Cascade SEIRD: Forecasting the Spread of COVID-19 with Dynamic Parameters Update

Yongliang Wen*, Jiangnan Xu, Yunming Ye, Xutao Li, Chuyao Luo, Tianlun Zhu
Department of Computer Science, Harbin Institute of Technology, Shenzhen, China
Email: yongliangwen.hit@gmail.com*

2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 978-1-7281-6215-7/20/$31.00 ©2020 IEEE DOI: 10.1109/BIBM49941.2020.9313525

***Abstract*—The SEIR model is widely used in simulating the spread of infectious diseases. COVID-19 virus is a very severe infectious disease. Some studies leverage the SEIR or SEIRD model to simulate the spread and estimate the number of infected and recovered people as time goes on. However, these models suffer from two key deficiencies: (i) conventional SEIRD does not update its model parameters w.r.t. time; (ii) it focuses on predicting the trend, instead of the actual number of infections in the future. In this paper, we propose a cascade SEIRD model. The model learns and updates its parameters every day. Moreover, it is able to predict the number of infection cases, recovered cases and deaths. Specifically, we leverage a machine learning like approach to dynamically estimate the parameters of infection rate, incubation rate, recovery rate and death rate, which can be updated by gradient descent algorithm. Once the nature of the parameters w.r.t. time is determined, ARIMA model is adopted to characterize the dynamics of the parameters and predict their future changes. To validate the effectiveness of the proposed cascade SEIRD model, we conduct experiments on five data sets of different scales of regions (China, Hubei, Wuhan, Shenzhen, US). Experimental results show that the proposed cascade SEIRD achieves the most accurate prediction and outperforms state-of-the-art techniques.**

***Index Terms*—SEIRD, Epidemic Model, Gradient Descent, ARIMA**

I. Introduction

Recently, the COVID-19 epidemics significantly affects the people's life and world economics. To better understand and predict how the epidemics would spread, many infection models have been utilized, such as Susceptible-Infectious-Susceptible model **(SIS)** [1], Susceptible-Exposed-Infectious-Recovered model **(SEIR)** [2] and Susceptible-Exposed-Infectious-Recovered-Dead model **(SEIRD)** [3].

However, the existing studies suffer from two key shortcomings. **(i)** The models can only analyze and predict the trends but cannot predict the number of infections and recovers accurately. **(ii)** Most of the studies [4]–[7] all use fixed parameters, namely the infection rate, incubation rate, recovery rate, and death rate are the same every day, which leads to the limited representation capability of the model. For example, SEIRD model adopts the fixed parameters $\beta, \alpha, \gamma, \theta$ in Fig. 1. In fact, the parameters vary greatly from one day to another, and also some subtle changes can lead to large variations in prediction.

To address the two deficiencies, we propose **cascade SEIRD model** which is inspired by the cascade structures in cascade recurrent convolution neural network (RCNN) [8] and cascade generative adverserial networks (GANs) [9]. These models have achieved state-of-the-art performance in computer vision tasks, e.g., object detection and image generation. Different from conventional SEIRD, which keeps its model parameters fixed, the proposed cascade SEIRD updates and learns the infection rate, incubation rate, recovery rate, and death rate every day. Previous studies show that a complex model incorporates more biological and epidemiological information about the epidemic and is more biologically realistic, but it requires more model parameters to be estimated compared to a simpler model [10]. The proposed method makes an excellent trade-off between the model complexity and dynamics. First, our method has fewer parameters than deep learning models. Second, as shown in Fig.1, in our cascade SEIRD, the parameters infection rate β_t, incubation rate α_t, recovery rate γ_t and death rate θ_t change from time to time and our model is more interpretable. By leveraging such a cascade structure, our method can better capture and characterize the spread dynamics.

