



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Contribution of machine learning approaches in response to SARS-CoV-2 infection

Mohammad Sadeq Mottaqi^a, Fatemeh Mohammadipناه^{a,*}, Hedieh Sajedi^{b,**}

^a Department of Microbial Biotechnology, School of Biology and Center of Excellence in Phylogeny of Living Organisms, College of Science, University of Tehran, 14155-6455, Tehran, Iran

^b Department of Computer Science, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, 14155-6455, Tehran, Iran

ARTICLE INFO

Keywords:

SARS-CoV-2
COVID-19
Coronavirus
Artificial intelligence
Machine learning
Respiratory infection
Deep learning

ABSTRACT

Problem: The lately emerged SARS-CoV-2 infection, which has put the whole world in an aberrant demanding situation, has generated an urgent need for developing effective responses through artificial intelligence (AI).

Aim: This paper aims to overview the recent applications of machine learning techniques contributing to prevention, diagnosis, monitoring, and treatment of coronavirus disease (SARS-CoV-2).

Methods: A progressive investigation of the recent publications up to November 2020, related to AI approaches towards managing the challenges of COVID-19 infection was made.

Results: For patient diagnosis and screening, Convolutional Neural Network (CNN) and Support Vector Machine (SVM) are broadly applied for classification purposes. Moreover, Deep Neural Network (DNN) and homology modeling are the most used SARS-CoV-2 drug repurposing models.

Conclusion: While the fields of diagnosis of the SARS-CoV-2 infection by medical image processing and its dissemination pattern through machine learning have been sufficiently studied, some areas such as treatment outcome in patients and drug development need to be further investigated using AI approaches.

1. Introduction to machine learning

Nowadays, the science of Artificial Intelligence (AI) and Machine Learning (ML) is used in most industrial activities and businesses to plan, make correct decisions, and increase the efficiency of the processes.

The most widely used branch of AI, ML, defines and explores the methods and algorithms to make computers and systems capable of learning. ML methods' ultimate aim is to gradually enhance the efficiency of the computational tasks in terms of correctness, automation, and predictions. ML has revolutionized industries such as medicine, healthcare, manufacturing, banking, and several others. Thus, it has become an essential part of modern industry.

Deep learning (DL) methods use neural network concepts for learning complex models from a massive amount of data. It is a subset of machine learning and therefore comprises supervised, unsupervised, and reinforcement methods.

Apart from clinical procedures, non-clinical techniques such as ML, data mining, deep learning, and other AI techniques afford numerous support in diagnosis, prognosis, and therapeutic motoring for several

infections with the aid of collected data. ML can be used to identify novel viruses and guess the nature of the virus through the world. Prediction for death of patients diagnosed with viruses, modeling the spread of viruses' infection, and determining their promising future impact are other ML algorithms' uses. ML techniques could be the key to finding high-quality predictive models. Compared to traditional prediction models, ML techniques have the gain of generalization.

ML techniques can be categorized into multivariate and univariate models. Multivariate models denote statistical models with more than one dependent or outcome variables. Each multivariate model can be reformulated as a univariate model by vectorizing the model. Amongst numerous ML techniques, deep learning is receiving additional attention due to the applicability to different datasets such as numerical and character. Deep learning models like Deep Feed Forward, Deep Convolutional Network, Recurrent Neural Network, Attention mechanism (Attention), and Long Short-Term Memory Networks (LSTM) can be employed for multivariate estimation.

The limitations of the current study are elaborated as follows: First, only the publications' scientific applications have been overviewed in

* Corresponding author.

** Corresponding author.

E-mail addresses: fmohammadipناه@ut.ac.ir (F. Mohammadipناه), hhsajedi@ut.ac.ir (H. Sajedi).

<https://doi.org/10.1016/j.imu.2021.100526>

Received 16 October 2020; Received in revised form 19 January 2021; Accepted 20 January 2021

Available online 24 January 2021

2352-9148/© 2021 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

this article. As a result, unreported models have not been discussed. Second, the details of the applied models have not been reviewed due to the paper's length limitation or details not being reported in the original publications. Finally, according to the fact that big data analysis has not yet been enough evolved, inferred analysis and performed learnings may not be quite accurate.

In the presented paper, we aimed to review the significant related works to investigate all aspects of ML models' contribution in response to SARS-CoV-2 infection. Our practical way of classifying various topics and subjects in this article is advantageous in contrast to other comparable review papers. Moreover, a summary of each application and model is provided, and the content is authored in a way to be more coherent and comprehensible for biologists.

ML algorithms are often categorized as supervised, unsupervised, and reinforcement [1], summarized in the following sections.

1.1. Supervised learning

Supervised ML algorithms can learn from the previous cases gathered in the past to predict future events. Starting the process from analyzing a known data set, the learning algorithm generates an abstract model to yield a prediction. The system can meet the goals of each new entry after adequate training. In the course of training, a learning algorithm can compare the model's predictions with the intended correct output (known from the previous cases) and find the model's errors to improve its performance. The problems that can be addressed by this method are divided into categories of Regression and Classification. In Regression problems, the output variable receives continuous values, and most of these problems are related to estimating or forecasting the response. Yet, the output variable in Classification problems is a class of labels that further identifies the groups or classes to which the data belongs [1].

1.2. Unsupervised learning

Unsupervised ML algorithms aim to model and use the information that is not classified and tagged. In unsupervised learning, the machine learns a function to express the hidden patterns of data. The unsupervised algorithms detect and learn patterns within the unlabeled data set. This approach places the data into different categories based on the density. Using this approach, high-dimensional data can be visualized. An example of this type of ML algorithm is Principal Component Analysis (PCA). K-means clustering is another sort of unsupervised training in which data is categorized into groups of the same type without labeling [1].

1.3. Reinforcement learning

Reinforcement learning algorithms are learning methods that communicate with their environment by generating actions and realizing penalties and rewards. These methods allow machines/agents to spontaneously determine the desired behavior in a particular field to boost their efficiency. Reward feedback is allocated for the machine/agent to determine what performs optimally.

Reinforcement learning is an emerging and popular ML algorithm used in various autonomous systems such as automobiles and industrial robotics. This algorithm tries to reach a result in a dynamic environment [2].

2. Current applications of ML in prevention of the SARS-CoV-2

2.1. Prevention of viral spread

Viral diseases spread rapidly and cannot be controlled by manual contact-tracing. An AI-based contact-tracing application for smartphones that keeps a memory of close contacts and instantly alerts them of any positive cases, if used adequately by individuals, can aid in the

management of the epidemic [3]. A similar application, running on Alipay and WeChat platforms, has been deployed in China with the same purpose. The database for the app which obtains the user information such as their medical records, exposure history, lifestyle choices, and the location of the person for tracking purposes is analyzed by an AI algorithm that gives each person a color code (green, yellow, or red) that demonstrates their health status [4].

2.2. Care management

Using AI tools to have a better quality of life during this challenging epidemic, and therefore notably controlling the spread of the infection, by designing smartphone applications or providing platforms for related data storage, is beneficial.

A team used AI to develop an app, which has the ability to recognize SARS-CoV-2 coughs from other kinds of coughs. This application (AI4COVID-19) has achieved promising accuracy of 95.6% and can reduce the load of common cold cases visiting clinical units [5].

In the paper [6], an AI and blockchain-based tracking system for SARS-CoV-2 infection is introduced as an app that requires the upload of test results. It can also inform authorities of the number of positive test results, and the GIS in the phones would facilitate tracking the infections patients.

Internet of things (IoT) has provided a platform that enables national public health agencies to approach data such as the number of daily active cases of the SARS-CoV-2 across districts and countries, intending to monitor the SARS-CoV-2 pandemic. Furthermore, some social media apps such as Twitter and Facebook are working with agencies to present real-time updates for the public [7].

2.3. Identifying epitopes for targeting vaccines

To develop the most effective immunogenic vaccine for the SARS-CoV-2, ML has been used by analyzing and predicting the suitable epitopes for vaccine development. The S protein of the SARS-CoV-2 has been extensively investigated in order to develop a stable multi-epitope vaccine. Researchers used molecular docking with TLR4, TLR2, and MHC receptors and *in-silico* cloning, codon optimization, and immune simulation to analyze the steady interactions between the predicted vaccine and MHC Toll-like receptors (TLRs) [8].

Another study analyzed the S protein to identify suitable recognizable epitopes by immune-informatics analysis and molecular docking, using the Immune Epitope Database (IEDB), SPARKS-X server, and PatchDock. They found 34, 29, and 8, linear B cell, MHC-type I, and MHC-type II epitopes, respectively [9].

Another research work [10] used a supervised ML-based inverse vaccinology application named Vaxign-ML to forecast SARS-CoV-2 vaccine candidates. Their research depicted six viral proteins as adhesive ones. They investigated the nsp3, despite spike protein, which has been frequently studied. The nsp3 was calculated to encompass 28 and 42, MHC-type I, and MHC-type II (T cell) epitopes, respectively, and have linear B cell epitopes on the protein's surface. Table 1 presents an overview of applied AI models intending to prevent SARS-CoV-2 infection.

