Applications of Artificial Intelligence in Combating Covid-19: A Systematic Review

Akpofure A. Enughwure, Isaac C. Febaide

Department of Electrical Engineering, Nigeria Maritime University, Okereenkoko, Delta State, Nigeria
Email: akpofure.enughwure@nmu.edu.ng, isaac.febaide@nmu.edu.ng

Abstract
Artificial Intelligence (AI), which has become widely accepted in the medical field over the years, have proven to be valuable in dealing with sicknesses of varying degrees. Following the spread of COVID-19, so many researches have been developed suggesting ways and models to combat the virus by exploring possible diagnosis, treatment, prevention and cure using AI. This research conducted a systematic literature review to unravel the applications of Artificial Intelligence (AI) to tackle the COVID-19 pandemic. Our data comprised of research papers on COVID-19 and AI or deep learning or machine learning, which are subsets of AI. We did not include review papers, as our research is an improvement on previous reviews on COVID-19 and AI. From our research, six papers examined the early detection and diagnosis of the infection, eight examined forecasting of the spread of the infection, one examined the development of drugs and vaccines and one examined the monitoring of patient treatment. All the researches showed the great potential of harnessing AI in combating the COVID-19 pandemic, when improved on.

Subject Areas
Artificial Intelligence

Keywords
Covid-19, Artificial Intelligence, Machine Learning, Deep Learning

1. Introduction
The novel coronavirus disease 2019 (COVID-19) which was first reported in Wuhan, China on 31st December 2019 [1], declared by the World Health Organization (WHO) [2] as a global pandemic due to its rapid spread across the globe has posed a serious threat to the global economy. The virus responsible for this
A disease is the Severe Acute Respiratory Syndrome Virus-2 (SARS COV-2) which has caused severe respiratory disorders and death in humans [2].

Common features or symptoms of COVID-19 are cold, cough, fever, headache, sore throat, fatigue, muscle pain and difficulty in breathing. As at the time of this study, researchers across the globe are working tirelessly to come up with a possible cure for COVID-19.

WHO approved testing technique currently in use for the diagnosis of COVID-19 is the Reverse transcription-polymerase Chain Reaction (RT-PCR) [3]. X-Ray and Chest Radiology imaging also finds importance in the early diagnosis and treatment of COVID-19.

Artificial Intelligence (AI) has proven to be a vital technological tool in the fight against COVID-19 [4]. Within this short period of the spread of COVID-19 across the globe, several studies have been made on possible treatment, cure and prevention of the spread of COVID-19 using AI.

AI as it relates to medicine could be broad in definition with concerning its applications. It is designed, through learning, knowledge and data processing to aid the practice of medicine. Branches of AI that finds application in medical practice include, machine learning, deep learning based on artificial neural network (ANN) which draws inspiration from the human nervous system [5].

Ben-Israel D. et al. [4] carried out a review work on the impact of machine learning on patient care. They carried out a systematic review in line with PRISMA guidelines using Medline(R), EBM Reviews, Embase, Psych Info, and Cochrane Databases, with a focus on human studies that used machine learning to directly address a clinical problem.


Tulin O. et al. [7] developed a new model for automatic COVID-19 detection using raw chest X-ray images. Their model is developed to provide accurate diagnostics for binary classification (COVID vs. No-Findings) and multi-class classification (COVID vs. No-Findings vs. Pneumonia). It produced a classification accuracy of 98.08% for binary classes and 87.02% for multi-class cases.

Raju V. et al. [8] identified seven significant applications of Artificial Intelligence for COVID-19 pandemic. Their work revealed an important role of AI in detecting cluster of cases and to predict where this virus will affect in future by collecting and analyzing all previous data.

