Journal of Network Intelligence ©2021 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **6**, Number **2**, May **2021**

A Method of Prediction of Congestion Degree in Electric Bus Based on Running Data

Wei-Dong Fang

Fujian Provincial Key Laboratory of Automotive Electronics and Electric Drive
Fujian University of Technology
No.33 Xueyuan Road, University Town, Minhou, Fuzhou, Fujian, 350118, China
wdfang@126.com

Sheng Wang

Fujian Provincial Key Laboratory of Automotive Electronics and Electric Drive
Fujian University of Technology
No.33 Xueyuan Road, University Town, Minhou, Fuzhou, Fujian, 350118, China
∗Corresponding Author,wsheng2021@163.com

Jeng-Shyang Pan

College of Computer Science and Engineering
Shandong University of Science and Technology
No.579 Qianwangang Road, Huangdao District, Qingdao, Shandong Province, 266590, China
Chaoyang University of Technology
No.168 Jifeng E. Rd., Wufeng District, Taichung, 413310, Taiwan
jengshyangpan@gmail.com

Han-Lin Chen

Fujian Provincial Key Laboratory of Automotive Electronics and Electric Drive
Fujian University of Technology
No.33 Xueyuan Road, University Town, Minhou, Fuzhou, Fujian, 350118, China
1164885850@qq.com

Chao-Da Xu

Fujian Provincial Key Laboratory of Automotive Electronics and Electric Drive
Fujian University of Technology
No.33 Xueyuan Road, University Town, Minhou, Fuzhou, Fujian, 350118, China
chiudaatcheui@163.com

Received February 2021; revised April 2021

Abstract. *In view of the current lack of real-time and accurate passenger flow data in the public transportation network, it is difficult for residents to obtain information on the congestion degree in the bus before traveling. A Method for predicting the congestion degree in the electric bus based on the RBF (Radical Basis Function) neural network algorithm is proposed. This method uses the RBF neural network algorithm to build a congestion prediction model, and relies on the vehicle running data obtained by the vehicle terminal (Telematics BOX,T-Box) equipped with electric bus. After segmenting the running data according to the GPS location coordinates of each bus station, it is matched with the data of the number of people in the vehicle collected on the spot, and the data obtained after the matching is used as the training sample and the test sample of the prediction model. The experimental results show that this method can accurately predict the congestion degree. Compared with the traditional method of estimating the congestion degree in the bus, it is more economical and efficient.*

Keywords: Running Data, RBF neural network, Electric bus, Congestion degree prediction

1. **Introduction.** Bus travel has become one of the main ways of travel for urban residents. The phenomenon of overcrowding in the bus when passengers take the bus can be seen everywhere due to the absence of real-time and accurate passenger flow information, leading to the experience of passengers taking the bus has been decreased [1]. The key to solving these problems is to accurately and reliably estimate the information on the degree of congestion in the bus, as well as to enhance the passenger experience based on the above problems.

At present, numerous domestic researchers have a lot of research results on the prediction of the congestion degree in the bus [2]. Chen et al. [3] carried out a photosensitive sensor network which is installed on both sides of the front and back doors of the bus to conduct passenger flow statistics. When passengers get on and off the bus, pulses are generated through the sensor network of the door. In the meanwhile, the number of passengers in the car is counted by counting the number of optical pulses. Haq et al. [4] proposed a counting system based on stereo vision. Through the shooting of the camera in the car, a reliable way to extract the three-dimensional information of the passengers in the car is constructed. Then, the number of passengers is statistically calculated. Zheng and Chen [5] proposed a new vision-based pedestrian detection method, including pedestrian detection method and sitting posture detection according to the different human motion information of passengers standing or sitting on seats in the car. The pedestrian detection algorithm is divided into static pedestrian detection and dynamic pedestrian detection, improving the accuracy of the number of people in the car.

The above studies are all completed by installing equipment in the bus in order to complete the counting of passengers and predict the degree of congestion. When passengers get on and off the bus, there are large errors in the counting with expensive equipment. With the development of the Internet of Things and application of communication technologies, and with the popularization of electric bus in major cities, the level of intelligence of electric bus is constantly improving [6]. Additionally, the vehicle terminal has also become the standard configuration of electric bus [7]. During the operation of the vehicle, the vehicle terminal uploads vehicle running data in real time, generating massive amounts of running data [8]. At present, there are few applications of employing electric bus running data to predict the degree of congestion in the vehicle.

