

Analysis and Design of Standard Knowledge Service System Based on Deep Learning

Yuzhong Zhou^{1,*}, Zhengping Lin¹, Liang Tu¹, Junkai Huang², Zifeng Zhang¹

¹Electric Power Research Institute, China Southern Power Grid Company (e-mail: yuzhong_zhou@hotmail.com, zhengping_lin@hotmail.com, Liang_Tu1@hotmail.com, and zifengzhangcsg@hotmail.com).

²Electric Power Research Institute of Guizhou Power Grid Co., Ltd (e-mail: junkaihuangcsg@hotmail.com).

Abstract

The development of information technology has changed the mode of communication of social information, and this change has put forward new requirements on the contents, methods and even objects of information science research. Knowledge service in the information service process can extract knowledge and information content from various explicit and implicit knowledge resources according to people's needs, build knowledge networks, and provide knowledge content or solutions for users' problems. Hence, it is very important to investigate how to analyze and design the advanced standard knowledge service system based on deep learning. To this end, we firstly introduce the typical deep learning networks of convolutional neural network (CNN) for the knowledge service system, and then employ the CNN to implement the knowledge classification based on deep learning. Finally, some simulation results on the knowledge service system are presented to validate the proposed studies in this paper.

Received on 28 August 2022; accepted on 09 October 2022; published on 12 October 2022

Keywords: Deep learning, standard knowledge service system, knowledge classification, convolutional neural network (CNN).

Copyright © 2022 Yuzhong Zhou *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi:10.4108/eetsis.v9i6.2637

1. Introduction

The development of information technology has changed the mode of communication of social information [1–3], and this change has put forward new requirements on the contents, methods and even objects of information science research [4, 5]. In terms of the specific forms of information, the diversified digital information resources have enriched the research objects of information science [6, 7]. Information research is no longer limited to academic literature. Information carriers such as web pages, books, patents, archives, and standard documents have begun to attract the attention from information researchers. Considering the perspective of information chain or basic concepts of information science, the object of information science research goes from information level to knowledge level. How to realize the orderly organization of knowledge in various information carriers and provide

effective knowledge services has become an important frontier of information science research.

As an important information source and knowledge carrier, the production, organization and utilization of standard knowledge in the digital network environment are also facing the problem of developing towards knowledge service. However, the current standard document service systems are mostly at the document granularity, which can not meet the knowledge needs of users. Based on the above considerations, it is important to combine optical character recognition, natural language processing, information visualization, information retrieval and other technologies to build a knowledge level oriented standard document information service system, which provides users with standard knowledge services such as knowledge extraction, knowledge map, and knowledge search [8–10]. To a certain extent, it improves the quality of standard knowledge information services and improves the user experience.

*Corresponding author. Email: yuzhong_zhou@hotmail.com

Generally speaking, the current standard knowledge service system still stays at the level of document retrieval based on keywords. The processing granularity of standard documents is relatively coarse, which fails to penetrate into the semantic knowledge units within standard documents and ignores the relationship between the standard document knowledge units. Therefore, it is urgent to conduct semantic organization and knowledge extraction on the standard knowledge content, and transform from the document service system to the knowledge service system. In this paper, we focus on using the deep learning model to mining the knowledge to construct the knowledge service system which can provide people with the recognition service of this knowledge.

2. System model of deep learning for standard knowledge service system

The method based on deep neural network model has been widely used in computer vision tasks, and has played a great role in the development of artificial intelligence [11–13]. Considering the excellent performance of deep network, it can be effectively used in the the construction of the standard knowledge service system. In this section, the deep learning model for the analysis and design of standard knowledgeservice system will be introduced and the network architecture is shown in Fig. 1. In this figure, we use the convolutional neural network (CNN) as the feature extractor to extract the knowledge from the data. CNN is a kind of feedforward neural networks which contains convolution calculation and has a deep structure [14]. It is one of the representative algorithms of deep learning. CNN has the ability of representation learning and can perform shift invariant classification of input information according to its hierarchical structure. Therefore, it is also called shift invariant artificial neural network. The CNN is constructed by imitating the visual perception mechanism of biology and can conduct supervised learning and unsupervised learning [15, 16]. The sharing of convolution kernel parameters in the hidden layer and the sparsity of inter layer connections enable the convolutional neural network to learn grid like topology features, such as pixels and audio stable effect and no additional feature engineering requirements for data.