3. Current applications of ML in diagnosis of the SARS-CoV-2

3.1. Diagnosis and patient screenings

In regions affected by the epidemic, a positive medical imaging measure and negative RT-PCR can be notable signs of COVID-19 [11]. AI can be utilized to detect the SARS-CoV-2 infection rapidly. If infection detection is accelerated, it will give healthcare professionals a better chance of controlling the pandemic.

The area under the Receiver Operating Characteristic (ROC) curve (AUC) is a test of how well a framework can classify two demonstrative

Table 1
An overview of applied novel AI models to prevent SARS-CoV-2 infection.

Method	Type	Purpose of analysis	Performance	Reference
CNN	DL	Identifying SARS-CoV-2 coughs from other kinds of coughs	96% sensitivity and 95% specificity	[5]
Cloning, Codon optimization, and Immune simulation	In-silico	Developing a stable multi-epitope vaccine	World population coverage of 95.78% for selected epitopes	[8]
Molecular docking	Immune-informatics analysis	Identifying suitable recognizable epitopes	Model "Z" score of -3.82 using ProSA server	[9]
Vaxign-ML	ML	Predicting SARS-CoV-2 vaccine candidates	Five supervised ML classification algorithms were employed	[10]

groups. The sensitivity of a test shows the effectiveness of the test in accurately recognizing the patients. Its specificity also indicates the point of the test inaccurately recognizing those without the disease [12].

An ML-based model using an Artificial Neural Network (ANN), Random Forest (RF), and shallow learning was advanced to recognize SARS-CoV-2 patients merely based on full blood counts at the early stages of the infection. The team used a dataset of standard full blood counts of SARS-CoV-2 patients at two Brazilian hospitals. Their developed independent models achieved an AUC of up to 0.86 in prediction [13].

There are a significant number of articles in the field of AI-based imaging processing of the SARS-CoV-2 infection. In this paper, merely some significant models with high detection accuracy are surveyed.

3.1.1. CT-based screening of the SARS-CoV-2 patients

According to the scarcity of time for individual radiologists to analyze medical images like CT, various AI models and algorithms such as CNN and SVM have been suggested to rapidly screen the patients' CT images and detect the SARS-CoV-2 infection with high accuracy.

An AI-based model applying CNN and conventional ML methods such as SVM, RF, and MultiLayer Perceptron (MLP) were developed to swiftly diagnose SARS-CoV-2 patients through chest CT images analysis and integrated with clinical characteristics, laboratory examination, and exposure history. CNN and ML methods were employed for diagnosis, and classification purposes, respectively. The AUC of the model in the testing set consisted of 279 patients was 0.92. The model achieved the same sensitivity as a senior radiologist [14].

CNN is a DL-based model, which is frequently used in investigating visual imagery. A CNN model, trained with a large dataset, was advanced to identify the SARS-CoV-2 patients by analyzing CT images rapidly. The sensitivity, specificity, and AUC, of the developed model were 90%, 96%, and 0.977 [15].

A DL-based system called COVID-19 Detection Neural Network (COVNet) was advanced to visually analyze chest CT images to detect the infection in a multi-center study. The sensitivity and specificity of the model at recognizing the infection in the test-set was 90% (114 out of 127) and 96% (294 out of 307), respectively, with an AUC of 0.96 [16].

A combination of DL algorithms and classification methods such as a hybrid method executing 3D classification on several crops at established resolution in an image, with a training set of 1280 international patients, was used to classify the SARS-CoV-2 patients by CT images through identifying parietal lung parenchyma. This model achieved

accuracy, sensitivity, and specificity of 90.8%, 84%, and 93% [17].

Another team of researchers advanced a DL-based method using "VB. Net" Neural Network (NN) to segment and quantify the SARS-CoV-2 infectious areas in the lung by CT images. The training set included data files of 249 patients, and the model achieved Dice similarity coefficients of 0.916 on the validation set [18].

3.1.2. X-ray based screening of SARS-CoV-2 patients

According to its availability and low cost, a chest X-ray (CXR) can be useful in detecting SARS-CoV-2 infection. A study developed an ACGAN-based model called CovidGAN as a method to produce synthetic CXR images to boost CNN performance in the detection of the infection. Applying this model, the accuracy of classification using CNN increased from 85% to 95% [19].

A DL-based model using a type of CNN (VGG16) was developed to detect SARS-CoV-2 infection by CXR images. The model was trained with 1000 CXR images of patients. They obtained 128 X-ray images from the Kaggle database and generated a larger dataset applying augmentation. The F-measure of the mentioned model was 95–99% [20].

Mohammed et al. created a Decision Matrix (DM) based on the MCDM method that combines twelve distinct machine learning SARS-CoV-2 diagnostic models and ten evaluation criteria in order to rank the models according to evaluation criteria with the amalgamation of TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) and Entropy methods. This article's evaluation criteria were precision, accuracy, recall, AUC point, time consumption, and four binary classification errors (TF, TN, FP, and FN). The analysis used 50 X-ray images. According to the several equations described and used in the mentioned study, three symbols has been calculated for each method, including Si^- , Si^+ , and Ci^+ , which represent the proximity of the model to the worst efficiency, the proximity of the model to the best efficiency, and the closeness coefficient, respectively. As a result, the SVM (linear) classifier demonstrated the best performance with the Ci^+ value of 0.9899, followed by Naive Bayes (NB) with the Ci^+ value of 0.9840, and Radial Basis Function network (RBF) the Ci^+ value of 0.9838 [21].

3.2. Predicting the probability of infection

By employing ML and AI, the probability of getting the SARS-CoV-2 infection and its transmission rate in each person can be predicted. The number of positive cases and recoveries across the world can also be forecasted using AI models.

LSTM is a regular RNN method with feedback connections, which is a useful technique to categorize, analyze, and forecast. Natural Language Processing (NLP) can also help predictions by gathering information related to the prevention and dissemination of the virus in the districts' news. An enhanced susceptible-infected (ISI) method, combined with LSTM and NLP, was developed to predict the infection rate of the SARS-CoV-2 accurately. They used text data and the number of daily-infected cases from the CDC (centers for disease control). Further, this hybrid-AI model depicted that the virus has a greater transmission rate in the period of 3–8 days after the beginning of infection in the body [22].

Support Vector Regression (SVR) in an algorithm, applied for categorization and regression challenges. A novel SVR method was introduced to analyze the SARS-CoV-2 pandemic. The accuracy of the technique in forecasting the number of positive cases, active cases, recoveries, and freshly found cases in the United States were 99.47%, 99.46%, 98.54%, and 92.1%, respectively. Based on Pearson's correlation, they noticed that the number of infected cases rises as the humidity declines in the city of Milan [23].

A team of researchers predicted the dissemination and scale of the SARS-CoV-2 pandemic using ML-based models (RS-SVM, and Bayesian Ridge) using the real-time patient dataset of Johns Hopkins Center. The data contained the symptoms and number of cases in each region. The accretion in the number of positive cases was forecasted, with 85%

accuracy. They also used Recurrent Neural Network (RNN), which saves the already computed data to calculate the recovery and mortality rate for each country with an overall accuracy of about 89% [24].

Three ML-based forecasting models have been built to estimate the number of new positive cases, deaths, and recoveries over the next ten days. This study accessed GitHub repository datasets, which were provided by Johns Hopkins University. The dataset shows the number of active and recovered cases in districts. The results demonstrate that the Exponential Smoothing (ES) model has higher performance at prediction than Linear Regression (LR) and LASSO. ES is a family of methods used for prediction through earlier courses of data. LR can identify the relationship between variables and be used in predicting. LASSO is an analysis technique for regression [25]. Another study predicted the same numbers but in the seven days ahead, using LSTM, the prophet algorithm (PA), and the ARIMA model. PA, as a prediction procedure used in R and Python, outperformed the other methods. The accuracy of their model in prediction was about 95% and 88% in Australia and Jordan, respectively [20].

A predictive model has been proposed using the ML-based FogBus structure extended on a cloud-computing platform to forecast the number of SARS-CoV-2 cases in the near future and the time it will end in various countries. They used the dataset of Our World in Data, which is daily updated from the World Health Organization (WHO). This team detected that Inverse Weibull (GIW) method fits the infection data files better than Gaussian or Beta model. A cloud platform would help us create an environment in which the institutions send their data, such as the number of daily SARS-CoV-2 infected patients [26].

A model employing data mining was advanced to forecast the recovery of patients infected by SARS-CoV-2 using a dataset of South Korean SARS-CoV-2 patients from KCDC. The dataset included data details such as age, gender, and the number of days between infection confirmation and their release from hospital or death. SVM, NB, the decision tree (DT), RF, and LR methods using python programming language were applied. DT method demonstrated the best accuracy, which was 99.85%. The model forecasted that the recovery process of SARS-CoV-2 patients takes 5–35 days. It also depicted that patients with the age of 65–85 are at high risk of not recovering from the infection, while patients with the age of 1–24, are recovered more quickly. Random forests (RF), as an ML method, proceeds by forming a large group of decision trees to handle classification or regression tasks [27]. Novel employed AI models in the field of diagnosis of SARS-CoV-2 are summarized in Table 2.

