Decision Making: Applications in Management and Engineering
Vol. 6, Issue 1, 2023, pp. 502-534.
ISSN: 2560-6018
eISSN: 2620-0104
DOI: https://doi.org/10.31181/dmame060129022023u

DEEP LEARNING BASED AN EFFICIENT HYBRID PREDICTION MODEL FOR COVID-19 CROSS-COUNTRY SPREAD AMONG E7 AND G7 COUNTRIES

Anil Utku[1]*

[1] Department of Computer Engineering, Munzur University, Tunceli, Turkey

Received: 28 September 2022;
Accepted: 25 February 2023;
Available online: 28 February 2023.

Original scientific paper

Abstract: *The COVID-19 pandemic has caused the death of many people around the world and has also caused economic problems for all countries in the world. In the literature, there are many studies to analyze and predict the spread of COVID-19 in cities and countries. However, there is no study to predict and analyze the cross-country spread in the world. In this study, a deep learning based hybrid model was developed to predict and analysis of COVID-19 cross-country spread and a case study was carried out for Emerging Seven (E7) and Group of Seven (G7) countries. It is aimed to reduce the workload of healthcare professionals and to make health plans by predicting the daily number of COVID-19 cases and deaths. Developed model was tested extensively using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R Squared (R^2). The experimental results showed that the developed model was more successful to predict and analysis of COVID-19 cross-country spread in E7 and G7 countries than Linear Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM). The developed model has R^2 value close to 0.9 in predicting the number of daily cases and deaths in the majority of E7 and G7 countries.*

Key words: *COVID-19, machine learning, deep learning, cross-country spread, CNN, RNN.*

1. Introduction

COVID-19, which was first seen in the Wuhan region of China in early December 2019, is a contagious virus that causes respiratory tract infection and can be passed from person to person (Jernigan, 2020). COVID-19 causes symptoms such as shortness of breath, high fever and cough (Ahmad et al., 2022).

The virus that causes COVID-19 spreads very quickly from person to person (Toğaçar et al., 2020). According to current information, the virus is transmitted to each other from people within a distance of about 2 meters (Klompas et al., 2021).

* Corresponding author.
E-mail address: anilutku@munzur.edu.tr (A. Utku)

This virus is transmitted by airborne droplets when a person sneezes, coughs, breathes and speaks (Schijven et al., 2021). These droplets can be transmitted by breathing or by getting into the mouth, nose and eyes. In some cases, the COVID-19 virus can be transmitted to people by exposure to airborne droplets. People can be exposed to the virus if they touch an infected area and then touch their mouth (Jarag et al., 2020).

It is known that people with certain diseases are more likely to become seriously ill and die if infected with the new coronavirus (Alakuş & Türkoğlu, 2020). People with advanced age or chronic diseases are among the risk groups for COVID-19. Hypertension and diabetes patients are in the risky group (Fanget al., 2020). Since COVID-19 is a type of virus that can affect the lungs, those with chronic or acute respiratory diseases such as COPD are also in the risk group. In addition, people with weakened immune systems, such as cancer patients, and those taking immunosuppressive drugs are advised to be more careful than other people (Minotti et al., 2020).

The concept of E7 means 7 developing countries. E7 countries consist of China, India, Brazil, Russia, Indonesia, Mexico and Turkey. These countries have made great progress in the last 20 years and have become one of the strongest economies in the world (Tong et al., 2020). The G7 concept means the seven largest economies in the world. The G7 countries consist of the USA, Germany, France, England, Italy, Japan and Canada (Lee et al., 2012). COVID-19 has affected all countries in the world negatively, as well as the E7 and G7 countries. COVID-19 is not only a health problem, but also a problem that affects the economies of countries (Mamun & Ullah, 2020). Therefore, in this study, it is aimed to predict the daily number of COVID-19 cases and deaths for E7 countries and G7 countries. It is also aimed to analyze the spread of COVID-19 between these countries.

PCR tests are used to detect COVID-19. However, the cost of these tests and the long-term delivery of the results of the tests cause the patients not to be able to start their treatment in a short time. For this reason, it has been inevitable to use intelligent systems to improve this process in the detection of COVID-19 disease. Artificial intelligence methods, especially deep learning algorithms, have become a technology that offers important solutions in the field of health in recent years. Artificial intelligence is the ability of machines to showcase human-specific skills such as reasoning, learning, planning, and creativity. Artificial intelligence allows machines to sense their environment, deal with what they perceive, solve problems, and act to achieve a specific goal. Artificial intelligence is widely used in the field of health as well as in many application areas in daily life. Artificial intelligence methods are used in medical applications such as early diagnosis, diagnosis, decision making, treatment, research and education. In addition, artificial intelligence can help healthcare professionals adopt a more comprehensive approach to disease management, and improve patient compliance by providing better coordination. Artificial intelligence can be used in application areas such as early diagnosis, monitoring of treatment processes, filiation and contact tracing and prediction of the number of cases/deaths in the COVID-19 pandemic. Many successful systems developed with deep learning algorithms will be useful in the fight against COVID-19 disease. Accurate prediction of the number of COVID-19 cases and deaths is important in terms of optimizing healthcare personnel, hospitals and equipment.

The motivation for this study is to predict the spread of COVID-19. During the spread of COVID-19, the healthcare system in most countries collapsed. Health workers worked in very difficult conditions. Millions of people were infected and died due to COVID-19. By predicting the spread of COVID-19, health systems can be optimized, processes in hospitals can be regulated, and measures can be taken to prevent transmission of the virus. Therefore, in this study, it was aimed to determine the spread of COVID-19 by predicting the daily number of COVID-19 cases and deaths. A hybrid deep learning model was developed and extensively compared with popular machine learning and deep learning models such as LR, RF, SVM, MLP, CNN, RNN and LSTM. In addition, based on the existing literature, it was seen that the spread characteristics of COVID-19 between countries were not analyzed in the literature.

Accurately detecting diseases requires years of medical education. Even after this education, diagnosing is a challenging and time-consuming task. In many areas of medicine, the demand for specialists exceeds supply, which puts doctors under stress and the diagnosis of diseases is delayed further. Machine learning, and especially deep learning methods, have made great progress in the automatic diagnosis of diseases in recent years. This makes the diagnostic process cheaper, easier and more accessible. Considering that pandemic diseases such as COVID-19 affect many people around the world, predicting the course of the pandemic is important in terms of optimizing the health systems of countries. In this section, studies in the literature using artificial intelligence methods for the diagnosis and prediction of the spread distribution of COVID-19 were evaluated.

Che Azemin et al. presented a ResNet-101-based model for detecting COVID-19 cases using chest radiography images (Che Azemin et al., 2020). The proposed model had 0.82 AUROC, 77.3% sensitivity, 71.8% specificity, and 71.9% accuracy.

Masum et al. presented a comparative analysis of LSTM, RF and SVM to predict the spread of COVID-19 in Bangladesh (Masum et al., 2020). Daily COVID-19 case, death and recovery data from May 2020 to June 15, 2020 were used as the dataset. Experimental results according to the RMSE metric showed that LSTM was more successful than other models.

Shahid et al. presented a comparative analysis of ARIMA, SVM, LSTM, Bi-LSTM, and GRU models to predict COVID-19 spread for Brazil, China, Germany, India, Israel, Italy, Russia, Spain, UK, and USA (Shahid et al., 2020). As a dataset, daily case, death and recovery numbers between January 2020 and June 2020 were used. Experimental studies based on MAE, RMSE and R^2 metrics showed that Bi-LSTM is more successful than other models.

Kırbaş et al. presented a comparative analysis of ARIMA, Nonlinear Autoregression Neural Network (NARNN), and LSTM using COVID-19 case data from 8 different European countries until May 2020 (Kırbaş et al., 2020). Predictions of the number of cases for 14 days were analyzed. Experimental results showed that the MAPE value of LSTM was more successful than other models.

Zeroual et al. presented a comparative analysis of the RNN, LSTM, Bi-LSTM, and Variational AutoEncoder (VAE) models to predict the number of COVID-19 cases and recoveries (Zeroual et al., 2020). In the study, COVID-19 data from Italy, Spain, France, China, USA and Australia from January 2020 to June 2020 were used as a dataset. Experimental results showed that VAE was more successful than other models.

Shastri et al. presented a comparative analysis of the LSTM, Stacked LSTM, Bi-directional LSTM and ConvLSTM models to predict the number of COVID-19 cases and deaths in India and the USA (Shastri et al., 2020). COVID-19 data from February 2020 to July 2020 were used as the dataset. Experimental results showed that ConvLSTM outperformed other models compared.

Satu et al. developed a forecasting model to forecast the seven-day cases in Bangladesh. In their study, 25-day cases dataset was used as the training dataset

(Satu et al., 2021). Experimental results showed that the prophet is more successful than LR, SVM and Multi-Layer Perceptron (MLP).

Abbasimehr and Paki developed a hybrid model for COVID-19 prediction using CNN, LSTM and Bayesian optimization algorithm (Abbasimehr & Paki, 2021). Bayesian optimization algorithm is used to optimize the hyperparameters of the models. Experimental results showed that the proposed model has a SMAPE value of 0.25 in short-term prediction and 2.59 in long-term prediction.

Ayoobi et al. presented a comparative analysis of the LSTM, GRU, ConvLSTM, Bi-LSTM, Bi-GRU, Bi-ConvLSTM models for predicting the number of cases and deaths in Australia and Iran (Ayoobi et al., 2021). Predictions were made for periods of 1, 3 and 7 days. Experimental results showed that Bi-GRU was more successful than other models compared in 1-day estimation of the number of deaths.

Marzouk et al. presented a comparative analysis of LSTM, CNN and MLP to predict the spread of COVID-19 in Egypt (Marzouk et al., 2021). COVID-19 data from February 2020 to August 2020 were used as the dataset. Experimental results showed that LSTM was more successful than other models compared.

Devaraj et al. presented a comparative analysis of the Stacked Long Short-Term Memory (SLSTM), ARIMA, LSTM, and Prophet models for predicting the number of COVID-19 cases (Devaraj et al., 2021). Predictions of the number of cases for 30, 60 and 90 days were made using the COVID-19 data from January 2020 to May 2020. Experimental results showed that SLSTM was more successful than other models.

Elsheikh et al. developed an LSTM-based model to predict the number of COVID-19 cases, deaths, and recoveries (Elsheikh et al., 2021). The developed model was compared with Nonlinear Autoregressive Artificial Neural Networks (NARANN) and ARIMA. COVID-19 data from Brazil, India, Saudi Arabia, South Africa, Spain and the USA were used as dataset. Experimental results showed that LSTM outperformed other models compared for all countries.

Alassafi et al. developed a model to predict the spread of COVID-19 in Morocco, Malaysia, and Saudi Arabia. Using the developed prediction model, the number of 7-day COVID-19 cases and deaths were predicted (Alassafi et al., 2022). Experimental studies using RNN and LSTM showed that LSTM has 98.58% precision and RNN has 93.45% precision.

Verma et al. presented a comparative analysis of LSTM variations, CNN and CNN+LSTM models for COVID-19 prediction (Verma et al., 2022). 7, 14 and 21 day predictions were made for the 4 cities with the highest number of cases in India. Experimental results have shown that the CNN+LSTM model is more successful than other models.

Ketu and Mishra presented a hybrid deep learning model for predicting the spread of COVID-19 for 29 states in India (Ketu & Mishra, 2022). In the study, the developed CNN-LSTM hybrid model was compared with ARIMA and LSTM. Experimental results showed that the developed hybrid model was more successful than other models. The developed model had R^2 of over 0.9 in almost all states.

