IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 January 2024, accepted 14 February 2024, date of publication 21 February 2024, date of current version 13 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3368854

RESEARCH ARTICLE

Artificial Intelligence-Enabled 5G Network Performance Evaluation With Fine Granularity and High Accuracy

QING ZHANG[1,2], TAOYE ZHANG[2], BIN CHEN[2], JI YAN[1], ZHONGYUAN ZHAO[1], (Member, IEEE), XIAOFEI QIN[2], CHAO CAI[2], AND XIANKUI LUO[2]

[1]School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]China Unicom Intelligent Network Innovation Center, Beijing 100048, China

Corresponding author: Zhongyuan Zhao (zyzhao@bupt.edu.cn)

This work was supported in part by Beijing Natural Science Foundation under Grant L223026 and in part by the 5G Evolution Wireless Air Interface Intelligent Research and Development and Verification Public Platform Project under Grant 2022-229-220.

ABSTRACT Network performance evaluation is crucial in ensuring the effective operation of 5G wireless networks, offering valuable insights into evaluating network status and user experience. However, the complexity of network conditions, characterized by high dynamics and diverse user requirements across various vertical applications, presents a significant challenge for generating accurate and detailed evaluation results using existing algorithms. To provide a feasible solution for this issue, an artificial intelligence-enabled 5G network performance evaluation scheme for private 5G networks is proposed. First, the network performance evaluation at different granularities is modeled with the deployment of network performance evaluation introduced. Furthermore, an intelligent network performance evaluation architecture based on residual networks with the attention mechanism is introduced, which can generate evaluation scores based on key performance indicators of reliability, accessibility, utilization, integrity, mobility and retainability. Additionally, the corresponding training strategies for the intelligent model, catering to different evaluation granularity, are thoroughly designed. Finally, to validate the effectiveness of the proposed scheme, comprehensive experiments are conducted using practical 5G network operation system data. The experimental results demonstrate the scheme's ability to achieve highly accurate evaluations with fine spatial granularity. These findings establish the feasibility and efficacy of the proposed artificial intelligence-enabled scheme in enhancing 5G network performance evaluation.

INDEX TERMS Network performance evaluation, 5G vertical applications, artificial intelligence, convolutional neural networks, attention mechanism.

I. INTRODUCTION

It is anticipated that the rapid development of the fifth-generation (5G) mobile communication systems facilitates various emerging network services [1], which can meet the demanding requirements of low latency, high throughput and massive connections [2]. Plentiful advanced technologies of 5G communication provide great convenience and benefits to support a diversity of creative applications, which include industrial automation, distributed energy control, cloud-based virtual reality (VR), network communication and so on [3]. For instance, the device-to-device (D2D) communication technologies of 5G communication enable terminal devices to circumvent the cellular base station, thereby facilitating direct information sharing with other targeted devices [4]. Moreover, as introduced in [5], robust and secure traffic management with lower collision probabilities can be achieved by the 5G advanced internet of vehicles (IoV) technologies. In [6], wearable technology is introduced as a promising 5G application, which can record different

The associate editor coordinating the review of this manuscript and approving it for publication was M. Venkateshkumar.

physiological signals continuously to help health monitoring and management [7]. In light of the multifarious advantages and enhanced convenience afforded by 5G technologies across diverse domains, it becomes imperative to evaluate the performance of 5G networks [8]. In particular, accurate network performance evaluation can facilitate monitoring real-time network operation and maintenance status so as to detect network anomalies and further adjust resource allocation, thereby ensuring the efficient network management and enhancing the user experiences. Moreover, synthetic network performance evaluation unifies multiple key performance indicators into single representative evaluation result, making it more distinctly and directly for network operators to observe the comprehensive situations of various network services and applications. Therefore, designing optimal network performance evaluation schemes of 5G networks has always been a significant and meaningful research field.

There have existed some related network performance evaluation methods, the main ideas of which can be roughly classified into two categories. On the one side, conventional evaluation algorithms based on linear weighting and regression scheme are applied. In [9], an evaluation method based on quality of service (QoS) performance is proposed, which linearly calculate importance weights of network nodes and links for performance measurement. In [10], a network performance evaluation approach is put forward based on objective weights determination, where the weight of each network performance parameter is calculated based on linear calculation. In [11], the concept of Quality of Monitoring (QoM) is introduced for the evaluation and enhancement of M-plane data within the context of 5G networks, which integrates a data-driven algorithm in tandem with a lossy-compression methodology. In [12], a real-time data measurement methodology based on hardware platform is proposed to evaluate and analyse the 5G network performance. In [13], the performance of 5G cognitive radio networks is assessed, where the numerical analysis of multiple network performance metrics is conducted to evaluate the Quality of Service (QoS) under different scenarios. On the other side, the technologies of artificial intelligence (AI) are utilized for network performance evaluation. In [14], the performance of 5G wireless sensor network is evaluated by utilizing the machine learning. In [15], a graph neural network is constructed for large-scale network performance evaluation, which is significantly less time-consuming than traditional methods. References [16] and [17] center their attention on the evaluation and prediction of wireless network traffic data utilizing deep learning networks. In [16], a residual network amalgamating multiple mechanisms is designed which can comprehensively capture the spatio-temporal correlation of the evaluated data. In [17], an enhanced multilayer perception deep neural network is designed, exhibiting heightened accuracy in the realms of data evaluation and prediction.