4. Current applications of ML in monitoring the SARS-CoV-2

4.1. Dissemination pattern

By utilizing AI in predictive modeling, researchers can forecast the dissemination pattern of the SARS-CoV-2 virus and its transmission across districts, states, and countries.

A team advanced an agent-based (ABM) computational imitation model (AceMod) to investigate the SARS-CoV-2 pandemic properties in Australia. The model was validated against real-time epidemic curves. They considered the factor of age and applied various strategies such as restrictions of international travels, patient isolation, and quarantine for intervening in the infection. They found out that the closure of schools merely generates a delay in the peak of the epidemic. Also, 90% compliance levels (cl) of social distancing (SD) would control the epidemic of the virus in Australia in 13 weeks, while less than 70% cl of SD fails in controlling the epidemic [28].

Hosseini et al. proposed a new SARS-CoV-2 Optimizer Algorithm (CVA) that imitates the spread of SARS-CoV-2 infection across several countries. In this model, transmission processes are different for any country according to its society’s behavior and activities. In the infection distribution process for the simulation, there are active cases and closed cases that cannot transmit the virus [29]. Broad simulations utilizing

Table 2
An overview of applied novel AI models in the field of diagnosis of the SARS-CoV-2 infection.

Method	Type	Purpose of analysis	Performance	Reference
ANN, RF, and Shallow learning	ML and DL	Identifying SARS-CoV-2 patients merely based on full blood counts	AUC of 86%	[13]
CNN, SVM, RF, and MLP	ML and DL	Identifying SARS-CoV-2 patients through chest CT images analysis	AUC of 92%, and sensitivity same as a senior radiologist	[14]
CNN	DL	Identifying SARS-CoV-2 patients through chest CT images analysis	AUC of 97.7%, sensitivity of 90%, and specificity of 96%	[15]
COVNet	DL	Detecting SARS-CoV-2 infection using CT images	AUC of 96%, sensitivity of 90%, and specificity of 96%	[16]
Combination of DL algorithms	DL	Classifying the SARS-CoV-2 patients by CT images	Accuracy of 90.8%, sensitivity of 84%, and specificity of 93%	[17]
"VB.Net" Neural Network	DL	Segmenting and quantifying the SARS-CoV-2 infected lung areas by CT images	Dice similarity coefficients of 0.916	[18]
CovidGAN	DL	Producing synthetic CXR images to boost CNN performance in the detection of the infection	CNN classification accuracy increased to 95%	[19]
VGG16	DL	Identifying SARS-CoV-2 patients by CXR images	F-measure of 95–99%	[20]
ISI, LSTM, and NLP	DL	Predicting the infection rate of the SARS-CoV-2 virus	ISI + NLP + LSTM achieved a more precise prediction than any other models	[22]
SVR	ML	Predicting the number of positive cases, active cases, recoveries, and freshly found cases in US	More than 92% overall	[23]
RS-SVM, Bayesian Ridge, and RNN	ML	Forecasting the number of positive cases	Accuracy of 85%	[24]
	DL	Predicting the recovery and mortality rate for each country	Accuracy of 89%	[24]
ES, LR, and LASSO	Regression analysis	Estimating the number of freshly positive cases, deaths, and recoveries over the next ten days	ES model achieved higher performance	[25]

(continued on next page)

Table 2 (continued)

Method	Type	Purpose of analysis	Performance	Reference
PA	Forecasting analysis	Estimating the number of new positive cases, deaths, and recoveries over the next seven days	Accuracy of 95% and 88% in Australia and Jordan, respectively	[20]
SVM, NB, DT, RF, and LR	ML	Forecasting the recovery chance of infected patients	DT achieved an accuracy of 99.85%	[27]

optimization issues depicted that this CVA model outperforms the Volcano Eruption Algorithm (VEA) [30] and Gray Wolf Optimizer (GWO) [31].

By applying RF method as a regressor to forecast Standardized Mortality Ratio (SMR) and Age-Standardized Positivity Ratio (SPR), and Canonical Correlation Analysis (CCA), on the epidemiological collected data from the Italian Civil Protection's data repository, it was revealed that among various factors, air pollution and fine-particulate (PM2.5) pollutant level correlates with the SARS-CoV-2 infectivity and mortality in Italy. According to Ref. [32], a rise of 5–10% in air pollution would result in 21–32% new cases and eventually leading to 4–14% more deaths.

4.2. Predicting the progress and severity of the disease

ML algorithms can analyze the progress and severity of the SARS-CoV-2 infection. If a patient in a critical health state is immediately recognized at the hospital admission, she/he can receive the proper treatment without any delay.

An individual SARS-CoV-2 patient's mortality probability could be estimated more than ten days ahead using an ML model based on three biomarkers, including LDH, lymphocytes, and hs-CRP. This prediction method can aid in the detection of high-risk cases and early interference. This model, named XGBoost, using a dataset of 485 patients, showed an accuracy of more than 90% [33].

According to the clinical symptoms of an individual patient at hospital admission, a DL-based Survival Cox model, which can efficiently recognize SARS-CoV-2 patients at risk of developing critical illnesses, has been developed. Ten hazardous characteristics, such as medical image abnormalities, comorbidities quantity, and age, were determined using the LASSO algorithm. This online calculation tool would assure that patients at major risk receive proper care as soon as possible. The concordance index (C-index) and AUC of this model on the internal validation set involving 1590 Chinese patients' information deriving from 575 medical centers were 0.894 and 0.911, respectively [34].

An ML-based model applying predictive analytics, SVM, and Decision tree algorithm has been developed in order to recognize SARS-CoV-2 patients with the potential to develop more critical illnesses such as acute respiratory distress syndrome (ARDS), based on clinical features such as a moderately increased alanine aminotransferase (ALT) level, the existence of myalgias (muscle aches), and an increased hemoglobin level at hospital admission. They used the medical record of 53 Chinese SARS-CoV-2 patients at two hospitals. This method's accuracy was about 70–80%, and it can be improved using a higher number of data to train [35].

A deep network, originating from Spatial Transformer Networks (STN) and by applying several AI methods, was developed to predict the severity of the SARS-CoV-2 infection by assigning a frame-based and video-based score for each person and recognizing areas encompassing pathological artifacts by automatically evaluating Lung Ultrasound (LUS) images. They developed this model using STN on 277 LUS videos

from the SARS-CoV-2 Lung Ultrasound DataBase from Italy (ICLUS-DB) to find the artifacts in the images [36].

According to the Viterbi algorithm and Hidden Markov model, an automatic program was introduced to examine and localize the pleural line in images and then evaluate the pathological severity of the infection setting score values for each image using SVM. They employed 58 videos from ICLUS-DB as the testing set. The accuracy in pleura detection for linear and convex probes was 94% and 84%, respectively. In addition, the SVM model's accuracy in classification for linear and convex probes was 94% and 88%, respectively [37].

Bai and his coworkers have discovered that the health status of many mild SARS-CoV-2 patients suddenly get worse in a brief time, and their condition deteriorates into more severe cases. They developed a DL-based model using MLP and LSTM to forecast malignant progression in infectious patients using analysis of their CT images. They acquired the best prediction AUC of 0.954 [38]. Table 3 summarizes the new AI models employed in the field of monitoring the SARS-CoV-2.

5. Current applications of ML in treatment of the SARS-CoV-2

5.1. Repurposing the existing drugs

Considering the 4–6 years needed from discovery to phase three clinical trial, repurposing the existing drugs is the fastest response for the treatment of the newly emerged viral or bacterial infections. There is an urgency to use AI-based models to rapidly recognize potential useful drugs in treating the infected patients by SARS-CoV-2.

Project IDentif. AI is a NN based model that is developed to quickly find effective medicine combinations according to correlations among mechanism of actions of drugs, their efficient doses for optimal drug synergy in the body, and safety, which does not need large data to train [39].

A research team that utilized AI-based BenevolentAI, incorporated bio-medical data deriving out of unstructured and structured authorities to examine certified medicines for SARS-CoV-2 and suggested "Baricitinib" as an inhibitor of AAK1 protein [40]. In Ref. [41], the AI-based knowledge graph of BenevolentAI was used to explore in approved drugs in order to recognize ones with both antiviral and anti-inflammatory characteristics and those that could prohibit viruses from entering the cells.

A combination of virtual drug screening, ML algorithms, and molecular docking, with the help of MOGON II supercomputer, was used to search for FDA-approved drug candidates and natural substances that might inhibit the SARS-CoV-2. Three proteins were selected as targets, including S protein, N protein, and 2'-O-ribose methyl transferase [42].

Another paper developed a DL-based medicine-target communication model, named MT-DTI, to distinguish accessible drugs that can interact with proteins of the SARS-CoV-2 virus, such as 3C-like protease and helicase, based on the scores of their binding affinity [43].