In the present study, the RBF neural network is used to match the running data of electric buses with the data on the number of people in the car collected on site, and use them as sample data for the prediction model to predict the degree of congestion in the electric bus. It is an innovative method for predicting the degree of congestion in the bus.

2. Related work.

2.1. RBF neural network. The RBF neural network was first proposed based on the corresponding characteristics of the human brain neuron cells. It can simulate the neural structure of the human brain that adjusts to each other and covers the receptive domain [9]. RBF neural network belongs to feedforward network in artificial neural network with its structure including three layers: input layer, hidden layer, and output layer. The structure is shown in the Figure 1.

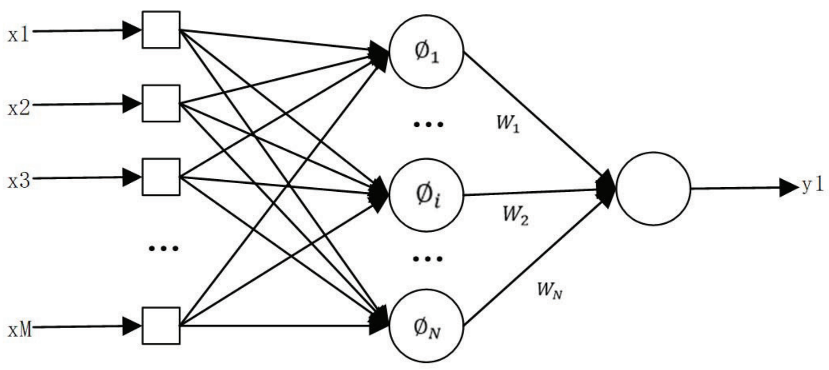


FIGURE 1. RBF neural network structure diagram

The input layer is composed of original input variables (signal source nodes). The number of neurons in the input layer is X= [x_1, x_2, ..., x_M], Each row in the X is used as the input to the input layer of the RBF neural network.

The radial basis function of the hidden layer in the RBF neural network has a lot of forms, generally we generally use Gaussian function as the radial basis function, as shown in Eq.(1).

$$\phi(x, c_i) = \exp(-\frac{\|x - c_i\|^2}{\delta^2}) \tag{1}$$

Where c_i denotes the center of the basis function, δ is the width of the basis function and the center and width of the Gaussian function represent important parameters of the radial basis function. The output of the hidden layer of the radial basis function neural network is as follows Eq.(2).

$$u_i(x) = R(\|x - c_i\|) = \exp(-\frac{\|x - c_i\|}{2\delta^2}) \tag{2}$$

The output layer linearly combines the output results of the hidden layers to obtain the final output value of the network. The number of neurons in the output layer is 1, which is the predicted value y. The mapping function is a linear function, which is a linear combination of the output results of each layer of the hidden layer by connecting the weights, as presented in Eq.(3).

$$y = \sum_{I=1}^{H} W_i U_i \tag{3}$$

The parameters of the RBF neural network are composed of center value, width and connection weight. The performance of the neural network is ultimately determined by the above three parameters, Therefore, it is of great importance to optimize the three parameters of the RBF neural network.

Regarding the RBF neural network, the model input layer only exerts the role of input data. Therefore, for the problem of predicting the degree of congestion in the bus, it is only necessary to select the appropriate data object as the input layer. The number of neurons in the output layer and the specific problems studied are associated with the function expression [11].

2.2. **Classification of congestion levels in bus compartments.** According to the specific parameters of the vehicle model, the size of the bus is 8450mm×2470mm×3180mm, the seat spacing is 650mm, and the width between front and rear door is above 650mm. The effective standing area in the bus can be calculated as 2.5m^2, according to the national standard. One square meter cannot exceed the data volume of 8 people. Therefore, the bus can accommodate up to 20 people. It is conceivable that the car can accommodate 20 standing passengers under ideal experimental conditions, which is difficult to achieve.