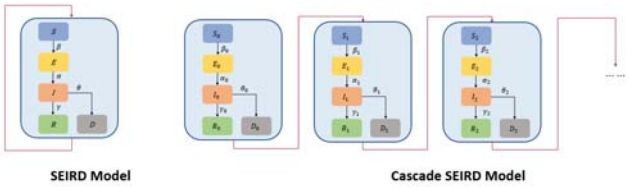

Fig. 1. **The architecture of cascade SEIRD Model. (a)** is the SEIRD model, and **(b)** is the cascade SEIRD model. S, E, I, R, D is the number of five states: Susceptibles, Exposed, Infectives, Resistances, Dead. **The biggest difference between (a) and (b) is that (b) has different parameters $\beta_t, \alpha_t, \gamma_t, \theta_t$ that change over time.**

Recently, a time-dependent SEIRD model is proposed by taking the varying parameters issue into account [11]. However, the model learns dynamic parameters via Markov chain Monte Carlo (MCMC), which is computationally expensive and takes hours to update according several days observations. Moreover, the MCMC is often hard to converge to the target posterior distribution in the presence of non-identifiable and extreme distribution skewness. [10]

In this paper, we develop a cascade SEIRD model. Different from the time-dependent SEIRD model, our model leverage a machine learning like method to dynamically estimate the parameters. Once the nature of the parameters w.r.t. time is determined, ARIMA model is adopted to characterize the

dynamics the parameters and predict their future changes by extrapolation. Due to the optimization is based on gradients, our model is more efficient.

In summary, this work has the following contributions:

- We propose a cascade SEIRD model to dynamically update the parameters in SEIRD. The proposed model can be trained in an end-to-end pipeline.
- The cascade SEIRD model can learn the time-dependent parameters infection rate, incubation rate, recovery rate and death rate efficiently.
- According to the parameters dynamics w.r.t. time, we build a ARIMA method to predict the future parameters, which are utilized to make more accurate recovery and death number predictions.

II. The Proposed Approach

A. Preliminary

The SEIRD model is widely used in the infectious disease transmission. The key idea of the model is dividing the disease spread into four phases, namely incubation, infection, cure, and death.. Correspondingly, by considering the transition between the phases, five states are designed, including Susceptible(S), Exposed(E), Infectious(I), Recovered(R) and Dead(D). The states may transit from one to another with a certain probability. For example, S transits to E with the probability of infection rate β, and E turns into I with the probability of disease incidence α. The designs of the two state because susceptible people may be infected as latent patients when they are exposed with patients. Patients will be cured with recovery rate γ, or they will die with the death rate θ. According to the probabilities of state transitions, we obtain the following equations to characterize the spread process:

$$\begin{aligned} N &= S_t + E_t + I_t + R_t + D_t \\ \frac{dS_t}{dt} &= -\beta * I_t * \frac{S_t}{N} \\ \frac{dE_t}{dt} &= \beta * I_t * \frac{S_t}{N} - \alpha * E_t \\ \frac{dI_t}{dt} &= \alpha * E_t - \gamma * I_t - \theta * I_t \\ \frac{dR_t}{dt} &= \gamma * I_t \\ \frac{dD_t}{dt} &= \theta * I_t \end{aligned} \tag{1}$$

Here N is a constant denoting the total population, which is equal to the sum of all states at any time. Though the equations nicely capture the transitions between different states. However, the transition probabilities between states are kept fixed as time goes on, which restrains the representation capacity of the model. Indeed, the transition probabilities change from time to time. By considering the state changes, we obtain the following equations for the states:

$$\begin{aligned} S_{t+1} &= S_t + \frac{dS_t}{dt} \\ E_{t+1} &= E_t + \frac{dE_t}{dt} \\ I_{t+1} &= I_t + \frac{dI_t}{dt} \\ R_{t+1} &= R_t + \frac{dR_t}{dt} \\ D_{t+1} &= D_t + \frac{dD_t}{dt} \end{aligned} \tag{2}$$

The states at time $t+1$ are update from the states at time t by the difference in equation (1).

