Chaos, Solitons and Fractals 140 (2020) 110214

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran

Peipei Wang[a], Xinqi Zheng[a,b,*], Gang Ai[a,*], Dongya Liu[a], Bangren Zhu[a]

[a] School of Information Engineering, China University of Geosciences, Beijing, China
[b] Technology Innovation Center for Territory Spatial Big-data, MNR of China, Beijing, China

ARTICLE INFO

Article history:
Received 13 July 2020
Accepted 17 August 2020
Available online 19 August 2020

Keywords:
Covid-19
LSTM
Rolling update mechanism
Modeling
Forecasting

ABSTRACT

The COVID-19 outbreak in late December 2019 is still spreading rapidly in many countries and regions around the world. It is thus urgent to predict the development and spread of the epidemic. In this paper, we have developed a forecasting model of COVID-19 by using a deep learning method with rolling update mechanism based on the epidemical data provided by Johns Hopkins University. First, as traditional epidemical models use the accumulative confirmed cases for training, it can only predict a rising trend of the epidemic and cannot predict when the epidemic will decline or end, an improved model is built based on long short-term memory (LSTM) with daily confirmed cases training set. Second, considering the existing forecasting model based on LSTM can only predict the epidemic trend within the next 30 days accurately, the rolling update mechanism is embedded with LSTM for long-term projections. Third, by introducing Diffusion Index (DI), the effectiveness of preventive measures like social isolation and lockdown on the spread of COVID-19 is analyzed in our novel research. The trends of the epidemic in 150 days ahead are modeled for Russia, Peru and Iran, three countries on different continents. Under our estimation, the current epidemic in Peru is predicted to continue until November 2020. The number of positive cases per day in Iran is expected to fall below 1000 by mid-November, with a gradual downward trend expected after several smaller peaks from July to September, while there will still be more than 2000 increase by early December in Russia. Moreover, our study highlights the importance of preventive measures which have been taken by the government, which shows that the strict controlment can significantly reduce the spread of COVID-19.

1. Introduction

Since the start of 2020, the COVID-19 epidemic is sweeping the world, infecting more than 12 million people and 569,000 deaths in more than 100 countries, posing a huge threat to human health [1]. The world health organization (WHO) declared an international emergency on 31 January, 2020. Unlike normal influenza, COVID-19 has a high R0 value (the basic reproduction number, representing viral infectivity) of 3.25–3.4 [2,3], characterized by strong human-to-human transmission through the air, which means the outbreak is not easily controlled. As is shown in Fig. 1, breathing and close contact are the main channel of COVID-19 transmission, while controlling the source of infection, cutting off the route of transmission and protection of vulnerable groups are the key to prevent the spreading of COVID-19. As a result, many countries have adopted unprecedented nationwide interventions, such as keep social distance, a curfew, closure of schools and businesses, bans on travel and others to prevent the further spread of the epidemic [4,5]. Thanks to the huge efforts of the government and people, these measures have yielded initial results in many countries, and the epidemic trends in these countries are approaching the middle and late stage.

Given the severity of the outbreak, predicting when the epidemic will end is particularly important for the production and life of affected countries. Many scholars have made efforts in the fight against the epidemic, and a number of predictive models based on mathematical model, infectious disease model and machine learning model are designed to forecast the trend of COVID-19 [6,7]. Since COVID-19 is time series data with daily records, the application of a sequence network extraction pattern is strongly recommended. Mathematical and machine learning technics has been quite successful and widely used in time series prediction, such as

* Corresponding authors at: School of Information Engineering, China University of Geosciences, No. 29 Xueyuan Road, Haidian District, Beijing, China.
E-mail addresses: zhengxq@cugb.edu.cn (X. Zheng), aigang@cugb.edu.cn (G. Ai).

https://doi.org/10.1016/j.chaos.2020.110214
0960-0779/

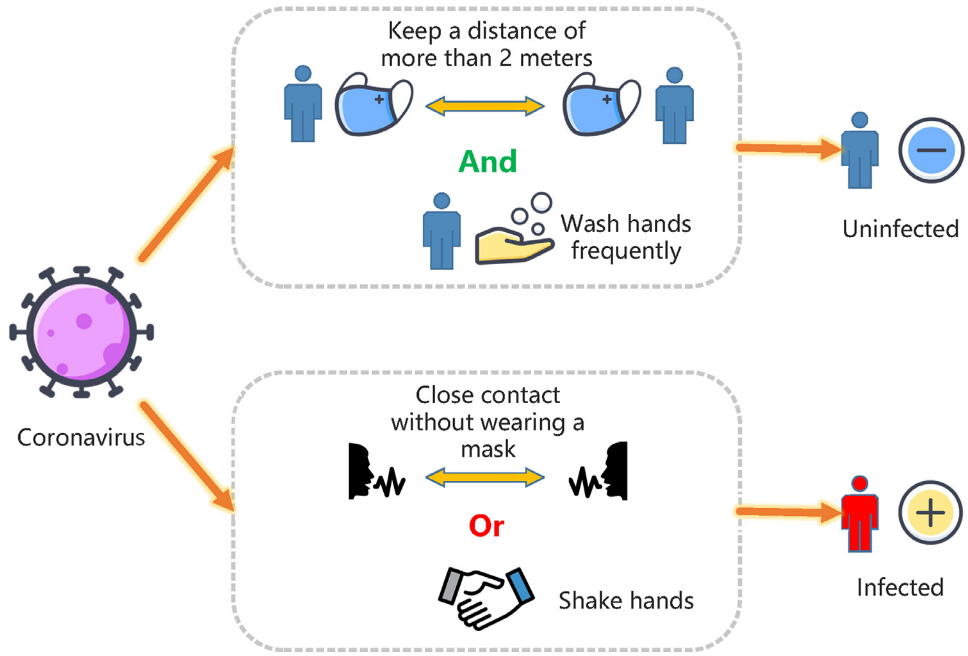

Fig. 1. Transmission of COVID-19.

