Information Processing and Management 60 (2023) 103171

Contents lists available at ScienceDirect

Information Processing and Management

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

Data information processing of traffic digital twins in smart cities using edge intelligent federation learning

Weixi Wang [a], Fan He [a], Yulei Li [b], Shengjun Tang [a], Xiaoming Li [a,*], Jizhe Xia [a], Zhihan Lv [c]

[a] *Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University and Shenzhen Key Laboratory of Spatial Smart Sensing and Services and MNR Technology Innovation Center of Territorial and Spatial Big Data and Guangdong–Hong Kong-Macau Joint Laboratory for Smart Cities, Shenzhen 518060, China*
[b] *Logistics Information Centre, Beijing 100842, China*
[c] *Department of Game Design, Faculty of Arts, Uppsala University, Sweden*

A R T I C L E I N F O

Keywords:
Digital twins cities
Deep learning
Traffic safety
Sign recognition
Edge computing

A B S T R A C T

The present work analyzes the application of deep learning in the context of digital twins (DTs) to promote the development of smart cities. According to the theoretical basis of DTs and the smart city construction, the five-dimensional DTs model is discussed to propose the conceptual framework of the DTs city. Then, edge computing technology is introduced to build an intelligent traffic perception system based on edge computing combined with DTs. Moreover, to improve the traffic scene recognition accuracy, the Single Shot MultiBox Detector (SSD) algorithm is optimized by the residual network, form the SSD-ResNet50 algorithm, and the DarkNet-53 is also improved. Finally, experiments are conducted to verify the effects of the improved algorithms and the data enhancement method. The experimental results indicate that the SSD-ResNet50 and the improved DarkNet-53 algorithm show fast training speed, high recognition accuracy, and favorable training effect. Compared with the original algorithms, the recognition time of the SSD-ResNet50 algorithm and the improved DarkNet-53 algorithm is reduced by 6.37ms and 4.25ms, respectively. The data enhancement method used in the present work is not only suitable for the algorithms reported here, but also has a good influence on other deep learning algorithms. Moreover, SSD-ResNet50 and improved DarkNet-53 algorithms have significant applicable advantages in the research of traffic sign target recognition. The rigorous research with appropriate methods and comprehensive results can offer effective reference for subsequent research on DTs cities.

1. Introduction

Nowadays, the new generation of information technology has become a critical support to social progression (Deng et al., 2021). At this stage, information technology is gradually revealing its significant value to society, and it has been applied to many fields such as the manufacturing industry, medical industry, and education industry (Khalid & Ameen, 2021). Information technology also plays a vital role in the construction and renewal of cities. Urban construction and progress are increasingly inseparable from digital

* Corresponding author.
E-mail address: lixming@szu.edu.cn (X. Li).

https://doi.org/10.1016/j.ipm.2022.103171
Received 30 July 2022; Received in revised form 8 November 2022; Accepted 13 November 2022
Available online 18 November 2022

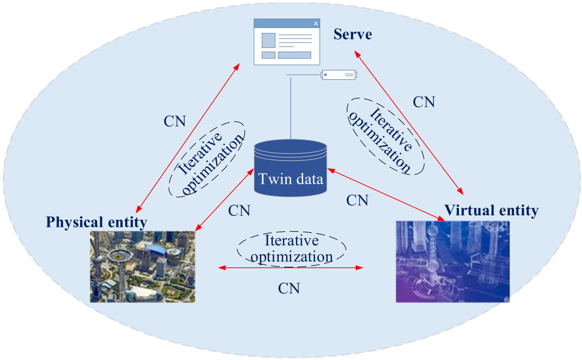

Fig. 1. Structure of the five-dimensional DTs model.

empowerment, and the construction of digitalized cities is the principal development direction of major cities in China and an inevitable requirement for enhancing the efficiency of urban governance (Camero & Alba, 2019). With the deepening of research and an increasingly extensive application, the concept of smart city is further improved and optimized. Digital twins (DTs) integrate a variety of technologies, such as machine learning, deep learning, and Internet of Things (IoT). In-depth discussion of these information technologies is of great research significance for the development and application of DTs (Fan et al., 2021). The smart city based on DTs has become the focus of researchers. Lu et al. (2020) took the West Cambridge site of the University of Cambridge as an example to illustrate the theoretical path of developing DTs at the city level. Francisco et al. (2020) explained the important role and significance of DTs in urban energy management. Elayan et al. (2021) discussed the application of DTs technology in the field of intelligent medical care in smart cities .

DTs city is a combination of DTs with urban construction, which builds a complex giant system of one-to-one correspondence, mutual mapping, and cooperative interaction between the physical world of the city and the virtual digital space. A twin city corresponding to the physical city is reconstructed in the virtual space to realize digitalization and virtualization of total urban elements, real-time and visualization of the whole state of the city, and coordination and intelligentization of urban management and decision-making (Pang et al., 2021). However, the research on DTs city is still in its infancy, so further research is of great theoretical value and practical significance. In the construction of DTs cities, the intelligent transportation system (ITS) is the most momentous part and attracts close attention (Yan et al., 2020), which realizes elementary functions such as dynamic discovery, automatic identification, scientific diagnosis, intelligent solution, and feedback optimization (Aujla et al., 2018). Besides, it provides a series of intelligent services, such as intelligent state extraction, intelligent pattern recognition, intelligent jump detection, intelligent situation diagnosis, intelligent collaborative optimization, efficiency balance regulation, spatio-temporal information aggregation, and fusion strategy, to improve the operation effect and intelligence of urban traffic. At present, studies on intelligent transportation generally focus on unmanned driving, intelligent road perception, and environment recognition (Aamir et al., 2019). In addition, edge intelligence provides services such as advanced data analysis, scene perception, real-time decision-making, self-organization and collaboration provided by edge nodes on the edge side, while edge computing can calculate, apply, store and communicate the data information collected on the edge side. This makes it possible to respond more quickly to the requests of edge devices in the process of Internet of Vehicles in the construction of smart city traffic (Jararweh et al., 2020). Therefore, current research should focus on optimizing the accuracy of intelligent road environment recognition and detection, so as to improve the efficiency of intelligent perception and effectively improve the information security performance of intelligent transportation networks (Berdik et al., 2021).

Based on the actual situation at the present stage, the present work principally studies environmental perception in the field of intelligent transportation in DTs cities. First, the theoretical framework of smart city and transportation model based on DTs is constructed based on the concept of DTs. In addition, the deep learning network is adopted and optimized, edge computing algorithm is introduced, and the data information on the edge side of the transportation network is collected and calculated in real time for intelligent identification of the transportation network. Moreover, experiments are carried out to verify the specific optimization effect and target recognition effect of the algorithm. Through the discussion of the theoretical model of DTs as well as related theories of smart cities and ITS, the present work proposes the conceptual models of technological DTs city and DTs transportation. Furthermore, the Single Shot MultiBox Detector (SSD) algorithm is optimized by the residual network, form the SSD-ResNet50 algorithm, and DarkNet-53 algorithm is also optimized. Then, the training effect, instantaneity performance, target recognition accuracy, and data enhancement effect are analyzed through experiments. The purpose of the present work is to provide scientific and effective reference for the subsequent research on DTs city and intelligent transportation model.

