Special Communication

Comparative study of machine learning methods for COVID-19 transmission forecasting

Abdelkader Dairi [a],[*], Fouzi Harrou [b],[*], Abdelhafid Zeroual [c],[d], Mohamad Mazen Hittawe [b], Ying Sun [b]

[a] University of Science and Technology of Oran-Mohamed Boudiaf (USTO-MB), Computer Science department Signal, Image and Speech Laboratory (SIMPA) Laboratory, El Mnaouar, BP 1505, Bir El Djir 31000, Oran, Algeria
[b] King Abdullah University of Science and Technology (KAUST) Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, Thuwal 23955-6900, Saudi Arabia
[c] Faculty of Technology, Department of electrical engineering, University of 20 August 1955, Skikda 21000, Algeria
[d] LAIG Laboratory, University of 08 May 1945, Guelma 24000, Algeria

ARTICLE INFO

ABSTRACT

Within the recent pandemic, scientists and clinicians are engaged in seeking new technology to stop or slow down the COVID-19 pandemic. The benefit of machine learning, as an essential aspect of artificial intelligence, on past epidemics offers a new line to tackle the novel Coronavirus outbreak. Accurate short-term forecasting of COVID-19 spread plays an essential role in improving the management of the overcrowding problem in hospitals and enables appropriate optimization of the available resources (i.e., materials and staff).This paper presents a comparative study of machine learning methods for COVID-19 transmission forecasting. We investigated the performances of deep learning methods, including the hybrid convolutional neural networks-Long short-term memory (LSTM-CNN), the hybrid gated recurrent unit-convolutional neural networks (GAN-GRU), GAN, CNN, LSTM, and Restricted Boltzmann Machine (RBM), as well as baseline machine learning methods, namely logistic regression (LR) and support vector regression (SVR). The employment of hybrid models (i.e., LSTM-CNN and GAN-GRU) is expected to eventually improve the forecasting accuracy of COVID-19 future trends. The performance of the investigated deep learning and machine learning models was tested using confirmed and recovered COVID-19 cases time-series data from seven impacted countries: Brazil, France, India, Mexico, Russia, Saudi Arabia, and the US. The results reveal that hybrid deep learning models can efficiently forecast COVID-19 cases. Also, results confirmed the superior performance of deep learning models compared to the two considered baseline machine learning models. Furthermore, results showed that LSTM-CNN achieved improved performances with an averaged mean absolute percentage error of 3.718%, among others.

1. Introduction

Humanity in the World has experienced various disease outbreaks in the past. World Health Organization (WHO) and numerous national authorities worldwide battle facing these pandemics to date. The coronavirus COVID-19 pandemic that has initially been identified in Wuhan, China, in December 2019 remains a significant problem facing our modern world [1]. As reported by the World Health Organization, 213 countries and territories worldwide are contaminated with the COVID-19. The COVID-19 virus can spread in diverse ways, such as direct body contact, infected money, air from contaminated persons when

coughing, or sneezing [2]. The incubation period of COVID-19 of a minimum of fourteen days plays a crucial in its propagation [3,4]. Thus, the accurate forecast of recovered and contaminated COVID-19 cases is indispensable in managing the outbreak and preparing efficient procedures to slow down the spread of COVID-19.

COVID-19 pandemic has had a serious effect on population health and prosperity. Research efforts of COVID-19 forecasting become critical; in particular, biomedical informatics is major and central for each research efforts to deliver healthcare for COVID-19 patients. With recent advancements in computer and software technologies, artificial intelligence (AI) becomes more prevalent in healthcare systems to detect

disease and perform clinical diagnosis. Moreover, this recent outbreak shows the necessity and possibility of using AI for outbreaks prediction. In other words, AI tools possess powerful feature extraction capabilities, which enables extracting relevant features from rapidly evolving data to support public health experts in decision-making. One of the key features of AI is its ability to rapidly extract pertinent information from multi-sources including health reports, social media, news, and media. In past outbreaks (e.g., Severe Acute Respiratory Syndrome (SARS) in China in 2003), limited data were available [5]. With easy access to different data sources, AI tools have become even more necessary in healthcare systems.

Note that artificial intelligence has already been employed in epidemiology. For example, the HealthMap group, founded by software developers and epidemiologists at Boston Children's Hospital founded in 2006, focuses on using online informal sources to monitor disease outbreaks and real-time surveillance of emerging public health threats [6]. HealthMap group considered monitoring the informal data sources in the COVID-19 outbreak. Also, there is Google Flu, which used search engine queries for improving the track of the flu epidemic. Remarkably, BlueDot, a Canadian software company, has signaled an alert to its customers about the COVID-19 outbreak on December 31 [5]. Hence, this company specializing in infectious disease epidemiology indicated an early detection nine days before the first WHO warning on January 9, 2020 [7]. This software company has also succeeded in predicting the transmission of the Zika virus from Brazil [8]. Hence, advances in AI methods to early detect epidemiologic risks will be necessary for enhancing the prediction and detection of future health risks.

Over the last decay, many advances have been made in AI methods. Machine learning and, in particular, deep learning is considered as an essential aspect of artificial intelligence that can automatically extract relevant features from complex data. These aspects are represented as a promising technology used by various healthcare providers and researchers [9–14]. Particularly, deep learning technology has achieved progress in multiple fields, such as computer vision, robotics, medical imaging, chemistry, and biology [15–18]. The deep learning-driven methods benefit other machine learning methods because they do not rely on feature engineering.

The ongoing development in machine learning and deep learning has significantly enhanced healthcare experts. Recently, COVID-19 has received remarkable consideration from the research community, the epidemiological experts, and local authorities [19,20,15]. Various studies have been reported in the literature for detecting and diagnosing COVID-19 diseases using X-ray imagery [21–26]. In [1], COVID-19 detection is addressed as a classification problem involving three classes: pneumonia, COVID-19 affected cases, and normal cases based on X-ray imagery. To this end, a support vector machine (SVM) is applied to the extracted features by CNN models (MobileNetV2 and SqueezeNet). However, the detection is performed in a supervised way, which demands data labeling. In [22], Goel et al. considered an optimized CNN (called OptCoNet) to automatically diagnosing of COVID-19 based on chest X-ray images. The Grey Wolf Optimizer has been used to optimize the hyperparameters of the CNN model during the training phase. They applied the OptCoNet approach for discriminating normal from pneumonia patients and showed satisfying results. In [25], Turkoglu Muammer proposed an expert-designed system for COVID-19 diagnosis based on X-ray images. Specifically, a pre-trained CNN-based AlexNet model has been used as a features extractor. Then, the SVM algorithm is applied to the extracted features to discriminate between healthy, COVID-19, and Pneumonia patients. This approach showed high classification accuracy of 99.18%. Further, in [27], Zebin and Rezvy applied three pre-trained convolutional backbones (ResNet50, VGG-16, and EfficientNetB0 pre-trained on ImageNet dataset) to detect COVID-19 and related infection based on chest X-ray images. Also, a generative adversarial framework is trained for the generation and augmentation of the minority COVID-19 class. It has been reported that even with a small number of images in the COVID-19 class, good discrimination

performance has been achieved by these pre-trained models. The method in [28] applied a deep CNN model called Decompose, Transfer, and Compose (DeTraC) to classify COVID-19 chest X-ray images. They showed that DeTraC could handle irregularities in the image dataset and detect COVID-19 cases.
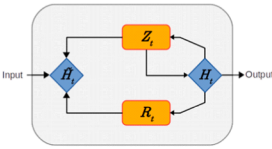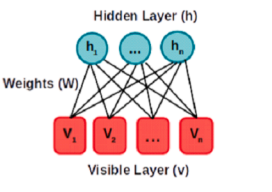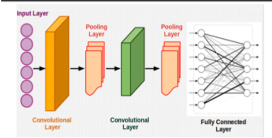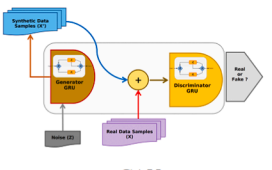
For understanding and managing this epidemic, numerous studies in the literature focused on modeling, predicting, and forecasting COVID-19 spread based on the recorded COVID-19 time series data [29]. In [30], Sarkar et al. proposed a mathematical model called $SARII_qS_q$ to model and forecast the transmission dynamics of COVID-19. This model is based on monitoring the dynamics of six behaviors named susceptible (S), asymptomatic (A), recovered (R), infected (I), isolated infected (Iq), and quarantined susceptible (Sq). It is an extended version of the traditional SEIR (Susceptible, Exposed, Infectious, Removed) model. In [31], Abbasi et al. presented an alternative version of the SEIR model called SQEIAR that considers Quarantined individuals (Q) and Asymptomatic (A) parameters to describe the COVID-19 epidemical dynamics. The method in [32] combines the benefits of the Ensemble of Kalman filter and the SIR (Susceptible, Infected, Recovered) model for modeling COVID-19 transmission. This coupled approach enables estimating the unknown transmission rates of the COVID-19 to reconstruct the basic reproduction number R_0. The forecasting performance of this approach has been verified based on real data from March 6 to May 17, 2020, of the current COVID-19 pandemic outbreak in Cameroon. In [33], Ribeiro et al. applied regression models for forecasting COVID-19 cases in Brazil. Specifically, they considered ARIMA, cubist regression, random forest, ridge regression, SVR, and stacking-ensemble learning for one, three, and six-step ahead forecasting. They showed that the SVM regression and stacking-ensemble learning achieved better forecasting than the other models. In [34], SVR with different nonlinearity structures is used to predict COVID-19 cases. It has been shown that the SVR with Gaussian kernel presented better prediction compared to SVR with other Kernel functions. In [35], for the global forecast of the COVID-19, a nonlinear machine learning approach is coupled with partial derivative regression to build an epidemiological prediction method. For more computation efficiency and better dataset parameters, the Progressive Partial Derivative Linear Regression is used.