In the literature, there are many studies on predicting the spread of COVID-19 in cities and countries and detecting COVID-19 from x-ray images. However, there is no study in the literature to predict and analyze the spread of COVID-19 among countries in the world. For this purpose, a hybrid deep learning model was developed to predict and analyze the spread of COVID-19 between countries. A case study was conducted for E7 and G7 countries. It is also aimed to effectively predict the daily number of cases and deaths in E7 and G7 countries.

The main contributions of this study to the literature can be summarized as follows:

- It is aimed to predict the daily number of COVID-19 cases and deaths in E7 and G7 countries.
- A CNN-RNN based hybrid deep learning prediction model was proposed and compared with popular machine learning and deep learning models such as LR, RF, SVM, MLP, CNN, RNN and LSTM.
- A comprehensive analysis was conducted to determine the countries with similar patterns for 5-day and 14-day incubation periods by determining the peak dates of the number of COVID-19 cases and deaths in E7 and G7 countries.

2. The Proposed Hybrid COVID-19 Prediction Model

In this study, a hybrid deep learning model was proposed to increase the prediction accuracy and efficiency. In the proposed model, the success of CNN in the feature extraction stage and the success of RNN in the learning and prediction stage in sequential data were used. With the proposed model, it is aimed to determine the daily number of COVID-19 cases and deaths. The architecture of the proposed model is shown in Figure 1.

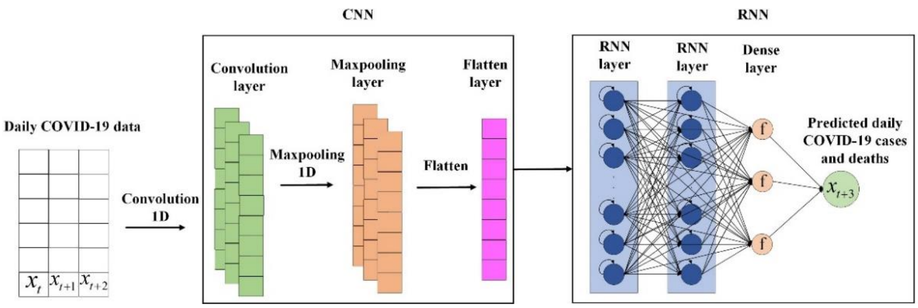

Figure 1. The architecture of the proposed model

In the proposed model, firstly, the daily number of cases and deaths along with their dates were extracted from the dataset. The model takes number of cases and deaths as input. The output of the model is the predicted number of daily cases and deaths. CNN is an efficient model for automatically extracting features and learning from one-dimensional series data such as univariate time series. In this study, CNN is used to interpret sub-sequences that are input to RNN.

In the developed hybrid model, CNN is used for feature extraction and RNN is used to analyze and predict features extracted by CNN. In order for the proposed model to be used in the prediction problem, the first step is divide the input sequences into sub-sequences that can be processed by the CNN. For this reason, using the sliding window method, univariate time series data is divided into input/output samples as 3 inputs and one output.

CNN interprets these sub-sequence samples and forwards them to RNN for processing as input. Here, the CNN has a one-dimensional convolutional layer with a kernel size of 1 and a number of filters of 64 to read substrings. After the convolutional layer there is a max pooling layer which is used to interpret the input feature. There is a dense layer to interpret the features extracted by the convolution layer. The flatten layer is used to reduce the 3D feature maps obtained from the convolution and pooling layers to a one-dimensional vector to be used as input to the RNN.

In order to apply the proposed model to the problem of predicting the spread of COVID-19, pre-processing was done on the dataset. The daily COVID-19 dataset used in this study is a time series dataset. In order for the time series data to be processed by machine learning and deep learning models, it is necessary to transform the time series data into supervised learning problem. For this purpose, the sliding window method is used.

In the sliding window method, the data is presented as an input to the window according to the specified window size, as seen in Figure 2. The value at the next time step is the output to be predicted. Thus, t_1, t_2 and t_3 are selected as inputs and t_4 as output. Min-max normalization is used for scaling the number of cases and deaths so that these fall in smaller range, such as 0 to 1(Henrikson et al., 2015).

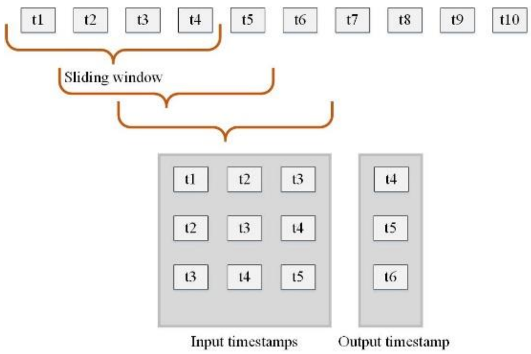

Figure 2. Sliding window method

In studies in the literature, 67% training and 33% testing rates are commonly used to divide training and test sets (Sun et al., 2021). In this study, these rates were used because lower error values were obtained at 67% training and 33% testing rates as a result of experimental studies. The training data was divided into 90% for training and %10 for validation. The validation data was used for optimization of model parameters. Model parameters were optimized using GridSearchCV from the Scikit Learn library so that the applied models could obtain the best prediction results.

3. Baseline Models

LR is a statistical method that summarizes the relationship between 2 quantitative data (Giambartolomei et al., 2018). It aims to predict a dependent variable using the values of the independent variables. LR predicts the coefficients of the linear equation using one or more independent variables that best predict the value of the dependent variable. LR creates a straight line that minimizes discrepancies between predicted values and actual values (El-Khaiary, 2008). The key point in LR is that the dependent variable is continuous. However, independent variables should be measurable on a categorical or continuous measurement scale. There are two types of linear regression models, simple regression and multiple regression. Simple linear regression is performed using an independent variable to predict a dependent variable (Ali et al., 2020). For example, by looking at the relationship between the age variable and the cigarette addiction variable, the cigarette use of a person of the relevant age can be predicted. When more than one independent variable is present, the process is called multiple linear regression (Mize, 2019). For example, to predict smoking addiction using age and gender.

RF is a machine learning method based on decision trees. RF is create a forest by combining the results from multiple decision trees. The final prediction is made by combining the prediction results obtained from each decision tree (Naghibi et al., 2016). Therefore, RF is an ensemble learning method. Instead of branching each node using the best branch among all the variables, RF branches each node using the best randomly selected variables at each node. Each dataset is generated by displacement from the original dataset. Trees are then developed using random feature selection (Speiser et al., 2019). Developed trees are not pruned.

SVM are machine learning algorithms based on convex optimization that work on the principle of structural risk minimization (Xu & Zhu, 2021). SVM are distribution-free learning algorithms as they do not need any combined distribution function knowledge of the data. SVM is a vector space based machine learning method that finds a decision boundary between the two most distant classes from any point in the training data determined on the dataset (Chamasemani & Singh , 2011). SVM maps training samples to points in space to maximize the width of the gap between the two classes (Ballabio & Sterlacchini, 2012). New instances are then mapped to the same space and guessed which class they belong to base on which side of the space they land on.

Artificial Neural Networks (ANN) are known as the most powerful and flexible machine learning methods (Samra et al, 2020). ANN was developed by being inspired by the working system of the human brain. The most important reason for the widespread use of ANN is that it provides an effective alternative for problems that are difficult to solve with classical methods (Salehi & Burgueño, 2018). MLP is the most widely used model with a simple structure and is popular among traditional feed-forward ANN (Mahmoudi et al., 2016). While single-layer perceptrons can predict linear events, MLP can be used to predict nonlinear events. MLP consists of several layers, including an input layer, one or more hidden layer(s), and an output layer (Bikku, 2020). In MLP, the input layer is the layer where the input parameters are presented. The values for the output layer parameters are determined according to the target values of the problem of interest (Huang et al., 2013).

CNN is a model that was successfully applied in image processing and classification problems (Basha et al., 2020). CNN basically consists of convolution and fully connected layer (Tiwari et al., 2022). Convolution is the first layer used to extract features from the input. Convolution creates an output matrix by performing matrix calculations (Hari et al., 2021). The pooling layer is used to reduce parameters and computations in the network. The fully connected layer vectorizes the matrix data passing through the convolutional layer and the pooling layer (Wang et al., 2018).

RNN are networks in which the connections between their units form a directed loop (Ramadevi et al., 2012). In RNN, predictions are created by associating the information coming to the neural networks with certain weight constants in the layers (Chen et al., 2015). If there is a margin of error when this predict is compared with the actual data, the weight constants are changed and the prediction is reconstructed.

LSTM is a kind of recurrent neural network architecture that avoids the vanishing gradient problem (Jozefowicz et al., 2015). LSTM allows the error to propagate backwards with a limited range. A basic LSTM unit consists of three basic gates: an input gate, an output gate, and a forget gate (Mirzaei et al., 2022). According to the status of the gates, it is determined which information should be protected and when the units will be accessed (Gao & Glowacka, 2016).

4. The Experimental Results

In this study, CNN-RNN based a hybrid model was proposed to predict the number of COVID-19 cases and deaths. The experimental results of the proposed model were extensively compared with LR, RF, SVM, MLP, CNN, RNN and LSTM. Experimental results of each model were compared according to MSE, RMSE and R^2.

4.1. Dataset

In this study, the COVID-19 dataset, which includes the daily number of cases and deaths, presented by the World Health Organization (WHO), was used. The dataset consists of eight features: 'Date_reported', 'Country_code', 'Country', 'WHO_region', 'New_cases', 'Cumulative_cases', 'New_deaths' and 'Cumulative_deaths'. "Date_reported" refers to the date of data reported to WHO. "Country_code" refers to the country code of type ISO Alpha-2. "Country" refers to the name of a country. "WHO_region" refers to WHO-assigned regional offices. "New_cases" refers to the number of new cases per day. "Cumulative_cases" refers to the total number of cases up to a given date. "New_deaths" refers to the number of new deaths. "Cumulative_deaths" refers to the total number of deaths up to a given date. In this study, daily COVID-19 data between January 03, 2020 and May 31, 2022 were used. The dataset used is publicly available via https://covid19.who.int/WHO-COVID-19-global-data.csv.

The E7 countries consist of China, India, Brazil, Russia, Indonesia, Mexico and Turkey. The G7 countries are the USA, Germany, France, England, Italy, Japan and Canada. Figure 3 shows the E7 countries and the total number of cases.

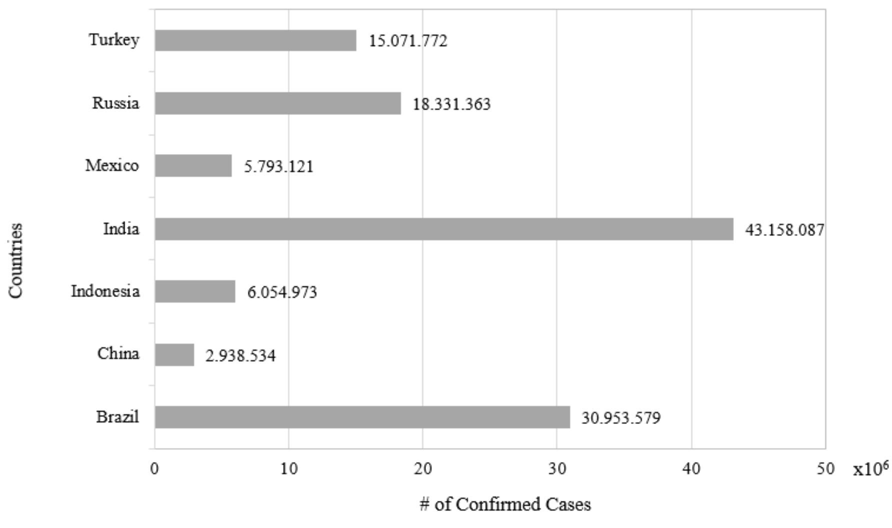

Figure 3. E7 countries and the total number of cases

As seen in the Figure 3, India has the highest number of cases at 43,158,097. After India, Brazil, Russia and Turkey have higher total number of cases than any other E7 countries. Figure 4 shows the total number of deaths in E7 countries.