2. Theoretical foundations and algorithm optimization scheme

2.1. Accuracy evaluation of communication big data node detection algorithm

DTs achieve interaction and feedback between physical systems in the real world and digital systems in the virtual world. Through data collection, mining, storage, and calculation, it ensures collaboration and adaptation between physical and digital systems in the

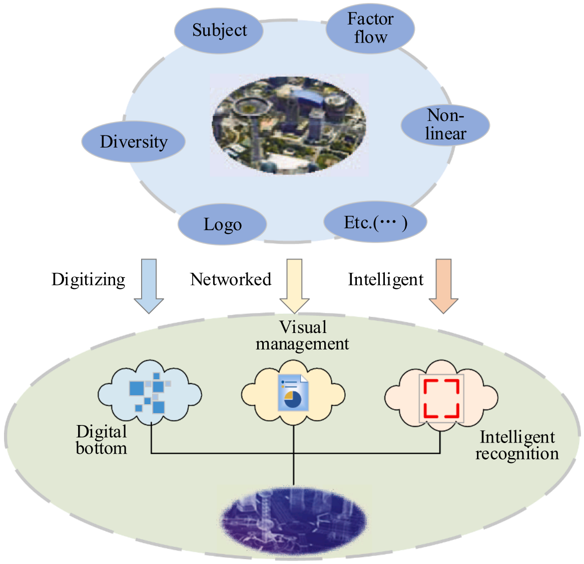

Fig. 2. Conceptual framework of the DTs city.

full life cycle. With technology advancement and extended application, the model structure of DTs is further deepened. For example, Tao et al. (2019) defined a five-dimensional DTs model, as shown in Eq. (1).

$$M_{DT} = (PE, VE, Ss, DD, CN) \tag{1}$$

In Eq. (1), *PE* denotes a physical entity, *VE* refers to a virtual entity, *Ss* signifies a service in the DTs model, and *CN* represents the connection among various components.

Fig. 1 displays the structure of the five-dimensional DTs model.

DTs primarily include PE, VE, twin data, services, and connections. PE is the basis and premise of DTs, which generally includes three levels, namely the Unit PE, System PE, and System of systems PE. VE can be expressed as Eq. (2).

$$VE = (G_v, P_v, B_v, R_v) \tag{2}$$

In Eq. (2), *Gv* stands for the geometric model, *Pv* signifies the physical model, *Bv* refers to the behavior model, and *Rv* indicates the rule model.

Service is classified into new function service and business service. Twin data is the data base of the DTs model, including physical entity data, virtual entity data,service data, knowledge data, and fusion derived data. The connectivity can realize the interconnection of the various parts of the DTs structure, as shown in Eq. (3).

$$\begin{aligned} CN = (&CN - PD, CN - PV, \\ &CN - PS, CN - VD, \\ &CN - VS, CN - SD) \end{aligned} \tag{3}$$

In Eq. (3), *CN-PD, CN-PV, CN-VD, CN-VS*, and *CN-SD* refer to the connection between PE and DTs data, PE and VE, PE and services, VE and DTs data, VE and services, or between DTs data and services, respectively.

The city is an extremely complex man-made system, and it is technically difficult to study the modeling of the urban system through traditional methods (Ham & Kim, 2020). The five-dimensional DTs model has obvious universality and has achieved excellent application effects in different fields and different objects. The five-dimensional DTs model can cooperate with IoT, big data, artificial intelligence, and other new-generation information technologies, which is highly consistent with the construction concept of smart cities (Wu et al., 2021). Therefore, a DTs city which is almost the same as the urban physical entity is established based on the five-dimensional DTs model and other related concepts, to realize the real-time connection and dynamics between the physical city and the virtual city (Qiao et al., 2019). The five-dimensional DTs model can analyze the unified data to track and identify urban dynamic changes, and make urban planning and management more in line with the law of urban development. The DTs city meets the needs of physical system integration, physical data fusion, virtual and real bidirectional connection, and interaction of information (Minerva et al., 2020).

Therefore, the conceptual framework of the smart city is designed based on DTs and the practical needs of smart city development, as presented in Fig. 2.

Fig. 2 indicates that the DTs city starts with the comprehensive digitalization of subject, diversity, and factor flow. It can realize the integration and sharing of urban government data resources and social data resources, form the organic unification of human production, living, and ecological data, and build the digital city mirror integrating humans, machines, and objects. Under the DTs city, digitalization will become the key basis of urban development, which will pay attention to the dynamic learning and adaptation of

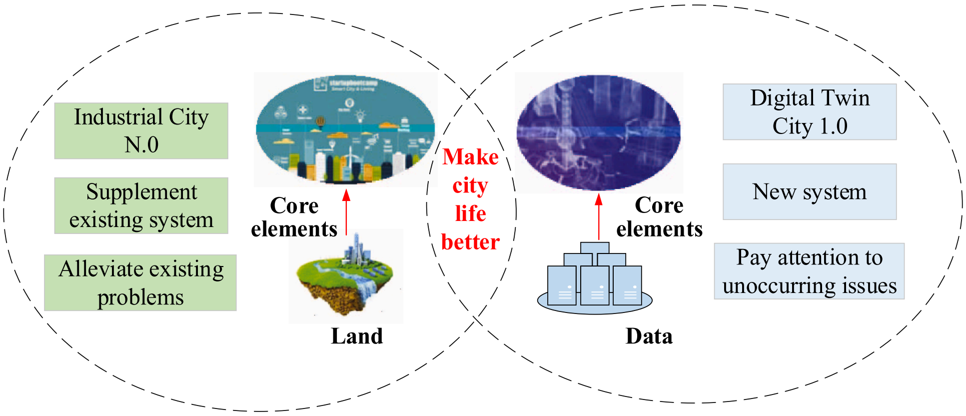

Fig. 3. The difference between DTs cities and traditional smart cities.

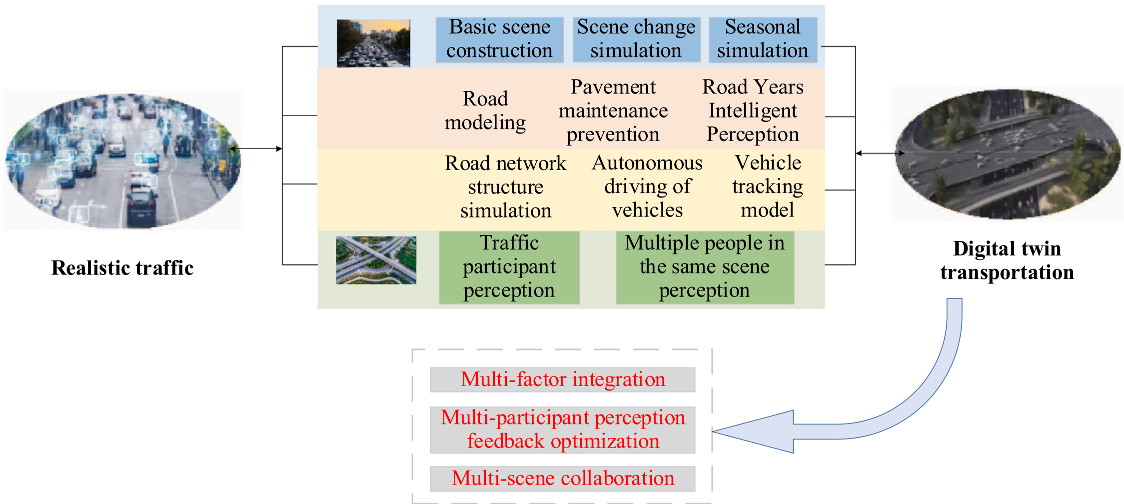

Fig. 4. Conceptual model of the traffic sensing system on DTs.

society and promote the intelligent coupling of various modules of cities.