electricity price prediction [8], stock price prediction [9] and air pollution forecast [10]. Lately, time series prediction methods such as Auto-Regressive Integrated Moving Average (ARIMA), Nonlinear Autoregression Neural Network (NARNN) and Long-Short Term Memory (LSTM) approaches are used to model the prediction of epidemic diseases. According to the results of this survey [11], LSTM was found the most accurate model.

However, there are some shortcomings in the existing LSTM prediction literatures. For example, in order to obtain the long-term prediction curve, Yang et al. [12] train the LSTM model based on the SARS data in 2003, since the time series data of SARS are complete. However, there are so many differences between SARS and COVID-19 in terms of latency rate, latency period, mortality rate and how developed the world is at the time of the outbreak that simulations can be unreliable. Some scholars have also designed prediction models with LSTM, using the cumulative number of confirmed cases as training set to train the model, however, it can only predict the rising trend of the epidemic within the next 30 days, and cannot predict when the epidemic will decline or end [13,14]. Since the cumulative confirmed data shows an overall upward trend and is a non-smooth sequence (Fig. 5), the LSTM is not very applicable with it. That is why those who used LSTM to make predictions only made short-term predictions, as the long-term predictions would continue to grow and the cap value could not be estimated, which was obviously illogical.

To deal with the aforementioned limitations, this study focuses on the modeling of the daily confirmed number of cases (a smooth sequence) in COVID-19 infection by LSTM method. Moreover, traditional forecasting methods mostly use current data such as historical confirmed cases to build models, while the real-time update of data is not considered. However, the existing predictions could play a greater role in the accuracy of the next prediction. Therefore, the rolling update mechanism is applied in this study to constantly update the input data of the model with the current prediction results, as previous study dose in short-term load forecasting [15]. To the best of our knowledge, this study includes the widest time interval (150 days ahead) ever made by LSTM networks. In order to predict the outbreak of integral curve, estimate DI and depict the evolution of the epidemic, experiments for Russia, Peru, and Iran are conducted in this paper. Modeling results show that under different policies in different continents, there is also a good fit between actually reported daily confirmed infections with our predicted ones. Our model can help public health providers and policy makers make the necessary arrangements to respond to the potential changes in COVID-19 trend. The experiments are based on a set of confirmed COVID-19 cases as of 7 July 2020.

The remainder of this paper is organized as follows. Section 2 introduces the data sources and proposes the framework of LSTM model with rolling update mechanism for COVID-19. Section 3 provides the experimental results and discussion based on the epidemic data of several typical countries in Asia, Europe and South America. The conclusion of this article is given in Section 4.

2. Methodology

2.1. Data description

In this study, we collected data on COVID-19 in country level from January 22, to July 7, 2020, based on real-time statistics from John Hopkins university. It is worth noting that the raw data were studied in the LSTM network through deep learning, and no change was caused to the infected database during the study. As of 10:33 PM on 7 July, 2020, a total of 11,662,574 COVID-19 cases have been reported worldwide, with 539,058 deaths and 6,336,160 survivors, with an overall case fatality rate of 4.62%. The most updated epidemical data of COVID-19 are provided by an ArcGIS platform published by Dong et al. [16].

2.2. LSTM based trend prediction model

In recent years, deep learning methods such as recurrent neural networks (RNNs) have been shown to be effective for prediction [17]. RNN is a neural network with a feedback structure whose inputs not only match the current, but also is related to the weight values of the network and the previous network structure, which automatically extracts relevant features from the training sample and combines the previous. The activation of a time step serves as