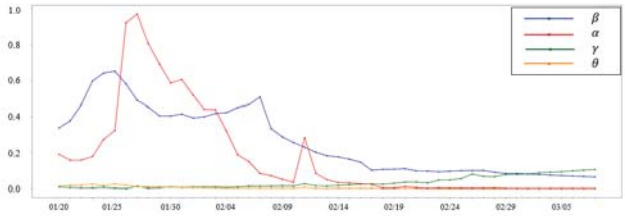

Fig. 2. **Parameters changing process.** The blue,red,green,orange lines refer to the parameters $\beta, \alpha, \gamma, \theta$ changing over time.

B. The Cascade SEIRD Model

In reality, the infection rate β, incubation rate α, recovery rate γ and death rate θ change every day. Fig.2 shows the fitted results of the parameters w.r.t time and an obvious dynamics can be found. Hence, we extend the SEIRD model into a temporal cascade one. Specifically, we assume that the dynamic parameters are denoted as β_t, α_t, γ_t, and θ_t, respectively. With the consideration of time dependence, the difference equation of cascade SEIRD is as follow:

$$\begin{aligned} \frac{dS_t}{dt} &= -\beta_t * I_t * \frac{S_t}{N} \\ \frac{dE_t}{dt} &= \beta_t * I_t * \frac{S_t}{N} - \alpha_t * E_t \\ \frac{dI_t}{dt} &= \alpha_t * E_t - \gamma_t * I_t - \theta_t * I_t \\ \frac{dR_t}{dt} &= \gamma_t * I_t \\ \frac{dD_t}{dt} &= \theta_t * I_t \end{aligned} \tag{3}$$

where N is the same as in equation (1).

Let us revisit the Fig.1. We can see that in our approach, the SEIRD-cell is a base module which utilizes S_t, E_t, I_t, R_t, D_t to predict $S_{t+1}, E_{t+1}, I_{t+1}, R_{t+1}, D_{t+1}$ following the equation (3). We build the cascade SEIRD model with SEIRD-cells and train the parameters in cells with the loss by gradient descent method. By cascading the SEIRD cells iteratively w.r.t time, we are able to better model the infectious disease