The DTs cities reported here are significantly different from the traditional smart cities. The specific differences are shown in Fig. 3.

The differences between DTs cities and traditional smart cities are as follows. Firstly, DTs city performs a unified digital record and presentation of material city and its economic and social characteristics, and meanwhile conforms to the law of urban adaptation and self-organization. Secondly, DTs city is the symbiosis of the virtual and the real, which aims to realize the digitalization of all elements of the city, the visualization of all states, and the intelligence of all activities through constructing a virtual city. Finally, DTs cities always pay attention to the continuous development and progress of the urban ecosystem, while traditional smart cities are more limited to a part of a city or a certain stage. In summary, there are significant differences between DTs cities and traditional smart cities in underlying logic, governance concepts, and technical solutions. However, the essential purpose of DTs city and smart city construction is the same, that is, to make the city better.

2.2. Design of intelligent traffic perception system based on edge computing combined with DTs

The transportation system is a critical component of urban life, which is related to the normal life of residents and the progression of the city (Bansal et al., 2020). The digitalization degree of the transportation system also reflects the digitalization process of cities to some extent. By applying edge computing to data detection in intelligent transportation, roadside sensors and on-board sensors can collect real-time information such as the surrounding environment and vehicle driving conditions, and transmit them to the edge server through the roadside micro cloud and base station. The edge server then maximizes the real-time traffic data stream it collects while offline. After the edge server completes the preliminary processing, the results of the preliminary processing are pushed to the centralized central platform in the 5G core network for caching or further intelligent processing. Finally, the prediction results will be fed back to the vehicle unit in time.

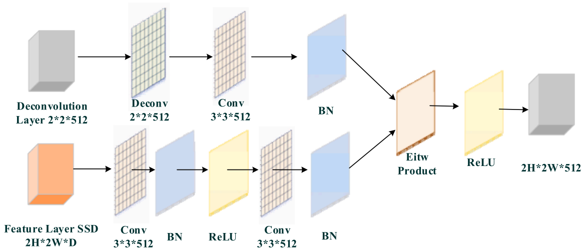

Fig. 5. Structure of the deconvolution module.

At the present stage, the research on DTs traffic model has raised attention, which focuses on the construction method of DTs dynamic traffic scenes (An & Wu, 2020). The DTs technology is further introduced into it, the DTs-based traffic sensing system mainly includes four aspects,. i.e., people, cars, roads, and the surrounding environment. Therefore, in the construction of DTs, it is essential to consider the influence of multiple factors rather than a single factor (Schintler & McNeely, 2020). Fig. 4 reveals the conceptual model of traffic scene based on DTs designed here.

According to Fig. 4, in the DTs urban traffic scene, people, vehicles, roads, environment, and other elements migrate from the physical world to the digital world, realizing the virtuality and reality combination of the traffic scene. This can be extensively used in remote operation and maintenance of transportation facilities, development of new transport vehicles, driverless virtual test, traffic problem analysis and experiment, etc. With the help of advanced information technology means such as deep learning, IoT, cloud computing, and positioning technology, DTs-based intelligent transportation produces a vitally significant application effect in solving practical problems.

2.3. *Traffic sign recognition via SSD-ResNet50 algorithm*

Here, the improved SSD algorithm is applied to the recognition of traffic signs. SSD algorithm belongs to one-stage detection algorithms. This algorithm can directly calculate the target position and category of the input image through the convolutional neural network (CNN) (Li & Zhou, 2017). The SSD algorithm can also generate a variety of feature maps to enhance the feature extraction capability of the algorithm (Nagrath et al., 2021).

The SSD algorithm primarily consists of the multi-layer deep convolutional neural network (DCNN) and multi-scale feature extraction network. The multi-layer DCNN is evolved from Visual Geometry Group Network 16 (VGG16), which is the basic network of the SSD algorithm (Miao et al., 2019). It is responsible for preliminarily extracting feature information and achieving target classification. The multi-scale feature extraction network is composed of many convolution layers, and the size of feature map obtained by these convolution layers decreases layer by layer. Therefore, multi-scale feature maps can be used to predict target position and category, to increase the accuracy and improve robustness to low resolution images (Wang et al., 2018). ResNet adopts the mode of jump connection to directly combine the input information with the output after the convolution operation as the overall output of this feature layer, and no new parameters are added in the direct connection process (Du et al., 2019). Consequently, the number of parameters in the network structure is greatly reduced, the feature extraction effect is improved, and the performance degradation with the increase in network layers is effectively solved.

To improve the depth of the convolutional network, the SSD algorithm is improved based on ResNet. Specifically, ResNet50 is introduced into the SSD algorithm to replace original VGG16, constituting the SSD-ResNet50 algorithm (Theckedath & Sedamkar, 2020). ResNet50 has four different residual layers, and each layer has a different number of residual units. Each residual unit contains a jump connection from input to output, and there is also an output obtained by the input after three convolution operations, which solves the problem that detection accuracy decreases instead of rising with the growth of network layers (Ali et al., 2019). Moreover, the parameters and filters in ResNet50 are less than those in VGG16, resulting in a more simple computation process (Nahar et al., 2018).

After the above improvements, deconvolution fusion operation is also carried out to solve the problem of poor network convergence of SSD-ResNet50 (Cheng et al., 2017). In the target detection, the lower the dimension of feature graph, the richer the semantic information of target feature. After deconvolution operations, the feature map can map low-dimensional features in the high layer to high-dimensional features in the low layer. This operation can strengthen the semantic connection between network feature layers and highlight the feature expression of network structure (He et al., 2018). Deconvolution is a kind of mathematical operation that maps low-dimensional local features to high-dimensional space through a zero-filling method. The operation process is shown in Eq. (4).

$$f_{(Deconvolution)} = [s(x-1) + n - 2pad]^2 \tag{4}$$

In Eq. (4), s denotes the step size, x refers to the size of input image pixel, n stands for the size of convolution kernel in deconvolution, and pad represents the value of padding.

Fig. 5 illustrates the structure of the deconvolution module.

Fig. 6 indicates the SSD-ResNet50 model after optimizing the SSD algorithm by ResNet50 and deconvolution module.

According to Fig. 6, the SSD-ResNet50 model adds five extra convolution layers on the basis of ResNet50 to extract feature images

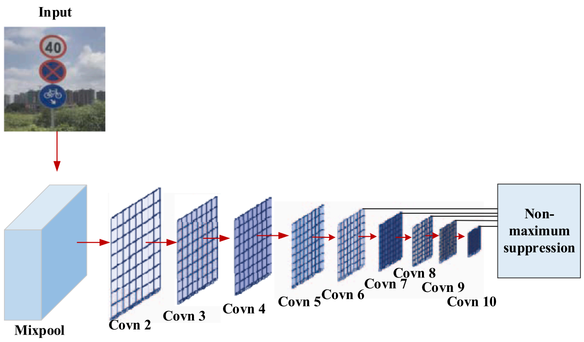

Fig. 6. SSD-ResNet50 model.