Let $I \in \mathbb{R}^{N \times H \times W \times C}$ be the input data for a certain layer of the CNN, and the output feature map from the convolutional layer can be obtained as [17, 18]

$$F = I * K, \quad (1)$$

$$F(m, n) = \sum_i \sum_j I(m-i, n-j)K(m, n), \quad (2)$$

where K represents the convolutional kernel of the neural network, m and n represent the spatial

coordination (m, n) of the input feature map, i and j range in the local area of (m, n) that relate to the size of convolution kernel K . The visualization of the convolution for a local area of the data is shown in Fig. 2. The feature representation generated by convolution of each local region is closely related to the selection of convolution kernel. For a given convolution kernel, one type of feature representation can be obtained, while different values of convolution kernels produce different local feature representations.

To ensure that the neural network can fit complex non-linear functions, the activation function needs to be added after convolution. The function of activation function is to map data features from one feature space to another through nonlinear transformation, so as to construct robust feature expression, which is suitable for dealing with complex visual analysis problems. There are so many activation functions that have been used in the deep neural network, and some of them are expressed as [19, 20]

$$f(x) = \frac{1}{1 + e^{-x}}, \quad (3)$$

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}. \quad (4)$$

Among the existing activation functions, there is a most widely used activation function, named as linear rectification function (ReLU), and it can be expressed as [21–23]

$$f(x) = \max(0, x). \quad (5)$$

It can be seen from the above formula that when the input takes a negative value, the function output is 0, and when the input takes a positive value, the output is the same as the input. The derivative of the ReLU activation function is as [24–26]

$$f'(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \quad (6)$$

It can be seen from the above formula that when deriving the ReLU activation function, the results are 0 or 1. The visualization of ReLU and its derivatives is shown in Fig. 3. In the interval $[0, +\infty]$, since the derivative of ReLU is 1, the error attenuation will not be caused by multiplying the derivative of the activation function during the back propagation of the network training, which alleviates the disappearance of the gradient and enables the network training to converge more stably and efficiently. Therefore, in this system model, ReLU is chosen. And the feature extraction of the model for one layer can be expressed as [27, 28]

$$y = \text{MaxP}(\text{ReLU}(\text{Cov}(I; W, b))), \quad (7)$$

where MaxP represents the max pooling operator, Cov represents the convolution operator, W is the parameter

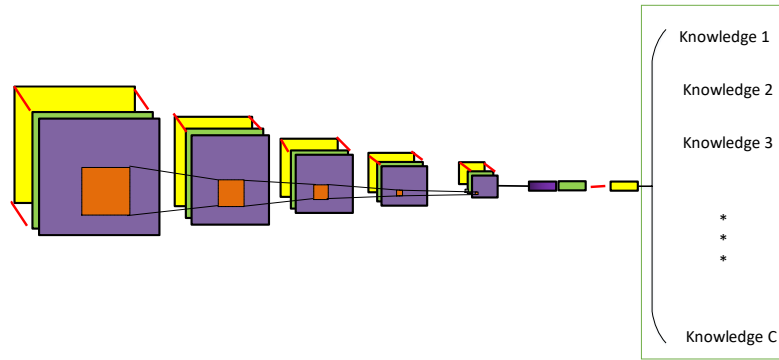


Figure 1. Deep learning model for the analysis and design of standard knowledge service system.