The bus studied in the present study is a 6-way electric bus operated by Quanzhou Passenger Transport Company, which is a TEG6851BEV model, at medium size. The specific parameters are shown in Table 1.

Table 1. Vehicle parameters

Model	BEV
Total mass (in kg)	8400
Outside dimensions (in mm)	8450×2470×3180
Approved number of passengers	35
Number of seats	15

This paper mainly refers to the judgment of the degree of congestion in the relevant literature for the classification of the degree of congestion in the car. Currently, the reference suggestions for determining the degree of congestion in the car put forward by relevant researchers are more detailed. After collating the various reference opinions [13], the standing area of the compartment occupied by the passengers in the car corresponds to the corresponding congestion level standard, which can be found in Table 2.

Table 2. Determining the congestion degree in the bus

Congestion degree	Passenger compartment area (in m^2)	Condition inside the car
Comfortable	Occupied area 0.5	Spacious interior space
Normal	Occupied area 0.35	The space in the car is not crowded
Comfortable	Occupied area 0.17	Slightly crowded in the car
Crowded	Occupied area 0.14	Passengers have a crowded psychology
Extremely crowded	Occupied area 0.12	Passengers cannot move and the environment is extremely poor

Based on the above classification criteria, when the bus model and internal seats remain known, only the real-time number of people in a certain bus can be known to determine the degree of congestion in the bus. Table 3 presents the crowding degree parameters corresponding to different levels of the number of people in the bus.

TABLE 3. Congestion level classification

Number of people	Crowdedness
<10	Loose
10-18	Normal
18-25	Mildly crowded
25-35	Heavy congestion

3. **Proposed method.** This paper proposes a method for predicting the degree of congestion in an electric bus based on the RBF neural network algorithm. Firstly, the actual bus running data of a specific line and the collected number of people in the bus are matched and preprocessed. In accordance with the GPS coordinate information of the station, the passenger data collected on the spot and the running data are matched as the input of the model, and the characteristic data with the passenger capacity is employed as the sample data to predict the real-time passenger capacity in the bus. The division relationship between the degree and the passenger capacity classifies the degree of congestion. This study uses the data of the number of people in the car collected on site for verification. At the same time, three standard errors of average absolute error MAE, average absolute percentage error MAPE, and root mean square error RMSE are used to measure the prediction results to evaluate the accuracy of the algorithm [14]. Figure 2 illustrates the specific steps of the method for predicting the degree of congestion in an electric bus.

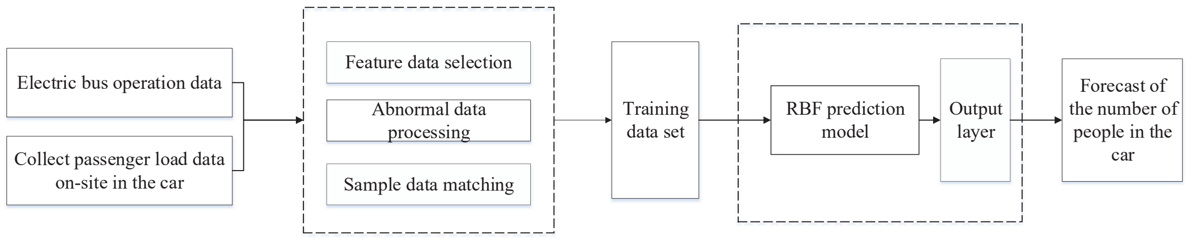
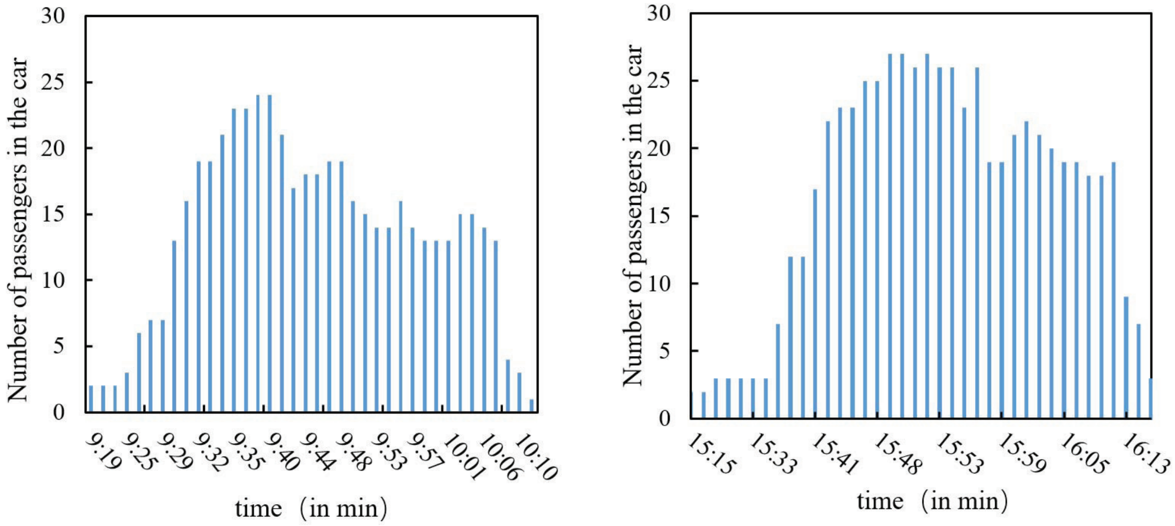