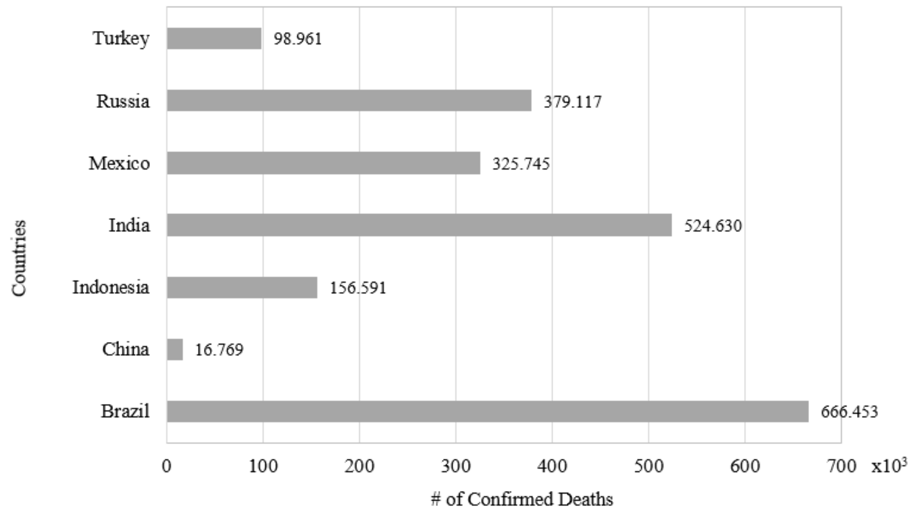

Figure 4. E7 countries and the total number of deaths

As seen in Figure 4, Brazil has the highest number of deaths at 666,453. After Brazil, India, Russia and Mexico have higher deaths than any other E7 countries. China has the lowest total number of deaths. Figure 5 shows the G7 countries and the total number of cases.

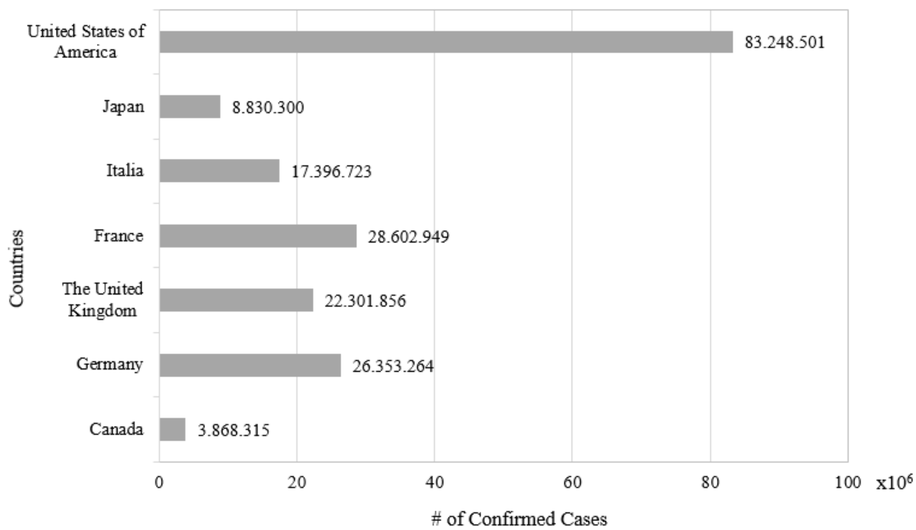

Figure 5. G7 countries and the total number of cases

As seen in the Figure 5, United States of America has the highest number of cases at 83,248,501. After France, Germany, The United Kingdom and Italia have higher total number of cases than any other G7 countries. Figure 6 shows the total number of deaths in G7 countries.

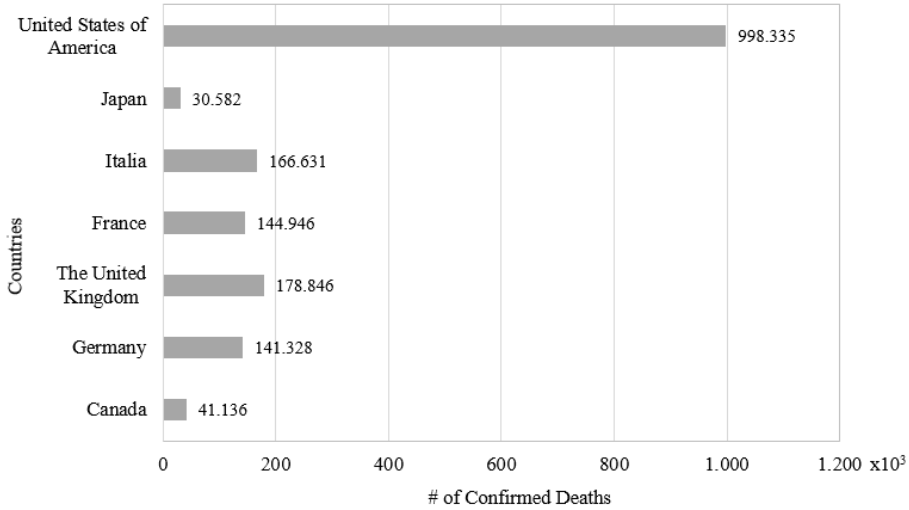

Figure 6. G7 countries and the total number of deaths

As seen in Figure 6, United States of America has the highest number of deaths at 998,335. After United States of America, The United Kingdom, Italia, France and Germany have higher deaths than any other G7 countries. Japan has the lowest total number of deaths.

4.2. The Evaluation Metrics

The RMSE is used to measure the distance of differences between actual values and predicted values. It is calculated by taking the square root of the mean of the squares of error. MSE is calculated using Eq. (1).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y-\hat{y})^2} \tag{1}$$

where, y is the actual value, $\hat{y}$ is the predicted value, and n is the total number of instances in the dataset.

The MAE is calculated by averaging the absolute values of the differences between the actual values and the predicted values, as seen in Eq. (2).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y-\hat{y}| \tag{2}$$

R^2 is a measure of variance and is a measure of how well the data fit the model used. R^2 is calculated using Eq. (3).

$$R^2 = 1 - \frac{\sum(y-\hat{y})^2}{(y-\bar{y})^2} \tag{3}$$

where, $\hat{y}$ is the predicted value and $\overline{y}$ represents the average of y.

4.3. The Prediction of COVID-19 Spread Between E7 and G7 Countries

In this study, it is aimed to predict the number of daily COVID-19 cases and deaths in E7 and G7 countries using the proposed hybrid prediction model. The proposed model was compared with LR, RF, SVM, MLP, CNN, RNN and LSTM using RMSE, MAE and R^2.

Table 1 shows the experimental results according to the RMSE for predicting the number of cases in E7 countries.

Table 1. The experimental results according to the RMSE for predicting the number of cases in E7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
BR	31035.54	39422.01	31068.81	30536.43	30894.43	30028.90	28992.99	23046.15
CN	6890.96	7601.69	13036.43	7058.25	7487.39	6045.49	5835.01	5605.66
ID	3400.57	3945.62	3405.81	3333.42	3281.38	3231.33	3178.54	2104.10
IN	8587.26	12021.30	8569.24	8639.98	9092.09	8698.31	8511.27	5210.78
MX	6690.18	10906.32	6968.31	3980.55	3947.55	3934.60	3820.71	3713.31
RU	4577.96	4417.21	8162.57	4286.38	5964.27	4217.77	4160.07	2276.99
TR	3882.42	11973.07	3883.98	3871.58	4000.07	3874.57	3851.69	1743.71

As seen in Table 1, the proposed model has a better prediction performance than the other models compared according to RMSE. After the proposed model, LSTM, RNN and MLP have more successful results than other models.

Table 2 shows the experimental results according to the MAE for predicting the number of cases in E7 countries.

Table 2. The experimental results according to the MAE for predicting the number of cases in E7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
BR	15722.68	19756.91	15227.46	14305.76	13800.51	15691.10	15645.47	11574.68
CN	2653.11	2951.15	5754.95	2700.66	2910.12	2267.26	2197.60	2177.04
ID	1492.61	1785.72	1605.80	1427.75	1431.07	1394.92	1377.16	931.92
IN	3890.91	5339.06	4328.65	3981.41	4142.12	3916.29	3823.47	2119.86
MX	3102.42	4002.35	2798.26	1876.30	1843.39	1855.33	1760.50	1759.80
RU	2367.32	2245.47	3951.41	2144.67	2726.59	2120.31	2074.65	937.47
TR	2339.70	5846.67	2444.73	2385.10	2379.08	2311.43	2310.87	1078.40

As seen in Table 2, the proposed model has a better prediction performance than the other models compared according to MAE. After the proposed model, LSTM, RNN and LR have more successful results than other models. Table 3 and Figure 7 show the experimental results according to the R^2 for predicting the number of cases in E7 countries.

Table 3. The experimental results according to the R^2 for predicting the number of cases in E7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
BR	0.635	0.411	0.634	0.646	0.638	0.659	0.682	0.754
CN	0.888	0.865	0.602	0.884	0.869	0.914	0.920	0.923
ID	0.940	0.920	0.940	0.943	0.945	0.946	0.948	0.968
IN	0.985	0.971	0.985	0.985	0.983	0.986	0.988	0.993
MX	0.804	0.479	0.787	0.930	0.931	0.932	0.935	0.939
RU	0.989	0.990	0.967	0.990	0.983	0.990	0.991	0.997
TR	0.979	0.942	0.979	0.979	0.978	0.979	0.979	0.998

As seen in Table 3 and Figure 7, the proposed model has a better prediction performance than the other models compared according to R^2. After the proposed model, LSTM, RNN, MLP, SVM and LR have more successful results than other models.

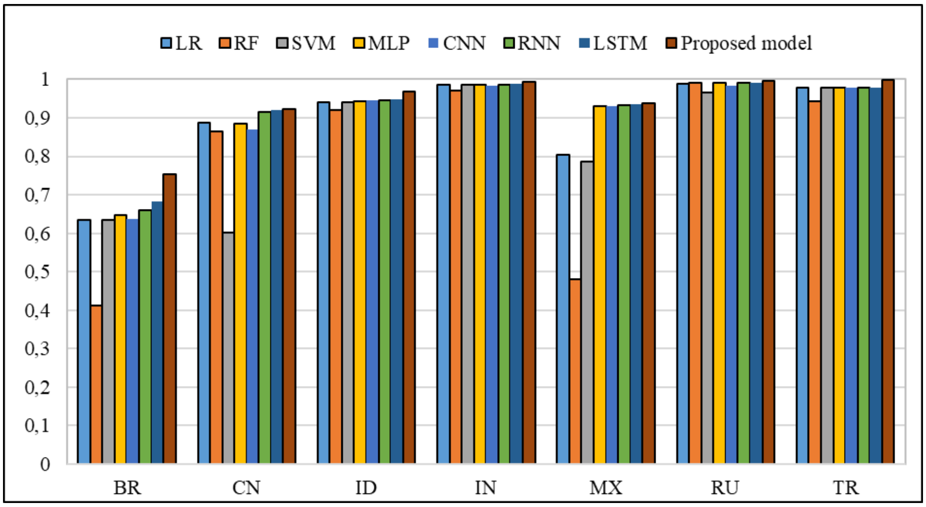

Figure 7. The experimental results according to the R^2 for predicting the number of cases in E7 countries

Table 4 shows the experimental results according to the RMSE for predicting the number of deaths in E7 countries.