Recently deep learning methods have gained particular attention in time-series modeling and analysis because of their outstanding generalization capability and superior nonlinear approximation [36,37]. Deep learning models can be obtained by concatenating several layers into the neural network structures [38]. The main feature of deep learning models consists of their capacity to automatically extracting relevant information from large data [39]. They have attracted attention in the machine-learning community and investigated in a wide range of applications [38,40,41,11]. For instance, in [42], an long short-term memory (LSTM) model is applied to forecast new confirmed cases of COVID-19 in Canadian. They used data from January 22 to March 31, 2020. Similarly, in [43], an improved LSTM model has been employed to predict the epidemic trends of COVID-19 in Russia, Peru, and Iran. In [44], support vector regression (SVR), LSTM, bidirectional LSTM (BiLSTM), and GRU are employed to forecast COVID-19 time-series data in ten countries most affected by COVID-19. Results reveal the superior performance of the BiLSTM based on COVID-19 data available until June 27, 2020.

The deep learning technology has realized advancement in different areas, such as computer vision [45,46], natural language processing [47], speech recognition [48], renewable energy forecasting [13], anomaly detection [49,11,50], and reinforcement learning [51].

This paper is aimed at presenting a comparative study of machine learning-driven methods for COVID-19 transmission forecasting. Notably, this study investigates the efficiency of deep learning methods to forecast recovered and confirmed COVID-19 time-series and assess their performance compared to the traditional forecasting methods. Essentially, six deep learning models, including the hybrid convolutional neural networks-Long short-term memory (CNN-LSTM), the

Table 1
The considered benchmark deep learning methods.

Model	Description	Key Points
GRU model	▪ The major demarcation of GRU from LSTM is that only one unit is used to control both the forgetting factor and the decision to update the state unit [41]. ▪ GRU contains only two gates, the update, and the reset gates. Gate mechanism is used to memorize historical data features. ▪ The GRU is trained in a supervised way. ▪ The GRU has been widely used in time-series, data sequence (e.g., speech and text processing), temporal features extraction, prediction, and forecasting.	**+Pros:** 👍 The attractive features of the GRU model are the shorter training time compared to the LSTM and the fewer parameters that the GRU model possesses compared to the LSTM [41]. 👍 It is suited for modeling time dependencies and sequences. **-Cons:** 🗨 GRU models have problems such as slow convergence rate and low learning efficiency, resulting in too long training time and even under-fitting.
Restricted Boltzmann machines	▪ RBMs are stochastic and generative neural networks [40] consisting of visible units, and hidden units, with no connections between visible-to-visible and hidden-to-hidden, however visible and hidden units are fully connected. ▪ RBMs are trained in an unsupervised way based on the contrastive divergence learning method. ▪ Contrastive divergence uses Gibbs sampling to compute the intractable negative phase.	**+Pros:** 👍 It is a simple architecture with two layers 👍 RBMs have good capability for data distribution approximation. 👍 They can be stacked to build deep learning models, such as DBN and DBM. **-Cons:** 🗨 They are slow to train due to Contrastive Divergence approach.
CNN	▪ CNNs are primarily designed for image classification. They are composed of a succession of convolution and pooling operations in order to learn features [38]. ▪ They are trained in a supervised way using backpropagation with gradient descent.	**+Pros:** 👍 They are efficient to learn relevant features from pixels (image) [39]. **-Cons:** 🗨 They require a large amount of labeled data.
GAN	▪ GAN is a composite generative model [42]. GAN uses two deep learning models called Generative (G) and Discrimitaive (D). The G model produces fake (or noisy) data to feed the D. The D model learns to discriminate fake and true data. ▪ The GAN training is achieved through alternately updating the parameters of generator G and a discriminator D. At the end of the GAN training, the data distribution learned by G should be similar to that learned by D. ▪ GAN-GRU and GAN-DNN are variants of GAN, with Generative and Discriminator as GRU for the first, and deep neural network for the second variant [41, 42].	**+Pros:** 👍 GANs have good capability for data distribution approximation. 👍 They have flexible architecture 👍 Generator and Discriminator can be any deep learning model, such as DNN, GRU, and LSTM. **-Cons:** 🗨 They suffers from the problem of convergence and Vanishing gradient.

hybrid gated recurrent unit-convolutional neural networks (GAN-GRU), GAN, CNN, LSTM, and Restricted Boltzmann Machine (RBM), are introduced and compared to the base-line methods, namely logistic regression (LR) and support vector regression (SVR), for forecasting confirmed and recovered COVID-19 cases time-series data. Here, the employment of hybrid models (i.e., LSTM-CNN and GAN-GRU) is expected to eventually improve the forecasting accuracy of COVID-19 future trends. For instance, in CNN-LSTM, this combination aims to integrate the capabilities of the LSTM in modeling time dependencies and the good ability of CNN in extracting features from complex data. As far as we know, this is the first time these two hybrid models were investigated to enhance the forecasting of COVID-19 time-series data. Confirmed and recovered COVID-19 cases time-series data from seven impacted countries, namely Brazil, France, India, Mexico, Russia, Saudi Arabia, and the US, are used to compare and verify the performance of the considered deep learning methods. The six deep learning methods and two commonly used machine learning methods, namely LR and SVR, were applied to daily recovered and confirmed COVID-19 time-

series datasets from January 22nd, 2020, to September 6th, 2020. Measurements of effectiveness including coefficient of determination ($R2$), Root Mean Square Error (RMSE), mean absolute error (MAE), Mean Squared Logarithmic Error (MLSE), and explained variance (EV) were adopted to assess the forecasting quality. Results confirmed the superior performance of deep learning models compared to the two considered shallow machine learning models, namely logistic regression and support vector regression. The results highlight that the hybrid LSTM-CNN model reaches the best forecasting performance, followed by the hybrid GRU-GAN model. Also, it has been shown that the considered hybrid models outperform the other considered non-hybrid models. Furthermore, this study has demonstrated that time-series modeling and forecasting are not restricted to only recurrent models like LSTM and GRU. Still, it can be extended to other generative composite models, such as GAN-DNN and GAN-GRU, or standalone models like RBM that showed a good forecasting ability with low computation cost and less number of parameters compared to the other considered models. Moreover, we showed that deep learning models provided satisfying

Table 2
Long Short-Term Memory gates and Memory equations.

	Component	Equation
σ	Input Gate	$\mathscr{I}_t = \sigma(\mathscr{X}_t \mathscr{W}_{xi} + \mathscr{H}_{t-1} \mathscr{W}_{hi} + \mathscr{B}_i)$
σ	Forget Gate	$\mathscr{F}_t = \sigma(\mathscr{X}_t \mathscr{W}_{xf} + \mathscr{H}_{t-1} \mathscr{W}_{hf} + \mathscr{B}_f)$,
σ	Output Gate	$\mathscr{O}_t = \sigma(\mathscr{X}_t \mathscr{W}_{xo} + \mathscr{H}_{t-1} \mathscr{W}_{ho} + \mathscr{B}_o)$,
c	Memory Cell	$\mathscr{C}_t = \mathscr{F}_t \odot \mathscr{C}_{t-1} + \mathscr{I}_t \odot \widetilde{\mathscr{C}_t}$.
c	Candidate Memory Cell	$\widetilde{\mathscr{C}_t} = \tanh(\mathscr{X}_t \mathscr{W}_{xc} + \mathscr{H}_{t-1} \mathscr{W}_{hc} + \mathscr{B}_c)$.
H	Hidden State	$\mathscr{H}_t = \mathscr{O}_t \odot \tanh(\mathscr{C}_t)$.

Table 3
Definition of measurements of effectiveness.

Metric	Definition
R^2	$\dfrac{\sum_{t=1}^n [(y_{t.} - \overline{y}) \cdot (\widehat{y}_t - \overline{y})]^2}{\sqrt{\sum_{t=1}^n (y_t - \overline{y})^2} \cdot \sqrt{\sum_{t=1}^n (\widehat{y}_t - \overline{y})^2}}$
RMSE	$\sqrt{\dfrac{1}{n}\sum_{t=1}^n (y_t - \widehat{y}_t)^2}$
MAE	$\dfrac{\sum_{t=1}^n |y_t - \widehat{y}_t|}{n}$
MAPE	$\dfrac{100}{n}\sum_{t=1}^n \left|\dfrac{y_t - \widehat{y}_t}{y_t}\right|\%$
EV	$1 - \dfrac{\text{Var}(\widehat{y} - y)}{\text{Var}(y)}$
MSLE	$\dfrac{1}{N}\sum_{t=0}^N (\log(y_t + 1) - \log(\widehat{y}_t + 1))^2$

forecasting results even with a relatively small amount of data. Thus, these methods can be extended to deal with different communicable diseases, such as SARS [52], EBOLA [53], HIV [54,55], and COVID-19, and non-communicable diseases including Cancer [56], Diabetic [57], and Heart disease [58] outbreak. In the next section, we provide a brief review of the essential features of DBN, RBM, CNN-GRU, and GAN-RNN models. In the third section, we present how they can be employed in forecasting. The fourth section offers the involved multivariate patient flow datasets and discusses the models forecasting results and comparisons. Lastly, in the fifth section, we conclude this study.