FIGURE 2. The concrete steps of the prediction method

As shown in Figure 2, the input vector of each sample and its target output value are defined. The input vector is the 7 types of feature data, and the target output value is the passenger capacity. Secondly, the first 4134 samples of the sample data are divided into training data, and the last 100 sample data are used as test data. The relative error result between the estimated value and the true value is displayed.

3.1. **Model.** This paper selects 6 bus lines in Quanzhou to predict the degree of congestion. Its operational characteristics are as follows: (1) The starting station is the front bus station and the last station is the Jiu ri shan bus station. There are totally 38 stations along the line. (2) The stations are densely populated, with mostly elderly passengers and students. It is a dedicated bus line that respects and loves the elderly, which starts and stops frequently.

The passenger capacity data in the car bases on the number of people between stations manually recorded on site. The total normal running time of 6 vehicles is 55 minutes, the entire bus line has 38 stations, and the collection time is one week. The maximum number of people in the car in the collected data is 28. Because the electric bus on the route is small, serious congestion in the car is more frequent, which is conductive to predicting the degree of congestion. According to the classification of the degree of congestion in Chapter 2, the number of people in the 4 types of congestion levels is collected. Here, we select two 6 bus lines to display the distribution of the passenger flow in the morning and the afternoon. Figure 3 shows the distribution of the number of people in the bus up and down.



(a) when the 6 bus goes up (b) when the 6 bus goes down

FIGURE 3. Passenger flow distribution characteristics of 6 bus line

As shown in Figure 3 above, the passenger flow distribution of the 6 bus lines is relatively regular, and the passenger flow is dense at the intermediate stops of the line. Based on the analysis of the actual road conditions, the 6 routes will be sent from the bus station in front of the departure station on the way for the reason that the bus departure station is relatively remote. The first few bus stops of the line are less crowded. After passing through scenic spots such as Fu wen Temple, Guan di Temple and West Lake Park, and passing through Pei yuan Middle School, this line is relatively busy with dense passenger flow. The collected passenger flow data is in consistence with the actual situation.

3.2. **Data preprocessing.** According to the 6 bus lines, the number of people collected on site is matched with the data during the driving of the vehicle, the bus station is identified according to the latitude and longitude information data, and the characteristic data between the stations is matched with the collected number of people. According to the position GPS coordinate information, the duplicate values and abnormal values that deviate from the bus line are eliminated. The input sample data of the prediction model total 4234, and some of the data are shown in Table 4.

