IJLAI Transactions on Science and Engineering (ISSN: 3006-3418) 2024, vol. 2, no. 1, pp: 24–31

Prediction of Disease Transmission Risk in Universities Based on SEIR and Multi-hidden Layer Back-propagation Neural Network Model

Jiangjiang Li[1] and Lijuan Feng[1]

School of Electronic and Electrical Engineering, Zhengzhou University of Science and Technology
Zhengzhou 450064 China
1124905128@qq.com;857003841@qq.com
Corresponding author:Lijuan Feng

Abstract. Against the background of regular epidemic prevention and control, in order to ensure the return of teachers to work, students to return to school and safe operation of schools, the risk of disease transmission is analyzed in key areas such as university canoons, auditoriums, teaching buildings and dormitories. The risk model of epidemic transmission in key regions of universities is established based on the improved SEIR model, considering the four groups of people, namely susceptible, latent, infected and displaced, and their mutual transformation relationship. After feature post-processing, the selected feature parameters are processed with monotone non-decreasing and smoothing, and used as noise-free samples of stacked sparse denoising automatic coding network to train the network. Then, the feature vectors after dimensionality reduction of the stacked sparse denoising automatic coding network are used as the input of the multi-hidden layer back-propagation neural network, and these features are used as tags to carry out fitting training for the network. The results show that the implementation of control measures can reduce the number of contacts between infected people and susceptible people, reduce the transmission rate of single contact, and reduce the peak number of infected people and latent people by 61% and 72% respectively, effectively controlling the disease spread in key regions of universities. Our method is able to accurately predict the number of infections.

Keywords: Disease transmission, SEIR model, Prediction, Stacked sparse denoising automatic coding network.

1. Introduction

The global epidemic of Corona Virus Disease 2019 is characterized by its high speed, wide range of infection and great difficulty in prevention and control. In June 29, 2020, there were 85,227 confirmed cases of Corona Virus Disease 2019 in China, there were 4648 deaths and 80055 cured cases. At present, the prevention and control of the epidemic situation has turned to normalization, the resumption of work and production has been carried out in an orderly manner, and the call for the resumption of school and classes in colleges and universities has become more frequent [1,2]. Different from primary and secondary schools, the key areas of colleges and universities (canteens, dormitories, teaching buildings, etc.) are typical places with high concentration of personnel, close contact and intensive risk points.If the risk management and control measures are not put in place, it may cause the virus to reappear, and then spread quickly, trigger the spread of large-scale epidemic [3,4]. The National Health Commission and the Ministry of Education jointly issued the technical plan for the prevention and control of the disease epidemic in colleges and universities, with emphasis on the prevention and control of the epidemic in key areas of schools. Therefore, to carry out risk analysis in the key areas of colleges and universities and formulate specific management and control programs will help to ensure the resumption of teachers, students and the safe operation of schools [5].

Risk assessment and risk transmission are two main aspects of the risk analysis and control of the epidemic situation of NCP. In the aspect of risk assessment, Ma et al. [6] used GIS spatial analysis technology, analytic hierarchy process and comprehensive weighting method to assess the risk of epidemic spread in Zhengzhou. Chauhan et al. [7] applied Delphi, AHP and other methods to evaluate the disease epidemic risk in schools. Gong et al. [8] used the combination model of EEMD-NAR neural network to forecast the public sentiment of emergent infectious diseases. Wex et al. [9] analyzed the application scenarios of each decision method of emergency action plan, and put forward a decision framework of emergency action plan and scheduling. Risk assessment is mostly static, so it is difficult to reveal the dynamic characteristics of epidemic spread, so it is necessary to carry out risk spread research. In the aspect of risk transmission, Shammi et al. [10] took the scenic spot as the object, analyzed the transmission route of the epidemic situation, and put forward risk control measures for four high-risk scenarios. Based on the cellular auto-method, Chen et al. [11] established a SIR epidemic model with mobility and population migration. Zakary et al. [12] used classic SIR model, differential recursive method and Logistic model to analyze