2. Methodology

Deep learning techniques, which could automatically learn relevant information from time-series data, are considered in this study to forecast confirmed and recovered COVID-19 time-series data. This section briefly describes the considered deep learning models, namely LSTM [59], CNN [60,61], RBM [62], Generative adversarial networks based on deep fully connected neural network (GAN-DNN) [63], GAN-GRU

[64,63], and LSTM-CNN. Table 1 offers the pros and cons of the five benchmark deep learning architectures.

2.1. The hybrid LSTM-CNN model

A hybrid deep learning architecture called LSTM-CNN has been introduced to merges the benefits of the LSTM, which learn and model temporal dependencies embedded in the data sequence, and the CNN model, which cam well process high-dimensional data. The LSTM is equipped with recurrent connections and a gating mechanism that help to memorize features. Table 2 lists the LSTM gates, memory cell, and hidden state equations. (see Table 3).

With $\mathscr{W}_{xi}, \mathscr{W}_{xf}, \mathscr{W}_{xo}, \mathscr{W}_{xc}, \mathscr{W}_{hc}, \mathscr{W}_{hi}, \mathscr{W}_{hf}$, and $\mathscr{W}_{ho}$ are the weight parameters and $\mathscr{B}_i, \mathscr{B}_f, \mathscr{B}_o$, and $\mathscr{B}_c$ are bias parameters, $\odot$ denotes the element-wise multiplication.

The second model integrated into the hybrid model is the well-known one-dimensional CNN that aims to extract features from the temporal feature space built by the nonlinear transformation performed by LSTM. Indeed CNN learns to extract features based on a series of nonlinear transformations via several Convolution and Pooling operations that usually end with a fully connected layer. CNN can extract features first and perform the internal features mapping based on one-dimensional (1-D) sequences data represented by COVID-19 cases recorded. On the other hand, the CNN-1D layer structure is simple and thus has a low computational cost. The mathematical formulation of each convolutional layer l is done as follows:

$$\mathscr{Y}_l = \sigma\left(\sum_{i=1}^{l-1} \mathscr{C}\left(\mathscr{W}_l, \mathscr{Y}_{l-1}\right) + b_l\right), \tag{1}$$

$\mathscr{C}(.)$ is 1-D convolution function and $\sigma(.)$ is the activation function, while $(\mathscr{W}, b)$ refers to the weight and bias of the layer l, respectively. Overall, in the LSTM-CNN approach, the LSTM process the input represented by the time-series data, and the CNN performs the forecasting. The basic flowchart of the LSTM-CNN is depicted in Fig. 1, where LSTM and CNN are stacked to form the deep hybrid architecture. The forecasting is done actually by the fully connected output layer of the CNN.

2.2. Forecasting model

In this study, we construct six deep learning models: non-hybrid models (i.e., LSTM, CNN, and RBM) and three hybrid models (i.e., GAN-DNN, and GAN-GRU, and LSTM-CNN). The COVID-19 time-series data is first normalized by min–max normalization within the interval [0, 1]. It should be noted that after the COVID-19 time-series forecasting using testing data, this normalization is reversed so that the forecasted data is comparable to the original testing time-series data. Also, the normalized data is preprocessed before the application of the deep learning models. Specifically, to train the six considered models, data preparation is performed, the normalized time series is transformed into a set of sequences $\mathscr{X}$ of fixed length using the window sliding methods,
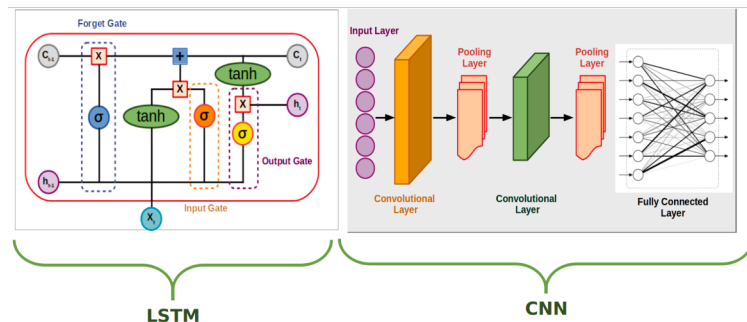

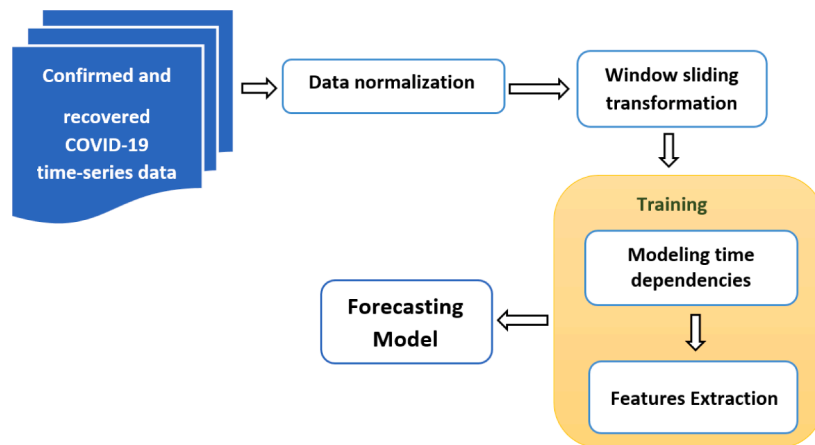**Fig. 1.** Flowchart of the LSTM-CNN approach.

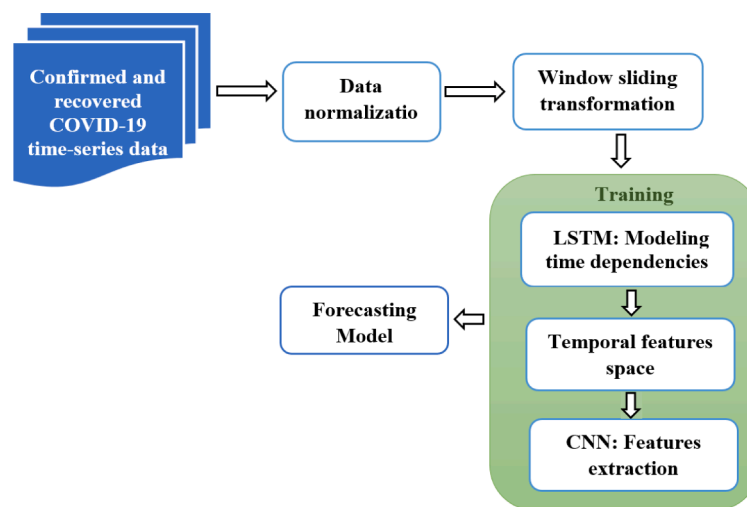**Fig. 2.** Flowchart of the proposed COVID-19 cases forecasting.

Fig. 3. Deep hybrid LSTM-CNN approach for COVID-19 cases forecasting.

where for each sequence $x \in \mathscr{X}$ is mapped y that represent the next value of the sequence x to form the mapping set of the subsequent values $\mathscr{Y}$. The general framework of the proposed deep learning-driven forecasting strategy is depicted in Fig. 2.

After normalizing and reshaping the COVID-19 time-series data, we will train the deep learning models. The training stage aims to learn complex features and construct the final forecasting model. It is worth pointing out that every model has its particularity regarding the learning approach, supervised or unsupervised.

It is worth noticing that four deep learning models are trained in an unsupervised way (i.e., LSTM, RBM, GAN-DNN, and GAN-GRU), and two models (i.e., CNN and LSTM-CNN) are trained in a supervised manner. For instance, GANs and RBM are generative models trained only with the historical data $\mathscr{X}$. At first, these two models try to learn the probability distribution without the need for $\mathscr{Y}$, and later a fine-tuning procedure is applied through supervised training to optimize the models' parameters via learning the mapping $\mathscr{X} \Rightarrow \mathscr{Y}$. During the fine-tuning step, a predictor layer with a single output to perform the forecasting is added. On the other hand, traditional LSTM and CNN are trained in a supervised way. Precisely, training a CNN model consists of extracting features from the input depending on the data structure like images, a 2D signal, or for sensory data of 1D. In contrast, LSTM belongs to gated recurrent neural networks designed to deal with time-series and data sequences to extract and discover temporal dependencies like text processing, where the objective is to predict the next word of a given

sentence (a series of words).

The training procedure of the deep hybrid LSTM-CNN is supervised. The two models LSTM, and CNN, are simultaneously trained (Fig. 3). The time-series data (input) is processed first by LSTM, resulting in temporal features space that feeds CNN. The forecasting is done through the last layer of CNN with a single output. The role of CNN is to learn the mapping of the temporal features to the next value in the original input (recovered or confirmed COVID-19 cases).

This study aims to compare deep learning methods to forecast COVID-19 time-series data. Towards this end, the present work uses two hybrid deep learning models (i.e., LSTM-CNN and GAN-GRU), four standalone deep learning models (i.e., GAN, CNN, LSTM, and RBM), and two baseline machine learning methods (i.e., LR, and SVR), and results are compared based on COVID-19 data from six well-impacted countries.