An AI-based model was developed using a Deep Neural Network to recognize available drugs, which can be used to treat SARS-CoV-2 by inhibiting the main protease of the virus. The results of feline coronavirus (FCoV) *in-vitro* assay were applied to the AI system again to enhance the model [44].

Zhang et al. designed a 3D model of 3C-like protease of the SARS-CoV-2 by analyzing RNA strings. Further, they developed a DL-based model named Dense Fully Convolutional Neural Network (DFCNN) based on neural networks for virtual drug screening and additionally used homology modeling to find suitable drugs and compounds interacting with this protease [45].

In another study, researchers gathered analytical data considering 65 human proteins, which are noted to interact with viral proteins of SARS-CoV-2. They used combined SVM models, applying Mean Absolute Error (MAE) and R, to predict FDA-approved drugs with suppressing activity on the proteins. Vapor pressure and toxicity of the forecasted drugs were also evaluated. They created a network of predicted drugs with their

Table 3

An overview of the employed novel AI models in the field of monitoring the SARS-CoV-2. Not specified.

Method	Type	Purpose of analysis	Performance	Reference
AceMod	Agent-based model (ABM)	Investigate the SARS-CoV-2 pandemic properties in Australia	NS	[28]
CVA	Optimizer algorithm	Imitating the spread of SARS-CoV-2 infection across several countries	Showed more efficiency than VEA, GWO, and PSO in several optimization problems	[29]
RF, and CCA	ML	Investigating the air pollution factor in SARS-CoV-2 infectivity	High accuracy of SMR ($R^2 = 0.95$ and RMSE = 28.9) and SPR ($R^2 = 0.93$ and RMSE = 20.3) values	10 New
XGBoost	ML	Investigating the probability of mortality of an individual SARS-CoV-2 patient	More than 90% Accuracy	[33]
Survival Cox	DL	Identifying SARS-CoV-2 patients at risk of developing critical illnesses	AUC of 91.1%, and C-index of 89.4%	[34]
SVM, and DT	ML	Recognizing SARS-CoV-2 patients with the potential to develop more critical illnesses	Accuracy of 70–80%	[35]
STN	DL	Predicting the severity of the SARS-CoV-2 infection by LUS images	F1-score of 65.1 for full model, and a pixel-wise accuracy of 96% for the segmentation model	[36]
Viterbi algorithm, Hidden Markov method, and SVM	ML	Evaluating the pathological severity of the infection by LUS images	Accuracy in pleura detection for linear and convex probes was 94% and 84%, respectively	[37]
MLP, and LSTM	ML	Employing CT images data to forecast malignant progression in infectious patients	AUC of 95.4%	[38]

associated human proteins [46].

A summary of the drugs whose efficiency in inhibiting the SARS-CoV-2 infection are predicted by artificial intelligence is provided in Table 4. As can be perceived from this table, several suitable drug repurposing candidates have been proposed by AI to combat the SARS-CoV-2 infection. In this regard, various ML and DL models such as RBF, SVM, NB, and DNN, as well as homology modeling, molecular dynamic simulations, and *in-vitro* antiviral assays were employed. Accordingly, these predictions might be advantageous for researchers and biotechnology companies, abating their research costs.

5.2. Viral protein conformation prediction

A number of 3D structures of viral proteins of the SARS-CoV-2 have

been predicted using ML techniques and molecular simulation to aid in the perception of the virus pathogenesis and its drug design.

In a study led by Heo et al., 3D structures of ten proteins derived from the SARS-CoV-2 genome were simulated. They utilized an ML method named trRosetta by applying inter-residue connection anticipations and Molecular Dynamics (MD) imitation-based processing through enhancing the protocol employed in CASP13 [49]. MD is a computer simulation method for the evaluation of the physical activity of atoms and molecules. They analyzed this model against other at-hand ones and state that their model presents more thorough coverage of sequences than the AlphaFold [50] one [51].

The critical protease of SARS-CoV-2 is called Mpro. Researchers took advantage of recent studies regarding the structure of this protein. They applied their AI-based model, integrating DL, and ML-based methods, to publish ten ideal 3D designs of Mpro in PDB format and also developed unique promising substances interacting with this protease. They used two concepts of pocket-based and ligand-based to develop new molecular conformations. They also applied 28 different ML models for structure prediction, and the structures were further optimized according to the reward function in reinforcement learning to obtain the best scores [52].

A DL-based (CNN) model has been introduced to estimate the localization of sub-cellular viral proteins of SARS-CoV-2. This model's regional accuracy is above 98%, which can effectively discover various-target drugs for viral infections [53].

5.3. Drug development

In the development of drugs to treat the newly emerged SARS-CoV-2 virus, time is a critical factor. Machine learning and AI can be beneficial for us by accelerating the pace of the process of drug development and clinical trials.

Researchers used a DL-based model evaluated by R^2 , Root Mean Square Error (RMSE), and MAE-Mean Square Error (MSE) metrics, for virtual evaluation and MD simulations to identify natural substances inhibiting the viral Mpro, which led to proposing the Sauchinone and Palmatine as the potential effective compounds among 1611 substances. These compounds further showed promising results based on the MMPBSA approach, which predicts the complex's stability [54].

Ton et al. built a new DL-based model called Deep Docking, which facilitated a structure-based virtual evaluation of 1.3 billion substances from the ZINC15 library to recognize suitable ones interacting with Mpro of the virus. They identified 1000 substances with the best score from Glide SP. Molecular structures of the recognized substances are made openly accessible [55].

Another paper developed a program to identify substances for the treatment of SARS-CoV-2, by applying virtual evaluation, structure-based design of drugs, high-throughput evaluation, and antiviral cell-based analysis against the Mpro of the SARS-CoV-2. Out of 10,000 pharmacological substances, six compounds demonstrated the ability to inhibit Mpro, and one (Ebselen) indicated outstanding results in FRET assays [56].

Fig. 1 illustrates the repurposed drugs and compounds that their inhibitory effect on SARS-CoV-2 was predicted in AI modeling and proved experimentally due to the *in-silico* prediction using ML approaches. As this figure demonstrates, each of these AI-predicted compounds has shown suppressive effects on the SARS-CoV-2 *in-vitro* assays. Therefore, it can be considered to assess the candidate drugs for the infection. The sharing and availability of data in this era are critical in order to accelerate the identification of the best treatments for diseases [57].

A review of the compounds whose efficiency in inhibiting the SARS-CoV-2 infection are predicted by artificial intelligence in recent related publications is presented in Table 5. As can be seen from this table, various ML and DL models such as NB, DFCNN, and homology modeling, molecular docking, and *in-silico* screening are applied to predict proper

Table 4
Repurposed drugs which their efficiency in inhibition of SARS-CoV-2 infection is predicted by artificial intelligence.