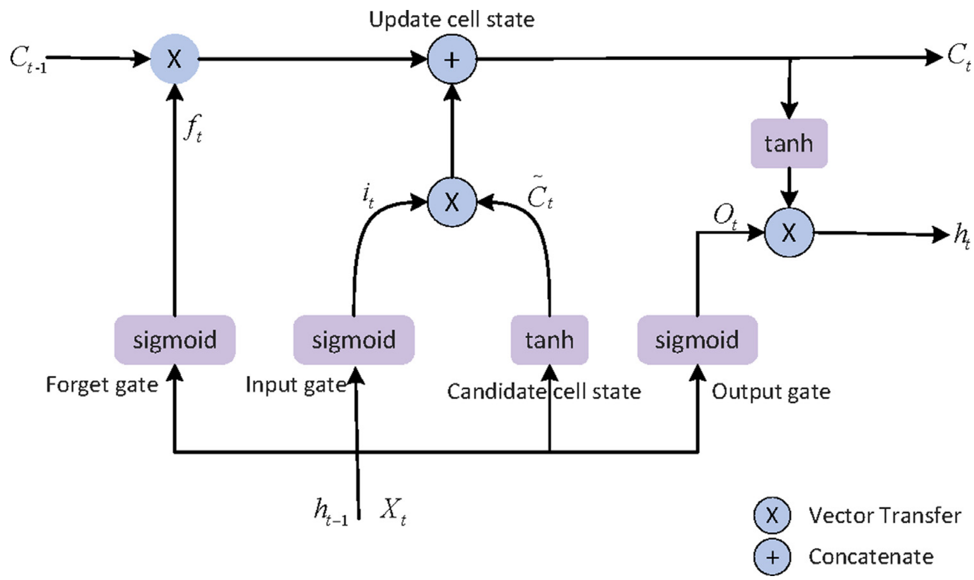

Fig. 2. The structure of LSTM consists of forget gate, input gate and output gate. The role of the forget gate is to forget the information in the cell state selectively. The input gate is to determine what new information is stored in the cell state. Finally, the output gate determines what value we want to output.

an input to the current time step and the network self-connection. Because RNNs store a large amount of historical information in their internal state, they show great potential for data processing [18]. However, the limitations are that there are problems of gradient disappearance and gradient explosion, and since RNNs can only remember part of the sequence, it is much less accurate than short sequences when performing on long sequences, resulting in reduced accuracy once the sequence is too long.

To address these limitations, The long short-term memory model (LSTM) was proposed by Hochreiter and Schmidhuber [19] and has improved and promoted by Alex Graves [20]. LSTM is a special structural type of RNN model, which adds three control units (cells) of input, output, and forgetting gates, and as information enters the model, the cells in the LSTM will judge the information, and the information that conforms to the rules will be left behind and the information that does not conform will be forgotten, and this principle can solve the long sequence dependency problem in neural networks by improving the structure of the hidden layer. In many issues, LSTM has been quite successful and widely used, such as highway trajectory prediction [21], stock price prediction [9] and air pollution forecast [10].

The structure of LSTM is illustrated in Fig. 2. Compared to traditional recurrent neural networks, which cannot handle long time series data due to their structure, LSTM networks introduce a number of gating structures, including forgetting gate, input gate and output gate. C_t, called cell sate, functions roughly the same as h_t in traditional neural networks.

The main role of the forget gate is to control whether to forget the hidden state of the previous layer with a certain probability or not, which can be expressed as follows:

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right) \tag{1}$$

where f_t ranges from 0 to 1 due to it is a hidden state with a certain probability of passing the previous level, and W_f is the weight of the network. Therefore, the activation function chooses Sigmoid function. When the current new information is received, the new input information is formed by adding part of the information at a certain probability. The function of the input gate in gated structure is to filter the current input information and determine how much of the current information is added to the current cell state. The calculation formula is as follows:

$$i_t = \sigma\left(W_u h_{t-1} + U_i x_t + b_i\right) \tag{2}$$

$$\widetilde{C}_t = tanh(W_C h_{t-1} + U_C x_t + b_C) \tag{3}$$

Since the input gate also filters input information with a certain probability, the value of i_t here is also between 0 and 1. The activation function selects Sigmoid function, and when receiving new information, it can updates to the new cell state by multiplying x_t and $\widetilde{C}_t$. The forget gate and the input gate change the current cell state C_t by probabilistic selection of the previous moment and current information. The process of cell state updating from the original C_{t-1} to the current C_t can be expressed by the follows:

$$C_t = C_{t-1} \odot f_t + i_t \odot \widetilde{C}_t \tag{4}$$

where $\odot$ is the Hadamard product of a matrix (the product of the elements in the same position of two matrices). The new cell state is an update of the current cell state from the filtered content of the new information at that time plus the information from the old cell state, which is transmitted with a certain probability. The output gate extracts information from the current cell state, and the extracted information is used to produce the hidden state h_t, which can be written as follows:

$$o_t = \sigma\left(W_o h_{t-1} + U_o x_t + b_o\right) \tag{5}$$

$$h_t = o_t \odot tanh(C_t) \tag{6}$$

The value of o_t in Eq. (5) ranges from 0 to 1, so Sigmoid function is selected to update the cell state to determine what will be left and lost, and then input the cell state into the tanh function for calculation, and finally leave the part of information we want. By combining the above expressions, it can be obtained that:

$$h_t = o_t \odot tanh(C_{t-1} \odot f_t + i_t \odot tanh(W_C h_{t-1} + U_C x_t + b_C)) \tag{7}$$

$$C_t = C_{t-1} \odot f_t + i_t \odot \widetilde{C}_t \tag{8}$$

As is derived from the formula we can see that the size of the h_t by the current state of cells on C_t and moment hidden state information contained in the combined impact of h_{t-1} and W_C has no effect on the current state of the cells in the calculation. Therefore, we can conclude that in the circulation of traditional neural network inside the W_C gradient is the main reason for the gradient disappearance, and here in LSTM, when the forgot gate f_t is opened, the gradient C_t can be effectively passed to the a moment of cell state C_{t-1}. Therefore, improving the traditional neural network by joining gate structure, can reduce the problems occurred in the training process of gradient disappeared.