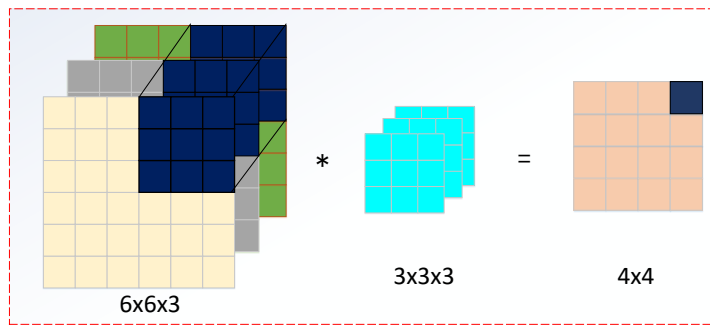


Figure 2. Visual schematic diagram of local area convolution operation in depth network for the standard knowledge service system.

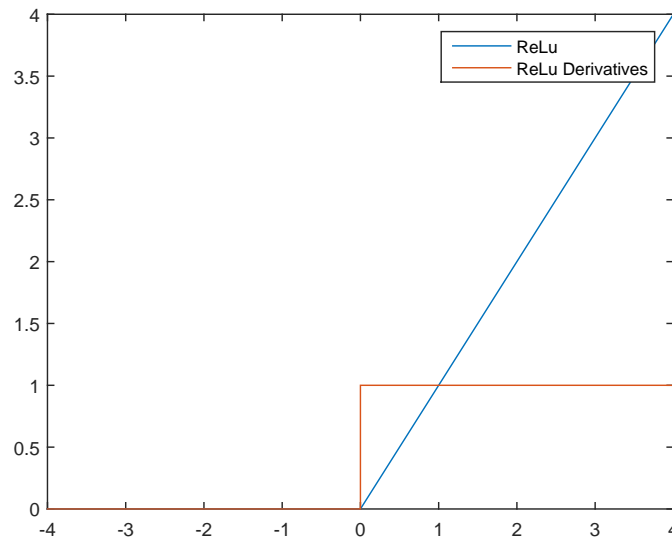


Figure 3. ReLu activation function and the derivatives for deep learning in the standard knowledge service system.

of the convolutional kernel, and b is the bias. After extracting the features by the convolutional layers, two linear layers are used to integrate the features and classify them into different types of knowledge. It can

be expressed as

$$p = \sigma(f_c(Y)), \tag{8}$$

where Y represents the output from the last convolutional layer of the network, p is the output probability

for the standard knowledge, and σ represents the linear classification layer that is expressed as [29, 30]

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^C e^{x_j}}. \quad (9)$$

The objective function of the network model is defined as

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C l_{ij} \log(p_{ij}), \quad (10)$$

where N is the number of sample, C is the number of standard knowledge, l_{ij} is the one-hot encode of the true label of the data that represents the i th sample belongs to the j th standard knowledge. p_{ij} is the predicted value of the output of the deep neural network, indicating the probability value of predicting the i th sample as the j th type of standard knowledge.

3. Results and Analysis

In this section, we will make some experiment to validate the effectiveness of the deep learning method for the analysis and design of standard knowledge service system.

3.1. Data illustration

As to the standard knowledge, 10,000 samples are collected in this work and they are divided into training set, validation set and test set according to the ratio of 3 : 1 : 1. The training set is used to train the deep neural network, the validation set is used to adjust the hyperparameters of the network so that the network can obtain a better performance. The test set is used to test the model and verify the generalization and validity of the model.

3.2. Experimental results and analysis

In this work, the codes are implemented based on Python 3.6.4 and Tensorflow 1.9.0 with GeForce RTX 2080 11GB for acceleration. The parameters of the networks are initialized by Xavier. In this section, the metrics of Precision, F1-score, and Accuracy (Acc) were used to measure the performance of the neural network, and the calculation of them are expressed as:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (11)$$

$$\text{F1-score} = \frac{2\text{Precision} * \text{Recal}}{\text{Precision} + \text{Recal}}, \quad (12)$$

$$\text{Acc} = \frac{TP + TN}{N}, \quad (13)$$

with

$$\text{Recal} = \frac{TP}{TP + FN}, \quad (14)$$

where TP is the number of samples that correctly predicted positive classes, TN represents the number of samples that correctly predicted negative classes, N is the number of the samples, FN is the number of positive cases that are classified into negative cases, and FP is the number of samples with negative class errors falling into positive classes. Considering that the initial form is used for two classification tasks, we used the macro-average method to extend them to multi classification tasks.