This paper is a concise review of the contributory role of AI in the fight against COVID-19 upon discovery to date. Its purpose is to analytically review the applications of AI in combating the novel COVID-19, showcase the various AI techniques for tackling specific aspect of the fight against the novel COVID-19.
2. Methodology

The primary objective of this systematic literature review is to unravel the applications of artificial intelligence (AI) to tackle the COVID-19 pandemic. The following research questions positioned us to better explore the employment of AI in the study:

**Question 1**
What artificial intelligence (AI) methods were used to tackle/manage the COVID-19 pandemic?

**Question 2**
What facet of the COVID-19 pandemic was addressed while applying these AI methods?

Furthermore, we intended to unsheathe some further insights from the research works, to complement the findings:
- The performance metrics of the models such as accuracy, specificity, sensitivity as well as the area under the receiver-operating characteristics curve (AUC).
- The type of data as well as the size of dataset used in the studies.

2.1. Data Collection

With the use of a certain set of queries on selected bibliographical databases and clear criteria, we managed to fetch a list of research articles within our area of interest. We queried Google Scholar and Science Direct. Google Scholar is a social networking platform for scientists, scholars and researchers to share papers as well as collaborate. Science Direct is one of the world-leading platform to source for scientific, technical and medical research. It hosts over 12 million items of content from over 3500 academic journals and 34,000 e-books.

2.2. Search Terms

During our search for relevant research work on these databases, our focus was on these terms: Artificial Intelligence and COVID-19. Given that machine learning, deep learning are subsets of AI, we interchanged AI with machine learning and deep learning during our search. Our search queries are in the following formats:

("COVID-19") AND ("Artificial Intelligence" or "Machine Learning" or “Deep Learning” or “AI”).

2.3. Study Selection

In this phase with a list of inclusion and exclusion benchmark, we were able to select the articles that were relevant to our work. At first glance, we scanned through each paper’s title and abstract to see if they contain relevant information. Publications with irrelevant content were excluded. Those classified in the inclusive set were read thoroughly (full text read) to ensure their significance to our work. This was achieved by considering the papers in the inclusion set with
our research questions in mind: to affirm that they provided sufficient information about analytical techniques, datatypes and stakeholders regarding the application of AI to combat COVID-19. We considered only research articles in our study hence we excluded review article with important information. We removed duplicates as well. Given the COVID-19 research area is novel in nature there was no need for time restrictions.

2.4. Data Analysis

For every research studies chosen for the literature assessment, we extracted data to answer our research questions and performed a mapping study to present the findings of each question using the area of application as a focal point.

3. Result and Discussion

3.1. Early Detection and Diagnosis of Infection

Artificial Intelligence has been used to combat Covid-19 spread by helping with the early detection of the virus. A good number of researchers have come up with ways to make this possible. Wu et al. [9] worked on a diagnosis model trained on the use of multi-view images of chest CT. The multi-view deep learning fusion model achieved the area under the receiver-operating characteristics curve (AUC) of 0.732, accuracy of 0.700, sensitivity of 0.730 and specificity of 0.615 in the validation set. In evaluating the model performance, AUC, accuracy, sensitivity and specificity of 0.819, 0.760, 0.811 and 0.615 respectively were recorded. This multi-view testing model holds some potential to improve the efficiency of diagnosis and reduce the demand placed on radiologists in times like these.

Vaishya Raju, Javaid Mohd, Khan Haleem & Abid Haleem [10] conducted a diagnostic research on the novel Coronavirus. They employed 10 convolutional neural network methods (AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, and Xception) to identify a COVID-19 case from a non-COVID-19 case (pneumonia and other viral infections) with a great precision. ResNet-101 and Xception performed best in the classification process. They also tested the performance of the models’ performance by their AUC score, Accuracy, Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value. The results of the testing showed that ResNet-101 could differentiate COVID-19 from non-COVID-19 cases with an AUC of 0.994 (sensitivity, 100%; specificity, 99.02%; accuracy, 99.51%). Xception achieved an AUC of 0.994 (sensitivity, 98.04%; specificity, 100%; accuracy, 99.02%). The researchers finally compared the results of ResNet-101 and Xception with the result of the radiology lab (0.873 sensitivity, 89.21%; specificity, 83.33%; accuracy, 86.27%). Given the Computer-aided design, the systems hold great promise with these findings from the research.