2.3. Rolling update mechanism for COVID-19 outbreaks and parameter settings

In order to achieve long-term prediction of COVID-19 outbreaks, a rolling update mechanism is adopted in this study. The rolling update mechanism is to update the training sample sequence according to the current prediction results at every moment so as to train the model. The so-called rolling optimization is to optimize the rolling feedback correction control in a limited time by using the predictive model and historical data to train the output model of the future moment iteratively. The learning process of LSTM based on update mechanism for COVID-19 outbreaks is depicted in Fig. 3.

There are several parameters in our study. The time step of input sequence is set to 3, which means the number of new cases in the following day is projected by the confirmed cases in the last three days. The number of hidden neuron units is set to 50, through which the hidden layer receives the data from the input layer. Our model selects the Adam optimizer, batch size of one and training for 500 rounds. To find the direction of gradient descent and estimate the error between true value and predict value, we apply Mean Square Error (MSE) as loss function, which can be calculated as follows:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{n} \tag{9}$$

3. Results and discussion

In this section, experiments are performed to analyze the effectiveness of our proposed model for predicting the curve of COVID-19 in Russia, Peru, and Iran. The epidemiological data we collected here is from January 22, 2020 to July 7, 2020, which have been used with Google Colaboratory environment.

3.1. Epidemic progression in Russia, Peru, and Iran

We study these three countries as they are among the top three states in Europe, South America and Asia in terms of the number of new confirmed cases per day. Because different continents implement different policies, the training of model in different continents can verify and analyze the robustness of our model better. Another reason is that the number of new infections per day in these countries does not follow a consistent upward trend, but both the increase also has a downward trend, which means the daily new infections in these countries curve is no longer a non-smooth sequence. Therefore, it is conducive to the construction of

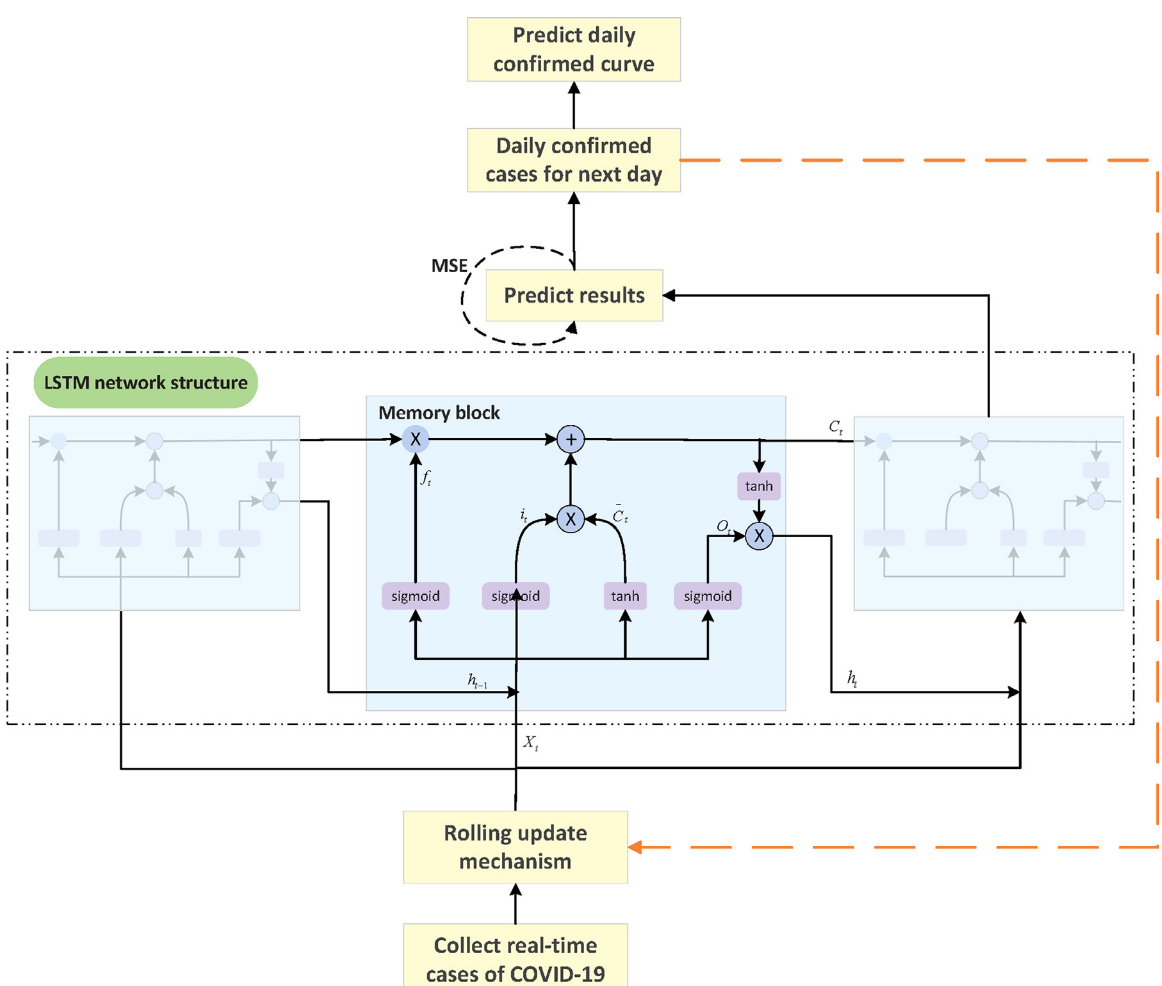

Fig. 3. Learning process of LSTM with rolling update mechanism for COVID-19 outbreaks.