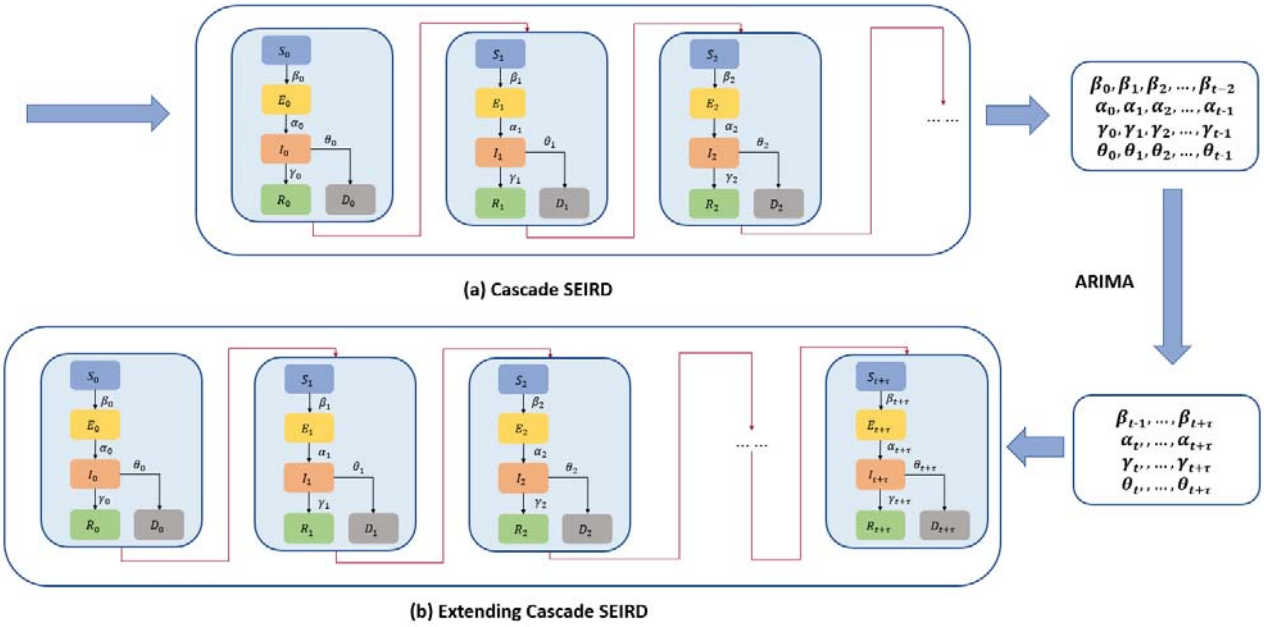

Fig. 3. **The pipeline of extending cascade SEIRD.** At time t, cascade SEIRD learn the historical parameters in **(a)**, and predict the future parameters by ARIMA. In **(b)**, use the predicted parameters to extending cascade SEIRD.

transitions. For example, cascade SEIRD model can capture mutations caused by emergencies, because the dynamic parameters changing over time can nicely characterize the current infection situation. Generally speaking, in the early stage of virus transmission, the virus has not spread widely; but in the middle stage, when the number of infected people tends to be saturated, the number of infected people will decrease. This process can be well expressed by our model, while the SEIRD model cannot capture the relationship between infection rate and time.

C. The Loss Function

In real applications, the numbers of confirmed, recovered and death cases, denoting the quantitative evaluations, are more important than the qualitative trends. As the numbers change every data. We design the following loss function:

$$L = L_C + L_I + w_R L_R + w_D L_D \tag{4}$$

where L_C, L_I, L_R, L_D denotes the numbers of confirmed cases, infected cases, recovered cases and death cases. In the data, there are only the number of confirmed cases, the number of recovered people and the number of deaths, and the number of confirmed cases is the sum of the number of infected, recovered and death cases. As the four numbers may have different magnitudes, we introduce the parameters w_R and w_D to balance the four losses: L_C, L_I, L_R, L_D. In our experiment, we set the two parameters as:

$$w_R = \frac{\sum_{i=0}^{t} I_i}{\sum_{j=0}^{t} R_j} \tag{5}$$

$$w_D = \frac{\sum_{i=0}^{t} I_i}{\sum_{j=0}^{t} D_j} \tag{6}$$

With the training data $I_{1\ldots t}, R_{1\ldots t}, D_{1\ldots t}$ from day 1 to day t, we can learn the parameters $\beta_{0\ldots t-2}, \alpha_{0\ldots t-1}, \gamma_{0\ldots t-1}, \theta_{0\ldots t-1}$ with gradient descent method.

D. Forecasting Parameters Method

Once the dynamic parameters are estimated from day 1 to day t, we able to know how the parameters change in the past. However, our objective is to predict the parameters in the future, so as to simulate the spread of disease according to the cascade SEIRD model. Hence, we need a forecasting model to predict the parameters in the future. Autoregressive Integrated Moving Average(ARIMA) is widely used in time series tasks. As shown in Fig.3, after model fitting the dynamic parameters($\beta, \alpha, \gamma, \theta$), we use ARIMA to forecast the furture parameters and feed the parameters into the cascade SEIRD-cell to forecast the future cases.

After differences operation, the parameters are stationary, because they have basic values and physical meanings. [12] For example, β is the infection rate. It is the virus itself that determines its basic infection risk. In addition, some factors

that people contact with each other may lead to high infections, or decision makers may block and isolate infected people and cause the virus infection rate to be lower. Hence, it is nature to apply the ARIMA model to predict the future parameters. Once the ARIMA prediction models $f_\beta(\cdot), f_\alpha(\cdot), f_\gamma(\cdot), f_\theta(\cdot)$ are estimated, we are able to forecast the future parameters as:

$$\begin{aligned} \beta_{t-1}, \beta_t, ..., \beta_{t+\tau} &= f_\beta(\beta_0, \beta_1, ..., \beta_{t-2}) \\ \alpha_t, \alpha_t, ..., \alpha_{t+\tau} &= f_\alpha(\alpha_0, \alpha_1, ..., \alpha_{t-1}) \\ \gamma_t, \gamma_t, ..., \gamma_{t+\tau} &= f_\gamma(\gamma_0, \gamma_1, ..., \gamma_{t-1}) \\ \theta_t, \theta_t, ..., \theta_{t+\tau} &= f_\theta(\theta_0, \theta_1, ..., \theta_{t-1}) \end{aligned} \tag{7}$$

By feeding the predicted parameters into the SEIRD cascade models, we are able to estimate the number of confirmed, recovered and death cases in the future.