Considering the limitation of the collected dataset, we design a simple network to deal with the problem of analysis and design of standard knowledge service system. According to practical conditions, we conducted experimental comparison on different simple neural networks, and the experimental results are shown in Table. 1. In this table, the method Two-MLP uses the network that has two full-connected layers and the method Three-MLP means a neural network that has three full-connected layers. In addition, the method Two-Conv uses a neural network that has two convolutional layers and two full-connected layers. The method One-Conv uses a neural network model that has one convolutional layer and two full-connected layers. From this table, we can see that the Two-MLP and Three-MLP methods achieve the similar performance, where the evaluation metrics are closed to each other correspondingly. Moreover, when using the Two-Conv, it can slightly improve the classification performance of standard knowledge. Additionally, the method One-Conv achieves a better performance improvement than Two-Conv method, including 1.98%, 1.94%, and 1.87% in Precision, F1-score, and Acc metrics, respectively. It means that exploiting more convolutional layers will degrade the classification performance of standard knowledge. Hence, we choose the method One-Conv to analysis and design the standard knowledge service system.

The figures from Fig. 4 to Fig. 7 shows the train process of the method One-Conv. In Fig. 4, it shows that the train loss goes down as the epoch goes up. When the epoch larger than 10, the rate of decline of the train loss decreases rapidly, and even changes little at last. Moreover, in Fig. 4, it shows that the accuracy of the train data goes up as epoch goes up. It means that the One-Conv model can efficiently fit the train data. In order to validate the effectiveness of the parameters of the model, the validation data are tested by this model during the training process, and the tested results are shown in Fig. 6 and 7. These two figures show that they have the same change trend with the train data, and it means that with the using of the validation data, a model with suitable parameters have been trained. Tabel 2 shows the validation results of the model. Compared with the test results in Tabel 1, they have the similar results among the three evaluation

Table 1. Numerical results of the deep model with different structures for the standard knowledge service system.

Methods	Precision	F1-score	Acc
Two-MLP	0.8961	0.8958	0.8984
Three-MLP	0.8990	0.8978	0.8997
Two-Conv	0.9093	0.9086	0.9105
One-Conv	0.9291	0.9280	0.9292

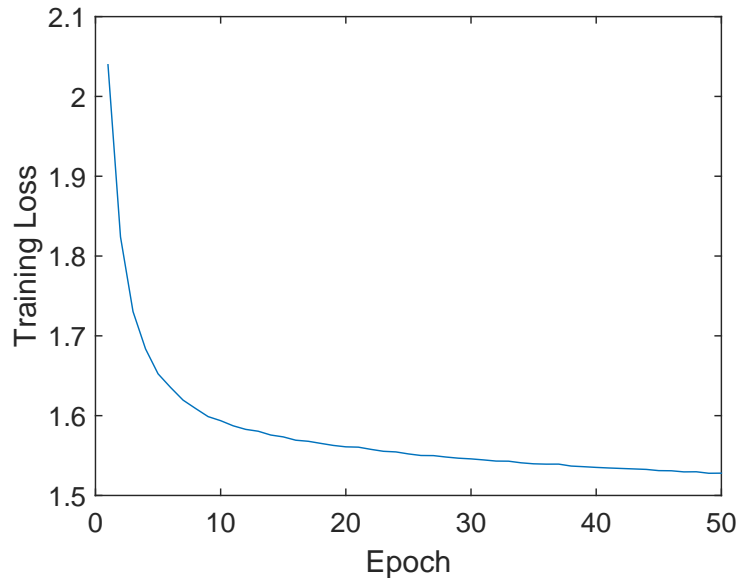


Figure 4. Train accuracy of the deep neural network model for the analysis and design of standard knowledge service system.