Table 4. The experimental results according to the RMSE for predicting the number of deaths in E7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
BR	259.44	244.71	244.50	233.44	236.81	227.54	226.70	160.27
CN	58.69	47.91	33.73	19.78	24.30	19.94	18.78	17.14
ID	57.36	89.94	57.31	57.39	64.32	53.98	51.72	40.35
IN	359.86	351.70	315.30	315.41	315.77	314.88	314.10	253.19
MX	25.51	27.74	26.30	25.86	28.84	26.11	25.14	17.19
RU	23.16	198.45	24.47	22.42	33.43	22.32	22.25	21.90
TR	17.66	17.71	17.35	17.37	17.71	17.44	17.13	10.47

As seen in Table 4, the proposed model has a better prediction performance than the other models compared according to RMSE. After the proposed model, LSTM, SVM, MLP and RNN have more successful results than other models.

Table 5 shows the experimental results according to the MAE for predicting the number of deaths in E7 countries.

Table 5. The experimental results according to the MAE for predicting the number of deaths in e7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
BR	172.95	143.40	151.76	136.74	139.34	168.29	137.90	99.94
CN	29.95	20.85	17.36	9.06	12.83	8.71	8.25	7.38
ID	24.69	36.35	25.13	24.57	39.85	24.53	22.69	15.54
IN	141.88	155.46	114.73	114.01	113.80	111.76	113.32	82.01
MX	17.12	19.39	18.59	16.74	19.59	17.47	16.44	10.41
RU	16.68	121.73	18.80	16.47	27.07	16.17	15.92	14.55
TR	13.19	13.21	13.04	13.05	13.19	12.67	12.51	6.78

As seen in Table 5, the proposed model has a better prediction performance than the other models compared according to MAE. After the proposed model, LSTM, RNN, SVM and MLP have more successful results than other models.

Table 6 and Figure 8 show the experimental results according to the R^2 for predicting the number of deaths in E7 countries.

Table 6. The experimental results according to the R^2 for predicting the number of deaths in E7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
BR	0.227	0.312	0.313	0.374	0.356	0.405	0.409	0.769
CN	0.370	0.580	0.791	0.928	0.892	0.927	0.935	0.946
ID	0.942	0.858	0.942	0.942	0.927	0.948	0.953	0.978
IN	0.079	0.121	0.293	0.293	0.291	0.295	0.298	0.532
MX	0.987	0.985	0.986	0.987	0.984	0.987	0.988	0.996
RU	0.995	0.680	0.995	0.995	0.991	0.995	0.996	0.998
TR	0.957	0.957	0.959	0.959	0.957	0.959	0.961	0.980

As seen in Table 6 and Figure 8, the proposed model has a better prediction performance than the other models compared according to R^2. After the proposed model, LSTM, RNN, MLP and SVM have more successful results than other models.

Table 7 shows the experimental results according to the RMSE for predicting the number of cases in G7 countries.

Table 7. The experimental results according to the RMSE for predicting the number of cases in G7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
CA	3339.03	3653.73	3316.97	3107.10	3141.32	3139.19	3065.45	1950.03
DE	29989.56	32179.93	34851.16	26899.95	27406.85	27975.53	24859.17	24092.27
GB	15275.47	15216.66	16040.86	14930.27	15891.24	14233.54	13945.96	12302.62
FR	55219.45	55607.95	62482.15	44741.03	55368.11	47202.93	45266.79	41146.59
IT	26688.93	25747.49	25986.84	23459.90	23606.33	22466.88	21213.23	14444.19
JP	12410.57	10033.74	10240.28	8848.66	10008.75	8285.96	7880.62	6888.87
US	102963.20	102982.41	108270.15	95577.58	102002.46	95174.88	94707.92	92571.18

As seen in Table 7, the proposed model has a better prediction performance than the other models compared according to RMSE. After the proposed model, LSTM, RNN and MLP have more successful results than other models.

Table 8 shows the experimental results according to the MAE for predicting the number of cases in G7 countries.

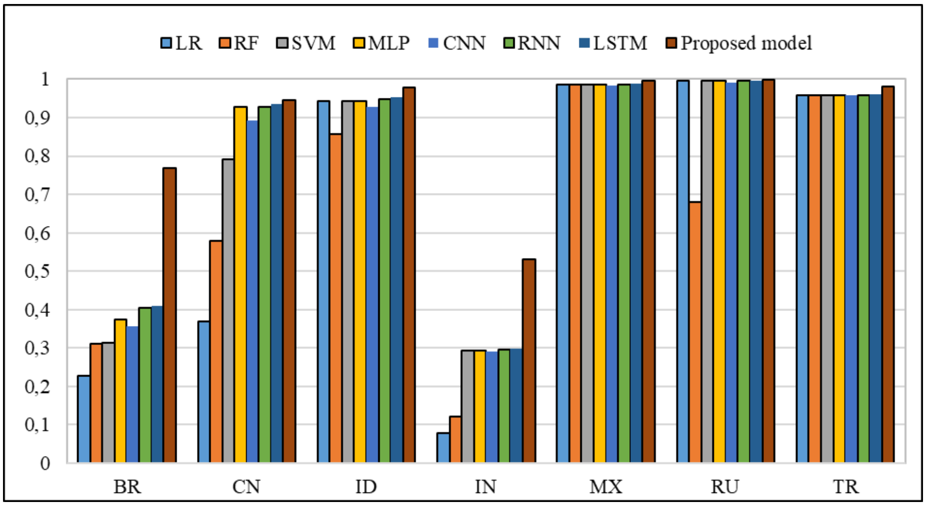

Figure 8. The experimental results according to the R^2 for predicting the number of cases in E7 countries

Table 8. The experimental results according to the MAE for predicting the number of cases in G7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
CA	1705.49	1864.68	1677.97	1582.26	1595.82	1611.13	1555.52	1103.46
DE	18859.71	19926.93	21835.54	14916.93	15855.05	15653.80	14435.26	14592.68
GB	8281.01	8127.54	8254.60	7896.22	7943.73	7758.94	7592.48	6996.61
FR	29069.05	29941.27	32222.20	21114.41	29301.26	21443.54	20432.93	20126.60
IT	14034.13	13250.54	12631.93	11701.93	11939.15	11839.78	10688.76	7393.03
JP	4849.84	4534.78	4996.07	3967.07	4605.53	4154.18	3841.77	2761.82
US	48875.60	48869.29	50468.84	47189.79	47653.17	45788.46	44499.82	42877.24

As seen in Table 8, the proposed model has a better prediction performance than the other models compared according to MAE. After the proposed model, LSTM, RNN, MLP and CNN have more successful results than other models.

Table 9 and Figure 9 show the experimental results according to the R^2 for predicting the number of cases in G7 countries.

Table 9. The experimental results according to the R2 for predicting the number of cases in G7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
CA	0.875	0.851	0.877	0.892	0.889	0.890	0.895	0.970
DE	0.856	0.835	0.806	0.884	0.880	0.875	0.901	0.907
GB	0.881	0.882	0.869	0.887	0.872	0.896	0.901	0.930
FR	0.702	0.699	0.619	0.804	0.701	0.781	0.800	0.835
IT	0.727	0.746	0.741	0.789	0.786	0.806	0.827	0.940
JP	0.808	0.876	0.869	0.902	0.876	0.914	0.922	0.949
US	0.745	0.746	0.718	0.780	0.750	0.782	0.784	0.794

As seen in Table 9 and Figure 9, the proposed model has a better prediction performance than the other models compared according to R^2. After the proposed model, LSTM, RNN, MLP and CNN have more successful results than other models.

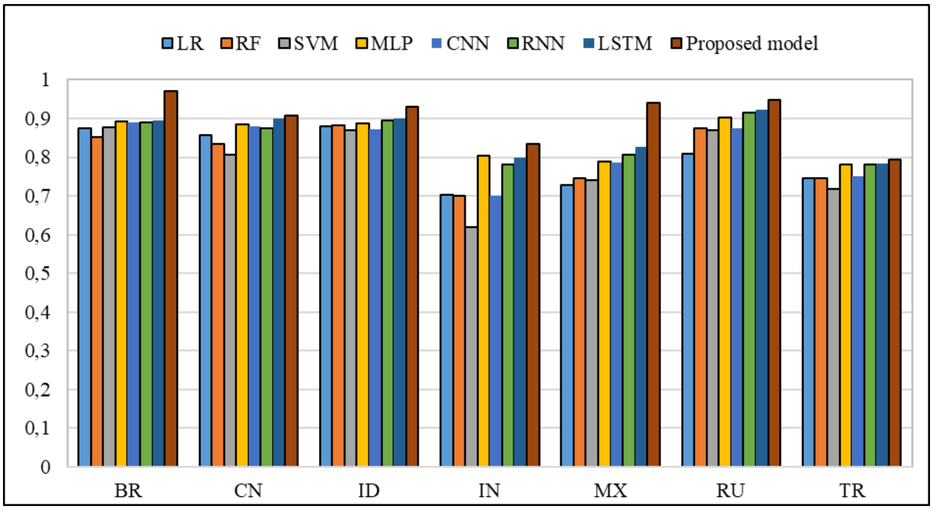

Figure 9. The experimental results according to the R^2 for predicting the number of cases in G7 countries

Table 10 shows the experimental results according to the RMSE for predicting the number of deaths in G7 countries.

Table 10. The experimental results according to the RMSE for predicting the number of deaths in G7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
CA	33.67	34.51	34.24	34.09	34.17	33.69	33.44	25.51
DE	61.67	60.40	63.57	51.48	65.01	49.49	47.71	38.55
FR	67.39	74.71	68.51	67.24	72.79	66.64	65.88	57.65
GB	19.14	18.57	19.35	18.71	18.92	18.62	18.55	11.01
IT	37.85	41.33	37.73	37.46	40.43	35.50	34.88	26.88
JP	30.87	46.34	31.08	28.27	29.78	27.90	27.71	22.93
US	722.28	818.05	733.04	722.84	724.87	721.08	712.65	605.12

As seen in Table 10, the proposed model has a better prediction performance than the other models compared according to RMSE. After the proposed model, LSTM, RNN, LR and MLP have more successful results than other models.

Table 11 shows the experimental results according to the MAE for predicting the number of deaths in G7 countries.

Table 11. The experimental results according to the MAE for predicting the number of deaths in G7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
CA	23.17	23.80	23.04	23.32	23.13	23.26	22.62	17.39
DE	43.74	43.96	43.04	37.16	48.55	35.16	33.66	29.01
FR	51.75	54.12	49.01	50.96	54.13	50.21	48.85	38.52
GB	15.29	14.59	15.50	14.75	15.02	14.85	14.57	8.05
IT	26.32	28.68	26.08	25.92	28.22	24.67	24.24	18.09
JP	15.34	22.82	14.90	13.68	14.02	13.74	13.97	11.52
US	513.10	628.67	502.76	513.06	514.05	502.15	496.77	451.98

As seen in Table 11, the proposed model has a better prediction performance than the other models compared according to MAE. After the proposed model, LSTM, RNN, SVM, LR and MLP have more successful results than other models.

Table 12 and Figure 10 show the experimental results according to the R^2 for predicting the number of deaths in G7 countries.