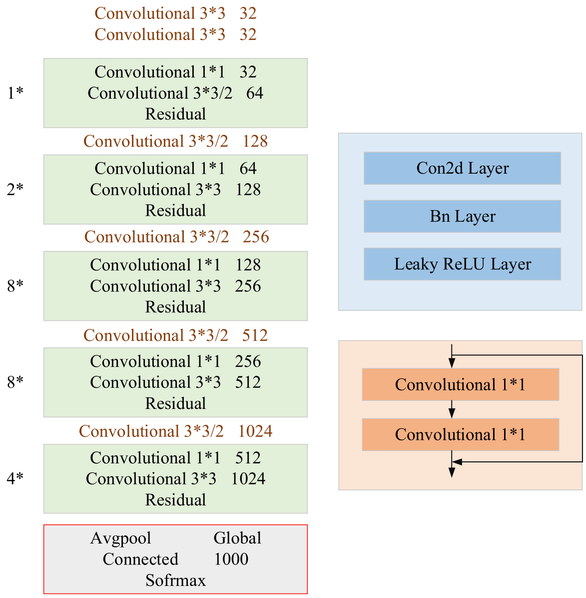

Fig. 7. The structure of original DarkNet-53.

of different sizes. The improved network has 65 layers, including 60 convolutional layers and 5 pooling layers. The size of the input data is 512*512, the size of output feature maps after the pooling layer of each layer is 1/2 of the size of the input image, and the size of the output data is 1*1. After each convolution layer in deconvolution module, there is an added batch normalization layer, and the learning-based deconvolution layer performs sampling. Besides, the superposition feature map is input into the ReLU activation layer by means of element dot product. Under this optimization scheme, the SSD-ResNet50 model can better extract target features and strengthen the fusion ability of target feature information.

2.4. Traffic sign recognition algorithm under improved DarkNet-53 model

The DarkNet-53 network model is common in target recognition. Fig. 7 shows the structure of original DarkNet-53.

In the original DarkNet-53, feature extraction is completed after the input data is processed by multi-layer 1*1 convolution, 3*3 convolution, reduction, and nonlinear transformation, etc. Then, the input enters into the module composed of average pooling (Avgpool) layer, full connection layer, and Softmax layer to realize image recognition and classification (Yang & Shi, 2021; Lawal, 2021).

Accuracy and instantaneity are two primary issues to be considered in the study of traffic sign recognition. Therefore, DarkNet-53 opptimized here, aiming at improving the accuracy of traffic sign recognition and enhancing the generalization ability of the model. The performance of the algorithm is improved by adjusting the classification module structure and loss function in the original DarkNet-53.

The structure of the classification module of DarkNet-53 is optimized first. Fig. 8 provides the structure of optimized DarkNet-53.

As shown in Fig. 8, a new classification module is designed, which is composed of the Avgpool layer and classifier. The Avgpool layer dimensionally reduces the features extracted by DarkNet-53, and the classifier consists of a linear layer, a one-dimensional

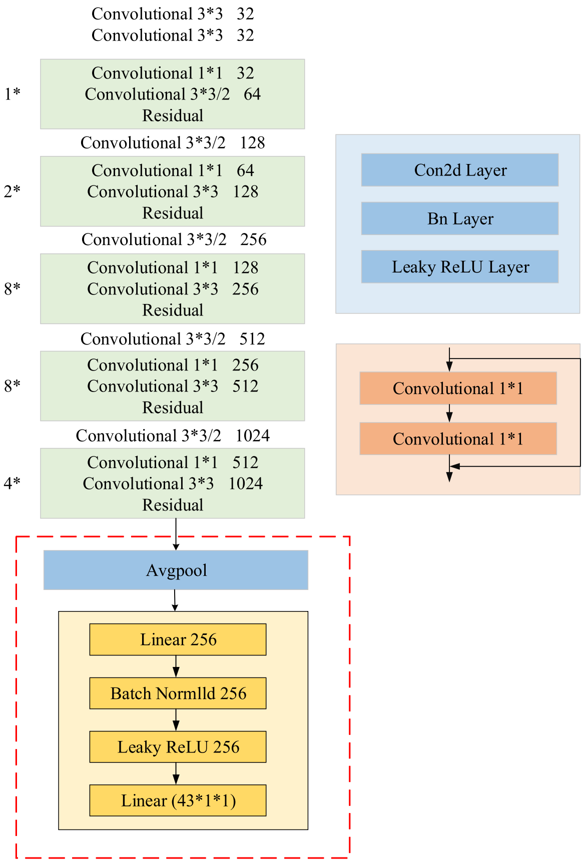

Fig. 8. Structure of optimized DarkNet-53.

BatchNormal layer, a LeakyReLU layer, and a linear layer.

In addition, the loss function is optimized. In the traditional two-class classification method, the conventional loss function is the cross-entropy function (CEF), which is expressed as Eq. (5).

$$L = -y\log y' - (1-y)\log(1-y') = \begin{cases} -\log y', y = 1 \\ -\log(1-y'), y = 0 \end{cases} \tag{5}$$

In Eq. (5), $y' \in \{0,1\}$ denotes the label of the real sample, and $y' \in [0,1]$ refers to the predicted output after the Sigmoid activation function.

According to Eq. (5), for positive samples, the larger the output probability value of the traditional cross entropy loss function, the smaller the loss. On the contrary, for negative samples, the smaller the output probability, the smaller the loss. Therefore, the loss function is slowly optimized or even unable to obtain the optimal solution in the iteration process of multitudes of samples.

Focal Loss function is utilized as loss function here, which can be written as:

$$L_F = \begin{cases} -(1-y')\log y', y = 1 \\ -y'^{a}\log(1-y'), y = 0 \end{cases} \tag{6}$$

$$L_F = \begin{cases} -b(1-y')\log y', y = 1 \\ -(1-b)y'^{a}\log(1-y'), y = 0 \end{cases} \tag{7}$$

Eqs. (6) and (7) show that Focal Loss function adds an impact factor a ($a > 0$) to the original CEF. If the impact factor reduces the loss of easy-to-classify categories, the loss of difficult-to-classify categories will increase. Therefore, the model parameters are modified in different ranges for different categories in each process of error transmission, to speed up network training and improve the overall recognition accuracy of the model. Besides, a balance factor b is added to deal with the imbalance of different types of data (Ross & Dollár, 2017; Sun et al., 2019).

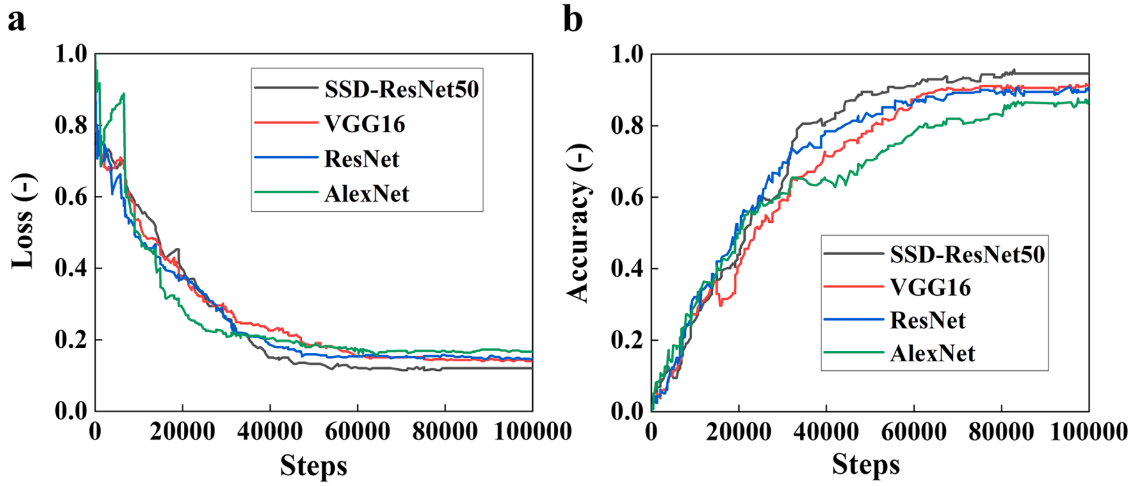

Fig. 9. Training results of SSD-ResNet50 algorithm.