the epidemic spread of new pneumonia in Wuhan before and after the city closure. Based on the Si-SIR model, Mohammadi et al. [13] carried out the simulation of the spread in a limited area, and put forward the corresponding suggestions. Kudryashov et al. [14] predicted the spread by improving the classic SIR epidemic model. Basing on the study of the classic SIR epidemic model, the infectious characteristics of the latent population in epidemic were not taken into account. However, the research based on SEIR model mentioned above is seldom aimed at the problem of returning to work and study. Therefore, we take the key areas of colleges and universities as the research object, consider that the latent population has the infectious ability, our method improves the standard SEIR model, carries on the research to the disease epidemic situation campus spread risk. In order to provide a reference for the prevention and control of epidemic situation in colleges and universities, the differences between management and control measures will be compared and analyzed.

2. Proposed Risk Model of Epidemic Transmission in Key Regions of Universities

2.1. SEIR Model

SEIR model is an improvement on the classic SIR model [15]. It divides the population into four categories: 1. Susceptible, it means a person who does not have a disease.But he lacks immunity and is susceptible to infection after contact infected person. 2. Exposed, people who have been infected but no signs of infection; 3. Infective, a person who is infected with an infectious disease and can infect it to susceptible people; 4. Removed, a person who has been cured or has died. So it is no longer involved in the process of infection. The transition between the four groups is shown in Figure 1.

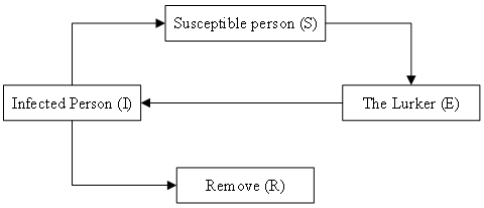

Fig. 1. Transformation relationships between populations in the SEIR model

3. Sparse Denoising Automatic Coding Network

Auto-Encoder (AE) networks are a kind of unsupervised learning field [16], which reconstructs input signals through the process of encoding and decoding to obtain higher-order representation of input signals. The input and output of the most common self-coding networks are exactly the same, but denoising construction has been added in this paper to make them denoising Auto-Encoder (DAE) [17]. The simplest DAE and AE are composed of input layer, hidden layer and output layer, and their network structure is shown in Figure 2.

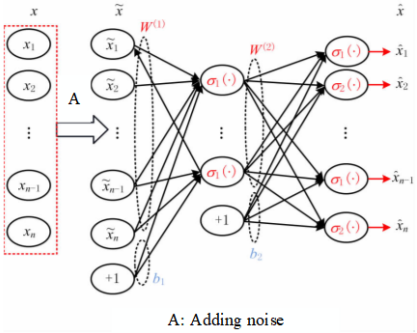

Fig. 2. Sparse denoising automatic coding network

The difference between DAE and AE is that DAE uses feature vector containing noise as input during training. This noise can be artificially added, or it can be embedded in the data itself. The output label is the original feature vector (without adding noise). Therefore, DAE must learn to remove noise to obtain input features that are not polluted by noise. This forces the coding network to learn more robust feature expression of the input signal, that is, to have stronger generalization ability.

For DAE, assuming that sample set $S = x_n|n = 1, 2, \cdots, n$ without noise. After adding noise, sample set $\tilde{S} = \tilde{x}_n|n = 1, 2, \cdots, n$. Encoding network's initial weight is $W^{(1)}$, bias is b_1. The initial weight of the decoding network is $W^{(2)}$ and the offset is b_2. The encoding process of the network can be expressed as:

$$h = g_{\theta_1}(\tilde{x}) = \sigma_1(W_1\tilde{x} + b_1). \tag{1}$$

The decoding process from the hidden layer to the output layer can be expressed as:

$$\tilde{x} = g_{\theta_2}(h) = \sigma_2(W_2\tilde{x} + b_2). \tag{2}$$

g_{θ_1} and g_{θ_2} are activation functions of encoding network and decoding network respectively, they are usually Sigmod function, identity function or Relu function. It is worth noting that DAE tries to make the output of noise-free samples even if the input is a noise-containing sample. Therefore, the loss function of DAE and AE is the same, which can be expressed as equation (3). It is the same gradient descent method that minimizes the error of x and $\tilde{x}$.

$$J_{DAE}(W, b) = \frac{1}{m}\sum_{r=1}^{m} 0.5||\hat{x}_{(r)} - x_{(r)}||^2. \tag{3}$$

Where m is the number of training samples. The training process of DAE is to constantly adjust the parameter set W_1, W_2, b_1, b_2 so that $J_{DAE}(W, b)$ can obtain the minimum value.