Name	Target	Prediction method	In vitro approved effect	Original clinical uses	References
Atazanavir	3C-like protease	DL (MT-DTI), ML (NLP, and NB), and AutoDock Vina	-	HIV	[43,47]
Astemizole	VCP, RHOA, TK2	ML (SVM, and RBF)	-	Antihistamine	[40,41,46]
Baricitinib	AAK1 protein Janus kinase (JAK)	BenevolentAI	-	Rheumatoid arthritis	[40,41,44]
Bazedoxifene	VCP, RHOA, ABCC1	ML (SVM, and RBF)	-	Postmenopausal osteoporosis	[44,46]
Bedaquiline	3C-like protease	DL (DNN), in-vitro antiviral assays	+	Tuberculosis	[44]
Brequinar	3C-like protease	DL (DNN), in-vitro antiviral assays	+	Immunosuppressant	[42,44]
Celecoxib	3C-like protease	DL (DNN), in-vitro antiviral assays	+	Anti-inflammatory	[43,44]
Clofazimine	Antiviral activities	DL (DNN), in-vitro antiviral assays	+	Leprosy	[41,44]
Convaptan	Nucleocapsid protein, 3C-like protease	Homology modeling, molecular docking, supervised ML (NN, and NB), DL (DNN), and in-vitro antiviral assays	+	Hyponatremia	[42,44,45]
Dabrafenib	BRD4, PSEN2, IDE	ML (SVM, and RBF)	-	Anti-Cancer	[44,46]
Darolutamide	TK2, TBK1, VCP	ML (SVM, and RBF)	-	Prostate cancer	[42,46]
Efavirenz	3C-like protease	DL (MT-DTI), ML (NLP), and AutoDock Vina	-	HIV	[43]
Entrectinib	IDE, MARK2, VCP	ML (SVM, and RBF)	-	Anti-Cancer	[42,46]
Etoricoxib	BRD4, PRKACA, DCTPP1	ML (SVM, and RBF)	-	Rheumatoid arthritis	[43,46]
Fedratinib	Janus kinase (JAK)	BenevolentAI	-	Myeloproliferative diseases	[41,43]
Ganciclovir	3C-like protease, RNA-dependent RNA polymerase	Homology modeling, DL (DFCNN, and MT-DTI), and AutoDock Vina	-	Cytomegalovirus (CMV) infections	[41,43,45]
Gemcitabine	Antiviral activities	DL (DNN), in-vitro antiviral assays	+	Anti-cancer	[42,44]
Grazoprevir	Spike protein	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	Hepatitis C	[42,44]
Ibrutinib	BRD4, IDE, TK2	ML (SVM, and RBF)	-	B cell cancer	[42,46]
Lapatinib	PSEN2, IDE, BRD4	ML (SVM, and RBF)	-	Breast cancer	[45,46]
Lasofoxifene	VCP, RHOA, ABHD12	ML (SVM, and RBF)	-	Osteoporosis	[44,46]
Lestaurtinib	BRD2, TBK1, MARK2	ML (SVM, and RBF)	-	Anti-Cancer	[42,46]
Lifitegrast	ITGB1, BRD4, PTGES2	ML (SVM, and RBF)	-	Dry eye	[44,46]
Lopinavir/Ritonavir	RNA Helicase	IDentif.AI, DL (MT-DTI), ML (NLP, and NB), and AutoDock Vina	-	HIV	[43,45,47,48]
Lumacaftor	IMPDH2, EIF4H, BRD2	ML (SVM, and RBF)	-	Cystic fibrosis	[45,46]
Midostaurin	TBK1, BRD2, MARK3	ML (SVM, and RBF)	-	Acute myeloid leukemia (AML)	[45,46]
Paritaprevir	Spike protein, 2'-o-ribose methyltransferase	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	Hepatitis C	[42,45,47]
Remdesivir	3C-like protease	IDentif.AI, DL (MT-DTI), ML (NLP), and AutoDock Vina	-	Antiviral	[42,43,48]
Ribociclib	PRKACA, ABCC1, HDAC2	ML (SVM, and RBF)	-	Breast cancer	[45,46]
Ritonavir	3C-like protease	IDentif.AI, DL (MT-DTI), ML (NLP, and NB), and AutoDock Vina	-	HIV	[43,45,47,48]
Ruxolitinib	Janus kinase (JAK)	BenevolentAI	-	Myelofibrosis	[41,42]
Simeprevir	Spike protein	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	Hepatitis C	[42]
Talazoparib	CSNK2A2, BRD4, BRD2	ML (SVM, and RBF)	-	Breast cancer	[46]
Telmisartan	NSD2, PRKACA, PTGES2	ML (SVM, and RBF)	-	Hypertension	[46]
Tolcapone	3C-like protease	DL (DNN), in-vitro antiviral assays	+	Parkinson's disease	[44]
Triazolam	CSNK2b, RIPK1, DCTPP1	ML (SVM, and RBF)	-	Insomnia	[46]
Velpatasvir	Spike protein	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	Hepatitis C	[42]
Vidarabine	3C-like protease	Homology modeling and DL (DFCNN)	-	Antiviral	[45]
Vismodegib	3C-like protease	DL (DNN), in-vitro antiviral assays	+	Basal-cell carcinoma (BCC)	[44]

drug candidates to control the SARS-CoV-2 pandemic. Although each of the publications has employed distinct parameters and data to predict, these predictions might come in handy in treating the infection if they show acceptable pharmacodynamics and pharmacokinetic properties in clinical trials.

5.4. Predicting the outcome of treatment

The disease's progress in the SARS-CoV-2 infection following the drug administration can be predicted using the ML approach.

An AI-based model (AIMDP) was developed to predict the response of SARS-CoV-2 patients to treatment. This model has two functions, including diagnosis (DM) using chest CT images and prediction (PM). The PM function utilizes CNN for the purpose of segmentation by applying the Whale Optimization Algorithm. PM function provides a probability of a patient's capability to respond to SARS-CoV-2 treatments according to the different input factors, including age, infection

level, treatment procedures, respiratory failure, and multi-organ failure [58]. AIMDP model trained by a dataset of 617 CT images to detect the infection, demonstrated higher accuracy and sensitivity than other similar methods such as COVNet [16].

In a study, 315 SARS-CoV-2 patients were investigated for two weeks using an AI-based radiomics method applying data mining algorithms. To anticipate the patient end-result concerning ICU acceptance, AUC for radiomics analysis and radiologists analysis were 0.75 and 0.68, respectively [59].

A technique that combines 3D analysis and 2D slice-level AI methods was developed to assess the progress of the infection in each patient over time, bringing about a score. They used a testing set of 157 patients. AUC, sensitivity, and specificity of this classification model was 0.99, 98.2%, and 92.2%, respectively [60]. Table 6.

The drug administration's efficacy can be evaluated from the analysis of the lung's opacification from CT scans using a measurable parameter of the image (QCT-PLO) applying a deep learning software.

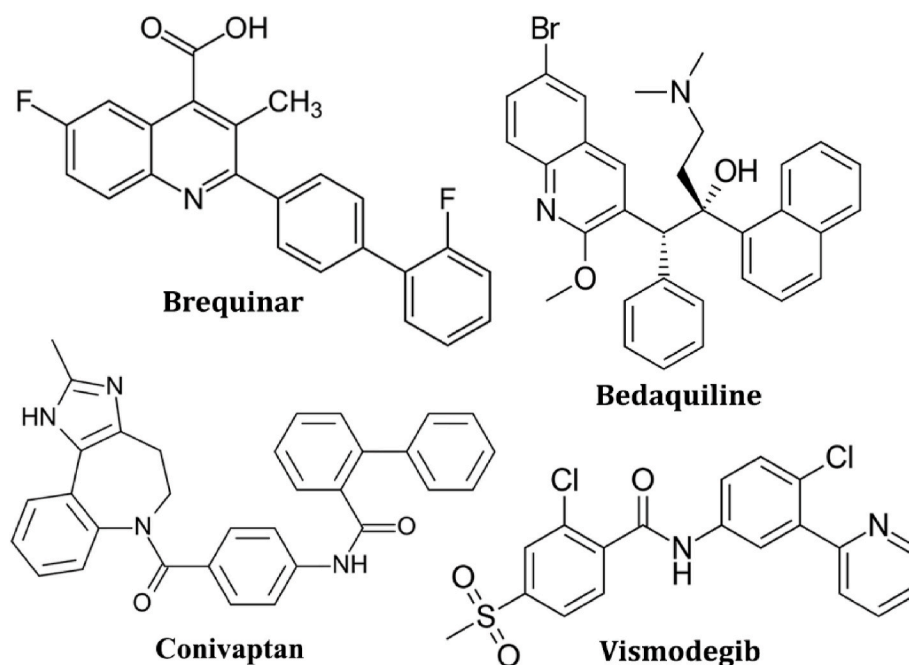


Fig. 1. AI-predicted FDA-approved drugs that have exhibited effectiveness in anti-SARS-CoV-2 assays.

Table 5

AI-predicted compounds with potential in the treatment of the SARS-CoV-2 infection. Not Available.

Name	Target	Prediction method	In vitro approved effect	Original clinical uses	References
Amyrin	Nucleocapsid protein, Spike protein, 2'-o-ribose methyltransferase	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	NA	[42]
Chlorobutanol	3C-like protease	Homology modeling and DL (DFCNN)	-	NA	[45]
D-Mannitol	3C-like protease	Homology modeling and DL (DFCNN)	-	NA	[45]
D-Sorbitol	3C-like protease	Homology modeling and DL (DFCNN)	-	NA	[45]
Ebselen	3C-like protease	"In-silico screening, FRET assay, in-vitro antiviral assays	+	NA	[47]
Ile + Lys + Pro	3C-like protease	Homology modeling and DL (DFCNN)	-	NA	[45]
Loniflavone	Spike protein	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	NA	[42]
Meglumine	3C-like protease	Homology modeling and DL (DFCNN)	-	NA	[45]
Palmatine	3C-like protease	DL(DeepScreening), AutoDock VINA, ADMET analysis, and MM-PBSA	-	NA	[54]
Sauchinone	3C-like protease	DL(DeepScreening), AutoDock VINA, ADMET analysis, and MM-PBSA	-	NA	[54]
Sodium_gluconate	3C-like protease	Homology modeling and DL (DFCNN)	-	NA	[45]
ZINC000008635575	Spike protein, 2'-o-ribose methyltransferase	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	NA	[42]
ZINC000027215582	Nucleocapsid protein, Spike protein, 2'-o-ribose methyltransferase	Homology modeling, AutoDock VINA, molecular docking, and supervised ML (NN, and NB)	-	NA	[42]

Follow-up CT images of 126 SARS-CoV-2 patients were analyzed. The outcome of this study indicates that pulmonary engagement in patients dramatically increases after the onset of symptoms, and it reaches its peak on the 13th day [61].

6. Future suggested applications of ML for management of SARS-CoV-2

The computational algorithms used by ML methods can extract the interpretations and solutions in a variety of ways, which are not possible by the conventional computational tools. The rate of dissemination and evolution of the SARS-CoV-2 demands rapid reactions and ML approaches attempt to limit its infection and different consequences relating to it.