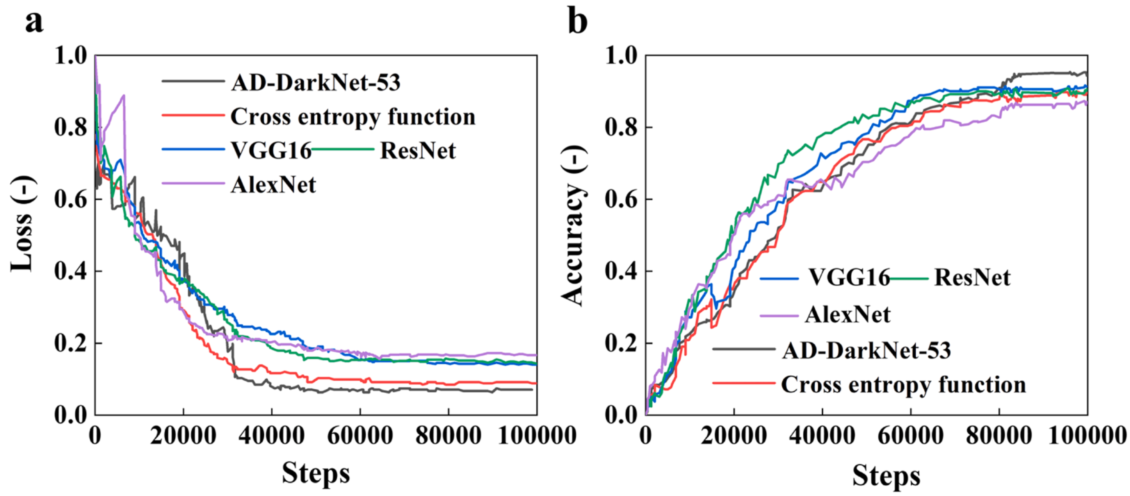

Fig. 10. Comparison of training results of different models.

2.5. Communication big data node detection algorithm based on machine learning

The data set used in this experiment is the German Traffic Sign Detection Benchmark (GTSDB), which is the most classic and the most recognized data set in traffic signal and detection (Ellahyani & El Ansari, 2017). The data set is classified into a training set and a test set. There are 600 images in the training set and 300 in the test set. Images in the data set are divided into three categories of prohibition, instruction, and warning. All images in this data set come from real scenes, including various traffic sign images of different application scenarios, different light scenarios, and different weather environments (Zhang et al., 2020).

The training with large-scale data can enable the CNN to extract more feature information of detection targets and avoid overfitting or under-fitting of the model due to insufficient data during model training. However, the GTSDB data set contains a small amount of images, which can't meet the needs for data of the model. Therefore, a data enhancement approach is carried out to provide sufficient amount of training data. The existing sample images are flipped, rotated, discolored, cropped, scaled, gray scaled, blurred, and adjusted for contrast and brightness, to generate new images and expand the traffic sign data set. After data enhancement, there are 45,000 images in the data set, which meets the needs of the experiment.

Moreover, the TensonFlow deep learning platform is adopted for design, improvement, and testing of algorithms. In the experiment of traffic sign detection and recognition under SSD-ResNet50 algorithm, the number of batch-siza is set to 8, the total number of training is 100,000. Besides, the output learning rate is set to 0.001, but it decreases to 0.0001 and 0.00001, respectively, at Step 30,000 and Step 60,000. Moreover, the Intersection over Union (IoU) threshold is 0.5. Samples with a threshold not less than 0.5 are regarded as positive samples, and those with a threshold less than 0.5 are regarded as negative samples.

The model performance is evaluated from two perspectives, namely accuracy and instantaneity. The evaluation criteria of accuracy

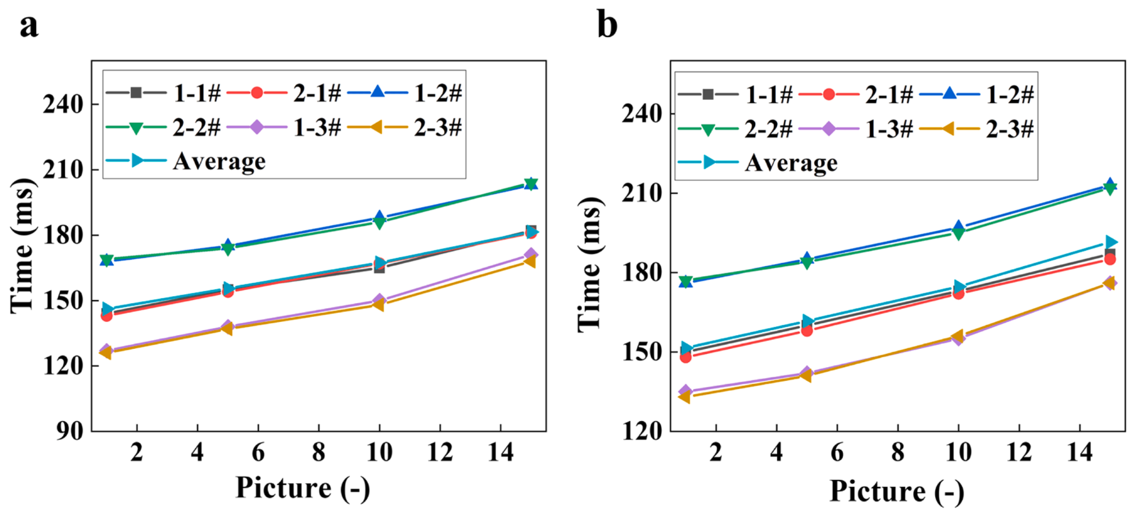

Fig. 11. Real-time analysis of SSD-ResNet50 algorithm.

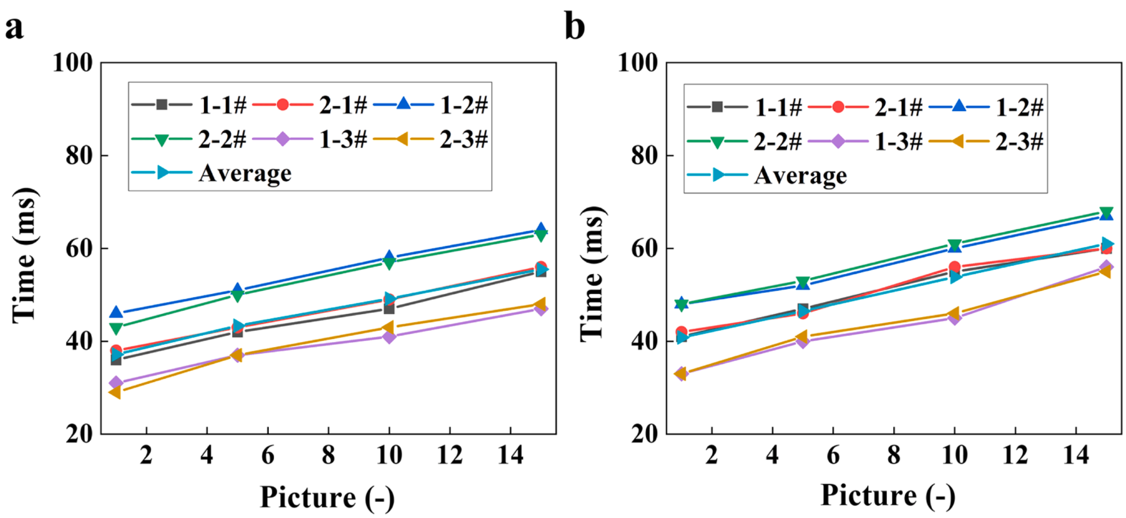

Fig. 12. Comparison of instantaneity performance of improved DarkNet-53 algorithm and traditional DarkNet-53 algorithm.

contains Accuracy, Precision, Recall, and Fi-value.