As can be seen from the DAE structure diagram in Figure 2, when the number of neurons in the hidden layer is smaller than the number of neurons in the input layer, the network can compress the dimensions of sample data. When the number of neurons in the hidden layer is larger than the number of neurons in the input layer, suppression of most neuron nodes in the hidden layer is required to achieve data dimension compression, which is also the original design intention of Sparse Auto-Encoder (SAE) [18].

The sparsity of sparse automatic coding networks can be understood as follows: assuming that the output range of neurons is 0-1 (Sigmod is used for activation function), when the output of neurons is 0 or close to 0, the neurons are considered to be inhibited. When the output of a neuron is 1, it is considered to be activated. Therefore, when most of the neurons in a certain layer of the neural network are in a suppressed state, the layer is considered to be sparse. On the basis of the three-layer DAE shown in Figure 2, this paper transforms it into Sparse Denoising Auto-Encoder (SDAE), and the average activation degree of neuron in the hidden layer.

$$\hat{\rho}_j = \frac{1}{m}\sum_{i=1}^{m}[a_j^{(2)} \cdot x^{(i)}]. \tag{4}$$

Where m is the number of training data sets. Assuming $\hat{\rho}_j = \rho$, under ideal conditions, when $\rho = 0$, the sparsity of the hidden layer will reach the best. However, in order to improve the efficiency of training, ρ is generally set as a small value, such as 0.05. To achieve this limitation, an additional penalty factor needs to be added to the reconstruction error loss function of DAE. The purpose of adding punishment factors is to punish those neurons with high activity, inhibit their expression, and finally achieve the sparse construction of the entire hidden layer. There are many options for punishment factors. This paper uses the most common KL divergence as punishment factor, as follows:

$$KL(\rho||\hat{\rho}_j) = \rho \times \log(\frac{\rho}{\hat{\rho}_j}) + (1 - \rho) \times \log(\frac{\rho}{\hat{\rho}_j}). \tag{5}$$

Finally, the reconstruction error loss function of SDAE can be expressed as:

$$J_{SDAE}(W, b) = J(W, b) + \beta\sum_{j=1}^{s_2} KL(\rho||\hat{\rho}_j). \tag{6}$$

3.1. Stack Self-coding Network

Generally, the self-coding network is a simple three-layer structure, belonging to the shallow neural network model, and its learning ability is limited when faced with the problem of dimensionality reduction of high-dimensional feature vectors. In order to increase the learning ability of the self-coding network, the way of deepening the level is generally adopted, but all shallow neural networks, including the ordinary self-coding network, adopt the back propagation of error. Therefore, the deeper the layer, the smaller the propagation error and the greater the possibility of the gradient disappearing. Stacked Auto-encoder (SA) can avoid this problem and directly use the feature values after dimension reduction for secondary training, which can be deepened at any level theoretically. Taking SA with two hidden layers as an example, its training process is as follows:

1. Train an auto-encoder to obtain the first-order feature of the original data representing $h^{(1)}$.
2. The feature $h^{(1)}$ output from the previous step is taken as the input of the next auto-encoder, and the second-order feature representation $h^{(2)}$ of the original data is obtained.
3. The feature $h^{(2)}$ output from the previous step is taken as the final dimension reduction vector, which can be used as the input feature vector of the neural network in the future, and the final classification or fitting task is completed.
4. The combination of these three layers forms a stacked self-coding network containing two hidden layers.

For this paper, the stack construction is applied to the SDAE model, and finally the SSDAE model is established for deep dimension reduction learning, so that better dimension reduction feature vectors can be obtained as the input of the subsequent multi-hidden layer BP neural network.