ML adds another dimension to how information is concluded by both continuous generations of data on SARS-CoV-2 and the improvement in computing power. For instance, the environmental stability of SARS-

CoV-2 on biotic and inanimate surfaces needs to be modeled by the ML approaches based on the variations in the nature of the surfaces and the surrounding physicochemical conditions.

Besides, ML approaches are able to analyze the pattern discovery of human fluids components such as a serum, mucosal liquid, urine, etc., following the infection by SARS-CoV-2. Furthermore, the empirical machine learning approach itself might be improved, which would lead to predictions with higher accuracy. Screening of the natural or synthetic compound libraries needs to be conducted in a more similar condition to the human body in order to rely on the observed interaction between the target protein and the drug candidate. The ML approaches could facilitate this simulation without the need to mimic the physiological condition.

Through proper documentation of the enormous data deriving from infected individuals in this coronavirus pandemic, patterns can be calculated by ML approaches to get us closer to personalized medicine in antiviral therapy. Complex prediction of the treatment-result in case of

Table 6

An outline of the employed novel AI models in the treatment of the SARS-CoV-2 infection.

Method	Type	Purpose of analysis	Performance	Reference
trRosetta	MD	Simulating the 3D structure of ten viral SARS-CoV-2 proteins	Less than 1.5 MolProbity score for the generated models	[51]
Several methods	ML and DL	Publishing ideal 3D designs of Mpro	28 ML models were employed to optimize the reward function	[52]
CNN	DL	Estimating the localization of sub-cellular viral proteins of SARS-CoV-2	Accuracy of more than 98%	[53]
AIMDP, and CNN	DL	Predicting the response of SARS-CoV-2 patients to treatment	Achieved more precision, accuracy, and sensitivity than other models such as COVNet and DeConNet	[35]
Radiomics	ML and DL	Anticipating the patient end-result concerning ICU acceptance	AUC of 75%	[59]
3D analysis and 2D slice-level methods	ML and DL	Assess the progress of the infection in each SARS-CoV-2 patients over time	AUC of 99.6%, sensitivity of 98.2%, and specificity of 92.2%	[60]

the simultaneous disease in the infected person needed to be performed by advances in data collections that could feed the machine learning platforms.

The use of resident stem cells, which can act as immunomodulatory and pro-osteogenic players in the local environment by taking advantage of AI progression to find treatment for SARS-CoV-2 infection, shall be further investigated in the near future. Some advancements in this field of study, with AI's help, would contribute to the improvement of personalized treatments with regenerated tissues by employing predictive modeling, which leads to reducing the complexity of the infection and abating the costs [62].

Mesenchymal Stem Cells (MSC) Assist human beings by having an immunomodulatory function in viral infections as well as efficiently lessening inflammation in acute lung injuries [63,64]. Various technologies have been applied in order to enhance the phenotype of cells developed from MSCs, which are employed as therapeutic supports in clinical applications [65]. Oral tissues are known to be valuable due to being rich in stem cells deriving out of various sources. Besides the Dental pulp stem cells (DPSCs), which are hugely investigated, hPCy-MSCs, stemming from wisdom teeth, has recently depicted promising characteristics such as vascular and neural phenotypes while maintaining the ability to proliferate [66,67]. Furthermore, the role of biomaterials or scaffolds as study models and the role of preclinical studies on biomaterials as part of tissue engineering protocols could be briefly improved with the help of AI. The robust effects of surface harshness on the mechanical intensity of samples, including Zirconia, have been identified [68]. The aforementioned study arranged two-layered zirconia-covered samples in order to do three-point bending evaluation tests. Due to the significance of the effect of surface roughness, researchers proposed employing Computer-Aided Design (CAD) or Computer-Aided Manufacturing (CAM) for the establishment of zirconia-based bridges and crowns.

Another proposed challenge in managing such a viral pandemic is scoring the performance of the health care or research and development staff in terms of following the biosafety level protocols. A personal reliable system can be developed using ML to estimate the reliability of the robustness of the individuals' performance in working with

infectious samples or patients to be recruited or kept in the high-risk working area. Some algorithms can be employed to train software programs to estimate the individuals' physiological susceptibility (cell-phone monitoring) to get symptomatic or severe infection by SARS-CoV-2 virus. In the long-term future, this estimation can turn into a more accurate prediction of the type of data for the machine learning changes, from physiological values to the individuals' genomic sequence.

Genomic computational tools can also increase data extraction by introducing new ML approaches for nucleotide sequence analysis. The rapid response needed by the highly contagious viral agents demands the development of ML-based decision-making, which needs to be developed for the policymakers. Another future upcoming advancement will be the addition of the ML analysis tools to the viruses' database to estimate some situations, which are queried from the database. The endurance testing process of the biosafety level-3 labs can also be improved using ML techniques to teach the machine the consequence of the flaws and predict the level of biosecurity in case of any deviations. Fig. 2.

The predominantly used approaches in training the machines also need to be enhanced by mathematical advances in developing new algorithms. In other words, it seems that in the data generation for the pandemic, the computational ability of the software is still not balanced with the capacity of the hardware.

7. Concluding remarks

The most recent publications regarding the advancement and use of ML and DL approaches in response to the SARS-CoV-2 infection are summarized in this paper to facilitate the pace of dealing with the pandemic around the globe. With the advancement of data recording and collection and their diverse assortment of applications, applying these models can prevail upon discovering the most effective and accurate approaches in managing and treating this lethal infection, abating the medical costs, and aiding the field of data analytics. We proposed a detailed overview through the aforementioned sections, intending to provide researchers with a thorough review of related and the most accurately developed models in various fields in managing this pandemic.

The majority of studies have been focused on drug repurposing using virtual screenings, analysis of dissemination patterns of the spread of the infection, and clinical diagnosis of the SARS-CoV-2. On the other hand, in areas such as drug development, clinical trials, and prediction of the treatment's end-result, more authentic and advanced machine learning-based models are still required to be developed.

In modeling the dissemination pattern of the spread of the infection and its transmission processes, SVR, LSTM, and optimization problems, among other models, have been frequently used to investigate the SARS-CoV-2 pandemic properties so far.

Machine learning methods including, CNN, SVM, and Generative Adversarial Network (GAN) are broadly applied to process the SARS-CoV-2 diagnosis for classification purposes in the field of medical image processing. These models often evaluate CT, X-ray, and LUS images of the patients. Several models in the aforementioned studies have been able to achieve more than 90% accuracy in detecting the infection.

Models including SVM, Decision tree (DT), and Spatial Transformer Networks (STN), have been dominantly employed by researchers in order to identify SARS-CoV-2 patients with the potential to develop more critical illnesses, based on evaluating LUS images of patients, or other clinical characteristics. According to the so far reports, Deep Neural Network (DNN), homology modeling, and molecular docking are the most used SARS-CoV-2 drug repurposing models in the rapid identification of potentially effective drugs for the treatment of the SARS-CoV-2 infection.

So far, all ML applications can be categorized in detection/monitoring, prevention/treatment, and pathogenesis analysis of this virus. More efforts have been focused on the drug repurposing, analysis of

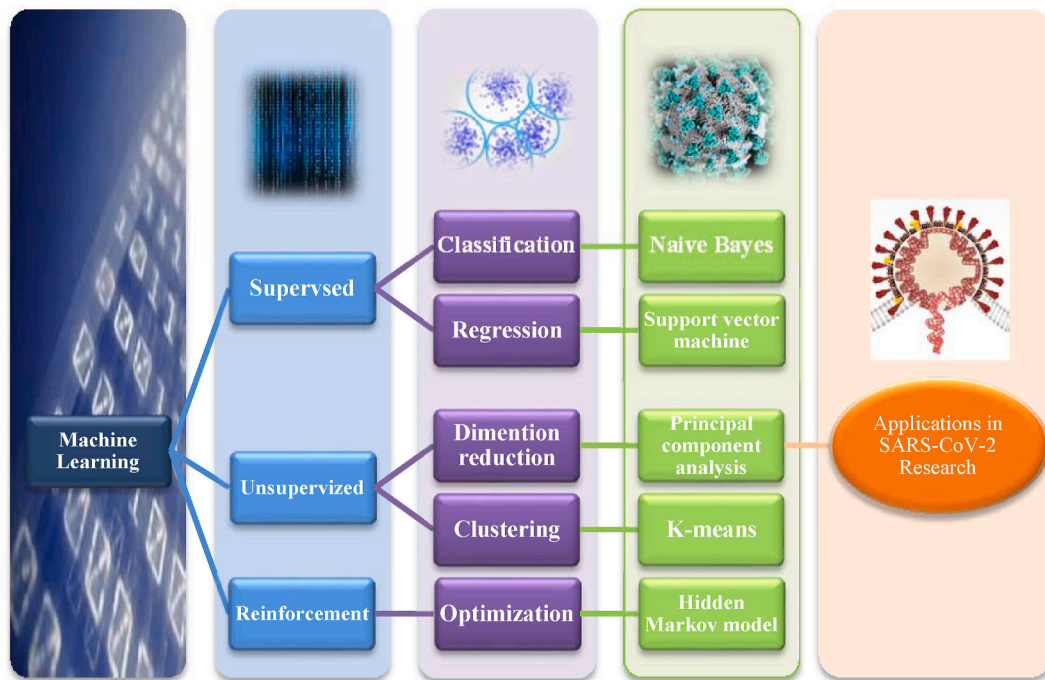


Fig. 2. Prevalent machine learning algorithms in SARS-CoV-2 research.

dissemination patterns, and clinical radiographs in terms of the reports' quantity. Nevertheless, the proper data record of the disease trend and individual status and response to the administrated antiviral drugs could lead to the best-recommended treatment with the existing drugs. However, in vast distributed clinical centers in world, there is not a proper recording of the cases that could feed the ML analysis for this purpose. Such inferred biomedical knowledge graphs are critically demanding for the appropriate management of the infected cases.