Table 12. The experimental results according to the R^2 for predicting the number of deaths in G7 countries

Code	LR	RF	SVM	MLP	CNN	RNN	LSTM	Proposed model
CA	0.443	0.415	0.424	0.429	0.426	0.443	0.451	0.704
DE	0.755	0.765	0.740	0.829	0.728	0.842	0.853	0.909
FR	0.493	0.376	0.475	0.495	0.408	0.504	0.515	0.598
GB	0.894	0.901	0.892	0.900	0.897	0.899	0.901	0.951
IT	0.861	0.834	0.861	0.863	0.841	0.877	0.882	0.946
JP	0.779	0.504	0.776	0.815	0.795	0.820	0.822	0.872
US	0.376	0.199	0.356	0.375	0.371	0.378	0.392	0.570

As seen in Table 12 and Figure 10, the proposed model has a better prediction performance than the other models compared according to R^2. After the proposed model, LSTM, RNN, LR, SVM and MLP have more successful results than other models.

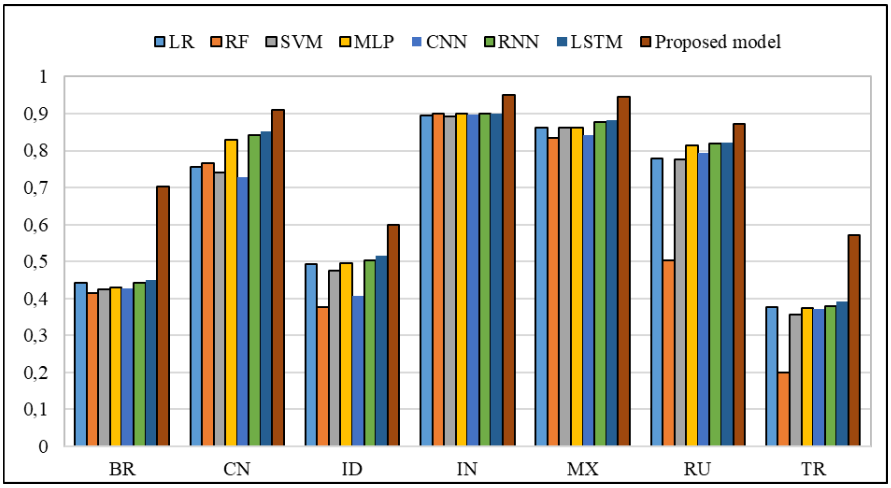

Figure 10. The experimental results according to the R^2 for predicting the number of deaths in G7 countries

The prediction graphs of the proposed model for E7 countries according to the number of cases and deaths are shown in Figure 11 and Figure 12.

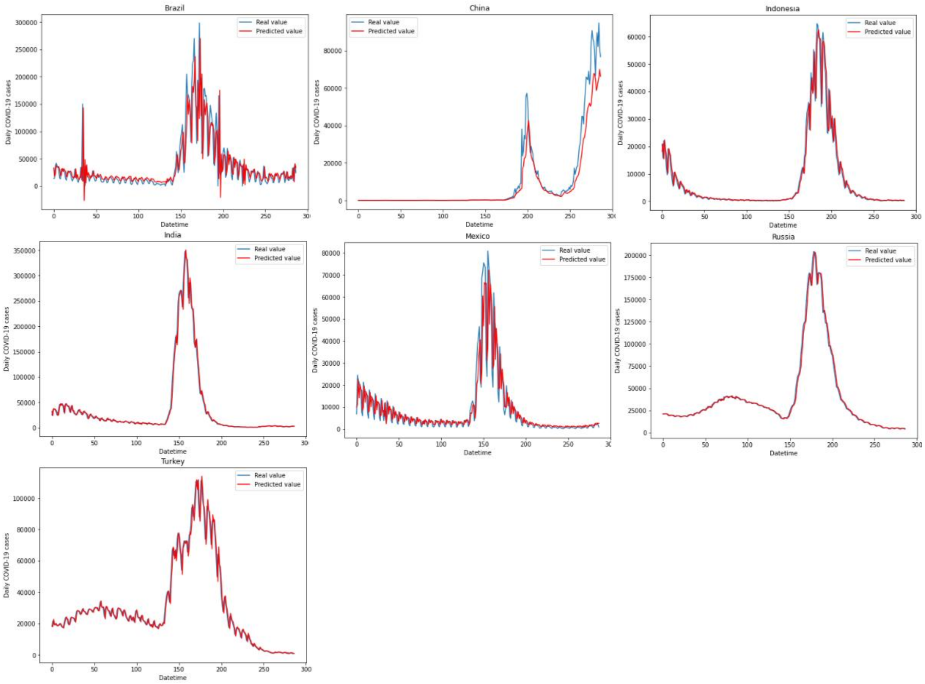

Figure 11. The prediction graphs of the proposed model for E7 countries according to the number of cases

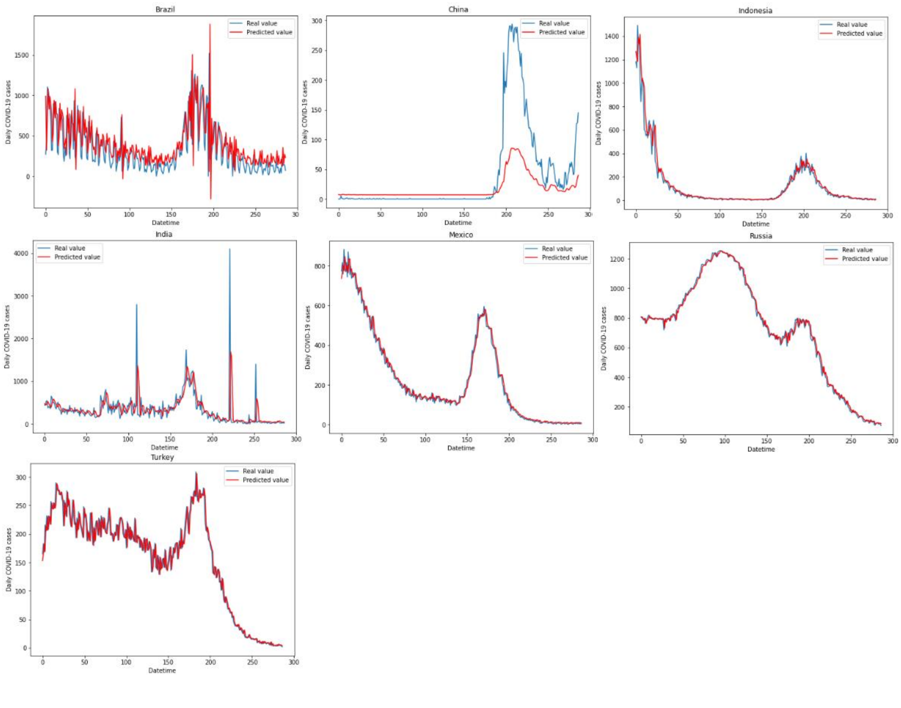

Figure 12. The prediction graphs of the proposed model for E7 countries according to the number of deaths

The prediction graphs of the proposed model for G7 countries according to the number of cases and deaths are shown in Figure 13 and Figure 14.

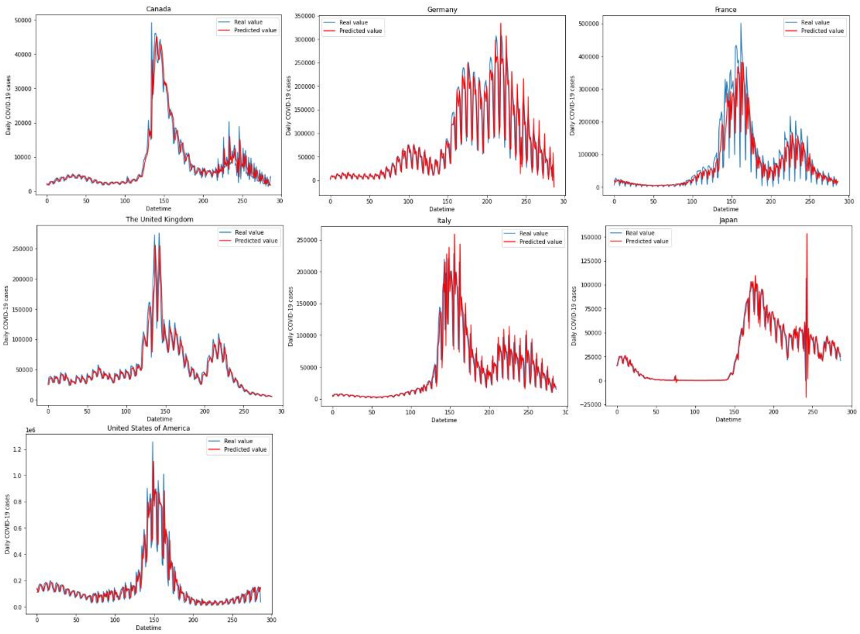

Figure 13. The prediction graphs of the proposed model for G7 countries according to the number of cases

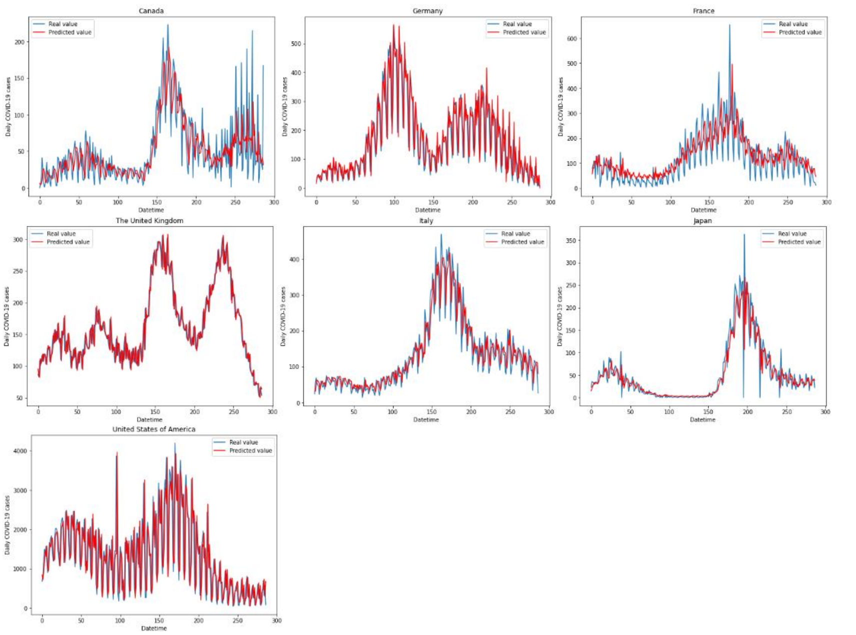

Figure 14. The prediction graphs of the proposed model for G7 countries according to the number of deaths

As seen in Figure 11-14, the proposed model has better results than the compared models for all error metrics and also the proposed model has successfully predicted the fluctuations in the number of cases and deaths.

4.4. Analysis of COVID-19 Cross-Country Spread in E7 and G7 Countries

Most people infected with COVID-19 commonly show symptoms of fever, cough, fatigue, and loss of taste or smell. It was found that these symptoms usually appear on the 5th day of the disease, but in different cases, they extend up to 14th day. For this reason, 5 days' and 14 days' incubation periods were used to analyze the spread of COVID-19 among E7 and G7 countries. For these analyses, peak dates of WHO-confirmed cases and deaths were determined for each country. In addition, chord diagrams were drawn for 5 days and 14 days the incubation periods to determine the spread of COVID-19 among the E7 and G7 countries. Chord diagrams provide a graphical representation of relationships between data. Relationships are shown with radial lines according to the frequency of relationships between data points. Chord diagrams can be drawn with the help of Python libraries or Websites. Chord diagrams used in this study were drawn using DataSmith (Data Smith, 2022).