3.2. Model Construction

For the simplification, the initial letters S, E, I and R of the four groups of people are used to represent the number of the four groups of people respectively (S is susceptible, E is latent, I is infected, and R is removed). The rules of the dynamic change of the number of the four groups of people over time can be expressed by the following ordinary differential equation:

$$\frac{dS}{dt} = -\beta_1 \frac{SI}{N} - \beta_2 \frac{SE}{N}. \tag{7}$$

$$\frac{dE}{dt} = \beta_1 \frac{SI}{N} + \beta_2 \frac{SE}{N} - \alpha E. \tag{8}$$

$$\frac{dI}{dt} = \alpha E - \gamma I. \tag{9}$$

$$\frac{dR}{dt} = \gamma I. \tag{10}$$

$$N = S + E + I + R. \tag{11}$$

4. Experimental Analysis

We select one university as the research object, and use MATLAB to calculate the parameters of the model and simulate the prediction. The number of this university is as follows: a total of 35,000 full-time undergraduates, 15,000 master students, 2,000 doctoral students, 1,500 foreign students, 5,000 faculty and staff. There are teaching buildings, dormitories, dining halls, auditorium and other typical crowded places.

1. No control measures. The number of people eating in the dining hall is 500 each time, and there are infected people among the dining crowd. The initial number of infected people is 2, and the number of latent people is 0. The virus transmission among the population in 120 days after eating in the dining hall without any control measures (T=120) is simulated, as shown in Figure 3. As can be seen from Figure 3, the peak of viral infection was around day 25, with the peak number of infected people exceeding 200.

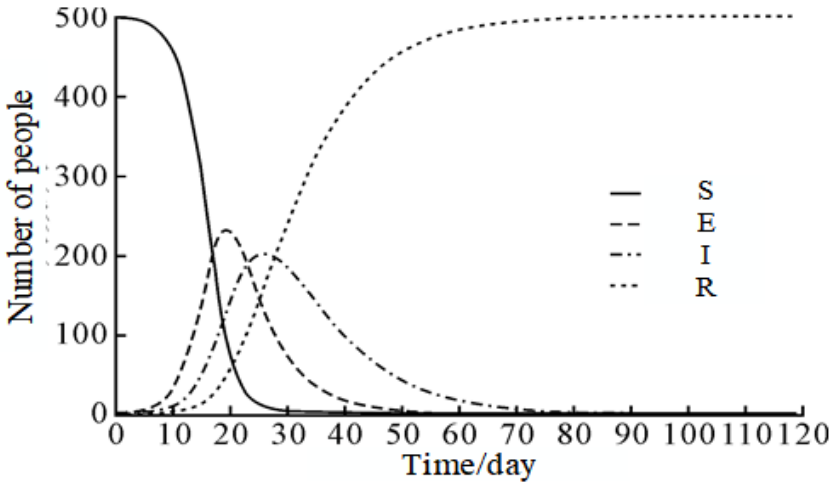

Fig. 3. The spread of the novel coronavirus without control measuresl

2. With control measures. Take measures such as cross-peak dining, one person, one dining position, increasing the distance in the dining queue, etc., to reduce the interaction and gathering among people and reduce the value. People entering and leaving high-risk areas must wear masks, wash their hands before and after meals, and canteen management staff should carry out regular sanitization of dining places every day to reduce the infection rate . Strict control measures can reduce the number of infected people and latent people who contact susceptible people every day to half of the original number, and reduce the of the infection rate by 1/3. As shown in Figure 4, the spread of the novel coronavirus when the control measures are adopted, it can be seen that the peak of the virus infection is around the 75th day, and the peak number of infected people drops to around 80.