Considering the need for early detection of the Covid-19 virus, Barstugan Mucahid, Ozkaya Umut & Ozturk Saban [11] used machine learning to develop a detection model. They used various feature-extraction methods (Grey Level
Co-occurrence Matrix (GLCM), Local Directional Pattern (LDP), Grey Level Run Length Matrix (GLRLM), Grey-Level Size Zone Matrix (GLSZM), and Discrete Wavelet Transform (DWT) algorithms) in this work after generating various sets of image data from the image set. In a bid to improve the performance of the model, 2-fold, 5-fold and 10-fold cross validation was employed in the classification process. The result showed that at 10-fold cross validation and application of GLSZM feature extraction method, the model performance was 99.68% accurate.

In this research work, Ozturk et al. [7] explored the possibility of automatically detecting COVID-19 cases using X-ray images. Their work was to identify COVID-19 cases from Non-COVID-19 cases (binary classification) and COVID-19 cases, Non-COVID-19 cases and pneumonic cases (multi-class classification). They applied DarkCovid deep learning to identify COVID-19 cases in the two scenarios: binary classification and multi-class classification. They employed a 5-fold cross validation technique to improve the model performance. With the use of a heat map, the model could easily distinguish between COVID-19 cases from pneumonia and ARDS cases. The developed system was able to perform binary and multi-class tasks with an accuracy of 98.08% and 87.02%, respectively. They mentioned one of the limitations of the work to be the low number of dataset (use of a limited number of COVID-19 X-ray images) used for training the model showing a need to see how the model will perform in a larger dataset.

Togaçar Mesut, Ergen Burhan & Zafer Comert [12] also developed a machine learning model to detect the Covid-19 virus in an earlier manner. They employed Support Vector Machine, a machine learning method, to distinguish COVID-19 cases from Non-COVID-19 and Pneumonia. Their work was distinct from other related work because they processed the x-ray image data using fuzzy colour technique and imaging stack techniques. The fuzzy colour technique was used to improve the input data (x-ray image files) by reducing the level of blur while the image stack technique was employed to eliminate the noises of the original files hence improving the quality of the image. Deep learning technique Squeezenet was used as well. The performance of the models was evaluated by F1-score, Sensitivity, Specificity, Precision and Accuracy. The three scenarios used were “Squeezenet + original data”, “Squeezenet + structured data (fuzzy technique)” and “Squeezenet + structured data (stacked technique)”. Their overall performance are as follows: 84.5%, 95.58% and 97.06% respectively.

Punn Narinder, Sonbhadra Kumar & Sonali Agarwal [13] also conducted a diagnostic research. This research evaluated the discrepancy of series of Computational Tomography (CT) manifestations in COVID-19 among different clinical types within the short-term follow-up periods, to help clinicians monitor and predict outcome and to make more accurate and effective clinical decisions. Seven patients who tested positive for Covid-19 were used for the study. The patients were subdivided into common, severe and critical groups for the sake of the study. All CT examinations were performed with a 128-row multi-detector
CT system. All data were statistically analyzed using IBM SPSS statistics software. On initial CT, the main abnormalities observed in common and severe cases respectively were pure ground-glass opacities (GGOs) and patchy consolidation surrounded by GGOs. Critical cases had multiple consolidation surrounded by wide range of GGOs distributed in the whole lung fields. The scope and density score in common (4.5 and 5), severe (9.5 and 9.5) and critical (19 and 12) cases were increased by gradient. On follow-up CT, common and severe types manifested as decreasing density of lesion, absorbed consolidation and GGOs. Critical cases showed progression of the disease. The extent and progression scores in common and severe patients were significantly decreased, while the range score of patients with critical disease reached the highest points, accompanied by an increase in the density score. This research showed that CT scanning could accurately assess the severity of COVID-19 and help to monitor disease transformation during follow-up among different clinical conditions.