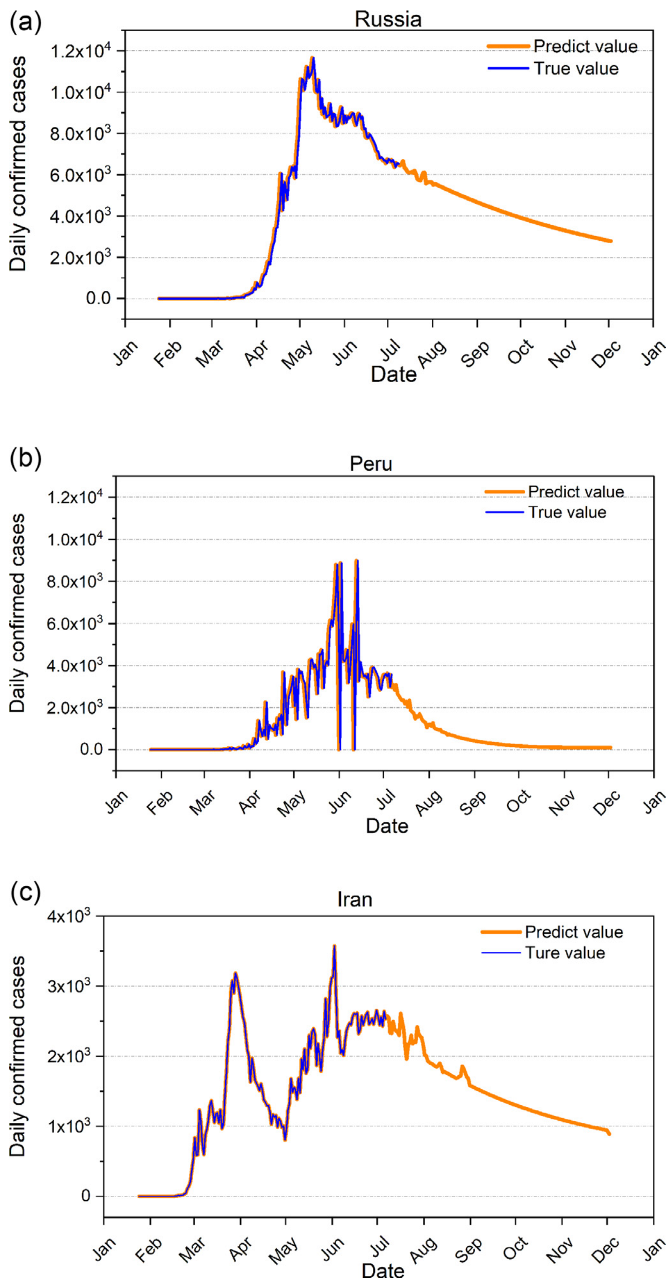

Fig. 4. The number of daily increased cases predicted by LSTM with rolling update mechanism for Russia (a), Peru (b) and Iran (c).

Table. 1
Comparison of reported and predicted daily confirmed cases.

Country	Date	Reported data	Predict data	Error (%)
Russia	8th July	6562	6517	−0.69
	9th July	6509	6482	−0.41
	10th July	6635	6440	−2.94
	11th July	6611	6501	−1.66
Peru	8th July	3575	3483	−2.57
	9th July	3633	3832	5.48
	10th July	3537	3084	−12.81
	11th July	3198	3118	−2.50
Iran	8th July	2691	2571	−4.46
	9th July	2079	2152	3.51
	10th July	2262	2334	3.18
	11th July	2397	2325	−3.00

our models and the increase of robustness to predict when the epidemic will end.

We have a rolling forecast of the number of new cases per day for the next 150 days in these three countries. The LSTM model predicts that although the outbreak in Russia has been under control recently and is on a downward trend, the number of new cases per day in Russia will still be more than 2000 by early December (Fig. 4(a)). This may be because the outbreak in Russia has grown beyond a certain scale, and it is difficult to fully control or return to the previous level.

Furthermore, the modeling study shows that initial victory in the fight against the epidemic has been achieved in Peru because of the effective measures have taken by the government, and the number of new confirmed COVID-19 cases per day is waning. Our model estimates that starting in mid-august, the daily confirmed cases in Peru will be less than 1000, from then on into a long tail die model (Fig. 4(b)). At the end of November and early December the outbreak will be fully controlled.

The model also indicates that the COVID-19 outbreak in Iran, which took strict measures in March and achieved results a month later, has rebounded in recent days after easing restrictions. The Iranian government announced that strict control of the outbreak has been restored since 5 July. Assuming that this control measure is as effective as previous ones, we have modeled how the outbreak will develop, and the results show that Iran, like Russia, still has a long way to go in combating COVID-19. The number of new positive cases per day in Iran is expected to fall below 1000 by mid-November, with a gradual downward trend expected after several smaller peaks from July to September, which is mapped in Fig. 4(c). We then plotted the number of daily new cases derived from LSTM and the actual reported data for these three countries. There was a remarkable fit between the actual number of new confirmed cases and the LSTM-predicted curve between January 22 and the July 7.