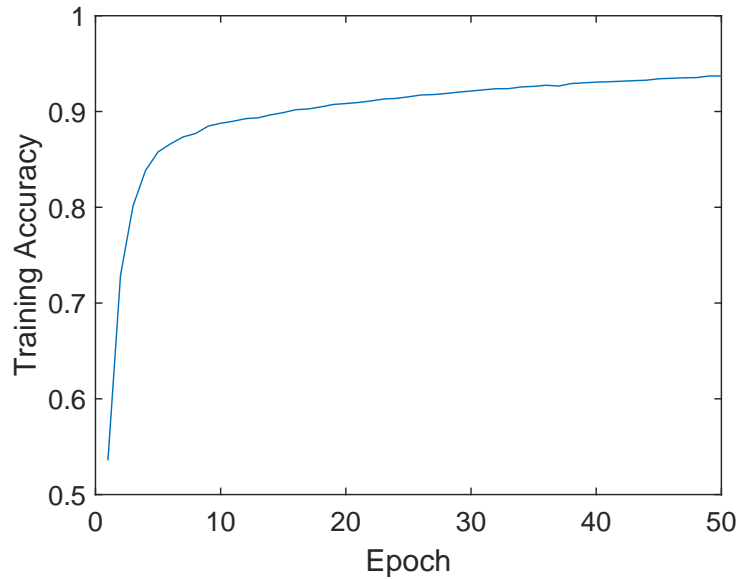


Figure 5. Train accuracy of the deep neural network model for the analysis and design of standard knowledge service system.

metrics. It means that the model can efficiently fit the test data without overfitting for the train data.

4. Conclusions

In this paper, we exploited the deep neural network model to analyze and design the standard knowledge

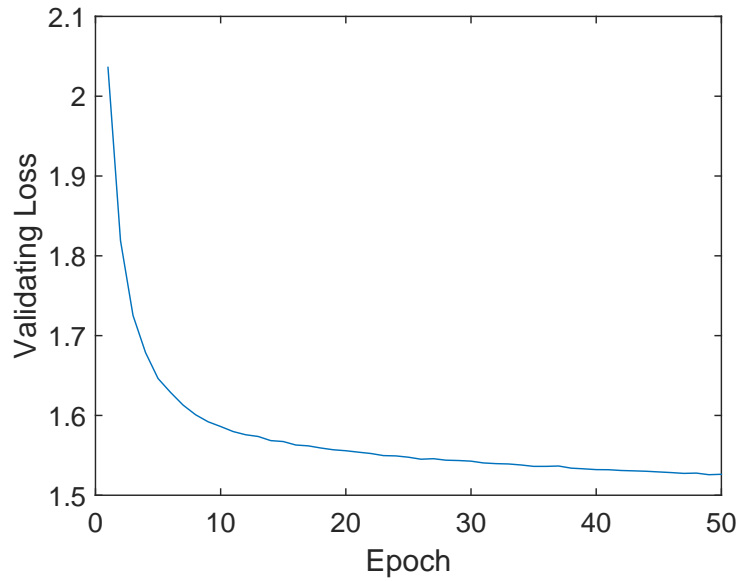


Figure 6. Validation loss of the deep neural network model for the analysis and design of standard knowledge service system.

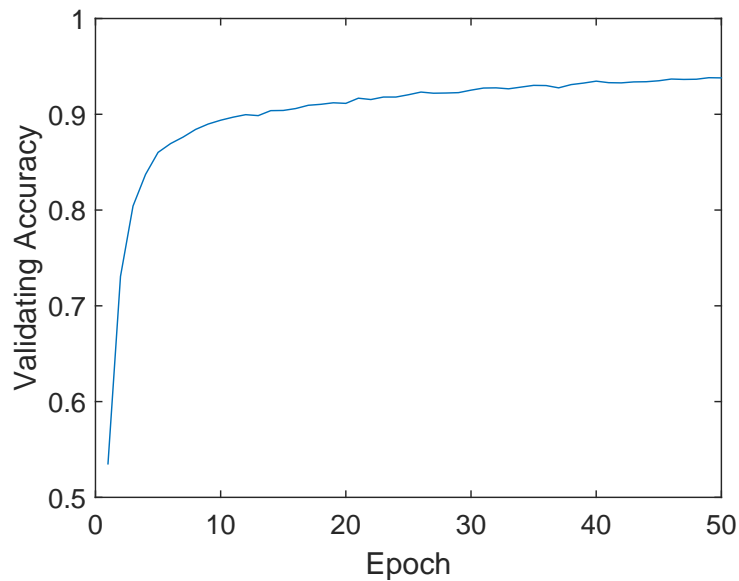


Figure 7. Validation accuracy of the deep neural network model for the analysis and design of standard knowledge service system.