The performance of the considered deep learning-driven methods is compared with traditional forecasting methods, SVR and LR, to validate their efficiency. The SVR and LR are deployed on a rolling window where the previous five days records are taken as covariates to forecast counts of one day ahead. Indeed, SVR showed good performance in various time series forecasting problems [65–67]. This is mainly because of its capacity to feel with the overfitting problem using kernel tricks and without the need for the specification of data distribution [68]. LR, which are usually incorporated into machine learning and deep learning models, can be viewed as a shallow neural network composed of one layer with a logistic
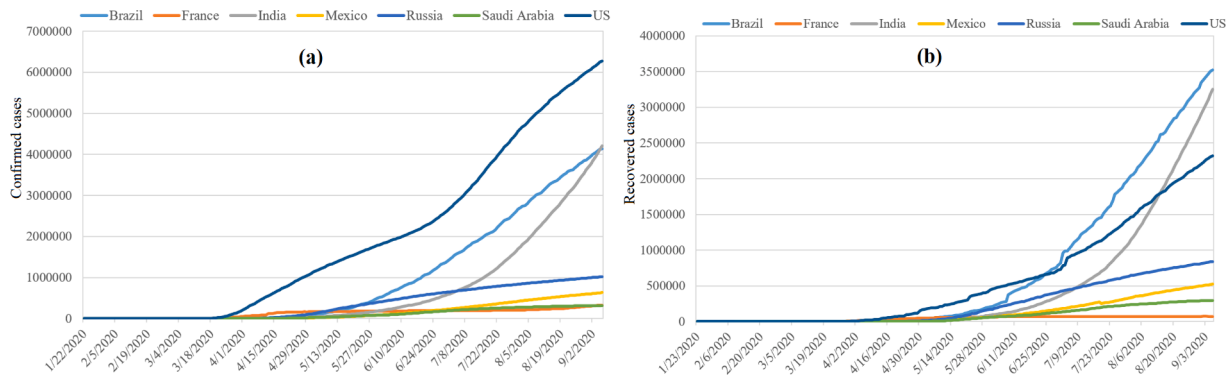
Fig. 4. Confirmed (a) and recovered (b) CODID-19 times series data recorded from January 22 to September 6, 2020.
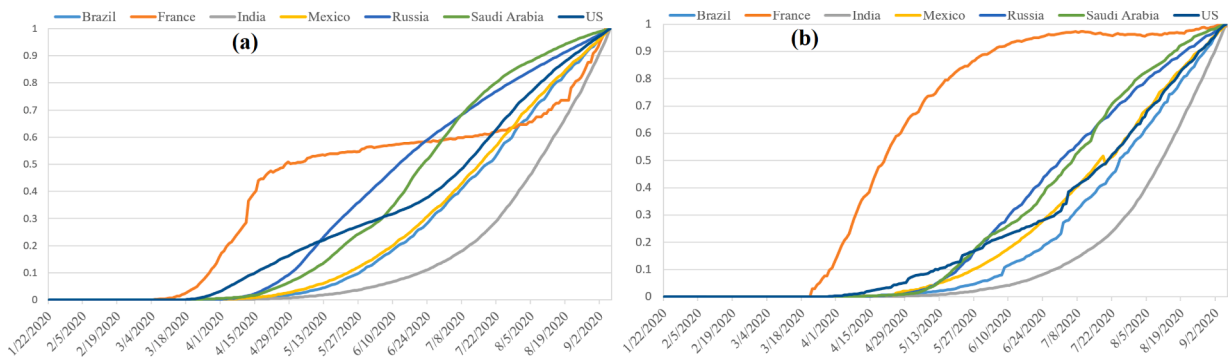

Fig. 5. Normalized confirmed (a) and recovered (b) CODID-19 times-series data.

function also called sigmoid formulated as $f(x) = \frac{1}{a+e^{-x}}$ [69]. LR has also been used in forecasting time-series data [69]. Here, a window sliding procedure is applied to reshape time-series data used by the SVR and LR models. Of course, SVR and LR are trained to learn to map a given sequence to its next value in the training phase, and then they are applied to unseen data to predict future observation (forecasted data).

2.3. Evaluation metrics

For performance comparisons, the following indexes have been adopted to quantify the forecasting of Confirmed and Recovered COVID-19: Coefficient of determination (R^2), Root Mean Square Error (RMSE), mean absolute error (MAE), Mean Squared Logarithmic Error (MSLE), and explained variance (EV).

With y_t is the number of Confirmed or Recovered patients, $\hat{y}_t$ is its corresponding forecasted values, and n is the number of data points.

3. Results and discussion

3.1. Data description

In this study, daily confirmed and recovered COVID-19 cases time

series data of seven different countries (Brazil, France, India, Mexico, Russia, Saudi Arabia, and the US) are used to assess the forecasting capability of the hybrid and non-hybrid deep learning models. The data are made accessible publicly by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (https://github.com/CSSEGISandData/COVID-19). Here, we consider the data from the appearance of the first cases in each country. Fig. 4 depicts the used COVID-19 time-series dataset in the considered countries recorded from January 22 to September 6, 2020.

Before applying the deep learning-based forecasting methods, the COVID-19 time series data displayed in Fig. 4 are first normalized via min–max normalization within the interval [0, 1], and the used for constructing the deep learning models.

$$\widetilde{x} = \frac{(x - x_{min})}{(x_{max} - x_{min})} \qquad (2)$$

where x_{min} and x_{max} denotes the minimum and maximum of the raw data, respectively. After the forecasting process, we allied a reverse operation so that the forecasted data correspond to the original confirmed and recovered COVID-19 time-series data. The normalized time series data (confirmed and recovered) are depicted in Fig. 5.

3.2. A comparison of methods for COVID-19 transmission forecasting

This study focus on the univariate daily forecasting of the number of confirmed and recovered COVID-19 cases from the seven considered countries (i.e., Brazil, France, India, Mexico, Russia, Saudi Arabia, and the US). Forecast models were developed for each time series. To do so, we split the normalized data into training and testing sub-datasets. The training data from January 22, 2020, to September 6, 2020, are used to construct the deep learning models: the standalone models (LSTM, CNN, and RBM) and the hybrid models (GAN-DNN, GAN-GRU, and LSTM-CNN), and baseline machine learning models: LR and SVR. The goal is

Table 4
Hyper-parameters using in the training of the investigated models.

Model	Parameters
LSTM	layers: 3, hidden units: [32,32,1] learning rate: 0.0001 epochs:200
GAN-GRU	layers: 3, hidden units: [32,1] learning rate: 0.0001 epochs:200
LSTM-CNN	layers: 5, hidden units: [16,32,32,1] learning rate: 0.0001 epochs:200
RBM	layers: 3, hidden units: [32,1] learning rate: 0.001 epochs:1000
GAN-DNN	layers: 3, hidden units: [16,1] learning rate: 0.0001 epochs:150
CNN	layers: 3, hidden units: [32,32,1] learning rate: 0.0001 epochs:200

Table 5
Measurements of effectiveness for confirmed cases COVID-19 forecasting.

Country	Model	RMSE	MAE	R2	EV	MSLE	MAPE
Brazil	GAN-GRU	1.02E + 05	8.30E + 04	0.993	0.996	0.005	6.485
	LSTM-CNN	4.49E + 04	3.39E + 04	0.999	0.999	0.002	2.891
	RBM	2.33E + 05	2.03E + 05	0.964	0.965	0.102	19.696
	GAN-DNN	1.25E + 05	9.48E + 04	0.99	0.993	0.020	8.358
	CNN	9.77E + 11	8.40E + 11	0.964	0.964	0.080	18.567
	LSTM	6.79E + 11	5.95E + 11	0.982	0.983	0.131	28.998
	SVM	1.85E + 12	1.67E + 12	0.87	0.878	0.455	83.117
	LR	2.13E + 12	1.84E + 12	0.828	0.865	0.586	103.288
France	GAN-GRU	8.21E + 03	7.72E + 03	0.945	0.993	0.002	3.780
	LSTM-CNN	2.75E + 03	1.51E + 03	0.994	0.994	0	0.628
	RBM	8.54E + 03	7.17E + 03	0.941	0.957	0.001	3.341
	GAN-DNN	6.28E + 03	4.16E + 03	0.968	0.969	0.001	1.815
	CNN	1.93E + 09	1.11E + 09	0.972	0.975	0.001	1.483
	LSTM	2.18E + 09	1.59E + 09	0.965	0.967	0.001	2.16
	SVM	6.56E + 09	5.36E + 09	0.679	0.892	0.007	7.231
	LR	5.21E + 09	3.20E + 09	0.798	0.81	0.004	4.12
India	GAN-GRU	8.36E + 04	6.69E + 04	0.995	0.997	0.008	8.495
	LSTM-CNN	8.31E + 04	6.09E + 04	0.995	0.998	0.005	6.021
	RBM	2.15E + 05	1.77E + 05	0.965	0.966	0.144	27.280
	GAN-DNN	1.07E + 05	7.83E + 04	0.991	0.993	0.075	15.954
	CNN	1.10E + 12	8.94E + 11	0.949	0.953	0.195	35.072
	LSTM	9.34E + 11	7.60E + 11	0.963	0.964	0.144	27.783
	SVM	2.45E + 12	2.34E + 12	0.744	0.841	1.439	256.757
	LR	3.04E + 12	2.75E + 12	0.608	0.661	1.672	303.742
Mexico	GAN-GRU	1.05E + 04	7.28E + 03	0.997	0.998	0.003	3.439
	LSTM-CNN	4.98E + 03	3.89E + 03	0.999	1	0.001	2.180
	RBM	3.58E + 04	3.18E + 04	0.964	0.965	0.097	19.655
	GAN-DNN	1.47E + 04	1.21E + 04	0.994	0.994	0.018	7.493
	CNN	2.36E + 10	2.02E + 10	0.961	0.965	0.082	18.292
	LSTM	1.67E + 10	1.27E + 10	0.981	0.987	0.03	12.785
	SVM	4.15E + 10	3.77E + 10	0.88	0.884	0.306	60.162
	LR	5.08E + 10	4.45E + 10	0.82	0.828	0.407	75.116
Russia	GAN-GRU	1.23E + 04	9.33E + 03	0.998	0.998	0.002	2.582
	LSTM-CNN	6.26E + 03	4.84E + 03	0.999	0.999	0	0.941
	RBM	5.88E + 04	5.19E + 04	0.944	0.944	0.052	12.213
	GAN-DNN	2.77E + 04	2.30E + 04	0.988	0.989	0.003	4.230
	CNN	6.35E + 10	5.53E + 10	0.938	0.939	0.054	12.933
	LSTM	5.74E + 10	5.00E + 10	0.949	0.953	0.042	11.658
	SVM	9.03E + 10	7.57E + 10	0.875	0.906	0.025	13.046
	LR	8.69E + 10	7.47E + 10	0.884	0.886	0.047	17.49
Saudi Arabia	GAN-GRU	6.56E + 03	5.50E + 03	0.995	0.997	0.006	5.661
	LSTM-CNN	2.85E + 03	2.51E + 03	0.999	0.999	0.001	2.250
	RBM	1.92E + 04	1.73E + 04	0.961	0.961	0.073	15.595
	GAN-DNN	9.43E + 03	7.99E + 03	0.99	0.99	0.004	4.969
	CNN	6.15E + 09	5.43E + 09	0.961	0.962	0.07	15.650
	LSTM	6.31E + 09	5.41E + 09	0.958	0.963	0.064	15.28
	SVM	1.03E + 10	8.91E + 09	0.889	0.904	0.069	22.314
	LR	1.41E + 10	1.26E + 10	0.792	0.813	0.13	33.386
US	GAN-GRU	6.18E + 04	4.70E + 04	0.998	0.999	0	1.534
	LSTM-CNN	5.69E + 04	4.75E + 04	0.999	0.999	0	1.632
	RBM	3.26E + 05	2.91E + 05	0.958	0.958	0.025	11.587
	GAN-DNN	1.61E + 05	1.16E + 05	0.99	0.991	0.001	3.018
	CNN	2.07E + 12	1.91E + 12	0.957	0.958	0.03	12.636
	LSTM	1.79E + 12	1.66E + 12	0.968	0.97	0.019	10.754
	SVM	3.65E + 12	3.32E + 12	0.867	0.868	0.044	19.841
	LR	4.06E + 12	3.64E + 12	0.835	0.839	0.065	23.974