The peak dates of the number of cases for E7 countries and 5 days' and 14 days' incubation period dates are shown in Table 13.

Table 13. The peak dates of the number of cases for E7 countries and incubation period dates

Code	14 days before peak date	5 days before peak date	Peak date	5 days after peak date	14 days after peak date
BR	2022/01/22-2022/02/04	2022/01/31-2022/02/04	2022/02/05	2022/02/06-2022/02/10	2022/02/06-2022/02/19
CN	2022/05/14-2022/05/27	2022/05/23-2022/05/27	2022/05/28	2022/05/29-2022/06/02	2022/05/29-2022/06/11
ID	2022/02/02-2022/02/15	2022/02/11-2022/02/15	2022/02/16	2022/02/17-2022/02/21	2022/02/17-2022/03/02
IN	2021/04/23-2021/05/06	2021/05/02-2021/05/06	2021/05/07	2021/05/08-2021/05/12	2021/05/08-2021/05/21
MX	2022/01/05-2022/01/18	2022/01/14-2022/01/18	2022/01/19	2022/01/20-2022/01/24	2022/01/20-2022/02/02
RU	2022/01/28-2022/02/10	2022/02/06-2022/02/10	2022/02/11	2022/02/12-2022/02/16	2022/02/12-2022/02/25
TR	2022/01/22-2022/02/04	2022/01/31-2022/02/04	2022/02/05	2022/02/06-2022/02/10	2022/02/06-2022/02/19

Figure 15 shows the spread of COVID-19 cases among E7 countries according to the 5-day incubation period.

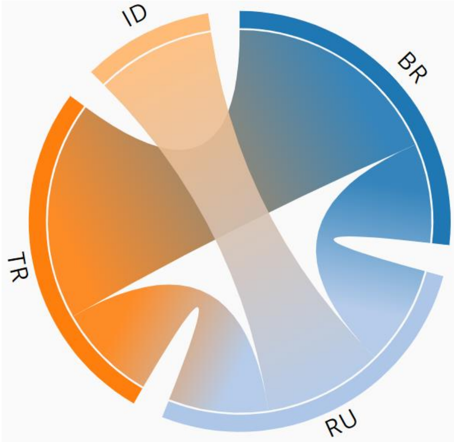

Figure 15. Cross-country spread of COVID-19 cases among E7 countries for 5-day incubation period

As it can be seen in Figure 15 and Table 13, for 5-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Brazil and Turkey, Brazil and Russia, Indonesia and Russia, Russia and Turkey.

Figure 16 shows the spread of COVID-19 cases among E7 countries according to the 14-day incubation period.

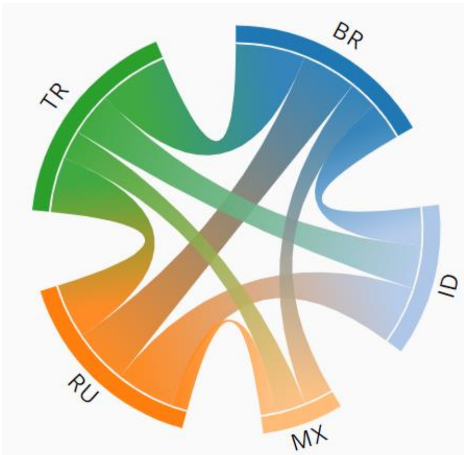

Figure 16. Cross-country spread of COVID-19 cases among E7 countries for 14-day incubation period

As it can be seen in Figure 16 and Table 13, for 14-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Brazil and Indonesia, Brazil and Mexico, Brazil and Russia, Brazil and Turkey, Brazil and Russia, Indonesia and Russia, Indonesia and Turkey, Mexico and Russia, Mexico and Turkey, Russia and Turkey.

The peak dates of the number of deaths for E7 countries and 5 days' and 14 days' incubation period dates are shown in Table 14.

Table 14. The peak dates of the number of deaths for E7 countries and incubation period dates

Code	14 days before peak date	5 days before peak date	Peak date	5 days after peak date	14 days after peak date
BR	2021/03/27-2021/04/09	2021/04/05-2021/04/09	2021/04/10	2021/04/11-2021/04/15	2021/04/11-2021/04/24
CN	2020/04/04-2020/04/17	2020/04/13-2020/04/17	2020/04/18	2021/04/19-2021/04/23	2021/04/19-2021/05/02
ID	2021/06/15-2021/06/28	2021/06/24-2021/06/28	2021/06/28	2021/06/29-2021/07/03	2021/06/29-2021/07/12
IN	2021/05/27-2021/06/09	2021/06/05-2021/06/09	2021/06/10	2021/06/11-2021/06/15	2021/06/11-2021/06/24
MX	2021/01/08-2021/01/21	2021/01/17-2021/01/21	2021/01/22	2021/01/23-2021/01/27	2021/01/23-2021/02/05
RU	2021/11/05-2021/11/18	2021/11/14-2021/11/18	2021/11/19	2021/11/20-2021/11/24	2021/11/20-2021/12/03
TR	2021/04/17-2021/04/30	2021/04/26-2021/04/30	2021/05/01	2021/05/02-2021/05/06	2021/05/02-2021/05/15

There is no similarity between the spread characteristic of COVID-19 deaths in E7 countries for the 5-day incubation period.

Figure 17 shows the spread of COVID-19 deaths among E7 countries according to the 14-day incubation period.

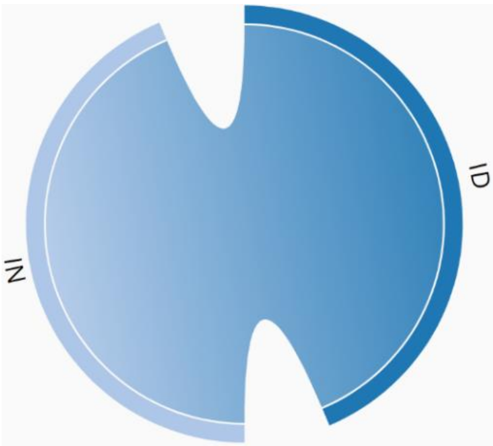

Figure 17. Cross-country spread of COVID-19 deaths among E7 countries for 14-day incubation period

As it can be seen in Figure 17 and Table 14, for 14-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Indonesia and India.

The peak dates of the number of cases for G7 countries and 5 days' and 14 days' incubation period dates are shown in Table 15.

Table 15. The peak dates of the number of cases for G7 countries and incubation period dates

Code	14 days before peak date	5 days before peak date	Peak date	5 days after peak date	14 days after peak date
CA	2021/12/15-2021/12/28	2021/12/24-2021/12/28	2021/12/29	2021/12/29-2022/01/02	2021/12/29-2022/01/11
DE	2022/03/10-2022/03/23	2022/03/19-2022/03/23	2022/03/24	2022/03/25-2022/03/29	2022/03/25-2022/04/08
FR	2022/01/12-2022/01/25	2022/01/21-2022/01/25	2022/01/26	2022/01/27-2022/01/31	2022/01/27-2022/02/09
GB	2021/12/24-2022/01/06	2022/01/02-2022/01/06	2022/01/06	2022/01/07-2022/01/11	2022/01/07-2022/01/20
IT	2022/01/05-2022/01/18	2022/01/14-2022/01/18	2022/01/19	2022/01/20-2022/01/24	2022/01/20-2022/02/02
JP	2022/04/02-2022/04/15	2022/04/11-2022/04/15	2022/04/16	2022/04/17-2022/04/21	2022/04/17-2022/04/30
US	2021/12/29-2022/01/11	2022/01/07-2022/01/11	2022/01/12	2022/01/13-2022/01/17	2022/01/13-2022/01/26

Figure 18 shows the spread of COVID-19 cases among G7 countries according to the 5-day incubation period.

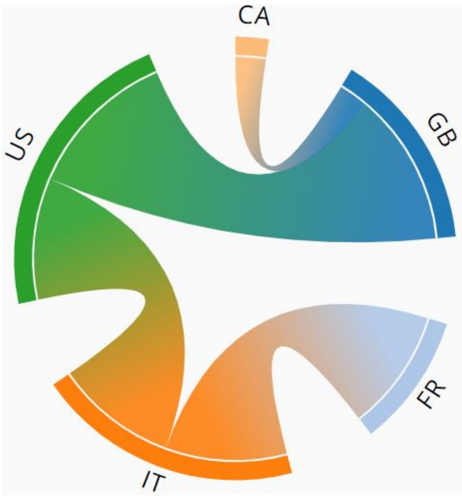

Figure 18. Cross-country spread of COVID-19 cases among G7 countries for 5-day incubation period

As it can be seen in Figure 18 and Table 15, for 5-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Canada and The United Kingdom, France and Italia, The United Kingdom and United States of America, Italia and United States of America.

Figure 19 shows the spread of COVID-19 cases among G7 countries according to the 14-day incubation period.

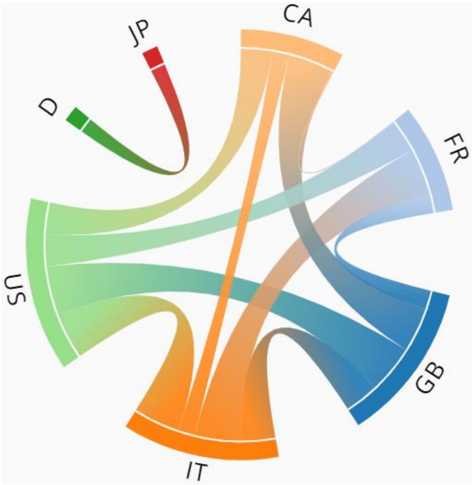

Figure 19. Cross-country spread of COVID-19 cases among G7 countries for 14-day incubation period

As it can be seen in Figure 19 and Table 15, for 14-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Canada and France, Canada and The United Kingdom, Canada and Italia, Canada and United States of America, Germany and Japan, France and The United Kingdom, France and Italy, France and United States of America, The United Kingdom and Italy, The United Kingdom and United States of America, Italy and United States of America.

The peak dates of the number of deaths for G7 countries and 5 days' and 14 days' incubation period dates are shown in Table 16.

Table 16. The peak dates of the number of deaths for G7 countries and incubation period dates

Code	14 days before peak date	5 days before peak date	Peak date	5 days after peak date	14 days after peak date
CA	2022/01/14-2022/01/27	2022/01/23-2022/01/27	2022/01/28	2022/01/29-2022/02/02	2022/01/29-2022/02/11
DE	2020/12/17-2020/12/30	2020/12/26-2020/12/30	2020/12/31	2021/01/01-2021/01/05	2021/01/01-2021/01/14
FR	2020/03/21-2020/04/03	2020/03/30-2020/04/03	2020/04/04	2020/04/05-2020/04/09	2020/04/05-2020/04/18
GB	2021/01/07-2021/01/20	2021/01/16-2021/01/20	2021/01/21	2021/01/22-2021/01/26	2021/01/22-2021/02/04
IT	2020/11/20-2020/12/03	2020/11/29-2020/12/03	2020/12/04	2020/12/05-2020/12/09	2020/12/05-2020/12/18
JP	2022/02/15-2022/02/28	2022/02/24-2022/02/28	2022/03/01	2022/03/02-2022/03/06	2022/03/02-2022/03/15
US	2021/01/31-2021/02/13	2021/02/09-2021/02/13	2021/02/14	2021/02/15-2021/02/19	2021/02/15-2021/02/28

There is no similarity between the spread characteristic of COVID-19 deaths in G7 countries for the 5-day incubation period. Figure 20 shows the spread of COVID-19 deaths among G7 countries according to the 14-day incubation period.