Table 2. Validation numerical results for the selected model in the standard knowledge system.

Methods	Precision	F1-score	Acc
One-Conv	0.9303	0.9290	0.9306

service system. To overcome the limitations of the collected dataset, the shallow network model was applied in this paper. For small data sets, the usage of shallow structural model could effectively reduce the complexity of the model and the risk of model overfitting. In order to effectively classify the standard

knowledge, numerical labels were exploited to label different types of knowledge. The experimental results showed that the model could efficiently classify the standard knowledge into different types and avoid overfitting, by using the CNN model.

4.1. Data Availability Statement

The data of this work can be obtained through the email to the authors: Yuzhong Zhou (yuzhong_zhou@hotmail.com), Zhengping Lin (zhengping_lin@hotmail.com), Liang Tu (Liang_Tu1@hotmail.com), Junkai Huang (junkaihuangcsg@hotmail.com), and Zifeng Zhang (zifengzhangcsg@hotmail.com). As to this work, the authors would like to sincerely thank the following researchers for the meaningful discussions: Xiazhi Lai (xiazhilai@hotmail.com), Bowen Lu (bwlu@ieee.org), and Yinghao Guo (yinghaoguo@ieee.org).

4.2. Copyright

The Copyright was licensed to EAI.

References

- [1] H. Wang and Z. Huang, "Guest editorial: WWWJ special issue of the 21th international conference on web information systems engineering (WISE 2020)," *World Wide Web*, vol. 25, no. 1, pp. 305–308, 2022.
- [2] H. Wang, J. Cao, and Y. Zhang, *Access Control Management in Cloud Environments*. Springer, 2020. [Online]. Available: <https://doi.org/10.1007/978-3-030-31729-4>
- [3] R. Zhao and M. Tang, "Profit maximization in cache-aided intelligent computing networks," *Physical Communication*, vol. PP, no. 99, pp. 1–10, 2022.
- [4] Q. H. Ngo, M. T. Kechadi, and N. Le-Khac, "Knowledge representation in digital agriculture: A step towards standardised model," *Comput. Electron. Agric.*, vol. 199, p. 107127, 2022.
- [5] N. Melluso, I. Grangel-González, and G. Fantoni, "Enhancing industry 4.0 standards interoperability via knowledge graphs with natural language processing," *Comput. Ind.*, vol. 140, p. 103676, 2022.
- [6] H. Wang, Y. Wang, T. Taleb, and X. Jiang, "Editorial: Special issue on security and privacy in network computing," *World Wide Web*, vol. 23, no. 2, pp. 951–957, 2020.
- [7] R. Zhao and M. Tang, "Impact of direct links on intelligent reflect surface-aided MEC networks," *Physical Communication*, vol. PP, no. 99, pp. 1–10, 2022.
- [8] K. Baghery, A. González, Z. Pindado, and C. Ràfols, "Signatures of knowledge for boolean circuits under standard assumptions," *Theor. Comput. Sci.*, vol. 916, pp. 86–110, 2022.
- [9] S. Basu, D. Rutstein, C. Tate, A. Rachmatullah, and H. Yang, "Standards-aligned instructional supports to promote computer science teachers' pedagogical content knowledge," in *SIGCSE 2022: The 53rd ACM Technical Symposium on Computer Science Education*, Providence, RI, USA, March 3-5, 2022, Volume 1, L. Merkle, M. Doyle, J. Sheard, L. Soh, and B. Dorn, Eds. ACM, 2022, pp. 404–410.
- [10] P. Lee, "Investigating the knowledge spillover and externality of technology standards based on patent data," *IEEE Trans. Engineering Management*, vol. 68, no. 4, pp. 1027–1041, 2021.
- [11] J. Lu and J. Xia, "Performance analysis for IRS-assisted MEC networks with unit selection," *Physical Communication*, vol. 2022, no. 8.
