgins, J. M. Castaldelli-Maia, and A. Ventriglio, "The outbreak of covid-19 coronavirus and its impact on global mental health," *International journal of social psychiatry*, vol. 66, no. 4, pp. 317–320, 2020.

[109] S. Li, Y. Wang, J. Xue, N. Zhao, and T. Zhu, "The impact of covid-19 epidemic declaration on psychological consequences: a study on active weibo users," *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, p. 2032, 2020.

[110] T. T. Nguyen, Q. V. H. Nguyen, D. T. Nguyen, E. B. Hsu, S. Yang, and P. Eklund, "Artificial intelligence in the battle against coronavirus (covid-19): a survey and future research directions," arXiv preprint arXiv:2008.07343, 2020, [Online]. https://arxiv.org/abs/2008.07343 (Accessed Date: March 17, 2020).

[111] P. L. Fung, M. A. Zaidan, O. Surakhi, S. Tarkoma, T. Petaj¨a,¨ and Hussein, "Data imputation in in situ-measured particle size dis-tributions by means of neural networks," *Atmospheric Measurement Techniques*, vol. 14, no. 8, pp. 5535–5554, 2021.

[112] A. Salameh and O. Surakhi, "An optimized convolutional neural network for handwritten digital recognition classification," *Journal of Theoretical and Applied Information Technology*, vol. 98, no. 21, pp. 3494–3503, 2020

[113] O. M. Surakhi, M. A. Zaidan, S. Serhan, I. Salah, and T. Hussein, "An optimal stacked ensemble deep learning model for predicting time-series data using a genetic algorithm—an application for aerosol particle number concentrations," *Computers*, vol. 9, no. 4, p. 89, 2020.

[114] O. Surakhi, S. Serhan, and I. Salah, "On the ensemble of recurrent neural network for air pollution forecasting: Issues and challenges," *Adv. Sci. Technol. Eng. Syst. J*, vol. 5, pp. 512–526, 2020

[115] Yeng, P. K., Nweke, L. O., Woldaregay, A. Z., Yang, B., & Snekkenes, E. A. (2021). Data-driven and artificial intelligence (AI)

approach for modelling and analyzing healthcare security practice: a systematic review. *In Intelligent Systems and Applications: Proceedings of the 2020 Intelligent Systems Conference (IntelliSys),* Vol. 1 (pp. 1-18). Springer International Publishing.

[116] O Tariq, N., Qamar, A., Asim, M., & Khan, F. A. (2020). Blockchain and smart healthcare security: a survey. *Procedia Computer Science*, vol. 175, 2020, pp. 615–20.

[117] M. A. Zaidan, O. Surakhi, P. L. Fung, and T. Hussein, "Sensitivity analysis for predicting sub-micron aerosol concentrations based on meteorological parameters," *Sensors*, vol. 20, no. 10, p. 2876, 2020.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- Walid Salameh, Ola Surakhi and Mohammad Khanafseh carried out the formal analysis, investigation, and methodology.
- Walid Salameh leads the planning and execution of research activities
- The original draft is created and prepared by Ola Surakhi
- Walid Salameh, Ola Surakhi and Mohammad Khanafseh reviewed and edited the final version

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

This research was funded by the Scientific Research Deanship at Princess Sumaya University for Technology.

Conflict of Interest

The authors have no conflicts of interest to declare.

APPENDIX

Table 2. ML and DL technologies used in the Diagnosis and Identification of COVID-19.