Among the ML algorithms, supervised and unsupervised ones such as SVM and NB, have been more functional for the analysis of coronavirus related data, and the reinforcement approach is employed more in clinical rather than population analysis.

The most continuous line of data generation is the daily deposition of the genome sequence of the SARS-CoV-2 clinical isolates worldwide. ML can aid in predicting pathogenesis, the evolution of the virus, and its drug susceptibility for the comparative sequence analysis. Improvement of technologies and their accreted accessibility is remarkably shaping the landscape of the ongoing research on managing treatment of Coronavirus infection using a machine learning approach.

With the aid of ML, intelligent systems can be developed to make autonomous decisions in such pandemic instead of clinicians and policymakers. Previous data samples can teach these algorithms through statistical analysis and pattern matching. Then, based on the trained data, they provide the predicted results. The exponential data generated by research on this virus accumulates daily, feeding the machine learning developments to get meaningful insights from this large collected data.

Funding

This research did not receive any specific grant from funding agencies.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgment

Not applicable.

References

- [1] Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006. ©20062006.
- [2] Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. *J Artif Intell Res* 1996;4(1):237–85.
- [3] Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, Parker M, Bonsall D, Fraser C. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* 2020;368(6491).
- [4] Paul Mozur RZ. Aaron Krolik, in coronavirus fight, China gives citizens a color code, with red flags. <https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html>; 2020.
- [5] Imran A, Posokhova I, Qureshi HN, Masood U, Riaz MS, Ali K, John CN, Hussain MDI, Nabeel M. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Informatics in Medicine Unlocked* 2020; 20:100378.
- [6] Mashamba-Thompson TP, Crayton ED. Blockchain and artificial intelligence technology for novel coronavirus disease-19 self-testing. *Diagnostics* 2020;10(4): 198.
- [7] Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and COVID-19. *Nat Med* 2020;26(4):459–61.
- [8] Kar T, Narsaria U, Basak S, Deb D, Castiglione F, Mueller DM, Srivastava AP. A candidate multi-epitope vaccine against SARS-CoV-2. *Sci Rep* 2020;10(1):10895.
- [9] Bhattacharya M, Sharma AR, Patra P, Ghosh P, Sharma G, Patra BC, Lee S-S, Chakraborty C. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach. *J Med Virol* 2020; 92(6):618–31.
- [10] Ong E, Wong MU, Huffman A, He Y. COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning, bioRxiv : the preprint server for biology. 2020.
- [11] Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* 2020;296(2):E32–40.
- [12] Trevelyan R. Sensitivity, specificity, and predictive values: foundations, pliabilitys, and pitfalls in research and practice. *Frontiers in Public Health* 2017;5 (307).
- [13] Banerjee A, Ray S, Vorselaars B, Kitson J, Mamalakis M, Weeks S, Baker M, Mackenzie LS. Use of machine learning and artificial intelligence to predict SARS-CoV-2 infection from full blood counts in a population. *Int Immunopharm* 2020;86: 106705.