TABLE 4. Model input data

Serial number	Motor input current (in A)	Motor speed (in rad)	Vehicle speed (in km/h)	Total battery voltage (in V)	Charge or discharge current (in A)	motor torque (in N·M)	SOC (in %)	Number of people
0	-0.005	0.334	0.333	0.569	0.02	0	0.56	0.23
1	0.099	0.268	0.245	0.473	0.123	0.313	0.56	0.23
2	-0.005	0.263	0.245	0.79	0.006	0.047	0.56	0.23
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
4232	-0.154	0.619	0.639	0.547	-0.156	0.292	0.35	0.615
4233	0.207	0.398	0.385	0.156	0.215	0.336	0.232	0.532

As shown in TABLE 4, when using the RBF radial basis function neural network to predict the degree of congestion in electric bus, the input variable is the running data generated during the operation of the electric bus [15], including the vehicle speed and the actual rotation of the motor. A total of 7 types of related characteristic data include torque, motor speed, lithium battery discharge current, motor input current, and SOC. The output variable is the real-time passenger capacity in the bus at different times. According to the above classification of the degree of congestion in the car, it will be predicted that the passenger capacity is converted to the degree of congestion. The input parameter is the sample value of the input layer, which is the main factor affecting the prediction of the number of people in the electric bus, that is, the determined vehicle data indicators generated by the above 7 electric bus during the operation process. Because the amount of data may be extremely large, it will affect the weight coefficients in the RBF neural network model [16,17]. Therefore, we need to standardize the input sample data. The standardization formula is shown in 4.

$$X = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

Where X is the normalized value, and the original value of X, refers to the maximum and minimum values of a certain characteristic parameter [22].

4. **Experiment.**

4.1. **Experimental model description.**

The standardized sample data is divided. The last 100 samples are used as test samples, and the data are extracted for offline training. The structure of the neural network is determined and the input node is 7, representing the dependent variable associated with the passenger load. The number of layer nodes is 20, and the number of output nodes is 1, which refers to the output value of passenger capacity. The initial values of the weights and centers in the RBF network model are arbitrary values [18]. Then, the learning algorithm of RBF network weights is used to obtain the optimized network weights and centers. Moreover, Matlab2018b programming is employed to get the error drop curve of RBF network.

The model setting error is 1e-6, the diffusion factor is 30, the maximum number of neurons is 100, and the function of creating a radial basis neural network is called. The number of neurons is increasing during training, and the training error is gradually reduced until the error requirements are satisfied, as shown in Figure 4.

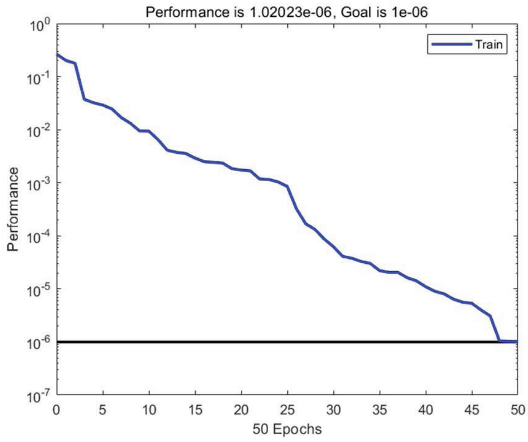


FIGURE 4. Error declining curve

From the above figure 4, it can be observed that when the training reaches the target value, only 50 steps are required, and the completion time of this process only takes 3s. Moreover, this well proves that RBF neural network has the advantages of larger convergence range and faster running speed.

4.2. **Experimental results.** Based on the RBF neural network model, a total of 100 sets of test data from the test group 4134 to 4234 are input into the model for prediction. The comparison between the prediction result and the real value is displayed in Figure 5.