3. Analysis of algorithm training results and target recognition results

3.1. Comparative analysis of algorithm training results

The specific parameters of SSD-ResNet50 algorithm in training are analyzed. To further analyze the effect of the model designed here, the training loss and accuracy of SSD-ResNet50 algorithm are compared with those of VGG16, ResNet, and AlexNet algorithm. The specific results are shown in Fig. 9, in which Fig. 9a reveals the training result of Loss, and Fig. 9b illustrates the training result of accuracy.

Through Fig. 9, after the training, the Loss of the SSD-ResNet50 model is basically around 0.12, and the accuracy is maintained at about 94.6%. Compared with the VGG16, ResNet, and AlexNet models, the SSD-ResNet50 model shows a better training effect. The optimization plan speeds up the training speed of network model, and elevates the recognition accuracy.

Then, the training effect of the improved DarkNet-53 model is compared with that of traditional CEF and improved Focal Loss function, and specific results are shown in Fig. 10.

According to Fig. 10, the DarkNet-53 algorithm improved by Focal Loss function can achieve faster model convergence and a lower total loss value. This result demonstrates that the optimization scheme reported here is reasonable and effective.

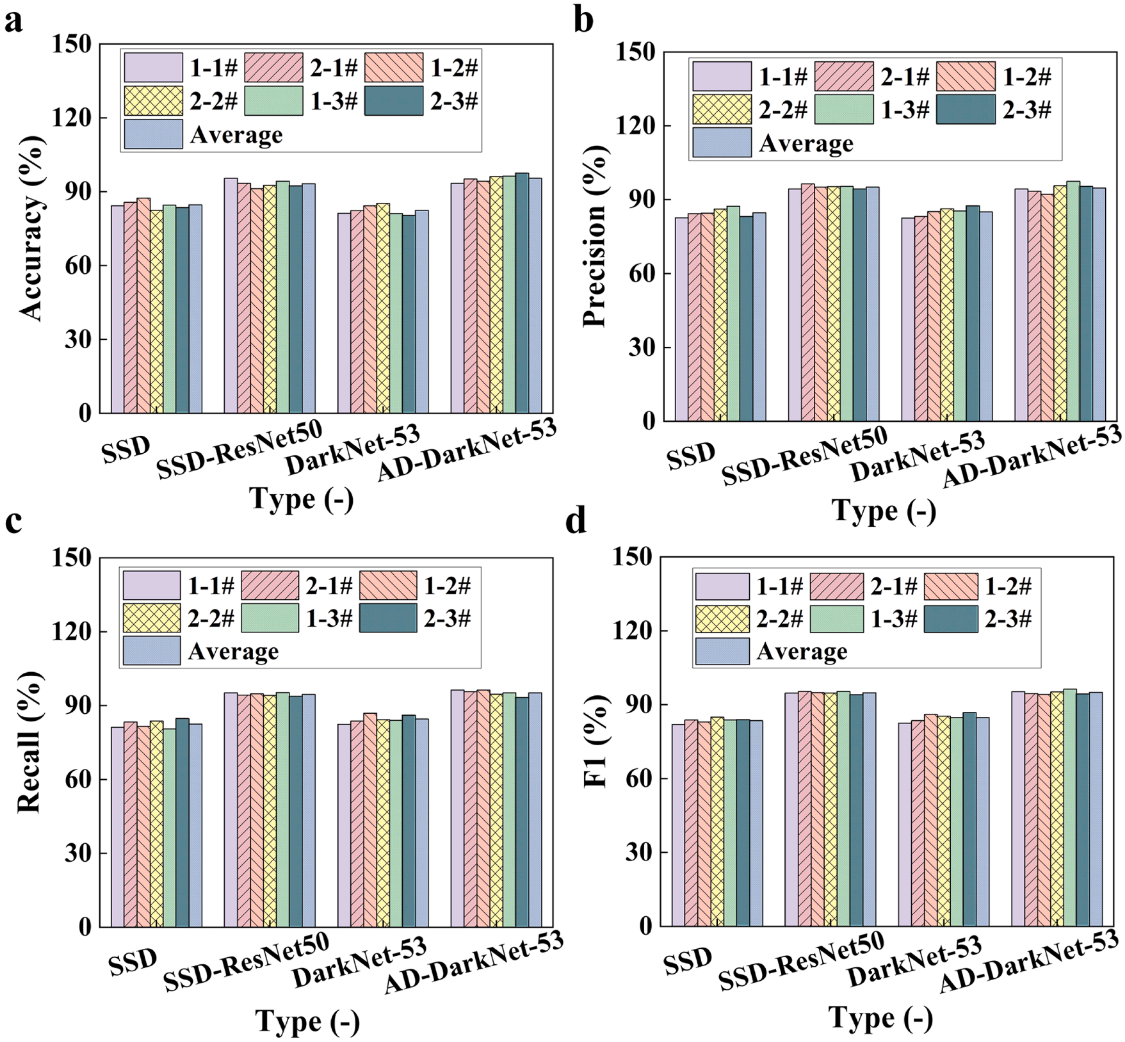

Fig. 13. Comparison of the target detection effect of algorithms.

3.2. Analysis of instantaneity of target recognition of optimized algorithm

The instantaneity performance of the SSD-ResNet50 algorithm during traffic sign recognition is analyzed in this experiment. Since the data set contains traffic indication data in the three categories of prohibition, instruction, and warning, they are coded as 1#, 2#, and 3#, respectively, and each group of categories is subjected to two tests. For example, the first test of 1# data set is recorded as 1-1#. Then, the instantaneity performance of the SSD-ResNet50 algorithm and the traditional SSD algorithm is compared. Fig. 11a illustrates the experimental result of the SSD-ResNet50 algorithm, and Fig. 11b shows that of the traditional SSD algorithm.

In Fig. 11a, when the number of recognized images is 1, 5, 10, and 15, the average recognition time of the SSD-ResNet50 algorithm is 146.167ms, 155.500ms, and 167.333ms, and 184.833ms, respectively. In Fig. 11b, under the same situation, the traditional SSD algorithm takes 151.500ms, 161.667ms, 174.667ms, and 191.500ms to recognize 1, 5, 10, and 15 images, respectively. The data shows that the recognition time of the SSD-ResNet50 algorithm is 6.37 ms shorter than that of traditional SSD.

The recognition time of improved DarkNet-53 and the traditional DarkNet-53 is compared using the same method as mentioned before. Fig. 12 shows the comparison.

From Fig. 12a, when the number of recognized images is 1, 5, 10, and 15, the average time of the improved DarkNet-53 algorithm is 37.167ms, 43.333ms, 49.167ms, and 55.500ms, respectively. Fig. 12b shows that in the same situation, the traditional DarkNet-53 algorithm takes 40.833ms, 46.500ms, 53.833ms, and 61.000ms, respectively. The recognition time of the improved DarkNet-53 algorithm is 4.25 ms less than that of the traditional DarkNet-53 algorithm, and the recognition time of the improved DarkNet-53 algorithm is less than that of SSD-ResNet50. In summary, the improved DarkNet-53 algorithm shows a faster speed in the recognition of traffic signs.