### 3.2. Forecasting of the Spread of the Infection

A major approach to curtail the spread of Covid-19 is the ability to make forecasts on the spread of the virus. This would ensure that certain preventive measures are applied. Different researchers from different countries have come up with models and suggestions to forecast the spread of COVID-19 using Artificial Intelligence.

A research conducted in India by Tiwari U. Kumar & Khan Rijwan [14] attempted to predict the spread of the COVID-19 infection in India using machine learning. With the use of COVID-19 related dataset, they were able to visualize the current spread of the viral infection across India. In a bid to predict the spread of the infection, the authors employed a machine-learning tool called Prophet. The prediction was done in a weekly fashion to estimate the number of deaths and recovery cases. The model was able to predict a rise of infection cases in India. Although the working principle of the Prophet tool was not discussed in the work, the work can be improved on so that performance can be accessed.

Zoabi Yazeed & Noam Shomron [15] conducted research that predicted the number of infected cases based on their symptoms and other demographic features in Israel. With the use of health records from the Ministry of Health in Israel, the authors built a predictive model using gradient boosting decision tree base learners. Their model was a classification in operation. The model predicted with 0.90 AUC (area under the receiver operating curve) with 95% CI: 0.892 - 0.905 possible working points are: 87.3% sensitivity and 72% specificity, or 85.7% sensitivity and 79% specificity. Symptoms like Fever and Cough were key features in predicting the contracting the virus. Close Contact with an infected person also played a vital role in contracting the disease as expected. They noted a flaw in the data that was used, which was 100% from the Israeli ministry of health that has limitations and bias. If this bias is removed, the AUC of their model will drop to 0.862. This has to be addressed.
Yu et al. [16] conducted a research based on an estimation of COVID-19 asymptomatic infection cases using a machine learning approach. They adopted a simulator called machine learning-based transmission simulator (MLSim) to predict COVID-19 asymptomatic infections. This simulator was built based on the transmission characteristics of COVID-19. In the simulator, the infection was classified into latent, quarantined, and confirmed patients with the latent patients can infect others. The disease progress is divided into three stages: incubation stage, quarantine stage, and confirmation stage. The performance of the models was evaluated by F1-score, Sensitivity, Specificity, Precision and Accuracy. The three scenarios used are “Squeezenet + original data”, “Squeezenet + structured data (fuzzy technique)” and “Squeezenet + structured data (stacked technique)”. Their overall performance are as follows: 84.5%, 95.58% and 97.06% respectively.

Punn N. Singh, Sonbhadra Kumar & Sonali Agarwal [13] conducted a research that was focused on the use of machine learning and deep learning models to understand the behavioural pattern of the virus and to predict the reachability of the Covid-19 virus across nations. Using a dataset that gave details of daily reported cases, this research was able to utilize mathematical and deep learning models for analysis. To analyze the transmission of the virus, mathematical models based on machine learning such as support vector regression (SVR) and polynomial regression (PR), and deep learning regression models such as a standard deep neural network (DNN) and recurrent neural networks (RNN) using long short-term memory (LSTM) cells were used. Machine learning and deep learning approaches were implemented using the python library “sklearn” and “Keras” to predict the total number of confirmed, recovered, and death cases worldwide. The machine learning and deep learning approaches used were able to predict the possible number of cases in the world for the next 10 days. Early prediction of cases using machine learning and deep learning remains one of the ways to control the transmission of the Covid-19 virus. The result of this research showed that polynomial regression (PR) yielded a minimum root mean square error (RMSE) score over other approaches in forecasting the COVID-19 transmission. If the spread follows the predicted trend of the PR model then it would lead to huge loss of lives as it presents the exponential growth of the transmission worldwide.