to investigate the performance of the six considered models in the presence of small-sized datasets and compare their performance with the traditional machine learning models (i.e., LR and SVR). The methods in this study provide one-step-ahead forecasting. The observation window is five, i.e., we are using the previous five days to predict the next day. The window size that maximizes the forecasting accuracy is chosen during the training phase using the grid search approach. For all deep learning models, we used the *cross-entropy* as loss function and *Rmsprop* as an optimizer in training. Table 4 lists the models' hyper-parameters used in the training stage. (see Table 5).

Reliable daily forecasting of the number of recovered and confirmed COVID-19 data offers pertinent information to assist medium-term planning(e.g., the assignment of rotas). This enables the hospital managers to improve decisions on staff that need to be contacted on call and getting prior knowledge on situational awareness. Overall, accurate forecasting is required for providing support for decisions. Now, we will explore and investigate the performance when applied to a relatively small-sized dataset. The testing data consists of confirmed and recovered COVID-19 cases recorded from June 06, 2020, to September 06, 2020, in the seven considered countries. Forecasting results from the trained
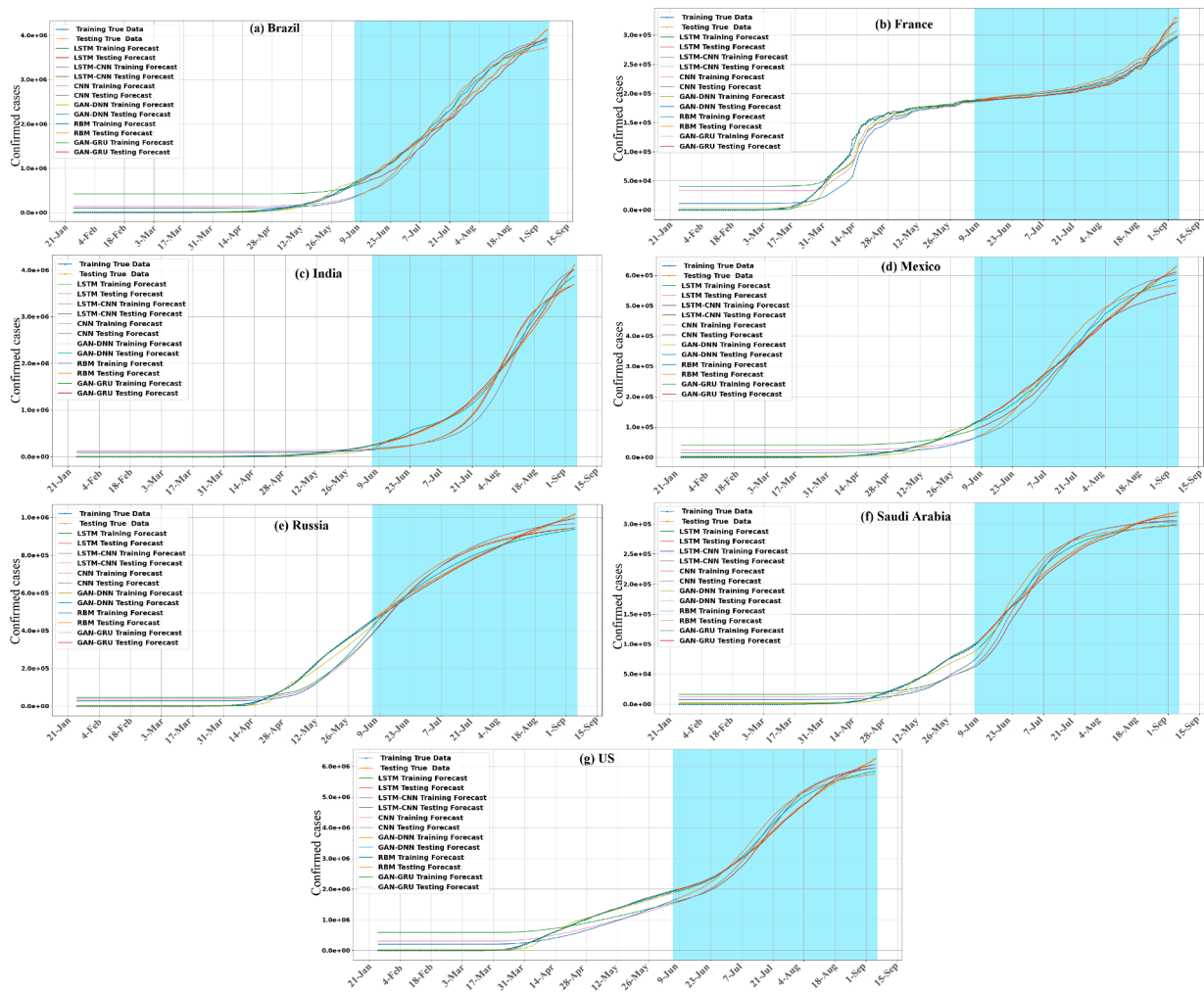
Fig. 6. Measured and forecasted confirmed COVID-19 cases from 6 June to 06 September 2020: (a) Brazil, (b) France, (c) India, (d) Mexico, (e) Russia, (f) Saudi Arabia, and (g) US.

deep learning models using testing data are depicted in Fig. 6. The shaded zone represents the testing period. From Fig. 6, we observe that these models are following the trend of recorded and confirmed COVID-19cases from the starting of the epidemic. After forecasting the COVID-19 time series using each trained model, to judge the quality of each model quantitatively, we calculate the evaluation metrics (R2, EV, RMSE, MAE, MAPE, and MSLE) from the testing datasets for the six deep learning models and two baseline machine learning models (i.e., LR and SVR). The obtained forecasting results are tabulated in Tables 6.

Results in Table 6 show that the quality of the forecast from the six trained models is promising. Table 6 indicates that deep learning models exhibited improved forecasting performance compared to the shallow methods (LR and SVR). We observe that the MAPE values of SVR and LR (data recorded from India and Brazil) are larger than 100%, which means that the forecasting errors are much greater than the actual values [70,71]. It is interesting to see better outcomes from the three hybrid models (GAN-DNN, GAN-GRU, and LSTM-CNN) than the other standalone models. We observe that the three hybrid models (GAN-DNN, GAN-GRU, and LSTM-CNN) can capture the most variability in confirmed time-series data by reaching values of EV and R^2 close to one. This indicates that the forecasted confirmed cases closely follow the trend of recorded confirmed cases. Also, the standalone models (LSTM, CNN, and RBM) can capture most of the variance in data with low forecasting error. By considering all metrics, LSTM-CNN is the best approach with high efficiency and satisfying forecasting accuracy. It could be attributed to its capacity to model time-dependent data and high features extraction capability.

Fig. 7 illustrates the forecasting results of six models based on the recovered COVID-19 time-series data. The forecasted data closely follow the trends of the recorded data. For shallow methods, the SVR model outperformed the LR forecasting method. It could be attributed to the SVR capacity in handling nonlinearity. Similarly to confirmed COVID-19 data, the hybrid models (GAN-DNN, GAN-GRU, and LSTM-CNN) record the highest score in terms of the statistical indicators (Table 6). In comparison, the non-hybrid models can achieve acceptable outcomes for all time-series. Again, the LSTM-CNN model keeps the best score in terms of the evaluation metrics and dominates the other models.