E. Implementation Details

The overall pipeline of our approach can be illustrated in Fig. 3, which includes three important stages: **training model, forecasting parameters, extending model**. After training the cascade SEIRD model, we obtain the historical parameters. Moreover, the calculation is very efficient. In general, it takes only several minutes to fit one month data.Once the parameters are estimated, we can feed them into the cascade SEIRD model to predict the spread. The core part is training the model with historical data. Here we summarize the training algorithm in Algorithm 1.

Algorithm 1: Training cascade SEIRD Model

Input: $I_0, I_1, ..., I_t$
$R_0, R_1, ..., R_t$
$D_0, D_1, ..., D_t$
Initialization:
$\hat{I}_0 \leftarrow I_0, \hat{R}_0 \leftarrow R_0, \hat{D}_0 \leftarrow D_0$
$E_0 = \omega_E * I_0$
$S_0 = N - E_0 - I_0 - R_0 - D_0$
$\hat{E}_0 \leftarrow E_0, \hat{S}_0 \leftarrow S_0$
$L = \infty$
while $L > \epsilon$ **do**
  for $i \leftarrow 0$ **to** $t-1$ **do**
    $\hat{S}_{i+1} = \hat{S}_i - \beta_i * I_i * \hat{S}_i / N$
    $\hat{E}_{i+1} = \hat{E}_i + \beta_i * I * \hat{S}_i / N - \alpha_i * \hat{E}_i$
    $\hat{I}_{i+1} = \hat{I}_i + \alpha_i * \hat{E}_i - (\gamma_i + \theta_i) * \hat{I}_i$
    $\hat{R}_{i+1} = \hat{R}_i + \gamma_i * \hat{I}_i$
    $\hat{D}_{i+1} = \hat{D}_i + \theta_i * \hat{I}_i$
  Compute loss:

$$L = L_C + L_I + w_R L_R + w_D L_D$$

  Compute the gradients of parameters $(\beta, \alpha, \gamma, \theta)$ with optimizer
  Update $\beta, \alpha, \gamma, \theta$

III. Experiments

A. Datasets and Metrics

To test the performance of the proposed cascade SEIRD model, we choose 5 regions (China, Hubei, Wuhan, Shenzhen and the United States of America) with large differences in population. The information about COVID-19 of time series confirmed cases, reported deaths and reported recoveries in China is collected from the website of China Center for Disease Control and Prevention. The US data comes from a variety public sources and is collated in the first instance via Johns Hopkins University on GitHub [13]. For reproducing the results, we release our codes and data sets on https://github.com/rssqzqyp/cascade_SEIRD_Model.

The data sets include the statistical information about the COVID-19, e.g., the numbers of confirmed cases, reported deaths and reported recoveries. We split the data sets into training and test data sets. The training sets of China, Hubei and Wuhan is from February 9 to March 2 and Shenzhen is from February 13 to March 2. The US training set is from February 29 to April 6. The test sets of China, Hubei, Wuhan and Shenzhen are from March 3 to March 9 and the US is from April 7 to April 13.

To evaluate the performance, we adopt the relative error rate, which is evaluated from the number of confirmed and cured. The relative error rate of C, R is E_C, E_R, which is defined as follows:

$$E_C = \frac{|\hat{C} - C|}{C}, E_R = \frac{|\hat{R} - R|}{R} \tag{8}$$

where $\hat{C}$ is the confirmed predictions of the model, $\hat{R}$ is the recovered predictions of the model and C and R is the ground truth observations. We calculate the mean of last week relative error at last.