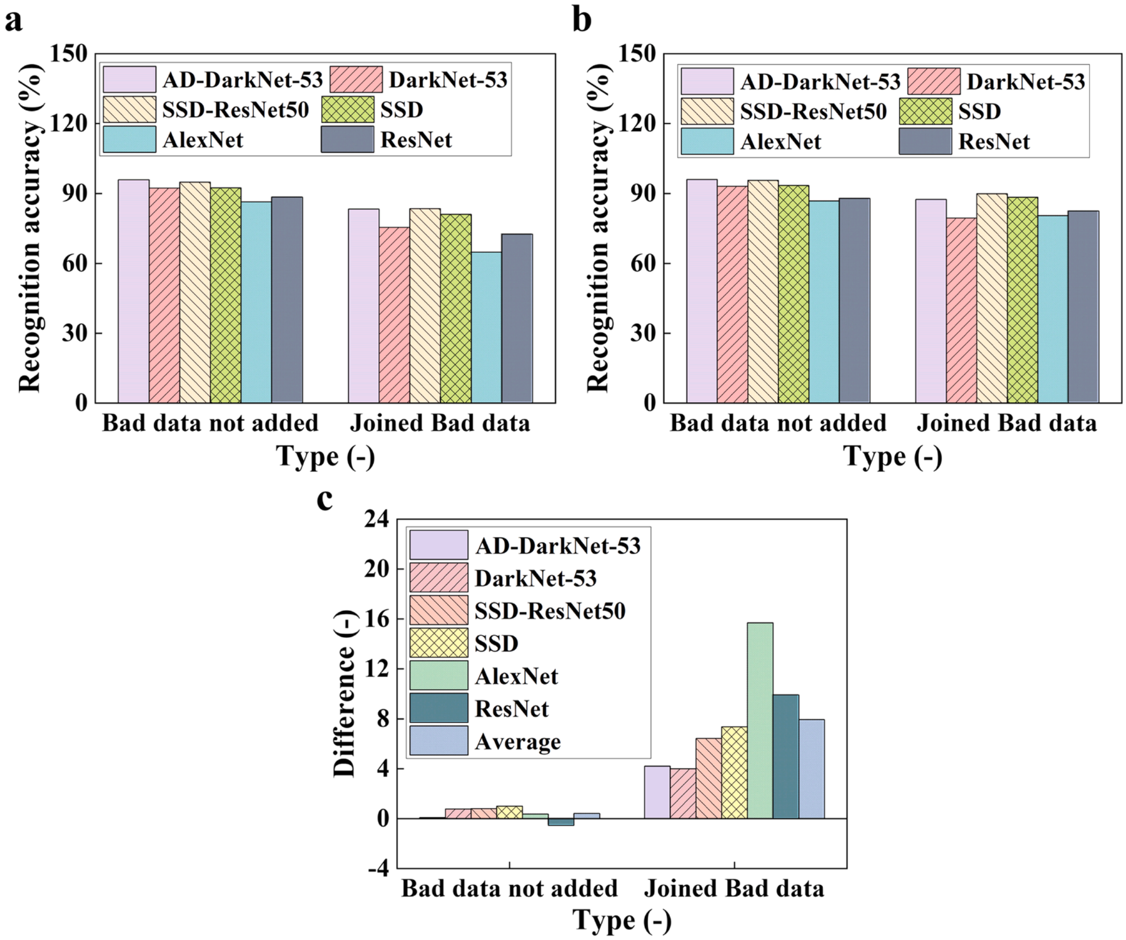

Fig. 14. Data enhancement performance.

3.3. Analysis of accuracy of algorithms

The Accuracy, Precision, Recall, and F1-value of the SSD-ResNet50, the traditional SSD, the improved DarkNet-53, and the traditional DarkNet-53 in recognizing three different traffic signs are compared, as presented in Fig. 13.

Fig. 13 shows that the two optimized algorithms reported here perform well in actual traffic sign target recognition. In the detection of three types of traffic signs, the average Accuracy of SSD-ResNet50, traditional SSD, traditional DarkNet-53, and improved DarkNet-53 is 93.227%, 84.665%, 82.437%, and 95.481%, respectively. Evidently, the optimization scheme can significantly improve the recognition effect of deep learning algorithms. The recognition accuracy of the improved DarkNet-53 is slightly better than that of the SSD-ResNet50 algorithm, but without considerable differences.

3.4. Analysis of effects of the data enhancement method

To verify the feasibility and specific performance of the data enhancement method, some bad data images with typical interference characteristics are added to the original data set, and traffic sign recognition tests are performed on different algorithms with and without data enhancement. The specific results are shown in Fig. 14, where Fig. 14a is the result before data enhancement, Fig. 14b is the result after data enhancement, and Fig. 14c shows the difference after data enhancement.

Through the results in Fig. 14, the performances of AD-DarkNet-53, DarkNet-53, SSD-ResNet50, SSD, AlexNet, and ResNet is improved greatly. Before adding bad data, the average recognition accuracy increases by 0.418%, and after adding bad data, the average recognition accuracy increases by 7.937%. Therefore, the data enhancement method not only optimizes the algorithms reported here, but also has a certain effect on other commonly used image recognition model methods.

4. Conclusion

The present work first discusses the theoretical model of DTs and theories related to smart cities and ITS. On this basis, the conceptual framework of DTs city and the intelligent traffic scene model based on edge computing combined with DTs are constructed. Then, the SSD algorithm and DarkNet-53 algorithm are optimized, developing the SSD-ResNet50 algorithm and AD-DarkNet-53 algorithm. Moreover, the algorithms reported here are tested from the perspectives of the training effect, instantaneity performance, and target recognition accuracy, and the data enhancement effect is also evaluated through experiments. The main conclusions are as follows. First, SSD-ResNet50 and AD-DarkNet-53 algorithms achieve fast model convergence speed and low total loss value. Secondly, the target recognition time of the two algorithms is short, indicating that the optimization scheme proposed here can effectively improve the training effect. Moreover, the recognition time of the improved DarkNet-53 algorithm is less than that of the SSD-ResNet50 algorithm. Third, the identification accuracy of SSD-ResNet50 and AD-DarkNet-53 algorithms after optimization is significantly higher than that of the original algorithms. Finally, the two algorithms reported here can effectively identify bad data, and the overall recognition effect is excellent. To sum up, the research content and research methods are fully considered. The combination of theoretical analysis and practical test can provide reference for subsequent research, and the present work has certain concrete application value. However, due to limited time, the present work just provides a theoretical framework of ITS, and there lacks a comprehensive exploration. The follow-up study will explore the whole process and full cycle of ITS to promote the study of ITS and smart cities.

CRediT authorship contribution statement

Weixi Wang: Conceptualization, Methodology. **Fan He:** Software, Investigation, Writing – original draft. **Yulei Li:** Investigation, Visualization. **Shengjun Tang:** Methodology, Writing – original draft. **Xiaoming Li:** Methodology, Visualization, Supervision. **Jizhe Xia:** Methodology, Software. **Zhihan Lv:** Methodology, Writing – review & editing.

Data availability

The data that has been used is confidential.