Table 7 presents the averaged statistical metric per model to evaluate the overall performance of the considered models. The forecasting results support that the hybrid models outperform the considered stand-alone models. The LSTM-CNN is dominating the other models by achieving an MAPE of 3.718%. It is followed by GAN-GRU and GAN-DNN with MAPE values of 5.254% and 11.105%, respectively.

In summary, this study highlights the highest forecasting accuracy of the hybrid models achieved by joining the two models in a single architecture, even with small-sized data. The results obtained base on hybrid and composite models are better than those obtained with the standalone models. Also, it can be noticed that the hybrid LSTM-CNN provides improved forecasting compared to the standalone LSTM and for CNN models. This could be due to the fact the double feature extraction performed by the deep hybrid model during the learning process. In other words, in the hybrid LSTM-CNN model, LSTM can
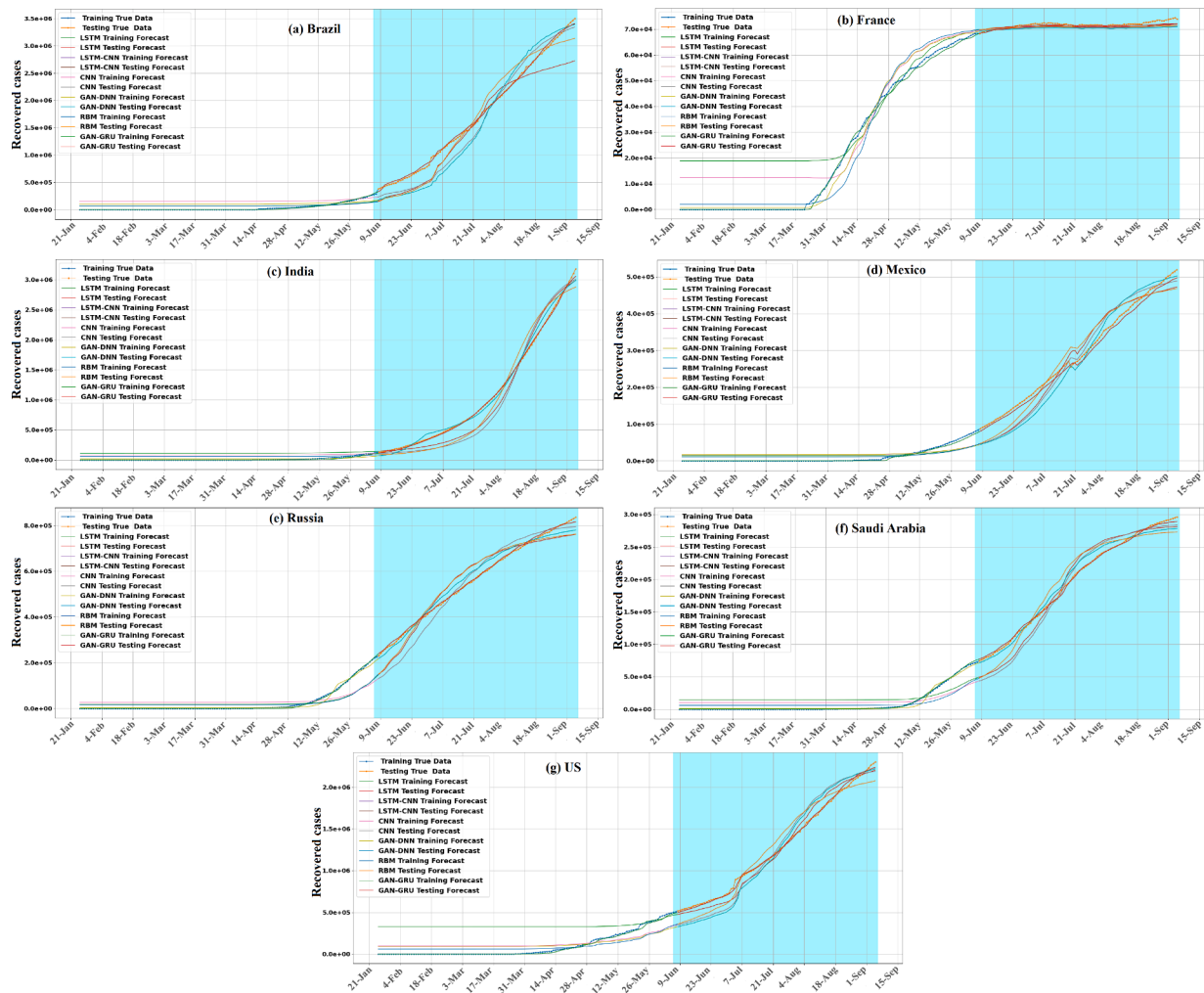
Fig. 7. Measured and forecasted recovered COVID-19 cases from 6 June to 06 September 2020(a) Brazil, (b) France, (c) India, (d) Mexico, (e) Russia, (f) Saudi Arabia, and (g) US.

capture time dependencies and the 1D feature extractor CNN predict the next value of a sequence of values. On the other hand, GANs (GAN-DNN and GAN-GRU) as a deep composite architecture (designed initially for image generation) demonstrate the potential of the generative models to fit COVID-19 time-series data and their ability to approximate the data distribution of time series data accurately. Also, we noticed that the RBM model obtain acceptable forecasting performance even with a straightforward architecture based on only two layers (i.e., visible and hidden layers), making it suitable for online forecasting.

This experimental study reveals that building a composite deep learning model can considerably improve the forecasting accuracy, where each element of the hybrid architecture plays an essential role in modeling temporal dependencies embed in the time-series data of the recorded COVID-19 cases. The proposed approach can be seen as hierarchical features extraction process resulting from a succession of nonlinear transformation performed in multi-stages (LSTM, CNN).

Furthermore, this study has demonstrated that time-series modeling and forecasting are not restricted to only recurrent models like LSTM and GRU. Still, it can be extended to other generative composite models, such as GAN-DNN and GAN-GRU, or standalone models like RBM that shown a good forecasting ability with low computation cost and less number of parameters compared to the other considered models. Moreover, we showed that deep learning models provided satisfying forecasting results even with a relatively small amount of data.

Deep learning models have outperformed the baseline approaches (i.

e., LR and SVR). Because deep learning models are efficient to model long temporal dependencies embed in time-series data (e.g., RNNs, LSTM-CNN, and GAN-GRU). The RNN-based models can capture short and long temporal dependencies in time-series data using memory and gate mechanisms. Also, RNN considers the past data to generate its output. This structure makes it appropriate to model time dependency than the traditional feed-forward that generates the output using only the current input.

4. Conclusion

Deep learning technology and artificial intelligence are promising techniques applied by different healthcare providers. The accurate COVID-19 spread forecasting models do not only support the management of the available resources in hospitals but also help slow down the progression of such diseases. In this study, short term forecasting of COVID-19 cases has been carried out through hybrid deep learning models to provide suitable forecasting accuracy. Towards this purpose, the LSTM-CNN model has been introduced for COVID-19 time series forecasting. Overall, we compared three hybrid models (i.e., LSTM-CNN, GAN-DNN, and GAN-GRU) and three standalone deep learning models (i.e., CNN, LSTM, and RBM). In other words, we examined the capability of advanced structures of deep learning methods in forecasting COVID-19 time-series data on datasets of limited size. The data from daily confirmed and recovered cases from seven of the most impacted

Table 6
Validation Metrics for Recovered cases COVID-19 forecasting.