B. Experimental Settings

Generally speaking,the future predictions are mainly related to the past few days. Hence, the cascade structure does not need be too complicated. In general, the cascade SEIRD model uses 20∼30 SEIRD-cells to achieve best performance. Therefore, our training data set also uses about 20 days to predict the number of confirmed and cured people in the next seven days. All computations are performed on a laptop with Intel(R) Core(TM) i5-8265U CPU@1.60GHz, 1800 Mhz, 16 GB RAM.

We adopt the Adam optimizer with a learning rate 0.001 to train and the initial N is the population of the whole region. The initial β is 0.2586, γ is 0.018, θ is 0.001, α is 0.2. According to our experience, the final results are insensitive to the initial parameters.

In real applications, we can save the trained parameters and adopt a lazy update scheme. Once the parameters are trained by cascade SEIRD model, we save and keep them fixed. When the data evolves to a new day, we only needs to train and update the parameters in the new day. By doing so, the cascade SEIRD model can be efficiently learned and updated.

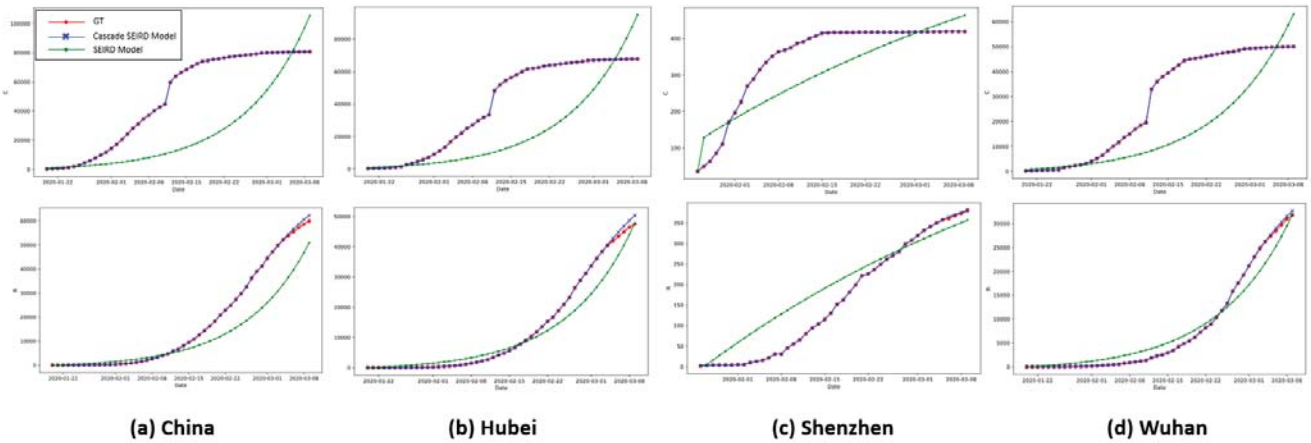

Fig. 4. **Comparison of fitting effect and prediction result.** The red line is ground truth, and the blue line is the fitting effect and prediction of cascade SEIRD model, and the green line is the fitting effect and prediction of SEIRD model. **(a)**,**(b)**,**(d)** are datasets China,Hubei,Wuhan from February 9 to March 2 and **c** is datasets Shenzhen from February 13 to March 2 for training. The last seven day from March 3 to March 9 is for testing.

TABLE I
PERFORMANCE OF CASCADE SEIRD

Method	China		Hubei		Shenzhen		Wuhan		US	
	$E_C(\%)$	$E_R(\%)$	$E_C(\%)$	$E_R(\%)$	$E_C(\%)$	$E_R(\%)$	$E_C(\%)$	$E_R(\%)$	$E_C(\%)$	$E_R(\%)$
SEIRD Model	3.50	27.80	11.18	13.61	2.43	10.71	6.42	6.51	4.88	33.4
Cascade SEIRD Model	**0.24**	**1.98**	**0.128**	**3.03**	**1.66**	**1.83**	**0.136**	**0.41**	**2.02**	**17.8**