The research by Abdollahia Amir & Rahbaralamb Maryam [17] used a machine learning approach to study the effect of temperature, humidity and wind speed on the number of infected people daily in some areas in Spain. They used the population of three different communities in 2019 and the number of confirmed COVID-19 cases until April 28th, 2020. These communities are located in different parts of Spain with varying conditions of climate. A temperature variation of up to about 16˚C is observed between these communities, highlighting the climate diversity of Spain. The Pearson’s correlation coefficient was used as a measure of the linear association between the predictors and the target. The linear regression, a supervised machine learning technique, to find the relationship between the most important predictors and the target was also used. The
Ordinary Least Squares (OLS) was used to estimate the unknown regression coefficients. The study showed that the average temperatures of up to 6 previous days had a moderate negative correlation with the daily increment of infected people. This result is relevant since the mean incubation period of COVID-19 is known to be around 5 - 6 days. The linear regression showed a decrease of about 200 daily-infected people per 1°C increase in average temperature.

This study by AFM Batista, JL Miraglia, Donato THR & ADP Chiavegatto [18] is based on the use of machine learning to predict the risk of positive Covid-19 diagnosis using only results from emergency care admission exams in Brazil as predictors. A total of 18 variables were used for training the algorithms which included age, sex, haemoglobin, platelets, red blood cells, mean corpuscular haemoglobin concentration (MCHC), mean corpuscular haemoglobin (MCH), red cell distribution width (RDW), mean corpuscular volume (MCV), leukocytes, lymphocytes, monocytes, basophils, eosinophils, platelets, red cell distribution width (RDW), mean corpuscular volume (MCV) and c-reactive protein (CRP). The predictive performance for positive diagnosis of COVID-19 was tested with machine learning algorithms: neural networks, gradient boosted trees, random forests, logistic regression and support vector machines. Predictive performance was gotten by calculating the area under the ROC curve (AUC), sensitivity, specificity, F1-score, Brier score, positive predictive value (PPV) and negative predictive value (NPV). All analyses were performed in Python using the scikit-learn library. By using only standard exams performed upon emergency care admission, machine-learning algorithms were able to predict with good performance the risk of each patient having a positive result for COVID-19. This study would be beneficial in allocating testing priorities in the case of a shortage of materials.

Still in a bid to build machine-learning models to forecast the spread of the virus, Tulia Shreshth, Tulic Shikhar, Tulib Rakesh & Gill S. Singh [19] built a model to predict the growth and trend of COVID-19 pandemic. They deployed this model on the cloud to enhance the prediction process using high-speed computations. In this work, they compared the distribution of the COVID-19 cases to that of SARS-Cov-1 cases distribution. They ruled out Gaussian or Exponential distributions. In an attempt to find the best fit distribution curve for the daily cases of COVID-19, they used five sets of global data and realized that the following five distribution type performed best: Inverse Weibull function, iterative versions of Gaussian, Beta (4-parameter), Fisher-Tippet (Extreme Value distribution), and Log Normal functions. Since every country had a unique approach to migrate the spread of the virus. With the use of their distributions, they were able to predict the new daily cases of different countries.

In this investigative study, Roy et al. [20] examined the effects of the COVID-19 pandemic on the world using machine learning. They used a regression model within a time series to predict the spread of the COVID-19 virus. They also developed a dynamic map that visualized the virus progress as it spreads from one location to another. They used an open source dataset from Kaggle where the Prophet prediction and non-linear trend models were employed. Plotly Python
Library was used to develop a Choropleth Dynamic Map to visualize how the pandemic virus is spreading all over the globe. They were able to predict the spread of the COVID-19 virus worldwide using machine learning and an additive regression model.