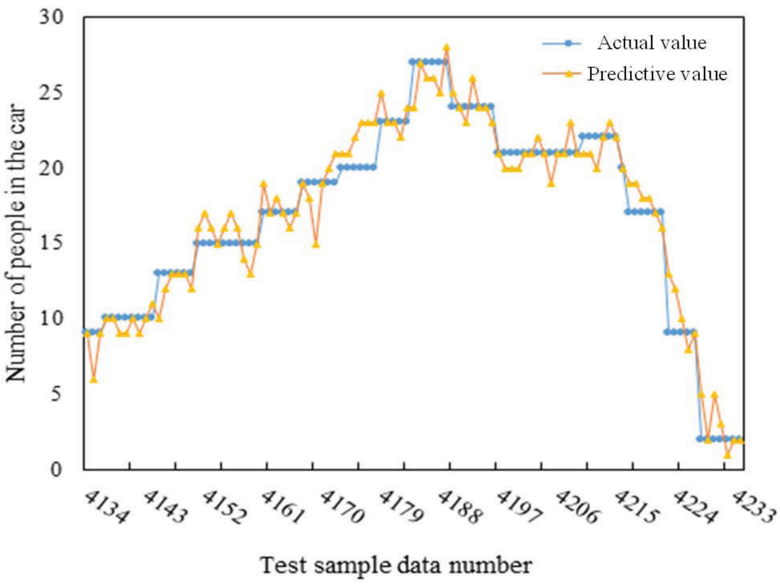


FIGURE 5. Comparison of predicted value with true value

It can be seen from Figure 5 that in the 4134-4234 test group data, the minimum number of real people in the car is 2 people, and the highest is 27 people. Based on the classification of the degree of congestion in the car presented in Chapter 2, each degree of congestion is involved. The number of people in the car predicted by the RBF network basically remains the same as the actual number of people in the car. The maximum error is 4 people, the average relative error is close to 10.13%, the root mean square error

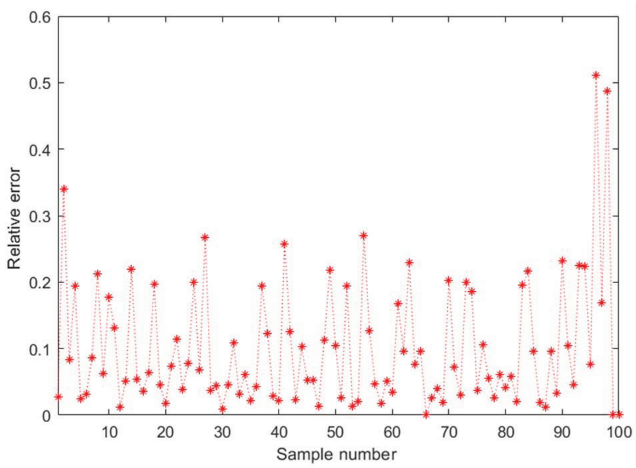


FIGURE 6. Relative error graph of predicted value and true value

(RSME) is 0.1463, the running time is 5.73s, and the relative error between the number of people collected and the predicted value, as shown in Figure 6.

The RBF radial basis function neural network is characterized by approximating a local network [19]. In the overall sample space of the model input, a small amount of parameters can be adjusted to achieve the accuracy requirements for samples in a local area. The neural network has a faster learning speed with better approximation than a nonlinear function and faster processing speed when the amount of data is large, and its generalization ability is good, which it can process data in parallel [20].

From the relative error of the predicted value and the real value predicted by the RBF neural network prediction model, the error is relatively large when the number of people in the bus is less than 10. Meanwhile, the error is relatively reduced when the number of people in the bus reaches 15 or more. From the average relative error, it can be found that the RBF neural network has a good prediction effect, which can basically meet the standard of high-precision prediction. According to the number of people in the bus predicted by the model, combined with the number of people in different congestion levels, it can be used in electric bus. The degree of congestion is estimated with high credibility.

4.3. **Evaluation.** For the number of people in the car predicted by the above RBF radial basis neural network prediction model, combined with the corresponding relationship between the number of people and the degree of congestion in the car, the predicted number of people is converted into a prediction of the degree of congestion, and the confusion matrix is applied to classify the degree of congestion as an evaluation. The specific operation of indicators is as follows [21].

By comparing the predicted results of 100 people with the actual number of people in the car, the predicted number of people and the actual number of people are transformed into the corresponding crowding level. There are four levels of crowding, respectively, loose, normal, lightly crowded, and severely crowded. Combined with the definition of the confusion matrix, the predicted number of people data is converted into the degree of congestion, and the above estimation model is evaluated. Table 5 is drawn as follows:

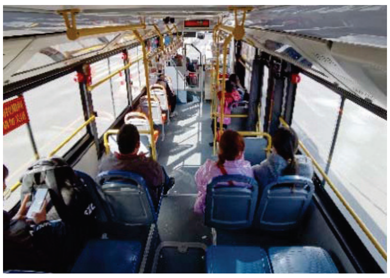
In a total of 100 samples, we have predicted 89 samples correctly. Obviously, the data are relatively concentrated on the diagonal. There is a confusion matrix accuracy rate formula to calculate the accuracy of the crowded prediction (Accuracy) of 89%. The corresponding samples of the four types of crowded prediction categories are selected, and the actual crowded conditions in the car are shown in Figure 8 below. The degree of