In Fig. 5, we plot confirmed COVID-19 cases and forecasted total cases in Russia, Peru and Iran over time. As can be seen from the figure, the cumulative number of confirmed cases in Peru will peak within 2020. Although the growth rate of Russia and Iran has slowed down, the cumulative number of cases is expected to show a linear trend. The above results provide that the epidemic of Peru will peak on early December, resulting in 398,991 total infections. We also predicted the epidemic size of Russia will reach 1,343,751 after 150 days, and the accumulative confirmed cases in Iran will be greater than 470 thousand by early December according to the results of our model. We plotted the actual reported values and the projections in a single line to depict the full curve of COVID-19. The results found that there is overall a good fit between our estimated and official data.

It is a challenging task to forecast the spread of outbreaks based on small data sets, limited by the small data available, from which much data and law cannot be deciphered if the initial data is almost a flat curve. In our experiment, we selected countries with a downward trend in the number of new confirmed cases per day, so as to learn more about the overall pattern of epidemic transmission. In this way, our forecast results have both several small peaks and a trend of continuous decline, which is more in line with the national conditions of the predicted countries.

For these data-driven projections, the data has been taken up to July 7, 2020. The comparison has also been made for total positive reported cases with predicted cases by our proposed method from July 8, to July 11, 2020 as shown in Table 1.

3.2. Effects of measures adopted to prevent the spread of COVID-19

Our study highlights the importance of disease control measures. Strict control can significantly release the spread of the epi-

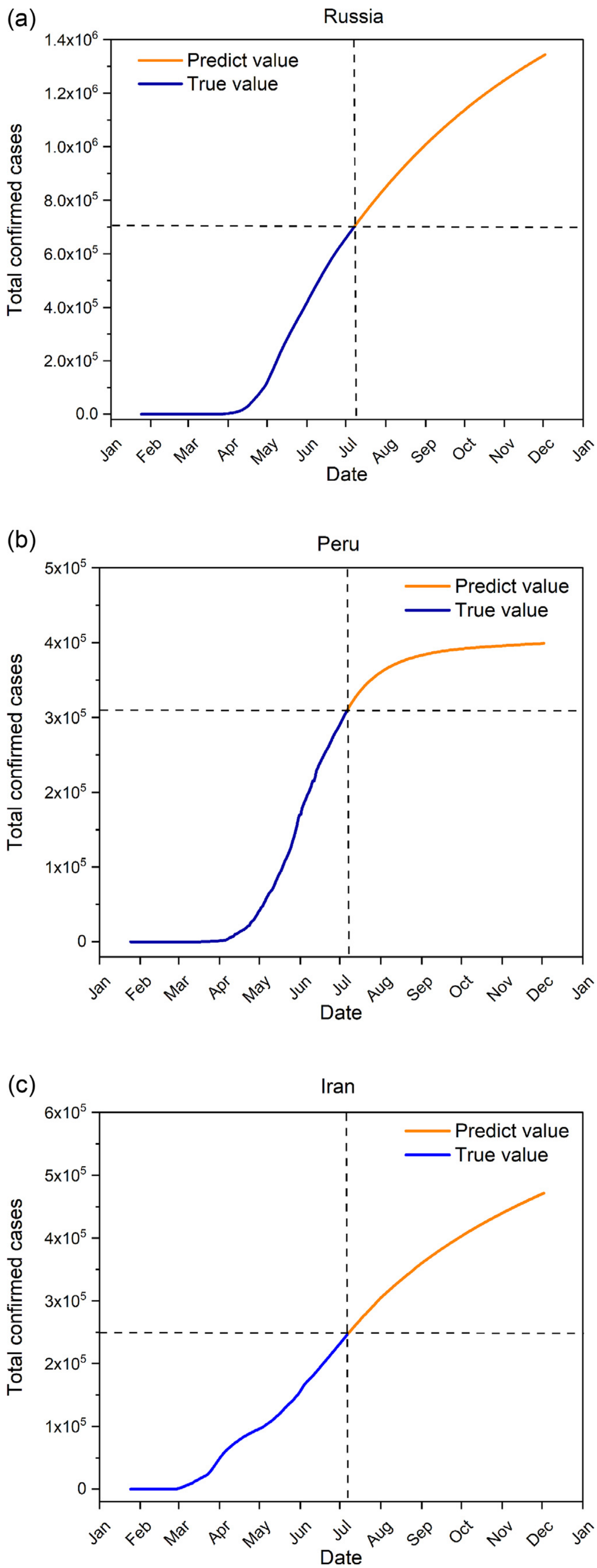

Fig. 5. The number of total confirmed cases predicted by LSTM with rolling update mechanism for Russia (a), Peru (b) and Iran (c).

demic, while relax control is likely to lead to the outbreak of epidemic rebound or secondary explosion. In order to describe the trend in epidemic control, we define the increase in the number of new confirmed cases on that day relative to the number of existing cases on the previous day as Diffusion Index (DI) and mapped DI dot line for each country (Fig. 6). After the first case of COVID-19 was reported by Russia, strict control measures have been implemented in early March, including travel ban, home office, etc., resulting the DI fell to 0.06 from 0.88 within a month. However, in order to resume normal production and life, the Russian government announced the end of national vacation and began to remove restrictions since May, the DI tends to a little rebound. Therefore, it is an important challenge for the affected countries to strike a balance between reasonable prevention and control of the epidemic and recovery of production and life.