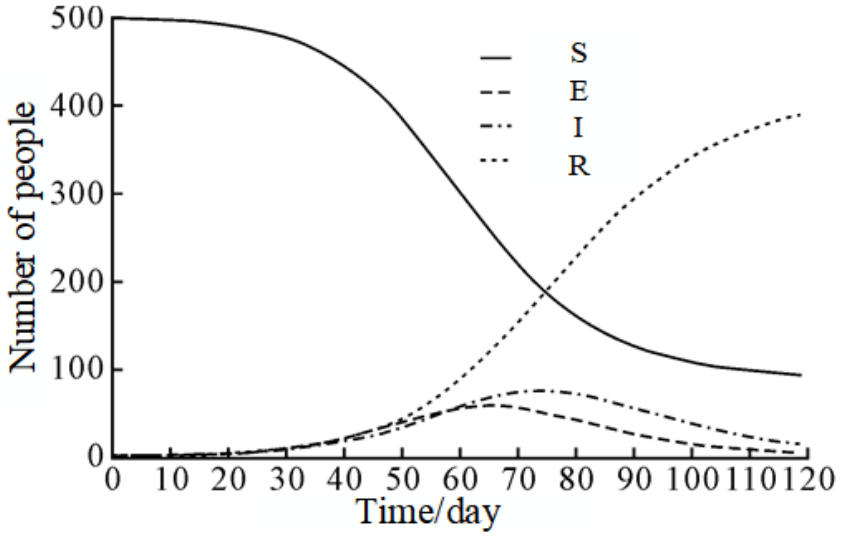

Fig. 4. The spread of the novel coronavirus with control measures

3. Comparison of control measures in key regions of colleges and universities. The simulation results of SEIR model in key regions of universities are shown in Table 1. It can be seen that without any control measures, the spread of the novel coronavirus develops exponentially. After the control measures are taken, the number of contacts between infected persons and susceptible persons is reduced, the transmission rate of a single contact is reduced, the spread of the virus is effectively controlled, and the peak number of infected persons and latent persons is reduced by 61% and 72% respectively (Without control measures and With control measures).

Table 1. Type Sizes for Camera-Ready Papers

Key areas of universities	Peak time of infection	Peak number of infections	Peak time of infection	Peak number of infections
Canteen (N=500)	Day 25(Without)	200(Without)	Day 75(With)	80(With)
Auditorium (N=3000)	Day 30(Without)	1200(Without)	Day 95(With)	400(With)
Teaching building (N=1500)	Day 25-30(Without)	600(Without)	Day 85(With)	$<$ 200(With)
Dormitory (N=1000)	Day 25-30(Without)	$>$ 400(Without)	Day 800(With)	150(With)

In this paper, the effect of SSDAE model with dimension reduction through deep learning is compared with Long Short Term Memory (LSTM) model, two traditional neural network models and two traditional machine learning models, which also used deep learning to predict disease.

Here, LSTM [19-21] adds door operation on the basis of RNN, alleviating the problem of gradient disappearance or gradient explosion that often occurs in RNN. Since the deep learning model can be used to process time series data, the collected vibration signals can be directly used as the input of the network, thus avoiding the trouble of artificial feature extraction [22]. However, the sampling frequency of the vibration sensor is as high as 50 KHz. In order to avoid the dimension of the input sample being too large, the tool walking data of each time is subsampled, so that the time step of LSTM is 100 and the data length of each moment is 10. In this example, the number of neurons in the hidden layer of LSTM is 42. In order to achieve the fitting function, it is necessary to add the fitting layer after it. Its complete network structure is shown in Equation (12):

$$input(100 \times 10) \rightarrow LSTM(42) \rightarrow regression(100 - 30 - 9 - 1). \tag{12}$$

The multi-hidden layer BPNN in the traditional neural network model has the same network structure as the BPNN part in SSDAE. The specific structure is 42-100-30-9-1. Another comparison object is Radial Basis Function Neural network (Radial Basis Function Neural network, 42-9-1. Because RBFN can contain only one hidden layer and the number of hidden layer neurons is generally not greater than that of the input layer, the RBFN network is 42-9-1 and uses the Gaussian function as its radial basis function.

Traditional Machine learning models include Linear Regression (LR) [23] and Least Squares Support Vector Machine (LSSVM), wherein LSSVM adopts Gaussian radial basis kernel function combined with particle swarm optimization algorithm to optimize parameter selection. In the experiment, 42-dimensional artificial features not compressed by SSDAE dimension are used as the input of traditional neural networks and traditional machine learning models to predict the infection of disease. At the same time, in order to explore the robustness of these models to face noise, the same Gaussian noise features are added as the input of each model, and the final prediction results are compared. The solution time of each model for the same data and RMSE for the test set are shown in Table 2, where the average solution time is the average solution time of 50 original inputs and 50 noise-added inputs.