TABLE 5. Model evaluation

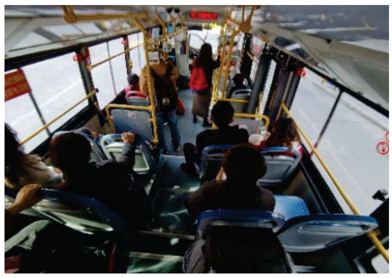
Confusion Matrix		Actual value			
		I	II	III	IV
Predictive value	I	12	4	0	0
	II	3	32	0	0
	III	0	2	40	1
	IV	0	0	1	5

congestion in the car predicted by the model is consistent with the degree of congestion in the bus collected on the spot, reflecting the superiority of the neural network in data prediction.

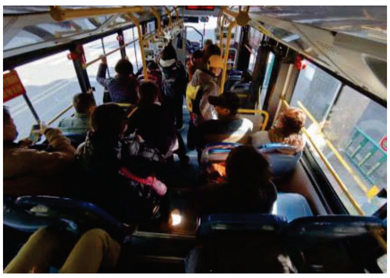
5. **Conclusion.** To conclude, in the present study, the RBF neural network algorithm is used to predict the degree of congestion in electric bus. Specifically, the degree of congestion in a specific bus is firstly classified. The running data of the vehicle are cleaned. Then, the data of the number of people in the vehicle collected on the spot are matched as the prediction sample data of the model. Using RBF neural network, a predictive model of the degree of congestion in the bus is established. Secondly, the information on the degree of congestion in the electric bus obtained by the proposed method is more economical and efficient than the traditional method of installing a counting device in the vehicle. Finally, the feasibility and effectiveness of the method are confirmed through actual predictions of specific bus lines.



(a) Dongmen Station



(b) Fuwenmiao Station



(c) Xinmen Street Station

(d) West Lake Park Station

FIGURE 7. Prediction of different levels of congestion in the car

Acknowledgment. The current work was supported in part by the Fujian Institute of Technology Fund GY-Z 13102 Research on Key Technologies of Intelligent Multi-function Display Terminals, the 2019 Ministry of Education Collaborative Education

Project 201901189009, the Development and application of Beidou/GPS dual-mode high-precision positioning technology 2013HZ0002-1, the Fujian Provincial Surveying and Mapping Geographic Information Industry Technology Public Service Platform 2014H2010, the Fujian New Energy Vehicle Supervision Service Platform GY-Z16014, the Development and Demonstration Application of New Energy Vehicle Supervision Service System in Fujian Province GY-Z160141, the Technical consultation on the integration of electric vehicles into modern power grids GY-H-17041, the Smart home related technical consultation GY-H-18133, the Technical consultation on the integration of electric vehicles into modern power grids GY-H-20115, the Fujian University of Technology and Fuyao Glass Industry Group Co., Ltd. Automotive Electronics and Electrical System Development Cooperation Project GY-H-20108.

REFERENCES

[1] G. G. Bunster, M. T. Torriti, and C. Oberli, Crowded pedestrian counting at bus stops from perspective transformations of foreground areas, *Iet Computer Vision* , vol. 6, no. 4, pp. 296–305, 2012.

[2] J. Z. Zhu, and Y. W. Lai, Research on the impact of road congestion on the reliability of bus travel time, *Traffic Engineering*, vol. 17, pp. 47–52, 2017.

[3] P. J. Chen, H. J. Chen, C. R. Zhang, and Z. Z. Tang, Real-time bus passenger flow statistics scheme based on light-sensitive wireless sensor network, *Chinese Control Conference*, vol. 35, pp. 59–63, 2016.

[4] E. U. Haq, H. R. Xu, X. H. Chen, and W. Q. Zhao, A fast hybrid computer vision technique for real-time embedded bus passenger flow calculation through camera, *Multimedia Tools and Applications*, vol. 79, no. 1, pp. 1007–1036, 2020.