Figure 20. Cross-country spread of COVID-19 deaths among G7 countries for 14-day incubation period

As it can be seen in Figure 20 and Table 16, for 14-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Germany and Italy, The United Kingdom and United States of America.

In addition, 5 days' and 14 days' chord diagrams of incubation periods were drawn to determine the spread of COVID-19 between E7 and G7 countries.

Figure 21 shows the spread of COVID-19 cases among between E7 and G7 countries according to the 5-day incubation period.

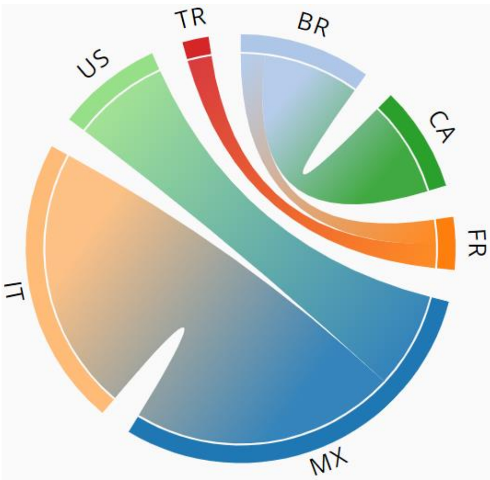

Figure 21. Cross-country spread of COVID-19 cases among between E7 and G7 countries for 5-day incubation period

As it can be seen in Figure 21, for 5-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Brazil and Canada, Brazil and France, Mexico and Italia, Mexico and United States of America, Turkey and France.

Figure 22 shows the spread of COVID-19 cases among between E7 and G7 countries according to the 14-day incubation period.

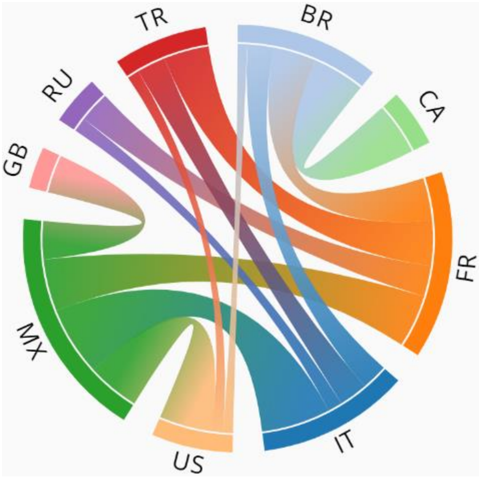

Figure 22. Cross-country spread of COVID-19 cases among between E7 and G7 countries for 14-day incubation period

As it can be seen in Figure 22, for 14-day incubation period, the spread characteristic of COVID-19 cases is quite similar between Brazil and Canada, Brazil and France, Brazil and Italia, Brazil and United States of America, Mexico and France, Mexico and The United Kingdom, Mexico and Italia, Mexico and United States of America, Russia and France, Russia and Italia, Turkey and France, Turkey and Italia, Turkey and United States of America.

Figure 23 shows the spread of COVID-19 deaths among between E7 and G7 countries according to the 5-day incubation period.

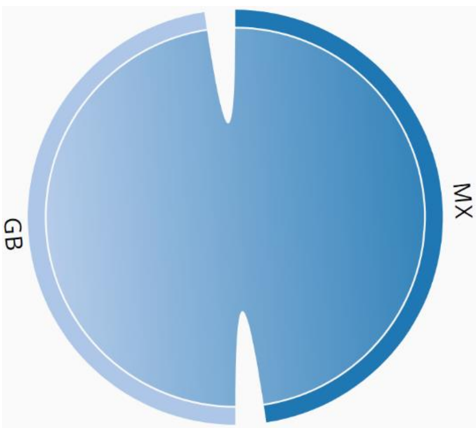

Figure 23. Cross-country spread of COVID-19 deaths among between E7 and G7 countries for 5-day incubation period

As it can be seen in Figure 23, for 5-day incubation period, the spread characteristic of COVID-19 deaths is quite similar between Mexico and The United Kingdom.

Figure 24 shows the spread of COVID-19 deaths among between E7 and G7 countries according to the 14-day incubation period.

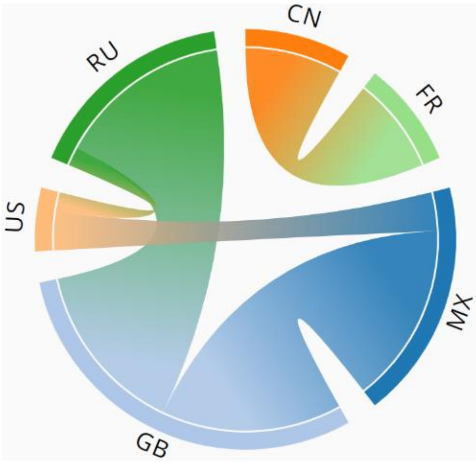

Figure 24. Cross-country spread of COVID-19 deaths among between E7 and G7 countries for 14-day incubation period

As it can be seen in Figure 24, for 14-day incubation period, the spread characteristic of COVID-19 deaths is quite similar between China and France, Mexico and The United Kingdom, Mexico and United States of America, Russia and The United Kingdom, Russia and United States of America.

Especially in E7 countries such as Brazil and India, the number of daily cases and deaths is also high, as the population density is high. In the remaining countries, the number of cases and deaths has increased within certain limits due to COVID-19 measures and lower population density. For this reason, the prediction results of the models were more consistent.

The prediction results of the G7 countries showed that the models did not have high success rates, especially in countries with a high population and case density. The reason for this situation is the high case density and indirectly the number of deaths. The preventive measures taken by countries against COVID-19 have affected case densities and prediction success.

Experimental results showed that SVM has better prediction performance than RF. RF works with a mixture of numerical and categorical features. RF is advantageous when features are of various scales. This allows RF to use the data as they are. SVM, on the other hand, maximizes the margin between different points and calculates the distance between points. Since the dataset used is a time series dataset, SVM was more successful than RF.

The fact that MLP, which is one of the neural network models, is more successful than RF and SVM, can be considered as RF SVM's more successful operation on tabular data such as voice, image and text data. Neural network models require scaling of features. However, features of greater importance will be considered more important in education. In this way, neurons will ensure that the training phase is more efficient.

The fact that LR is more successful than other machine learning models has to do with the characteristic of time series data. SVM and RF work better on categorical data. LR makes predictions with the help of a function that models the relationship between data points.

CNN's prediction performance close to MLP can be interpreted as both models are neural network architectures. MLP takes vectors as input and CNN takes tensor as input.

The fact that RNN is more successful than LR, RF, SVM, MLP and CNN can be interpreted as CNN and RNN having different architectures. CNNs are feedforward neural networks that use filters and pooling layers. RNN feeds the results back to the network. In CNN, the size of the input and the output are fixed. That is, a CNN takes inputs of a fixed size and scales them to the appropriate level along with the confidence level of its prediction. In RNN, the size of the input and the resulting output can vary. The feedback structure of RNN enabled the past features to be remembered and presented to the network as an input again, thus providing a more successful result.

The success of LSTM over other models compared is that LSTM's architecture includes special units in addition to the standard units found in the RNN. LSTM units contain a memory cell that can hold information in memory for long periods of time. A set of gates is used to control when information enters memory, when it leaves, and when it is forgotten. This architecture enables learning of longer term dependencies.

The most successful prediction performance of the proposed hybrid model can be interpreted by using the prominent features of CNN and LSTM together. The proposed model takes advantage of the success of CNN in the feature extraction stage and the success of LSTM in the learning and prediction stage.

5. Conclusions

Artificial intelligence is widely used in the field of health as well as in many aspects of daily life. Artificial intelligence can help clinicians adopt a more comprehensive approach to disease management and improve patient compliance by providing better coordination. By using artificial intelligence methods, measures can be taken for the spread of the disease by predicting the number of daily COVID-19 positive cases and deaths.

In this study, CNN-RNN based hybrid deep learning model was proposed to predict the number of cases and deaths and analyze their cross-country spread in E7 and G7 countries. The proposed model was extensively compared with LR, RF, SVM, MLP, CNN, RNN and LSTM. The dataset includes WHO-confirmed cases and deaths up to May 31, 2022. The compared models were evaluated using RMSE, MAE, and R^2. The experimental results showed that the developed LSTM-based model has better prediction performance than other according to RMSE, MAE and R^2 metrics.

WHO has found that the symptoms of COVID-19 usually appear on day 5 of illness but persist up to day 14. For this reason, incubation periods of 5 days and 14 days were used for the analysis of the spread of COVID-19 among E7 and G7 countries. Chord diagrams were drawn by making analyses for 5 days and 14 days by determining the peak number of cases and deaths in the countries. Experimental results show that the spread characteristic of COVID-19 cases is quite similar between Brazil and Indonesia, Brazil and Mexico, Brazil and Russia, Brazil and Turkey, Brazil and Russia, Indonesia and Russia, Indonesia and Turkey, Mexico and

Russia, Mexico and Turkey, Russia and Turkey, Canada and France, Canada and The United Kingdom, Canada and Italia, Canada and United States of America, Germany and Japan, France and The United Kingdom, France and Italy, France and United States of America, The United Kingdom and Italy, The United Kingdom and United States of America, Italy and United States of America, Brazil and Canada, Brazil and France, Brazil and Italia, Brazil and United States of America, Mexico and France, Mexico and The United Kingdom, Mexico and Italia, Mexico and United States of America, Russia and France, Russia and Italia, Turkey and France, Turkey and Italia, Turkey and United States of America.

The spread characteristic of COVID-19 deaths is quite similar between Canada and The United Kingdom, France and Italia, The United Kingdom and United States of America, Italia and United States of America, Germany and Italy, The United Kingdom and United States of America, China and France, Mexico and The United Kingdom, Mexico and United States of America, Russia and The United Kingdom, Russia and United States of America.

In the future works, it is planned to apply different hybrid deep learning models for more countries. In this way, it is aimed to predict the number of cases and deaths of countries with lower error rates, and thus to model the spread of COVID-19 between countries more effectively.

Author Contributions: Conceptualization, A.U.; methodology, A.U.; validation, A.U.; formal analysis, A.U.; investigation, A.U.; resources, A.U.; writing—original draft preparation, A.U.; writing—review and editing, A.U.; visualization, A.U.; supervision, A.U.; The author has read and agreed to the published version of the manuscript.

Funding*:* This research received no external funding.

Data Availability Statement*:* The dataset used is accessible via https://covid19.who.int/WHO-COVID-19-global-data.csv.