C. Experimental Results

In Fig.4, we depict the the ground truth, the fitting results of our cascade SEIRD model and the estimation of the conventional SEIRD model. We can see that the cascade SEIRD model is nicely in line with the ground-truth, which suggest that the model fits the data excellently. To compare the two models on test sets, we summarize the results in Table I. We can see that the cascade SEIRD model significantly outperforms the SEIRD model on the five data sets. The reason is that a statistical parameter setting on SEIRD cannot nicely model the daily dynamics of COVID-19 spread. Hence, our cascade SEIRD, which learns the parameters sequences with the historical data. With the learned sequences, we can better characterize their dynamics and predict the future. As a result, the cascade SEIRD can simulate the spread with more accurate parameters in each day, which thus yields better performance.

IV. CONCLUSION

In this paper, we propose a novel architecture named cascade SEIRD to estimate the infections, recoveries, and deaths from the disease. Different from conventional SEIRD method, the cascade SEIRD learns the dynamic parameters based on the training data, instead of utilizing a fixed parameter setting. An ARIMA model is built upon the learned parameter sequences to predict the parameters in the future. The parameters are then fed into the cascade SEIRD to simulate the spread of COVID-19. Experimental results on five data sets demonstrate the effectiveness and the superiority of the proposed method.

ACKNOWLEDGEMENT

This research was supported in part by NSFC under Grant Nos. U1836107, 61972111, 61602132 and 61572158.

REFERENCES

[1] P. van den Driessche and J. Watmough, "A simple sis epidemic model with a backward bifurcation," *Journal of Mathematical Biology*, vol. 40, no. 6, pp. 525–540, 2000.

[2] M. Y. Li, H. L. Smith, and L. Wang, "Global dynamics of an seir epidemic model with vertical transmission," *SIAM Journal on Applied Mathematics*, vol. 62, no. 1, pp. 58–69, 2001.

[3] I. Korolev, "Identification and estimation of the seird epidemic model for covid-19," *Binghamton University. http://dx. doi. org/10.2139/ssrn*, vol. 3569367, 2020.

[4] Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, Z. Mai *et al.*, "Modified seir and ai prediction of the epidemics trend of covid-19 in china under public health interventions," *Journal of Thoracic Disease*, vol. 12, no. 3, p. 165, 2020.

[5] C. Hou, J. Chen, Y. Zhou, L. Hua, J. Yuan, S. He, Y. Guo, S. Zhang, Q. Jia, C. Zhao *et al.*, "The effectiveness of quarantine of wuhan city against the corona virus disease 2019 (covid-19): A well-mixed seir model analysis," *Journal of medical virology*, 2020.

[6] S. He, Y. Peng, and K. Sun, "Seir modeling of the covid-19 and its dynamics," *Nonlinear Dynamics*, pp. 1–14, 2020.

[7] L. López and X. Rodo, "A modified seir model to predict the covid-19 outbreak in spain and italy: simulating control scenarios and multi-scale epidemics," *Available at SSRN 3576802*, 2020.

[8] Z. Cai and N. Vasconcelos, "Cascade r-cnn: High quality object detection and instance segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1–1, 2019. [Online]. Available: http://dx.doi.org/10.1109/tpami.2019.2956516

[9] M. Salvaris, D. Dean, and W. H. Tok, "Generative adversarial networks," *Deep Learning with Azure*, p. 187–208, 2018. [Online]. Available: http://dx.doi.org/10.1007/978-1-4842-3679-6_8

[10] W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li, "Why is it difficult to accurately predict the covid-19 epidemic?" *Infectious Disease Modelling*, 2020.

[11] T. Rapolu, B. Nutakki, T. S. Rani, and S. D. Bhavani, "A time-dependent seird model for forecasting the covid-19 transmission dynamics," *medRxiv*, 2020.
[12] J. D. Cryer and K.-S. Chan, *Time series analysis: with applications in R*. Springer Science & Business Media, 2008.
[13] I. M. Anuar Ustayev, "Novel Coronavirus 2019 Data on Datahub," accessed 2020-04-14. [Online]. Available: https://datahub.io/core/covid-19