Country	Model	RMSE	MAE	R2	EV	MSLE	MAPE
Brazil	GAN-GRU	5.53E + 04	3.91E + 04	0.997	0.998	0.077	8.630
	LSTM-CNN	4.72E + 04	3.93E + 04	0.998	0.999	0.004	5.395
	RBM	1.94E + 05	1.61E + 05	0.967	0.968	0.13	22.837
	GAN-DNN	2.26E + 05	1.87E + 05	0.956	0.959	0.162	29.272
	CNN	6.16E + 11	5.22E + 11	0.974	0.975	0.156	33.177
	LSTM	8.27E + 11	6.00E + 11	0.952	0.967	0.107	21.868
	SVM	1.53E + 12	1.42E + 12	0.836	0.863	1.074	191.317
	LR	1.75E + 12	1.53E + 12	0.787	0.855	1.281	229.856
France	GAN-GRU	1.10E + 03	8.94E + 02	0.961	0.961	0	1.334
	LSTM-CNN	9.37E + 02	8.12E + 02	0.971	0.992	0	1.180
	RBM	2.84E + 03	2.06E + 03	0.736	0.753	0.002	3.264
	GAN-DNN	2.53E + 03	2.07E + 03	0.791	0.791	0.002	3.199
	CNN	1.86E + 08	1.44E + 08	0.796	0.8	0.002	3.034
	LSTM	1.10E + 08	8.28E + 07	0.928	0.934	0.001	1.722
	SVM	8.24E + 08	792008938.63	0.692	0.692	0.029	15.206
	LR	3.17E + 08	2.97E + 08	0.41	0.686	0.004	5.78
India	GAN-GRU	2.71E + 04	1.57E + 04	0.999	0.999	0.126	10.203
	LSTM-CNN	3.24E + 04	2.49E + 04	0.999	0.999	0.514	16.113
	GAN-DNN	7.35E + 04	5.30E + 04	0.993	0.994	0.092	17.493
	RBM	1.68E + 05	1.34E + 05	0.964	0.965	0.277	46.168
	CNN	6.00E + 11	4.80E + 11	0.957	0.959	0.389	63.179
	LSTM	4.70E + 11	3.92E + 11	0.973	0.974	0.454	80.68
	SVM	1.57E + 12	1.50E + 12	0.703	0.835	2.539	559.762
	LR	1.71E + 12	1.56E + 12	0.646	0.803	2.822	645.671
Mexico	GAN-GRU	6.38E + 03	3.99E + 03	0.998	0.999	0.001	2.006
	LSTM-CNN	6.29E + 03	5.41E + 03	0.998	0.999	0.002	3.593
	RBM	2.99E + 04	2.63E + 04	0.964	0.964	0.109	20.695
	GAN-DNN	3.34E + 04	2.80E + 04	0.955	0.962	0.12	21.328
	CNN	1.61E + 10	1.38E + 10	0.962	0.964	0.103	19.994
	LSTM	1.66E + 10	1.44E + 10	0.959	0.962	0.116	21.06
	SVM	2.98E + 10	2.70E + 10	0.868	0.876	0.426	78.665
	LR	3.47E + 10	3.03E + 10	0.821	0.863	0.563	99.751
Russia	GAN-GRU	2.52E + 04	2.32E + 04	0.99	0.997	0.016	10.266
	LSTM-CNN	1.03E + 04	9.03E + 03	0.998	0.999	0.001	2.937
	RBM	5.05E + 04	4.42E + 04	0.961	0.961	0.112	18.116
	GAN-DNN	2.59E + 04	2.12E + 04	0.99	0.991	0.058	9.709
	CNN	4.44E + 10	3.69E + 10	0.957	0.961	0.1	18.594
	LSTM	4.28E + 10	3.72E + 10	0.96	0.961	0.103	17.906
	SVM	7.01E + 10	6.07E + 10	0.893	0.893	0.353	66.182
	LR	8.03E + 10	6.82E + 10	0.86	0.863	0.445	81.487
Saudi Arabia	GAN-GRU	4.63E + 03	4.27E + 03	0.998	0.999	0.007	6.035
	LSTM-CNN	3.34E + 03	2.85E + 03	0.999	0.999	0.004	4.303
	RBM	1.84E + 04	1.63E + 04	0.962	0.963	0.108	19.045
	GAN-DNN	9.08E + 03	7.25E + 03	0.991	0.992	0.05	8.948
	CNN	5.45E + 09	4.58E + 09	0.962	0.967	0.104	20.453
	LSTM	4.99E + 09	4.36E + 09	0.968	0.969	0.084	20.799
	SVM	9.34E + 09	8.17E + 09	0.889	0.89	0.382	73.476
	LR	1.23E + 10	1.08E + 10	0.809	0.809	0.492	93.661
US	GAN-GRU	3.20E + 04	2.47E + 04	0.997	0.998	0.001	3.111
	LSTM-CNN	2.41E + 04	1.83E + 04	0.999	0.999	0.001	1.987
	RBM	1.31E + 05	1.21E + 05	0.958	0.959	0.076	18.832
	GAN-DNN	1.42E + 05	1.28E + 05	0.951	0.955	0.082	19.680
	CNN	3.15E + 11	2.82E + 11	0.954	0.956	0.059	17.619
	LSTM	1.91E + 11	1.62E + 11	0.983	0.984	0.027	12.061
	SVM	5.50E + 11	5.09E + 11	0.861	0.867	0.162	40.553
	LR	6.38E + 11	5.59E + 11	0.813	0.819	0.21	46.899

countries, namely Brazil, France, India, Mexico, Russia, Saudi Arabia, and the US, are modeled using six advanced data-driven models. The number of daily confirmed and recovered COVID-19 cases recorded from January 22nd, 2020, to September 6th, 2020, are used in this study. We applied the six deep learning models and two baseline machine learning models (i.e., LR and SVR) to forecast COVID-19 confirmed and recovered cases for a 92-day forecast horizon from June 6th, 2020. Six measurements of effectiveness are adopted to assess and compare the forecasting performance: R2, RMSE, MAE, MAPE, EV, and MSLE. Results show the domination of the LSTM-CNN model due to its capability to learn higher-level features that permit good forecasting precision.

The main challenges of training deep learning models are the availability of a sufficient amount of data to build an accurate model. Unfortunately, till now, there are no exact rules or methods to check these criteria. Essentially, we investigate the performance of deep learning models when applied to a small amount of time-series data and compare their forecasting accuracy with the traditional time-series models. Results in this study showed that deep learning models provide satisfying results, even with relatively small-sized data. The second challenge resides in tuning the hyper-parameter model used in the design of deep neural networks and the way to learn like the number of layers, the number of neural units by layer, the activation function, the optimizer, the learning rate, the loss function, number of epoch and

Table 7
Averaged statistical metric per method.

Model	RMSE	MAE	R2	EV	MSLE	MAPE
CNN	4.17E + 11	3.62E + 11	0.947	0.950	0.102	20.763
GAN-DNN	6.88E + 04	5.45E + 04	0.967	0.969	0.049	11.105
GAN-GRU	3.12E + 04	2.42E + 04	0.990	0.995	0.018	5.254
LSTM	3.60E + 11	3.07E + 11	0.964	0.967	0.095	20.394
LSTM-CNN	2.33E + 04	1.83E + 04	0.996	0.998	0.038	3.718
RBM	1.06E + 05	9.17E + 04	0.944	0.946	0.093	18.452
LR	9.72E + 11	8.65741E + 11	0.765	0.814	0.623	126
SVR	8.46852E + 11	7.83997E + 11	0.824	0.863	0.522	106

batch size. To mitigate this challenge several approaches are reported in the literature, such as grid search, random search, and Bayesian optimizating, to help optimize the set of hyper-parameters.

A direction for future improvement is to incorporate explanatory variables, such as meteorological measurements and lockdown indicators, in constructing the deep learning models. Further, it will be interesting to conduct comparative studies to investigate the impacts of measures established during this pandemic event and to extend this method to deal with different communicable and non-communicable diseases.

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.

Acknowledgement

References

[1] M. Toğaçar, B. Ergen, Z. Cömert, "COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches," Computers in Biology and Medicine, p. 103805, 2020.

[2] F. Khan, A. Saeed, S. Ali, Modelling and forecasting of new cases, deaths and recover cases of covid-19 by using vector autoregressive model in pakistan, Chaos, Solitons & Fractals 140 (2020) 110189.

[3] F. He, Y. Deng, W. Li, Coronavirus disease 2019: What we know? Journal of medical virology 92 (7) (2020) 719–725.

[4] Z.J. Cheng and J. Shan, "2019 novel coronavirus: where we are and what we know," Infection, pp. 1–9, 2020.

[5] J.B. Long and J.M. Ehrenfeld, "The role of augmented intelligence (ai) in detecting and preventing the spread of novel coronavirus," 2020.

[6] "About Healthmap (accessed on march 21, 2021)." [Online]. Available: http://www.diseasedaily.or/about.

[7] " Niller E. An AI Epidemiologist Sent the First Warnings of the Wuhan Virus: WIRED. (accessed on march 21, 2021)." [Online]. Available: https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/.

[8] I.I. Bogoch, O.J. Brady, M.U. Kraemer, M. German, M.I. Creatore, M.A. Kulkarni, J.S. Brownstein, S.R. Mekaru, S.I. Hay, E. Groot, et al., Anticipating the international spread of zika virus from brazil, The Lancet 387 (10016) (2016) 335–336.

[9] T. Davenport, R. Kalakota, The potential for artificial intelligence in healthcare, Future healthcare journal 6 (2) (2019) 94.

[10] L. Liu, J. Xu, Y. Huan, Z. Zou, S.-C. Yeh, L.-R. Zheng, A smart dental health-iot platform based on intelligent hardware, deep learning, and mobile terminal, IEEE journal of biomedical and health informatics 24 (3) (2019) 898–906.

[11] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early detection of parkinson's disease using deep learning and machine learning", IEEE Access, vol. 8, pp. 147 635–147 646, 2020.

[12] N.M. Khan, N. Abraham, and M. Hon, "Transfer learning with intelligent training data selection for prediction of alzheimer's disease", IEEE Access, vol. 7, pp. 72 726–72 735, 2019.

[13] F. Harrou, A. Dairi, F. Kadri, Y. Sun, Forecasting emergency department overcrowding: A deep learning framework, Chaos, Solitons & Fractals 139 (2020) 110247.

[14] W.-J. Chang, L.-B. Chen, C.-H. Hsu, C.-P. Lin, and T.-C. Yang, "A deep learning-based intelligent medicine recognition system for chronic patients," IEEE Access, vol. 7, pp. 44 441–44 458, 2019.

[15] S. Lalmuanawma, J. Hussain, L. Chhakchhuak, Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review, Chaos, Solitons & Fractals (2020) 110059.

[16] W. Xue, Q. Li, Q. Xue, Text detection and recognition for images of medical laboratory reports with a deep learning approach, IEEE Access 8 (2019) 407–416.

[17] W.-J. Chang, L.-B. Chen, M.-C. Chen, Y.-C. Chiu, and J.-Y. Lin, "ScalpEye: A Deep Learning-Based Scalp Hair Inspection and Diagnosis System for Scalp Health," IEEE Access, vol. 8, pp. 134 826–134 837, 2020.

[18] F. Harrou, A. Dairi, Y. Sun, M. Senouci, Statistical monitoring of a wastewater treatment plant: A case study, Journal of environmental management 223 (2018) 807–814.