[5] G. Zheng and Y. Chen, A review on vision-based pedestrian detection, *Global High Tech Congress on Electronics*, pp. 49–54, 2013.

[6] S. Lee, W. Choi, H. J. Jo, and D. H. Lee, T-Box: A Forensics-Enabled Trusted Automotive Data Recording Method, *IEEE Access*, vol. 7, pp. 49738–49755, 2019.

[7] F. Q. Zhang, T. Y. Wu, Y. Wang, R. Xiong, G. Y. Ding, P. Mei, and L. Y. Liu, Application of quantum genetic optimization of LVQ neural network in smart city traffic network prediction, *IEEE Access*, vol. 8, pp. 104555–104564, 2020.

[8] Y. Qu, L. Xu, J. H. Wang, X. K. Fang, and S. S. Gu, RBF neural network theory and its application in control, *World Congress on Intelligent Control and Automation*, vol. 20, pp. 40–42, 2006.

[9] D. J. Du, M. R. Fei, and L. X. Li, Radial-basis-function neural network based on fast recursive algorithm and its application, *Control Theory & Applications*, vol. 25, no. 5, pp. 827–830, 2008.

[10] G. M. Fetene, S. Kaplan, S. L. Mabit, A. F. Jensen, and C. G. Prato, Big data processing and analysis in the field of electric vehicles, *Transportation Research Part D: Transport and Environment*, vol. 54, pp. 1–11, 2017.

[11] C. Q. Zhang, R. N. Qi, and Z. W. Qiu, Comparing BP and RBF Neural Network for Forecasting the Resident Consumer Level by MATLAB, *International Conference on Computer and Electrical Engineering*, pp. 169–172, 2008.

[12] J. H. Li, Y. A. Tian, and X. J. Xu, Design and implementation of bus lane video image monitor system based on FPGA, *Chinese Control And Decision Conference (CCDC)*, pp. 169–172, 2008.

[13] Z. G. Ren, J. C. Chen, L. W. Ye, C. Wang, Y. Liu, and W. H. Zhou, Application of RBF Neural Network Optimized Based on K-Means Cluster Algorithm in Fault Diagnosis, *International Conference on Electrical Machines and Systems (ICEMS)*, pp. 2492–2496, 2018.

[14] G. S. Babu, X. L. Li, and S. Suresh, Meta-cognitive Regression Neural Network for function approximation: Application to Remaining Useful Life estimation, *International Joint Conference on Neural Networks (IJCNN)*, pp. 2492–2496, 2016.

[15] M. Gan, X. Y. Peng, and H. Peng, Two hybrid optimization algorithms for RBF neural network parameter estimation, *Control and Decision*, vol. 24, no. 8, pp. 1172–1176, 2009.

[16] J. J. Zhong, Z. Hong, and Z. Y. Tong, Research on controller parameter optimization design based on RBF neural network, *Missile and Space Vehicle Technology*, vol.374, no. 3, pp. 80–84, 2020.

[17] K. A. Zhang and K. Zhao, Research on greenhouse temperature prediction based on improved CFA PSO-RBF neural network, *Computer Applications and Software*, vol. 37, no. 6, pp. 95–99, 2020.

[18] G. Tong, J. l. Lan, and Y. F. Li, Network traffic prediction based on quantum adaptive particle swarm optimization radial basis function neural network, *Journal of Electronics and Information Technology*, vol. 35, no. 9, pp. 2220–2226, 2013.
[19] F. Lin, S. J. Liu, H. C. Chao, and J. S. Pan, Short-Term Household Load Forecasting Model Based on Variational Mode Decomposition and Gated Recurrent Unit with Attention Mechanism, *Journal of Network Intelligence*, vol. 6, no. 1, pp. 143–153, 2021.
[20] G. X. Zhou and J. J. Tang, Forecast of Urban Rail Transit Passenger Flow in Holidays Based on Support Vector Machine Model, *International Conference on Electromechanical Control Technology and Transportation (ICECTT)*, pp. 585–589, 2020.
[21] C. Chen, X. J. Ji, and Z. Chen, Performance evaluation of traffic video background reconstruction based on low-rank matrix recovery, *Computer Engineering and Design*, vol. 38, no. 5, pp. 1301–1308, 2017.