Acknowledgment

This work has been partly supported by the National Natural Science Foundation of China (Grant No. 41971341, 41971354, 42171400), General Project of the National Natural Science Foundation of Guangdong Province (Grant No. 2019A1515010748, 2019A1515011872, 2021A1515011324), the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, MNR (No. KF-2020-05-032), the financial support of the Guangdong Science and Technology Strategic Innovation Fund (the Guangdong–Hong Kong-Macau Joint Laboratory Program (Project No. 2020B1212030009).

References

Deng, M., Menassa, C. C., & Kamat, V. R. (2021). From BIM to digital twins: A systematic review of the evolution of intelligent building representations in the AEC-FM industry. *Journal of Information Technology in Construction (ITcon), 26*(5), 58–83.

Khalid, L. F., & Ameen, S. Y. (2021). Secure Iot integration in daily lives: A review. *Journal of Information Technology and Informatics, 1*(1), 6–12.

Camero, A., & Alba, E. (2019). Smart City and information technology: A review. *Cities, 93*, 84–94.

Fan, C., Zhang, C., Yahja, A., & Mostafavi, A. (2021). Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management, 56*, Article 102049.

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., et al. (2020). Developing a digital twin at building and city levels: Case study of West Cambridge campus. *Journal of Management in Engineering, 36*(3), Article 05020004.

Francisco, A., Mohammadi, N., & Taylor, J. E. (2020). Smart city digital twin-enabled energy management: Toward real-time urban building energy benchmarking. *Journal of Management in Engineering, 36*(2), Article 04019045.

Elayan, H., Aloqaily, M., & Guizani, M. (2021). Digital twin for intelligent context-aware IoT healthcare systems. *IEEE Internet of Things Journal, 8*(23), 16749–16757.

Pang, J., Huang, Y., Xie, Z., Li, J., & Cai, Z. (2021). Collaborative city digital twin for the COVID-19 pandemic: A federated learning solution. *Tsinghua Science and Technology, 26*(5), 759–771.

Yan, J., Liu, J., & Tseng, F.. M. (2020). An evaluation system based on the self-organizing system framework of smart cities: A case study of smart transportation systems in China. *Technological Forecasting and Social Change, 153*, Article 119371.

Aujla, G. S., Jindal, A., & Kumar, N. (2018). EVaaS: Electric vehicle-as-a-service for energy trading in SDN-enabled smart transportation system. *Computer Networks, 143*, 247–262.

Aamir, M., Masroor, S., Ali, Z. A., & Ting, B. T. (2019). Sustainable framework for smart transportation system: A case study of karachi. *Wireless Personal Communications, 106*(1), 27–40.

Jararweh, Y., Otoum, S., & Al Ridhawi, I. (2020). Trustworthy and sustainable smart city services at the edge. *Sustainable Cities and Society, 62*, Article 102394.

Berdik, D., Otoum, S., Schmidt, N., Porter, D., Jararweh, Y., et al. (2021). A survey on blockchain for information systems management and security. *Information Processing & Management, 58*(1), Article 102397.

Tao, F., Liu, W., Zhang, M., Hu, T., Qi, Q., Zhang, H., et al. (2019). Five- dimension digtal twin model and its ten app;ications. *Computer Integrated Manufacturing Systems, 25*(1), 1–18.

Ham, Y., & Kim, J. (2020). Participatory Sensing and digital twin city: Updating virtual city models for enhanced risk-informed decision-making. *Journal of Management in Engineering, 36*(3), Article 04020005.

Wu, C., Zhou, Y., Pessôa, M. V. P., Peng, Q., & Tan, R. (2021). Conceptual digital twin modeling based on an integrated five-dimensional framework and TRIZ function model. *Journal of Manufacturing Systems, 58*, 79–93.

Qiao, Q., Wang, J., Ye, L., & Gao, R. X. (2019). Digital Twin for machining tool condition prediction. *Procedia CIRP, 81*, 1388–1393.

Minerva, R., Lee, G. M., & Crespi, N. (2020). Digital twin in the IoT context: a survey on technical features, scenarios, and architectural models. *Proceedings of the IEEE, 108*(10), 1785–1824.

Bansal, K., Mittal, K., Ahuja, G., Singh, A., & Gill, S. S. (2020). DeepBus: Machine learning based real time pothole detection system for smart transportation using IoT. *Internet Technology Letters, 3*(3), e156.

An, C., & Wu, C. (2020). Traffic big data assisted V2X communications toward smart transportation. *Wireless Networks, 26*(3), 1601–1610.

Schintler, L. A., & McNeely, C. L. (2020). Mobilizing a culture of health in the era of smart transportation and automation. *World Medical & Health Policy, 12*(2), 137–162.

Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P., & Hemanth, J. (2021). SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustainable Cities and Society, 66*, Article 102692.

Miao, X., Liu, X., Chen, J., Zhuang, S., Fan, J., & Jiang, H. (2019). Insulator detection in aerial images for transmission line inspection using single shot multibox detector. *IEEE Access, 7*, 9945–9956.

Wang, Y., Wang, C., & Zhang, H. (2018). Combining a single shot multibox detector with transfer learning for ship detection using sentinel-1 SAR images. *Remote Sensing Letters, 9*(8), 780–788.

Du, L., Li, L., Wei, D., & Mao, J. (2019). Saliency-guided single shot multibox detector for target detection in SAR images. *IEEE Transactions on Geoscience and Remote Sensing, 58*(5), 3366–3376.

Theckedath, D., & Sedamkar, R. (2020). Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science, 1*(2), 1–7.

Ali, N., Quansah, E., Köhler, K., Meyer, T., Schmitt, M., Popp, J., et al. (2019). Automatic label-free detection of breast cancer using nonlinear multimodal imaging and the convolutional neural network ResNet50. *Translational Biophotonics, 1*(1-2), Article e201900003.

Nahar, P., Tanwani, S., & Chaudhari, N. S. (2018). Fingerprint classification using deep neural network model resnet50. *International journal of research and analytical reviews, 5*(4), 1521–1537.

Cheng, Y., Cai, R., Li, Z., Zhao, X., & Huang, K. (2017). Locality-sensitive deconvolution networks with gated fusion for rgb-d indoor semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

He, X., Yu, Z., Wang, T., Lei, B., & Shi, Y. (2018). Dense deconvolution net: Multi path fusion and dense deconvolution for high resolution skin lesion segmentation. *Technology and Health Care, 26*(S1), 307–316.

Yang, M., & Shi, X. (2021). A deep learning model S-Darknet suitable for small target detection. In *Proceedings of the journal of physics: conference series*. IOP Publishing.

Lawal, M. O. (2021). Tomato detection based on modified YOLOv3 framework. *Scientific Reports, 11*(1), 1–11.

Ross, T. Y, & Dollár, G. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Sun, X., Dong, K., Ma, L., Sutcliffe, R., He, F., Chen, S., et al. (2019). Drug-drug interaction extraction via recurrent hybrid convolutional neural networks with an improved focal loss. *Entropy, 21*(1), 37.

Ellahyani, A., & El Ansari, M. (2017). Mean shift and log-polar transform for road sign detection. *Multimedia Tools and Applications, 76*(22), 24495–24513.

Zhang, H., Qin, L., Li, J., Guo, Y., Zhou, Y., Zhang, J., et al. (2020). Real-time detection method for small traffic signs based on yolov3. *IEEE Access, 8*, 64145–64156.

Li, Z., Zhou, F. (2017). FSSD: feature fusion single shot multibox detector. Computer Vision and Pattern Recognition, arxiv:1712.00960.