[19] J. Shuja, E. Alanazi, W. Alasmary, A. Alashaikh, COVID-19 open source data sets: a comprehensive survey, Applied Intelligence (2020) 1–30.

[20] C. Butt, J. Gill, D. Chun, and B.A. Babu, "Deep learning system to screen coronavirus disease 2019 pneumonia," Applied Intelligence, p. 1, 2020.

[21] Y. Mohamadou, A. Halidou, and P.T. Kapen, "A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of covid-19," Applied Intelligence, pp. 1–13, 2020.

[22] T. Goel, R. Murugan, S. Mirjalili, and D.K. Chakrabartty, "Optconet: an optimized convolutional neural network for an automatic diagnosis of covid-19," Applied Intelligence, pp. 1–16, 2020.

[23] L. Brunese, F. Mercaldo, A. Reginelli, A. Santone, Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays, Comput. Methods Programs Biomed. 196 (2020) 105608.

[24] A.I. Khan, J.L. Shah, and M.M. Bhat, "Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," Computer Methods and Programs in Biomedicine, p. 105581, 2020.

[25] M. Turkoglu, COVIDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble, Applied Intelligence (2020) 1–14.

[26] F. Ucar and D. Korkmaz, "COVIDiagnosis-Net: Deep Bayes-SqueezeNet based Diagnostic of the Coronavirus Disease 2019 (COVID-19) from X-Ray Images," Medical Hypotheses, p. 109761, 2020.

[27] T. Zebin and S. Rezvy, "COVID-19 detection and disease progression visualization: Deep learning on chest X-rays for classification and coarse localization," Applied Intelligence, pp. 1–12, 2020.

[28] A. Abbas, M.M. Abdelsamea, and M.M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," arXiv preprint arXiv:2003.13815, 2020.

[29] A. Zeroual, F. Harrou, A. Dairi, Y. Sun, Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study, Chaos, Solitons & Fractals 140 (2020) 110121.

[30] K. Sarkar, S. Khajanchi, J.J. Nieto, Modeling and forecasting the covid-19 pandemic in india, Chaos, Solitons & Fractals 139 (2020) 110049.

[31] Z. Abbasi, I. Zamani, A.H.A. Mehra, M. Shafieirad, A. Ibeas, Optimal control design of impulsive sqeiar epidemic models with application to covid-19, Chaos, Solitons & Fractals 139 (2020) 110054.

[32] C.H. Nkwayep, S. Bowong, J. Tewa, J. Kurths, Short-term forecasts of the covid-19 pandemic: a study case of cameroon, Chaos, Solitons & Fractals 140 (2020) 110106.

[33] M.H.D.M. Ribeiro, R.G. da Silva, V.C. Mariani, L. dos Santos Coelho, Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil, Chaos, Solitons & Fractals (2020) 109853.

[34] Y. Peng, M.H. Nagata, An empirical overview of nonlinearity and overfitting in machine learning using covid-19 data, Chaos, Solitons & Fractals 139 (2020) 110055.

[35] D.P. Kavadi, R. Patan, M. Ramachandran, A.H. Gandomi, Partial derivative nonlinear global pandemic machine learning prediction of covid 19, Chaos, Solitons & Fractals 139 (2020) 110056.

[36] F. Harrou, A. Dairi, Y. Sun, F. Kadri, Detecting abnormal ozone measurements with a deep learning-based strategy, IEEE Sens. J. 18 (17) (2018) 7222–7232.

[37] T. Cheng, F. Harrou, F. Kadri, Y. Sun, and T. Leiknes, "Forecasting of wastewater treatment plant key features using deep learning-based models: A case study," IEEE Access, 2020.

[38] F. Harrou, Y. Sun, A.S. Hering, M. Madakyaru, et al., Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches: Theory and Practical Applications, Elsevier, 2020.

[39] A. Dairi, F. Harrou, Y. Sun, M. Senouci, Obstacle detection for intelligent transportation systems using deep stacked autoencoder and k-nearest neighbor scheme, IEEE Sens. J. 18 (12) (2018) 5122–5132.

[40] F. Harrou, M.M. Hittawe, Y. Sun, O. Beya, Malicious attacks detection in crowded areas using deep learning-based approach, IEEE Instrumentation & Measurement Magazine 23 (5) (2020) 57–62.

[41] A. Dairi, T. Cheng, F. Harrou, Y. Sun, T. Leiknes, Deep learning approach for sustainable WWTP operation: A case study on data-driven influent conditions monitoring, Sustainable Cities and Society 50 (2019) 101670.

[42] V.K.R. Chimmula, L. Zhang, Time series forecasting of covid-19 transmission in canada using lstm networks, Chaos, Solitons & Fractals (2020) 109864.

[43] P. Wang, X. Zheng, G. Ai, D. Liu, B. Zhu, Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: case studies in Russia, Peru and Iran, Chaos, Solitons & Fractals (2020) 110214.

[44] F. Shahid, A. Zameer, M. Muneeb, Predictions for covid-19 with deep learning models of lstm, gru and bi-lstm, Chaos, Solitons & Fractals 140 (2020) 110212.

[45] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing

Systems 25, F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, Eds. Curran Associates Inc, 2012, pp. 1097–1105.

[46] M.M. Hittawe, D. Sidibé, F. Mériaudeau, Bag of words representation and svm classifier for timber knots detection on color images, 2015 14th IAPR international conference on machine vision applications (MVA). IEEE (2015) 287–290.

[47] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," 2014.

[48] A. Graves, A. rahman Mohamed, and G.E. Hinton, "Speech recognition with deep recurrent neural networks," CoRR, vol. abs/ 1303.5778, 2013. [Online]. Available: http://arxiv.org/abs/1303.5778.

[49] M.M. Hittawe, S. Afzal, T. Jamil, H. Snoussi, I. Hoteit, O. Knio, Abnormal events detection using deep neural networks: application to extreme sea surface temperature detection in the red sea, J. Electron. Imaging 28 (2) (2019) 021012.

[50] M.M. Hittawe, D. Sidibé, and F. Mériaudeau, "A machine vision based approach for timber knots detection," in Twelfth International Conference on Quality Control by Artificial Vision 2015, vol. 9534. International Society for Optics and Photonics, 2015, p. 95340L.

[51] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of go with deep neural networks and tree search, Nature 529 (2016) 484–503.

[52] M.M. Barbat, C. Wesche, A.V. Werhli, M.M. Mata, An adaptive machine learning approach to improve automatic iceberg detection from SAR images, ISPRS Journal of Photogrammetry and Remote Sensing 156 (2019) 247–259.

[53] A. Colubri, M.-A. Hartley, M. Siakor, V. Wolfman, A. Felix, T. Sesay, J.G. Shaffer, R. F. Garry, D.S. Grant, A.C. Levine, et al., Machine-learning prognostic models from the 2014–16 Ebola outbreak: data-harmonization challenges, validation strategies, and mHealth applications, EClinicalMedicine 11 (2019) 54–64.

[54] G. Nápoles, I. Grau, R. Bello, R. Grau, Two-steps learning of Fuzzy Cognitive Maps for prediction and knowledge discovery on the HIV-1 drug resistance, Expert Syst. Appl. 41 (3) (2014) 821–830.

[55] U. Chockanathan, A.M. DSouza, A.Z. Abidin, G. Schifitto, A. Wismüller, Automated diagnosis of HIV-associated neurocognitive disorders using large-scale Granger causality analysis of resting-state functional MRI, Comput. Biol. Med. 106 (2019) 24–30.

[56] S. Saxena, M. Gyanchandani, Machine learning methods for computer-aided breast cancer diagnosis using histopathology: A narrative review, Journal of Medical Imaging and Radiation Sciences 51 (1) (2020) 182–193.

[57] T. Nazir, A. Irtaza, Z. Shabbir, A. Javed, U. Akram, M.T. Mahmood, Diabetic retinopathy detection through novel tetragonal local octa patterns and extreme learning machines, Artificial intelligence in medicine 99 (2019) 101695.

[58] P. Sharma, K. Choudhary, K. Gupta, R. Chawla, D. Gupta, A. Sharma, Artificial plant optimization algorithm to detect heart rate & presence of heart disease using machine learning, Artif. Intell. Med. 102 (2020) 101752.

[59] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (8) (1997) 1735–1780.

[60] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.

[61] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436–444.

[62] P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory; cu-cs-321-86," 1986.

[63] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672–2680.

[64] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

[65] L.-J. Cao, F.E.H. Tay, Support vector machine with adaptive parameters in financial time series forecasting, IEEE Transactions on neural networks 14 (6) (2003) 1506–1518.

[66] B.-Z. Yao, C.-Y. Yang, J.-B. Yao, J. Sun, Tunnel surrounding rock displacement prediction using support vector machine, International Journal of Computational Intelligence Systems 3 (6) (2010) 843–852.

[67] Z. Mingheng, Z. Yaobao, H. Ganglong, C. Gang, Accurate multisteps traffic flow prediction based on svm, Mathematical Problems in Engineering 2013 (2013).

[68] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Statistics and computing 14 (3) (2004) 199–222.

[69] F.-L. Chu, Using a logistic growth regression model to forecast the demand for tourism in las vegas, Tourism Management Perspectives 12 (2014) 62–67.

[70] R.J. Hyndman, A.B. Koehler, Another look at measures of forecast accuracy, International journal of forecasting 22 (4) (2006) 679–688.

[71] S. Kim, H. Kim, A new metric of absolute percentage error for intermittent demand forecasts, Int. J. Forecast. 32 (3) (2016) 669–679.