3.3. Development of Drugs and Vaccines

With the urgent need to produce a vaccine to eliminate the novel coronavirus, Ong Edison, Wong Mei, Huffman Anthony & Hethese Yonghun [21] conducted a study to help the development of a vaccine using reverse vaccinology and machine learning. The Vaxign RV and Vaxign-ML approaches were used to predict COVID-19 protein candidates for vaccine development. They identified six possible adhesins, including the structural S protein and five other non-structural proteins, and three of them (S, nsp3, and nsp8 proteins) were predicted to induce high protective immunity. The S protein was predicted to have the highest protective antigenicity score and it has been extensively studied as the target of coronavirus vaccines by other researchers. They identified six possible adhesins, including the structural S protein and five other non-structural proteins, and three of them (S, nsp3, and nsp8 proteins) were predicted to induce high protective immunity. The S protein was predicted to have the highest protective antigenicity score and it has been extensively studied as the target of coronavirus vaccines by other researchers. They selected nsp3 protein as an alternative vaccine candidate, which was predicted to have the second-highest protective antigenicity score yet, but has not been considered in any vaccine studies. Furthermore, they investigated the sequence conservation and immunogenicity of the multi-domain nsp3 protein as a vaccine candidate. The results indicated that for the first time many non-structural proteins could be used as potential vaccine candidates. Following the various protein samples used as candidates, it is recommended that instead of using a single protein as the vaccine antigen, there should be a development of a “cocktail vaccine” which includes more than antigen to cover different aspects of protection as an effective strategy for COVID-19 vaccine development.

3.4. Monitoring Patient Treatment

Another critical aspect to consider on COVID-19 is the monitoring of patient treatment. Nemati Mohammadreza, Ansary Jamal & Nemati Nazafarin [22] conducted a study on the use of machine learning for survival analysis and predicting patient discharge time for COVID-19 cases. They also considered the features that stood out for the survival characteristics. In this work, they used a dataset that contained epidemiological information such as various features about the surveyed cases, including case ID, age, gender, the onset date of symptoms, date of hospitalization, infection confirmation date, death or discharge time, death or discharge status, symptoms, chronic disease history, travel history, and location. The data was sourced from national health reports. Survival analysis is a form of regression where the predicted value is continuous in nature. The following ML model was used in this study: Accelerated Time Failure, Stage-wise Gradient Boosting, Component-wise Gradient Boosting, Support Vector Machines while Kaplan-Meier estimator, Cox Proportional Hazard and CoxNet Regression model were the statistical packages. Among all the ML models employed in this study, IPCRidge performed the worst accuracy; however, it showed
a strong correlation between the discharge time and features in this algorithm. The Stagewise GB algorithm was most accurate in terms of predicting discharge time. The benefit of the ensemble method is to take advantage of a collection of decision trees instead of one predictor, so it tends to yield the best results in this case study. Estimating the discharge date of a patient using KM estimator, it was observed before the 15th day in the hospital, men had a higher recovery rate than women however after the 15th day, and women tend to recover faster than men do.

4. Conclusion

Artificial Intelligence is a collection of tools and techniques that can greatly influence the impact of COVID-19 on humanity. The call to integrate AI into the medical front to combat COVID-19 is imperative. In this study, the authors spotted out the various areas; AI can help mitigate and manage the COVID-19 pandemic. Firstly with the big data available in various repositories, the application of AI will facilitate the research on this deadly virus. This study shows AI can greatly improve the rate at which early infections due to coronavirus is detected. Most studies showed that the accuracy of an AI application in detecting early infection is better than the conventional method. It holds an unlimited bundle of promises when it comes to developing a vaccine, treatment regimens and drugs to tackle this disease.

Acknowledgements

We cannot express enough thanks to Dr Ruopa Agbadede, the Head of Department, Electrical Engineering, Nigeria Maritime University, Okerenkoko for his continued support and encouragement. Dr Martins Obaseki, thank you for your leadership.

Our completion of this project could not have been accomplished without the support of our family, Mr and Mrs Lucky Enughwure & Mr and Mrs Collins Febaide:

The countless times you kept speaking words of faith into our lives will not be forgotten.

Finally, to the love of our lives, our deepest gratitude. Your encouragement when the times got rough are much appreciated and duly noted. Our heartfelt thanks.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References