Table 2. Comparison of different algorithms

Method	Average solution time/s	RMSE
SSDAE	0.45	1.35
SSDAE with noise	0.45	3.95
LSTM	0.62	3.48
LSTM with noise	0.62	9.29
BPNN	0.29	5.24
BPNN with noise	0.29	12.1
RBFN	0.16	8.38
RBFN with noise	0.16	14.5
LR	0.08	32.7
LR with noise	0.08	54.9
LSSVM	0.19	3.18
LSSVM with noise	0.19	10.01

As can be seen from Table 2, the prediction accuracy of the SSDAE model proposed in this paper is higher than that of LSTM network, traditional neural network and traditional machine learning algorithm with or without noise. In particular, SSDAE shows its robustness in the face of noise when facing noisy data. By comparing the prediction accuracy of each model before and after adding noise, it can be found that the RMSE of other algorithm models increased significantly after adding noise, while the RMSE of SSDAE model only increased

slightly. However, by comparing the solving time of the same problem, it can be found that SSDAE and LSTM models based on deep learning are obviously slower, but this is inevitable, because the number of parameters and operational complexity of deep learning models are much higher than other models. In conclusion, it can be proved that the SSDAE prediction model established in this paper not only has high accuracy, but also has certain robustness in the face of noise.

5. Conclusion

In view of the resumption of work and school in universities under the background of normalized epidemic prevention and control, the risk of epidemic transmission in key areas of universities was studied based on the improved SEIR model based on the canteen, auditorium, teaching building, dormitory and other key areas, and the control measures for key areas were proposed. (1) Considering the infectivity of the sleeper, the standard SEIR model was improved to increase the transformation relationship between the sleeper and the susceptible, and the risk model of epidemic transmission in key regions of universities was established based on the improved SEIR. (2) A denoising structure was added to the deep auto-encoding network, and the feature post-processing method was used to obtain theoretically smooth feature vectors, which were used as the output of SSDAE, forcing the network to learn how to remove noise and obtain more robust feature representation of the input signal. However, because the disease infection is affected by many factors, the prediction accuracy is still somewhat decreased. In future work, we will adopt more advanced deep learning methods to solve these problems.

6. Conflict of Interest

The authors declare that there are no conflict of interests, we do not have any possible conflicts of interest.

Acknowledgments. This paper was supported by Key research Project of higher education institutions in Henan Province (Project: Name: A Study on Students' concentration in Class Based on Deep Multi-task Learning Framework; Project No. 23B413004) and the Science and Technology Project No. 222102310222.