- [14] Mei X, Lee H-C, Diao K-y, Huang M, Lin B, Liu C, Xie Z, Ma Y, Robson PM, Chung M, Bernheim A, Mani V, Calcagno C, Li K, Li S, Shan H, Lv J, Zhao T, Xia J, Long Q, Steinberger S, Jacobi A, Deyer T, Luksza M, Liu F, Little BP, Fayad ZA, Yang Y. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat Med* 2020;26(8):1224-8.
- [15] Chen W, Cao Y, Xu Z, Zhang X, Deng L, Zheng C, Zhou J, Shi H, Feng J. Development and evaluation of an AI system for COVID-19. 2020.
- [16] Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology* 2020;296(2):E65-71.
- [17] Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, Yang D, Myronenko A, Anderson V, Amalou A, Blain M, Khassin M, Long D, Varble N, Walker SM, Bagci U, Ierardi AM, Stellato E, Plensich GG, Franceschelli G, Girlando C, Irmici G, Labella D, Mahmoud D, Malayeri A, Jones E, Summers RM, Choyke PL, Xu D, Flores M, Tamura K, Obinata H, Mori H, Patella F, Cariati M, Carrafiello G, An P, Wood BJ, Turkbey B. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun* 2020;11(1):4080.
- [18] Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D, Shi Y. Lung infection quantification of COVID-19 in CT images with deep learning. 2020.
- [19] Waheed A, Goyal M, Gupta A, Khanna M, Long D, Varble N, Walker SM, Bagci U, Ierardi AM, Stellato E, Plensich GG, Franceschelli G, Girlando C, Irmici G, Labella D, Mahmoud D, Malayeri A, Jones E, Summers RM, Choyke PL, Xu D, Flores M, Tamura K, Obinata H, Mori H, Patella F, Cariati M, Carrafiello G, An P, Wood BJ, Turkbey B. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun* 2020;11(1):4080.
- [20] Alazab M, Awajan A, Mesleh A, Abraham A, Jatana V, Alhyari S. COVID-19 prediction and detection using deep learning. *International Journal of Computer Information Systems and Industrial Management Applications* 2020;12:168-81.
- [21] Mohammed MA, Abdulkareem KH, Al-Waisy AS, Mostafa SA, Al-Fahdawi S, Dinar AM, Alhakami W, Al-Mhiqani MN, Alhakami H, Arbaiy N, Maashi MS, Mutlag AA, García-Zapirain B, Díez ID. Benchmarking methodology for selection of optimal COVID-19 diagnostic model based on Entropy and TOPSIS methods. *IEEE Access* 2020;8:99115-31.
- [22] Zheng N, Du S, Wang J, Zhang H, Cui W, Kang Z, Yang T, Lou B, Chi Y, Long H, Ma M, Yuan Q, Zhang S, Zhang D, Ye F, Xin J. Predicting COVID-19 in China using hybrid AI model. *IEEE Trans Cybern* 2020;50(7):2891-904.
- [23] Yadav M, Perumal M, Srinivas M. Analysis on novel coronavirus (COVID-19) using machine learning methods. *Chaos, Solit Fractals* 2020;139:110050.
- [24] Arun SS, Iyer GN. On the analysis of COVID19 - novel corona viral disease pandemic spread data using machine learning techniques. 4th international conference on intelligent computing and control systems (ICICCS), 2020, 2020. p. 1222-7.
- [25] Rustam F, Reshi AA, Mehmood A, Ullah S, On B, Aslam W, Choi GS. COVID-19 future forecasting using supervised machine learning models. *IEEE Access* 2020;8:101489-99.
- [26] Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. *Internet of Things* 2020;11:100222.
- [27] Muhammad LJ, Islam MM, Usman SS, Ayon SI. Predictive data mining models for novel coronavirus (COVID-19) infected patients' recovery. *SN Computer Science* 2020;1(4):206.
- [28] Chang S, Harding N, Zachreson C, Cliff O, Prokopenko M. Modelling transmission and control of the COVID-19 pandemic in Australia. 2020.
- [29] Hosseini E, Ghafoor K, Sadiq A, Guizani M, Emrouznejad A. COVID-19 optimizer algorithm, modeling and controlling of coronavirus distribution process. *IEEE Journal of Biomedical and Health Informatics* 2020:1.
- [30] Hosseini E, Sadiq A, Ghafoor K, Rawat DB, Saif M, Yang X. Volcano eruption algorithm for solving optimization problems, *Neural Computing and Applications*. 2020.
- [31] Mirjalili S, Mirjalili SM, Lewis A. Grey Wolf optimizer. *Adv Eng Software* 2014;69:46-61.
- [32] Cazzolla Gatti R, Velichevskaya A, Tateo A, Amoroso N, Monaco A. Machine learning reveals that prolonged exposure to air pollution is associated with SARS-CoV-2 mortality and infectivity in Italy. *Environ Pollut* 2020;267:115471.
- [33] Yan L, Zhang H-T, Goncalves J, Xiao Y, Wang M, Guo Y, Sun C, Tang X, Jing L, Zhang M, Huang X, Xiao Y, Cao H, Chen Y, Ren T, Wang F, Xiao Y, Huang S, Tan X, Huang N, Jiao B, Cheng C, Zhang Y, Luo A, Mombaerts L, Jin J, Cao Z, Li S, Xu H, Yuan Y. An interpretable mortality prediction model for COVID-19 patients. *Nature Machine Intelligence* 2020;2(5):283-8.
- [34] Liang W, Yao J, Chen A, Lv Q, Zanin M, Liu J, Wong S, Li Y, Lu J, Liang H, Chen G, Guo H, Guo J, Zhou R, Ou L, Zhou N, Chen H, Yang F, Han X, Huan W, Tang W, Guan W, Chen Z, Zhao Y, Sang L, Xu Y, Wang W, Li S, Lu L, Zhang N, Zhong N, Huang J, He J. Early triage of critically ill COVID-19 patients using deep learning. *Nat Commun* 2020;11(1):3543.
- [35] Hassanien A, Elghamrawy S. Diagnosis and prediction model for COVID-19 patients response to treatment based on convolutional neural networks and Whale optimization algorithm using CT images. 2020.
- [36] Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, Huijben I, Chennakeshava N, Mento F, Sentelli A, Peschiera E, Trevisan R, Maschietto G, Torri E, Inchingolo R, Smargiassi A, Soldati G, Rota P, Passerini A, Sloun RJGV, Ricci E, Demi L. Deep learning for classification and localization of COVID-19 markers in point-of-care lung Ultrasound. *IEEE Trans Med Imag* 2020;39(8):2676-87.
- [37] Carrer L, Donini E, Marinelli D, Zanetti M, Mento F, Torri E, Smargiassi A, Inchingolo R, Soldati G, Demi L, Bovolo F, Bruzzone L. Automatic pleural line extraction and COVID-19 scoring from lung Ultrasound data. *IEEE Trans Ultrason Ferroelectrics Freq Contr* 2020:1.
- [38] Bai X, Fang C, Zhou Y, Bai S, Liu Z, Chen Q, et al. Predicting COVID-19 malignant progression with AI techniques. *medRxiv* 2020. <https://doi.org/10.1101/2020.03.20.20037325>.
- [39] Ho D. Addressing COVID-19 drug development with artificial intelligence. *Advanced Intelligent Systems* 2020;2(5):2000070.
- [40] Richardson P, Griffin I, Tucker C, Smith D, Oechsle O, Phelan A, Rawling M, Savory E, Stebbing J. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *Lancet* 2020;395(10223):e30-1.
- [41] Stebbing J, Phelan A, Griffin I, Tucker C, Oechsle O, Smith D, Richardson P. COVID-19: combining antiviral and anti-inflammatory treatments. *Lancet Infect Dis* 2020;20(4):400-2.
- [42] Kadioglu O, Saeed M, Greten H, Efferth T. Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. 2020.
- [43] Beck BR, Shin B, Choi Y, Park S, Kang K. Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Comput Struct Biotechnol J* 2020;18:784-90.
- [44] Ke Y-Y, Peng T-T, Yeh T-K, Huang W-Z, Chang S-E, Wu S-H, et al. Artificial intelligence approach fighting COVID-19 with repurposing drugs. *Biomed J* 2020;43(4):355-62.
- [45] Zhang H, Saravanan KM, Yang Y, Hossain MT, Li J, Ren X, Pan Y, Wei Y. Deep learning based drug screening for novel coronavirus 2019-nCoV. *Interdiscip Sci* 2020;12(3):368-76.
- [46] Kowalewski J, Ray A. Predicting novel drugs for SARS-CoV-2 using machine learning from a >10 million chemical space. *Heliyon* 2020;6(8):e04639.
- [47] Mahapatra S, Nath P, Chatterjee M, Das N, Kalita D, Roy P, et al. Repurposing therapeutics for COVID-19: rapid prediction of commercially available drugs through machine learning and docking. *medRxiv* 2020;15(11):e0241543.
- [48] Blasiak A, Lim J, Seah S, Kee T, Remus A, Chye D, Wong P, Hooi L, Truong A, Le N, Chan C, Desai R, Ding X, Hanson B, Chow E, Ho D, IDentif AI. Artificial intelligence pinpoints remdesivir in combination with ritonavir and lopinavir as an optimal regimen against severe acute respiratory syndrome coronavirus 2. 2020 (SARS-CoV-2).
- [49] Heo L, Arbour CF, Feig M. Driven to near-experimental accuracy by refinement via molecular dynamics simulations. *Proteins: Structure, Function, and Bioinformatics* 2019;87(12):1263-75.
- [50] Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Židek A, Nelson AWR, Bridgland A, Penedones H, Petersen S, Simonyan K, Crossan S, Kohli P, Jones DT, Silver D, Kavukcuoglu K, Hassabis D. Improved protein structure prediction using potentials from deep learning. *Nature* 2020;577(7792):706-10.
- [51] Heo L, Feig M. Modeling of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins by machine learning and physics-based refinement, *bioRxiv* : the preprint server for biology. 2020.
- [52] Zhavoronkov A, Aladinskiy V, Zhebrak A, Zagribelnny B, Terentiev V, Bezrukov D, Polykovskiy D, Shayakhmetov R, Filimonov A, Orekhov P, Yan Y, Popova O, Vanhaelen Q, Aliper A, Ivanenkov Y. Potential 2019-nCoV 3C-like protease inhibitors designed using generative deep learning approaches. 2020.
- [53] Shao Y, Chou K-C. pLoc-Deep-mVirus: a CNN model for predicting subcellular localization of virus proteins by deep learning. *Nat Sci* 2020;12:388-99.
- [54] Joshi T, Joshi T, Pundir H, Sharma P, Mathpal S, Chandra S. Predictive modeling by deep learning, virtual screening and molecular dynamics study of natural compounds against SARS-CoV-2 main protease. *J Biomol Struct Dyn* 2020:1-19.
- [55] Ton A-T, Gentile F, Hsing M, Ban F, Cherkasov A. Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3Billion compounds. *Molecular Informatics* 2020;39(8):2000028.
- [56] Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, Duan Y, Yu J, Wang L, Yang K, Liu F, Jiang R, Yang X, You T, Liu X, Yang X, Bai F, Liu H, Liu X, Guddatt LW, Xu W, Xiao G, Qin C, Shi Z, Jiang H, Rao Z, Yang H. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* 2020;582(7811):289-93.
- [57] Cosgriff CV, Ebner DK, Celi LA. Data sharing in the era of COVID-19. *Lancet Digit Health* 2020;2(5):e224.
- [58] Elghamrawy SM, Hassanien AE. Diagnosis and prediction model for COVID19 patients response to treatment based on convolutional neural networks and Whale optimization algorithm using CT images. *medRxiv* 2020. <https://doi.org/10.1101/2020.04.16.20063990>.
- [59] Homayounieh F, Ebrahimian S, Babaei R, Mobin HK, Zhang E, Bizzo BC, Mohseni I, Digumarthy SR, Kalra MK. CT radiomics, radiologists and clinical information in predicting outcome of patients with COVID-19 pneumonia. *Radiology: Cardiothoracic Imaging* 2020;2(4):e200322.
- [60] Gozes O, Frid-Adar M, Greenspan H, Browning P, Zhang H, Ji W, Bernheim A. Rapid AI development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis. 2020.
- [61] Huang L, Han R, Ai T, Yu P, Kang H, Tao Q, Xia L. Serial quantitative chest CT assessment of COVID-19: deep-learning approach. *Radiology: Cardiothoracic Imaging* 2020;2(2):e200075.
- [62] Sniecinski I, Seghatchian J. Artificial intelligence: a joint narrative on potential use in pediatric stem and immune cell therapies and regenerative medicine. *Transfus Apher Sci* 2018;57(3):422-4.
- [63] Sadeghi S, Soudi S, Shafiee A, Hashemi SM. Mesenchymal stem cell therapies for COVID-19: current status and mechanism of action. *Life Sci* 2020;262:118493.
- [64] Shetty R, Murugeswari P, Chakrabarty K, Jayadev C, Matalia H, Ghosh A, et al. Stem cell therapy in coronavirus disease 2019: current evidence and future potential. *Cytotherapy* 2020;S1465-3249(20):30932-4.

- [65] Ballini A, Scacco S, Coletti D, Pluchino S, Tatullo M. Mesenchymal stem cells as promoters, enhancers, and playmakers of the translational regenerative medicine. *Stem Cell Int* 2017;2017:3292810.
- [66] Spagnuolo G, Codispoti B, Marrelli M, Rengo C, Rengo S, Tatullo M. Commitment of oral-derived stem cells in dental and maxillofacial applications. *Dent J* 2018;6 (4):72.
- [67] Tatullo M, Codispoti B, Paduano F, Nuzzolese M, Makeeva I. Strategic tools in regenerative and translational dentistry. *Int J Mol Sci* 2019;20:1879.
- [68] Marrelli M, Maletta C, Inchingolo F, Alfano M, Tatullo M. Three-point bending tests of zirconia core/veneer ceramics for dental restorations. *International Journal of Dentistry* 2013;2013:831976.