The DI curves for Peru and Iran also signify that control measures taken in early March were effective, leading to an overall decline in the rate of growth around mid-May. In addition, our model predicts that after July, the DI of these three countries will not exceed 0.03, which means that the government and the public have had sufficient understanding of COVID-19, the epidemic prevention policy will remain strict, and the daily epidemic prevention of the public will become normal. People should continue to cooperate with the prevention and control of the epidemic, so as to bring an early end to the epidemic.

While more data is needed to make more detailed predictions, these models can help predict future confirmed cases if the spread of the virus does not change more than expected. It is well known that this is a new virus with severe transmission capacity. This feature may affect all of our projections, but as far as we know at the time of writing, the proposed model has proven to be effective.

4. Conclusion and future work

In view of the global COVID-19 pandemic, this study focuses on modeling the long-term epidemic trend of COVID-19 by using LSTM networks and rolling update mechanism by feeding new forecasting results into model training for next iteration for Russia, Peru and Iran. By estimating the number of daily increased cases, the full curve of total confirmed cases in next 150 days is thus derived. Then, we introduce DI to evaluate the effectiveness of preventive measures implemented by the government. The above results provide that the epidemic of Peru will peak on early December, resulting in 398,991 total infections. The number of positive cases per day in Iran is expected to fall below 1000 by mid-November, with a gradual downward trend expected after several smaller peaks from July to September, while there will still be more than 2000 increase by early December in Russia. Moreover, our study highlights the importance of preventive measures have taken by each government, which shows that the strict controlment can significantly reduce the spread of COVID-19.

Our study, as with all other modeling researches, has several assumptions and limitations. For example, as is shown in Section 3, all of our experiments are based on the assumption that the present policies will not change in 150 days and people will cooperate with the preventive measures, while there may be fluctuation or rebound in the real world due to the change of government policies and the degree of public cooperation. Aside from policy changes, the impact of imported cases and the spatial influence between countries are not taken into account in our model.

To deal with the aforementioned limitations, the following two aspects are worthy for future study. First, considering that intervention and control measures and people's attitude towards the epidemic are closely related to the spread of the virus, Natural Language Processing (NLP) technology can be used to extract semantic features from news reports related to people's awareness

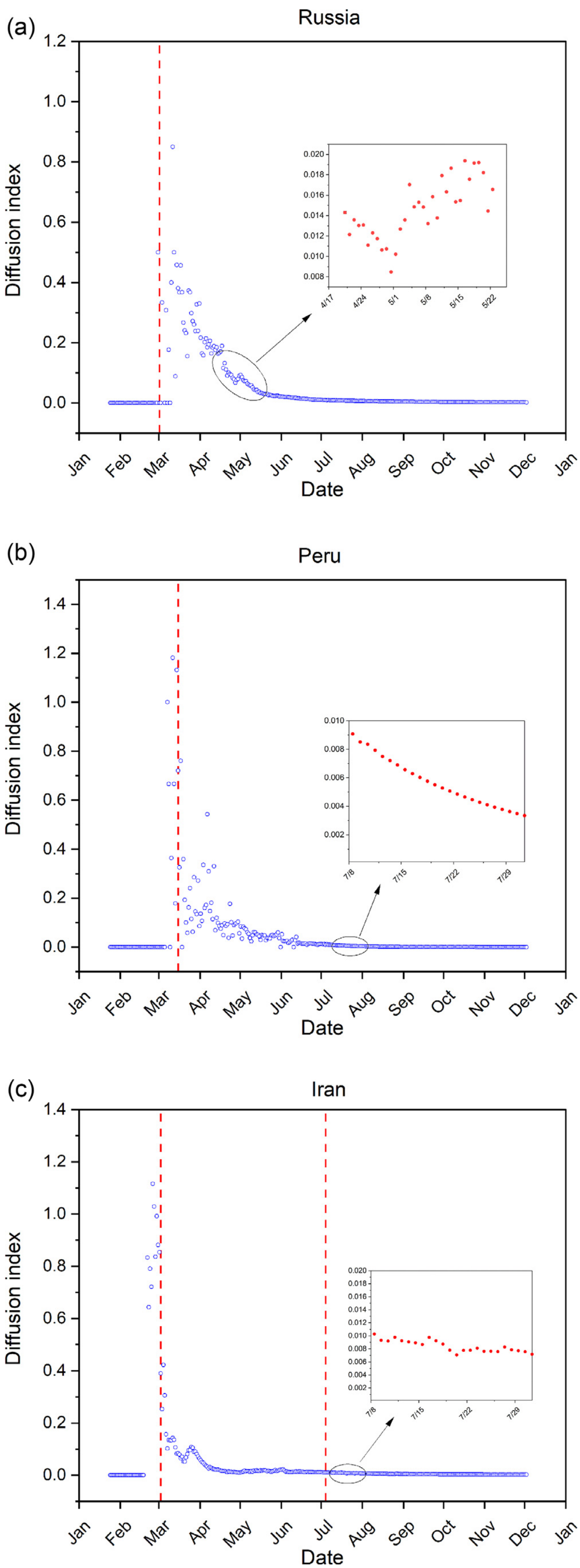

Fig. 6. The diffusion index in Russia (a), Peru (b) and Iran (c). The dotted red line represents the time of the government's announcement for strict control. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of epidemic prevention and control measures, which could be embedded with LSTM to realize dynamic prediction of epidemic trend under real-time status according to the actual situation. Second, the spatio-temporal residual network can be used to predict the number of imported cases in each region of the country and to model the spatial characteristics of epidemic trends to forecast the trend of the epidemic under the influence of imported cases from other countries.