Conflicts of Interest*:* The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abbasimehr, H., & Paki, R. (2021). Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. Chaos, Solitons & Fractals, 142. doi:https://doi.org/10.1016/j.chaos.2020.110511

Ahmad, M., Sadiq, S., Alluhaidan, A., & Umer, M. (2022). Industry 4.0 technologies and their applications in fighting COVID-19 pandemic using deep learning techniques. Computers in Biology and Medicine, 145. doi:https://doi.org/10.1016/j.dsx.2020.04.032

Alakuş, T., & Türkoğlu, İ. (2020). Comparison of deep learning approaches to predict COVID-19 infection. Chaos, Solitons & Fractals, 140. doi:https://doi.org/10.1016/j.chaos.2020.110120

Alassafi, M., Jarrah, M., & Alotaibi, R. (2022). Time series predicting of COVID-19 based on deep learning. Neurocomputing, 468, 335-344. doi:https://doi.org/10.1016/j.neucom.2021.10.035

Ali, M., Prasad, R., Xiang, Y., & Deo, R. (2020). Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms. Renewable and Sustainable Energy Reviews, 132. doi:https://doi.org/10.1016/j.rser.2020.110003

Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J., Moosaei, H., & Mosavi, A. (2021). Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results in Physics, 27. doi:https://doi.org/10.1016/j.rinp.2021.104495

Ballabio, C., & Sterlacchini, S. (2012). Support vector machines for landslide susceptibility mapping: The Staffora River Basin case study, Italy. Mathematical geosciences, 44(1), 47-70. doi:https://doi.org/10.1007/s11004-011-9379-9

Basha, S., Dubey, S., Pulabaigari, V., & Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. Neurocomputing, 378, 112-119. doi:https://doi.org/10.1016/j.neucom.2019.10.008

Bikku, T. (2020). Multi-layered deep learning perceptron approach for health risk prediction. Journal of Big Data, 7(1), 1-14. doi:https://doi.org/10.1186/s40537-020-00316-7

Chamasemani, F., & Singh , Y. (2011). Multi-class support vector machine (SVM) classifiers--an application in hypothyroid detection and classification. 2011 sixth international conference on bio-inspired computing: theories and applications, (s. 351-356). doi:https://doi.org/10.1109/BIC-TA.2011.51

Che Azemin, M., Hassan, R., & Mohd Tamrin, M. (2020). COVID-19 deep learning prediction model using publicly available radiologist-adjudicated chest X-ray images as training data: preliminary findings. International Journal of Biomedical Imaging. doi:https://doi.org/10.1155/2020/8828855

Chen, X., Liu, X., Gales, M., & Woodland, P. (2015). Recurrent neural network language model training with noise contrastive estimation for speech recognition. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). doi:https://doi.org/10.1109/ICASSP.2015.7179005

Data Smith. (2022, Novamber 06). Experiments in data visualization: http://www.datasmith.org/2018/06/02/a-bold-chord-diagram-generator/ adresinden alındı

Devaraj, J., Elavarasan, R., Pugazhendhi, R., Shafiullah, G., Ganesan, S., Jeysree, A., & Hossain, E. (2021). Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? Results in Physics, 21. doi:https://doi.org/10.1016/j.rinp.2021.103817

El-Khaiary, M. (2008). Least-squares regression of adsorption equilibrium data: comparing the options. Journal of Hazardous Materials, 158(1), 73-87. doi:https://doi.org/10.1016/j.jhazmat.2008.01.052

Elsheikh, A., Saba, A., Abd Elaziz, M., Lu, S., Shanmugan, S., Muthuramalingam, T., & Shehabeldeen, T. (2021). Deep learning-based forecasting model for COVID-19 outbreak in Saudi Arabia. Process Safety and Environmental Protection, 149, 223-233. doi:https://doi.org/10.1016/j.psep.2020.10.048

Fang, L., Karakiulakis, G., & Roth, M. (2020). Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? The lancet respiratory medicine, 8(4), e21. doi:https://doi.org/10.1016/S2213-2600(20)30116-8

Gao, Y., & Glowacka, D. (2016). Deep gate recurrent neural network. In Asian conference on machine learning, 63, s. 350-365.

Giambartolomei, C., Zhenli, L., Zhang, W., Hauberg, M., Shi, H., Boocock, J., & Roussos, P. (2018). A Bayesian framework for multiple trait colocalization from summary association statistics. Bioinformatics, 2538-2545. doi:https://doi.org/10.1093/bioinformatics/bty147

Hari, S., Sullivan, M., Tsai, T., & Keckler, S. (2021). Making convolutions resilient via algorithm-based error detection techniques. IEEE Transactions on Dependable and Secure Computing, 19(4), 2546-2558. doi:https://doi.org/10.1109/TDSC.2021.3063083

Henrikson, N., Opel, D., Grothaus, L., Nelson, J., Scrol, A., Dunn, J., & Grossman, D. (2015). Physician communication training and parental vaccine hesitancy: a randomized trial. Pediatrics, 136(1), 70-79. doi:https://doi.org/10.1542/peds.2014-3199

Huang, J., Li, J., Yu, D., Deng, L., & Gong, Y. (2013). Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, (s. 7304-7308). doi:https://doi.org/10.1109/ICASSP.2013.6639081

Jarag, P., Kengar, M., Jadhav, R., Shinde, A., Koli, S., & Honmane, P. (2020). On overview-Corona virus and Hanta virus Disease. Asian Journal of Research in Pharmaceutical Science, 10(3), 178-182.

Jernigan, D. (2020). Update: public health response to the coronavirus disease 2019 outbreak—United States, February 24, 2020. Morbidity and mortality weekly report.

Jozefowicz, R., Zaremba, W., & Sutskever, I. (2015). An empirical exploration of recurrent network architectures. International conference on machine learning, (s. 2342-2350). doi:https://doi.org/10.5555/3045118.3045367

Ketu, S., & Mishra, P. (2022). India perspective: CNN-LSTM hybrid deep learning model-based COVID-19 prediction and current status of medical resource availability. Soft Computing, 26(2), 645-664. doi:https://doi.org/10.1007/s00500-021-06490-x

Kırbaş, İ., Sözen, A., Tuncer, A., & Kazancıoğlu, F. (2020). Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. Chaos, Solitons & Fractals, 138. doi:https://doi.org/10.1016/j.chaos.2020.110015

Klompas, M., Milton, D., Rhee, C., Baker, M., & Leekha, S. (2021). Current insights into respiratory virus transmission and potential implications for infection control programs: a narrative review. Annals of Internal Medicine, 174(12), 1710-1718. doi:https://doi.org/10.7326/M21-2780

Lee, B., Yang, C., & Huang, B. (2012). Oil price movements and stock markets revisited: A case of sector stock price indexes in the G-7 countries. Energy Economics, 34(5), 1284-1300. doi:https://doi.org/10.1016/j.eneco.2012.06.004

Mahmoudi, J., Arjomand, M., Rezaei, M., & Mohammadi, M. (2016). Predicting the earthquake magnitude using the multilayer perceptron neural network with two hidden layers. Civil engineering journal, 2(1), 1-12. doi:https://doi.org/10.28991/cej-2016-00000008

Mamun, M., & Ullah, I. (2020). COVID-19 suicides in Pakistan, dying off not COVID-19 fear but poverty?–The forthcoming economic challenges for a developing country. Brain, behavior, and immunity, 87, 163-166. doi:https://doi.org/10.1016/j.bbi.2020.05.028

Marzouk, M., Elshaboury, N., Abdel-Latif, A., & Azab, S. (2021). Deep learning model for forecasting COVID-19 outbreak in Egypt. Process Safety and Environmental Protection, 153, 363-375. doi:https://doi.org/10.1016/j.psep.2021.07.034

Masum, A., Khushbu, S., Keya, M., Abujar, S., & Hossain, S. (2020). COVID-19 in Bangladesh: a deeper outlook into the forecast with prediction of upcoming per day cases using time series. Procedia Computer Science, 178, 291-300. doi:https://doi.org/10.1016/j.procs.2020.11.031

Minotti, C., Tirelli, F., Barbieri, E., Giaquinto, C., & Donà, D. (2020). How is immunosuppressive status affecting children and adults in SARS-CoV-2 infection? A systematic review. Journal of Infection, 81(1), 61-66. doi:https://doi.org/10.1016/j.jinf.2020.04.026

Mirzaei, S., Kang, J., & Chu, K. (2022). A comparative study on long short-term memory and gated recurrent unit neural networks in fault diagnosis for chemical processes using visualization. Journal of the Taiwan Institute of Chemical Engineers, 130. doi:https://doi.org/10.1016/j.jtice.2021.08.016

Mize, T. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. Sociological Science, 6, 81-117. doi:https://doi.org/10.15195/v6.a4

Naghibi, S., Pourghasemi, H., & Dixon, B. (2016). GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environmental monitoring and assessment, 188(1), 1-27. doi:https://doi.org/10.1007/s10661-015-5049-6

Ramadevi, R., Sheela Rani, B., & Prakash, V. (2012). Role of hidden neurons in an elman recurrent neural network in classification of cavitation signals. International Journal of Computer Applications, 37(7), 9-13. doi:https://doi.org/10.5120/4618-6626

Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural engineering. Engineering structures, 171, 170-189. doi:https://doi.org/10.1016/j.engstruct.2018.05.084

Samra, M., Abed, B., Zaqout, H., & Abu-Naser, S. (2020). ANN Model for Predicting Protein Localization Sites in Cells. International Journal of Academic and Applied Research (IJAAR), 4(9), 43-50.

Satu, M., Howlader, K., Mahmud, M., Kaiser, M., Shariful Islam, S., Quinn, J., & Moni, M. (2021). Short-term prediction of COVID-19 cases using machine learning models. Applied Sciences, 11(9). doi:https://doi.org/10.3390/app11094266

Schijven, J., Vermeulen, L., Swart, A., Meijer, A., Duizer, E., & de Roda Husman, A. (2021). Quantitative microbial risk assessment for airborne transmission of SARS-CoV-2 via breathing, speaking, singing, coughing, and sneezing. Environmental health perspectives, 129(4). doi:https://doi.org/10.1289/EHP7886

Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos, Solitons & Fractals, 140. doi:https://doi.org/10.1016/j.chaos.2020.110212

Shastri, S., Singh, K., Kumar, S., Kour, P., & Mansotra, V. (2020). Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. Chaos, Solitons & Fractals, 140. doi:https://doi.org/10.1016/j.chaos.2020.110227

Speiser, J., Miller, M., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. Expert systems with applications, 134, 93-101. doi:https://doi.org/10.1016/j.eswa.2019.05.028

Sun, H., Burton, H., & Huang, H. (2021). Machine learning applications for building structural design and performance assessment: State-of-the-art review. Journal of Building Engineering, 33. doi:https://doi.org/10.1016/j.jobe.2020.101816

Tiwari, P., Pant, B., Elarabawy, M., Abd-Elnaby, M., Mohd, N., Dhiman, G., & Sharma, S. (2022). Cnn based multiclass brain tumor detection using medical imaging. Computational Intelligence and Neuroscience. doi:https://doi.org/10.1155/2022/1830010

Toğaçar, M., Ergen, B., & Cömert, Z. (2020). 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, 121. doi:https://doi.org/10.1016/j.compbiomed.2020.103805

Tong, T., Ortiz, J., Xu, C., & Li, F. (2020). Economic growth, energy consumption, and carbon dioxide emissions in the E7 countries: a bootstrap ARDL bound test. Energy, Sustainability and Society, 10(1), 1-17. doi:https://doi. org/10.1186/s13705-020-00253-6

Verma, H., Mandal, S., & Gupta, A. (2022). Temporal deep learning architecture for prediction of COVID-19 cases in India. Expert Systems with Applications, 195. doi:https://doi.org/10.1016/j.eswa.2022.116611

Wang, S., Lv, Y., Sui, Y., Liu, S., Wang, S., & Zhang, Y. (2018). Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling. Journal of medical systems, 42(1), 1-11. doi:https://doi.org/10.1007/s10916-017-0845-x

Xu, X., & Zhu, D. (2021). New method for solving Ivanov regularization-based support vector machine learning. Computers & Operations Research, 136. doi:https://doi.org/10.1016/j.cor.2021.105504

Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. Chaos, Solitons & Fractals, 140. doi:https://doi.org/10.1016/j.chaos.2020.110121