References

1. Alazab M, Awajan A, Mesleh A, et al, "COVID-19 prediction and detection using deep learning", International Journal of Computer Information Systems and Industrial Management Applications, Vol. 12(June), pp. 168-181, 2020.
2. Santosh K C, "COVID-19 prediction models and unexploited data", Journal of medical systems, Vol. 44, No. 9, pp. 1-4, 2020.
3. Mandal M, Jana S, Nandi S K, et al. A model based study on the dynamics of COVID-19: Prediction and control[J]. Chaos, Solitons & Fractals, 2020, 136: 109889.
4. Qingwu Shi, Shoulin Yin, Kun Wang, Lin Teng and Hang Li. Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation. Evolving Systems, Vol. 13, No. 4, pp. 535-549, 2022. DOI: 10.1007/s12530-021-09392-3
5. Xiaowei Wang, Shoulin Yin, Muhammad Shafiq, Asif Ali Laghari, Shahid Karim, Omar Cheikhrouhou, Wajdi Alhakami, Habib Hamam, "A New V-Net Convolutional Neural Network Based on Four-Dimensional Hyperchaotic System for Medical Image Encryption", Security and Communication Networks, vol. 2022, Article ID 4260804, 14 pages, 2022. https://doi.org/10.1155/2022/4260804
6. Ma Y, Lyu D, Sun K, et al. Spatiotemporal Analysis and War Impact Assessment of Agricultural Land in Ukraine Using RS and GIS Technology[J]. Land, 2022, 11(10): 1810.
7. Chauhan C, Akram M U, Gaur D. Technology-driven responsiveness in times of COVID-19: A fuzzy Delphi and fuzzy AHP-based approach[J]. International Journal of Global Business and Competitiveness, 2021, 16(1): 48-61.
8. Gong M. Analysis of internet public opinion popularity trend based on a deep neural network[J]. Computational Intelligence and Neuroscience, 2022, 2022.
9. Wex F, Schryen G, Feuerriegel S, et al. Emergency response in natural disaster management: Allocation and scheduling of rescue units[J]. European Journal of Operational Research, 2014, 235(3): 697-708.
10. Shammi M, Bodrud-Doza M, Islam A R M, et al. Strategic assessment of COVID-19 pandemic in Bangladesh: comparative lockdown scenario analysis, public perception, and management for sustainability[J]. Environment, Development and Sustainability, 2021, 23(4): 6148-6191.
11. Chen M, Li M, Hao Y, et al. The introduction of population migration to SEIAR for COVID-19 epidemic modeling with an efficient intervention strategy[J]. Information Fusion, 2020, 64: 252-258.
12. Zakary O, Rachik M, Elmouki I. On the analysis of a multi-regions discrete SIR epidemic model: an optimal control approach[J]. International Journal of Dynamics and Control, 2017, 5(3): 917-930.

13. Mohammadi V, Rahmani A M, Darwesh A M, et al. Trust-based recommendation systems in Internet of Things: a systematic literature review[J]. Human-centric Computing and Information Sciences, 2019, 9(1): 1-61.
14. Kudryashov N A, Chmykhov M A, Vigdorowitsch M. Analytical features of the SIR model and their applications to COVID-19[J]. Applied Mathematical Modelling, 2021, 90: 466-473.
15. Zare Z, Vasegh N. Modeling and analysis of the spread of the COVID-19 pandemic using the classical SIR model[J]. Journal of Control, 2021, 14(5): 89-96.
16. Shoulin Yin, Hang Li, Asif Ali Laghari, et al. A Bagging Strategy-Based Kernel Extreme Learning Machine for Complex Network Intrusion Detection[J]. EAI Endorsed Transactions on Scalable Information Systems. 21(33), e8, 2021. http://dx.doi.org/10.4108/eai.6-10-2021.171247.
17. Shen W. A Novel Conditional Generative Adversarial Network Based On Graph Attention Network For Moving Image Denoising[J]. Journal of Applied Science and Engineering, 2022, 26(6): 831-841.
18. Gao S. A two-channel attention mechanism-based MobileNetV2 and bidirectional long short memory network for multimodal dimension dance emotion recognition[J]. Journal of Applied Science and Engineering, 2022, 26(4): 455-464.
19. Chen W. A novel long short-term memory network model for multimodal music emotion analysis in affective computing[J]. Journal of Applied Science and Engineering, 2022, 26(3): 367-376.
20. Meng, X., Wang, X., Yin, S. et al. Few-shot image classification algorithm based on attention mechanism and weight fusion. Journal of Engineering and Applied Science. 70, 14 (2023). https://doi.org/10.1186/s44147-023-00186-9.
21. Orlandi M, Escudero-Casao M, Licini G. Nucleophilicity prediction via multivariate linear regression analysis[J]. The Journal of organic chemistry, 2021, 86(4): 3555-3564.
22. Teng L. Brief Review of Medical Image Segmentation Based on Deep Learning[J]. IJLAI Transactions on Science and Engineering, 2023, 1(2): 01-08.
23. Zhu B, Ye S, Wang P, et al. Forecasting carbon price using a multi-objective least squares support vector machine with mixture kernels[J]. Journal of Forecasting, 2022, 41(1): 100-117.

Biography

Jiangjiang Li is with the School of Electronic and Electrical Engineering, Zhengzhou University of Science and Technology. Research direction is cloud computing, computer application and AI.

Lijuan Feng is with the School of Electronic and Electrical Engineering, Zhengzhou University of Science and Technology. Research direction is computer application and AI.