Related Works	ML/DL Methods	Features	Datasets	Country	Results
[69]	Multilayer perceptron artificial neural networks and decision trees	Patient's medical history and symptoms and Patient's clinical outcome	Mexican Federal Health Secretary, General Director of Epidemiology	Mexican	The model achieved up to 80% prediction accuracy for the dataset used
[70]	logistic regression (LR), random forest (RF), and extreme gradient boosting (XGB)	Clinical features (20)	287 COVID-19 samples of patients from the King Fahad University Hospital	Saudi Arabia	The RF outperformed the other classifiers with an accuracy of 0.95
[71]	Convolutional Neural Network (CNN)	Chest CT scans	4563 three-dimensional (3D) volumetric chest CT scans from 3506 patients acquired at six medical centers between August 16, 2016, and February 17	China	A deep learning method was able to identify coronavirus disease 2019 on chest CT scans
[72]	3D CNN model	Chest CT scans	498 CT scans from 151 positive COVID-19 subjects and 497 CT scans from different subjects with various types of pneumonia	China	Achieved a moderate diagnostic ability overall (Area under the curve (AUC) of 0.70 with 99% Confidential interval (0.56–0.85)
[73]	Using SVM (Support Vector Machine), CNN (Conventional Neural Networks), ResNet50, InceptionResNetV2, Xception, VGGNet16	X-ray images	5857 Chest X-rays and 767 Chest CTs for COVID-19-positive cases	-	Achieve 84% and 75% classification accuracy
[74]	deep learning model	cough, breathing, and speech	Coswara dataset	-	AUC of 96.4% and an accuracy of 96%
[75]	convolutional neural network	Cough, voice, and breath	sounds dataset from Cambridge University	UK	The proposed approach improves system performance to diagnose COVID-19 disease and provides better results on the COVID-19 respiratory sound dataset.
[76]	VGG19 and U-Net	X-ray images	BIMCV-COVID19, , BIMCV-COVID, and Spain Pre-COVID era dataset	Spain	Achieved 97% of accuracy
[77]	Regression Linear Regression model	COVID-19 confirmed, and mortal cases	Egyptian Ministry of health and population	Egypt	The proposed model was beneficial for the Egyptian government in managing the COVID-19 outbreak for the following months.
[78]	Classification Logistic Regression	Patient health record	United States health systems 197 patients	USA	The proposed algorithm can accurately identify 16% more patients than a widely used scoring system while minimizing false-positive results

Table 3. ML and DL technologies used in the Forecasting of COVID-19

Related Works	ML/DL Methods	Features	Datasets	Country	Results
[88]	linear regression (LR), least absolute shrinkage, selection operator (LASSO), support vector machine (SVM), and exponential smoothing (ES)	The number of new positive cases, the number of deaths, and the number of recoveries.	GitHub repository provided by the Center for Systems Science and Engineering, Johns Hopkins University [89]	International	The ES performs best among all the used models.
[90]	Long short-term memory (LSTM) network	Meteorological and mobility data	Google Cloud online COVID-19	Japan	The proposed framework provided more accurate and consistent estimations than that offered by Google Cloud
[91]	encoding–decoding LSTM	Confirmed positive cases, death cases, and recovery cases	Saudi Ministry of Health website, the existing interactive dashboards, and the available application program interface (API)	International	The proposed model generated high accuracy with less error rate
[92]	Bayesian regression neural network, cubist regression, k-nearest neighbors, quantile random forest, variational mode decomposition, and support vector regression	cumulative COVID-19 cases and exogenous variables such as daily temperature and precipitation	COVID-19 Data Repository [2] And API (Application Program Interface) in 27 Brazilian State Health Offices	Brazilian and American states	It was observed that climatic variables, such as temperature and precipitation, indeed influence increasing the accuracy when predicting COVID-19 cases, and the adopted models can be recommended as a promising model for forecasting
[93]	Machine learning algorithms along with SIR and SIR-F models	Patient information, mobility, number of cases	John Hopkins University dataset	Saudi Arabia	The results show that government lockdowns and isolation of individuals are not enough to stop the pandemic
[94]	supervised machine learning algorithms	human mobility data, number of cases	Collected	USA	tree-based classifiers performed best on the forecasting task. Gradient Boosting had the highest classification accuracy.
[95]	A hybrid machine learning method of adaptive network-based fuzzy inference system (ANFIS) and multi-layered perceptron-imperialist competitive algorithm (MLP-ICA)	Number of cases and deaths	The statistical reports of COVID-19 cases and mortality rate of Hungary	Hungary	suggests machine learning as a potential technology to be considered to model the outbreak
[96]	SEIR model and Regression model	confirmed cases	John Hopkins University dataset	India	The model gave a short-term prediction (two weeks) to help the Government and doctors prepare their plans
[97]	linear regression, Multilayer perceptron, and Vector autoregression method	confirmed, death, and recovered cases	COVID-19 Kaggle data	India	The MLP method gave better prediction results than that of the LR and VAR method
[98]	Logistic model and FbProphet model	confirmed cases, recovered cases, and death cases	John Hopkins University dataset	International	The model can significantly improve estimates of the number of infections.
[99]	Long short-term memory (LSTM) and Gated Recurrent Unit (GRU)	confirmed, negative, released, deceased cases	kaggle	-	The proposed approach helps generate suitable results based on the critical disease outbreak
[100]	quasi-Poisson regression	the trends of the daily incident diagnosed cases, deaths, and intensive care unit admissions	The websites of the Italian and Spanish Ministries of Health	Italy and Spain	the positive signs already shown by the decreasing trend slopes after a more restrictive lockdown in Italy and Spain