In summary, based on our proposed model, the rolling update mechanism with LSTM is embedded for long-term projections. The forecasting results of the model are highly consistent with reported daily positive cases, which demonstrate that the proposed method can accurately analyze the trend of epidemic. The effects of preventive measures, such as social isolation and isolation, have also been observed in this article, suggesting that the spread of the virus can be significantly reduced through these preventive measures. Modeling analysis of COVID-19 outbreaks may help national health agencies to develop response plans. To the best of our knowledge, in this study includes the widest time interval ever made based LSTM networks.

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.

Acknowledgments

This research has been jointly supported by the National Key R & D Program of China under grant no. 2016YFC0600510 and the Fundamental Research Funds for the Central Universities under grant nos. 2652020002 and 2652020004.

References

[1] Jacobsen KH. Will COVID-19 generate global preparedness. Lancet 2020;395:1013–14.
[2] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New Engl J Med 2020;382:1199–207.
[3] Shao, N., Cheng, J., Chen, W., 2020. The reproductive number R0 of COVID-19 based on estimate of a statistical time delay dynamical system. medRxiv.
[4] Kraemer MUG, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM, Plessis Ldu, Faria NR, Li R, Hanage WP, Brownstein JS, Layan M, Vespignani A, Tian H, Dye C, Pybus OG, Scarpino SV. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 2020;368:493–7.
[5] Ozaslan M, Safdar M, Kilic IH, Khailany RA. Practical measures to prevent COVID-19: a mini-review. J Biol Sci 2020;20:100–2.
[6] Wang P, Zheng X, Li J, Zhu B. Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. Chaos Solitons Fractals 2020;139:110058. doi:10.1016/j.chaos.2020.110058.
[7] Zheng N, Du S, Wang J, Zhang H, Cui W, Kang Z, Yang T, Lou B, Chi Y, Long H, Ma M, Yuan Q, Zhang S, Zhang D, Ye F, Xin J. Predicting COVID-19 in China using hybrid AI model. IEEE Trans Cybern 2020:1–14. doi:10.1109/TCYB.2020.2990162.
[8] Contreras J, Espinola R, Nogales FJ, Conejo AJ. ARIMA models to predict next–day electricity prices. IEEE Trans Power Syst 2002;18:1014–20.
[9] Persio, L.D., Honchar, O., 2016. Artificial neural networks architectures for stock price prediction: Comparisons and applications.
[10] Q. Zhang, J.C.K. Lam, V.O.K. Li, Y. Han, 2020. Deep-AIR: A Hybrid CNN-LSTM Framework forFine-Grained Air Pollution Forecast.
[11] Kırbaş İ, Sözen A, Tuncer AD, Kazancıoğlu FŞ. Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. Chaos Solitons Fractals 2020;138:110015. doi:10.1016/j.chaos.2020.110015.
[12] Yang Z, Zeng Z, Wang K, Wong S-S, Liang W, Zanin M, Liu P, Cao X, Gao Z, Mai Z, Liang J, Liu X, Li S, Li Y, Ye F, Guan W, Yang Y, Li F, Luo S, Xie Y, Liu B, Wang Z, Zhang S, Wang Y, Zhong N, He J. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thorac Dis 2020;12:165–74. doi:10.21037/jtd.2020.02.64.

[13] Azarafza M, Azarafza M, Tanha J. COVID-19 infection forecasting based on deep learning in Iran (preprint). Epidemiology 2020. doi:10.1101/2020.05.16.20104182.
[14] Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos Solitons Fractals 2020;135:109864. doi:10.1016/j.chaos.2020.109864.
[15] Wang S, Wang X, Wang S, Wang D. Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting. Int J Electr Power Energy Syst 2019;109:470–9. doi:10.1016/j.ijepes.2019.02.022.
[16] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020;20:533–4.
[17] Wang J, Zhang C. Software reliability prediction using a deep learning model based on the RNN encoder–decoder. Reliab Eng Syst Saf 2018;170:73–82. doi:10.1016/j.ress.2017.10.019.
[18] Connor, J., Martin, R., Atlas, L., 1994. Recurrent neural networks and robust time series prediction 5, 240–254. 10.1109/72.279188
[19] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735–80. doi:10.1162/neco.1997.9.8.1735.
[20] Deutsch Supervised sequence labelling with recurrent neural networks. Berlin Heidelberg: Springer Springer Verlag; 2012. doi:101007/978-3-642-24797-2.
[21] Altché F, Arnaud DLF. An LSTM network for highway trajectory prediction. 2017 IEEE 20th international conference on intelligent transportation systems (ITSC); 2018. doi:101109/ITSC20178317913.