- [12] X. Hu, J. Wang, and C. Zhong, "Statistical CSI based design for intelligent reflecting surface assisted MISO systems," *Science China: Information Science*, vol. 63, no. 12, p. 222303, 2020.
- [13] X. Hu, C. Zhong, Y. Zhu, X. Chen, and Z. Zhang, "Programmable metasurface-based multicast systems: Design and analysis," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1763–1776, 2020.
- [14] L. Zhang and C. Gao, "Deep reinforcement learning based IRS-assisted mobile edge computing under physical-layer security," *Physical Communication*, vol. PP, no. 99, pp. 1–10, 2022.
- [15] D. Cai, P. Fan, Q. Zou, Y. Xu, Z. Ding, and Z. Liu, "Active device detection and performance analysis of massive non-orthogonal transmissions in cellular internet of things," *Science China information sciences*, vol. 5, no. 8, pp. 182301:1–182301:18, 2022.
- [16] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.
- [17] S. Tang and X. Lei, "Collaborative cache-aided relaying networks: Performance evaluation and system optimization," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–12, 2022.
- [18] Y. Wu and C. Gao, "Task offloading for vehicular edge computing with imperfect CSI: A deep reinforcement approach," *Physical Communication*, vol. PP, no. 99, pp. 1–10, 2022.
- [19] X. Lai, "Outdated access point selection for mobile edge computing with cochannel interference," *IEEE Trans. Vehic. Tech.*, vol. 71, no. 7, pp. 7445–7455, 2022.
- [20] K. He and Y. Deng, "Efficient memory-bounded optimal detection for GSM-MIMO systems," *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4359–4372, 2022.
- [21] S. Tang, "Dilated convolution based CSI feedback compression for massive MIMO systems," *IEEE Trans. Vehic. Tech.*, vol. 71, no. 5, pp. 211–216, 2022.
- [22] S. Tang and L. Chen, "Computational intelligence and deep learning for next-generation edge-enabled industrial IoT," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 3, pp. 105–117, 2022.
- [23] L. Chen, "Physical-layer security on mobile edge computing for emerging cyber physical systems," *Computer Communications*, vol. PP, no. 99, pp. 1–12, 2022.
- [24] J. Sun, X. Wang, Y. Fang, X. Tian, M. Zhu, J. Ou, and C. Fan, "Security performance analysis of relay networks based on-shadowed channels with rhis and cees," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [25] X. Deng, S. Zeng, L. Chang, Y. Wang, X. Wu, J. Liang, J. Ou, and C. Fan, "An ant colony optimization-based routing algorithm for load balancing in leo satellite networks," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.

- [26] C. Wang, W. Yu, F. Zhu, J. Ou, C. Fan, J. Ou, and D. Fan, "Uav-aided multiuser mobile edge computing networks with energy harvesting," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [27] J. Chen, Y. Wang, J. Ou, C. Fan, X. Lu, C. Liao, X. Huang, and H. Zhang, "Albrl: Automatic load-balancing architecture based on reinforcement learning in software-defined networking," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [28] C. Ge, Y. Rao, J. Ou, C. Fan, J. Ou, and D. Fan, "Joint offloading design and bandwidth allocation for ris-aided multiuser mec networks," *Physical Communication*, p. 101752, 2022.
- [29] L. Chen and X. Lei, "Relay-assisted federated edge learning: Performance analysis and system optimization," *IEEE Transactions on Communications*, vol. PP, no. 99, pp. 1–12, 2022.
- [30] W. Zhou and X. Lei, "Priority-aware resource scheduling for uav-mounted mobile edge computing networks," *IEEE Trans. Vehic. Tech.*, vol. PP, no. 99, pp. 1–6, 2023.