Although previous research on evaluating the performance of 5G networks has offered valuable insights, there is a lack of precise and applicable evaluation methods with high accuracy across various 5G scenarios. Traditional linear evaluation methods fall short in capturing complex correlation features between network data and results. Moreover, targeted AI-based evaluation methods for emerging 5G scenarios are underexplored. To overcome the above issues, it is feasible to exploit the powerful feature learning ability of AI technologies [18] to generate accurate and geographically precise evaluation results directly based on network indicator data. And as one of the most important and widely studied technologies, AI has made impressive achievements in many feature-learning areas [19], including time series prediction [20], signal identification [21], [22], image classification [23], [24], object recognition [25] and so on. In the research fields of wireless communication, AI is also regarded as an efficient way to address conventional communication problems. Reference [26] suggests applying AI to cellular networks and puts forward an AI-empowering architecture, where an AI controller is introduced as an independent network entity that can communicate with core networks (CNs) and radio access network (RAN). References [27], [28], and [29] utilize the AI technologies to help improve the efficiency of network routing rules. In [27], a 3-layer deep neural network is constructed to classify routing node degree. In [28], a deep belief network is proposed to determine the next routing node and it is also employed to construct a software defined router. In [29], tensors are utilized to represent weights, hidden layers and biases in deep belief networks, which obtain better routing performance. Reference [30] and [31] investigate network scheduling based on deep learning technologies. In [30], a deep Q learning-powered scheduling mechanism is put forward, which aims to decrease the energy consumption in 5G real-time systems. In [31], deep Q leaning is used for scheduling in roadside communication networks.

Since massive research work has confirmed the powerful capacity of AI applied in the field of communication, it is appropriate to utilize its splendid learning ability to evaluate the performance of 5G networks. And compared with other types of deep learning networks, convolutional neural network (CNN) can achieve higher learning efficiency and has more powerful feature learning ability in the face of massive network data. Therefore, CNN is widely explored and utilized for network-level and user-level data analysis [32]. In [33], a CNN-based scheme is proposed for network-level traffic classification, which aims to recognize specific protocols or services based on traffic in network. In [34], CNN is employed to forecast the mobile traffic. In [35] and [36], CNNs are exploited for efficient mobile health data analysis and medical data analysis, respectively. In particular, the applicability of CNNs is explored in the field of synthetic network evaluation and management. In [37], a hybrid CNN-based model is proposed to achieve real-time evaluation and anomaly detection of network data in central clouds, and the simulation results show that the proposed model exhibits an improvement in terms of accuracy. In [38], CNN is utilized to construct deep Q-network, which evaluates the user QoS demand data and energy consumption data to facilitate further network resource optimization. Moreover, recent research begins to focus on the introduction of attention mechanism to CNN to further enhance the model performance when processing large amounts of data [39], which has been explored in various scenarios including image caption generation [40], [41], machine translation [42], [43], speech recognition [44], [45], etc. Reference [47] introduces the attention mechanism to residual convolutional neural networks, which is called residual attention network and the attention mechanism is achieved by stacking attention modules that can generate attention-aware features. Reference [46] achieves the attention mechanism of CNN by adding a squeeze-and-excitation (SE) block, which explicitly models inter-dependencies between different channels to adaptively learn channel-wise features and proves to have a great performance enhancement of learning multi-channel data features. The attention mechanism-based CNNs have more powerful capabilities of data analysis and feature learning, thereby making it highly suitable for the performance evaluation of 5G networks.

Despite the great success of applying AI technologies in the research field of communication, there still exist several key challenges of evaluating the performance of 5G networks, which are elaborated as follows: Firstly, highly adaptive and practically applicable algorithms of evaluating the performance 5G networks are lacking, due to the unavailability of practical 5G network data in different 5G private network scenarios. Therefore, it is hard to train network performance evaluation models with high evaluation accuracy based on the practical 5G network data. Secondly, data labels of collected multimodal 5G network data are lacking, which poses a great obstacle to the specific design of the supervised task during the training of network performance evaluation models constructed based on CNN. Thirdly, current algorithms focus on the performance evaluation of cellular-scale network areas, which can not achieve geographically precise network performance evaluation, i.e, the performance evaluation of network areas with much smaller coverage regions and finer evaluation granularity.

In this paper, to more accurately and precisely evaluate the performance of 5G networks, an AI-enabled network performance evaluation scheme which utilizes attention mechanism-based CNN is proposed. And the attention mechanism in the proposed CNN is achieved by SE module as introduced in [46]. In particular, the proposed scheme supports cellular-scale and fine-grained network performance evaluations based on training different types of multimodal network data, which is elaborated in the following sections. The main contribution of this paper is summarized as follows:

- Firstly, both the cellular-scale and fine-grained network performance evaluation are modeled and the deployment of network performance evaluation in 5G network architecture is introduced to facilitate the achievement of high-performance network evaluation based on flexible and efficient 5G network functions.
- Secondly, an intelligent network performance evaluation scheme is proposed, which takes network performance data as input, to generate both cellular-scale and fine-grained network performance evaluation results with high accuracy. In particular, attention mechanism with multiple residual blocks is adopted in the proposed intelligent model to enhance the overall model performance. Moreover, training strategies are designed, where a weighted sum approach (WSA)-based method is provided to address the problem of lacking data labels and loss functions are also designed for network performance evaluations with different evaluation granularity.
- Finally, experimental results obtained based on practical 5G network operation system data are presented to verify the evaluation performance of the proposed CNN-based model. In particular, evaluation results generated by the fine-grained intelligent model are transferred into visible heat maps, which illustrate high geographical precision of the corresponding evaluation results. And correlation analysis between output results and input data is carried out for both the cellular-scale and fine-grained intelligent models, which shows the high accuracy of proposed evaluation scheme.

II. SYSTEM MODEL AND DEPLOYMENT OF 5G NETWORK PERFORMANCE EVALUATION

A. ANALYTICAL SYSTEM MODEL OF CELLULAR-SCALE AND FINE-GRAINED NETWORK PERFORMANCE EVALUATION

Conventional network performance evaluation generally focuses on the overall performance, namely the synthetic situations of network operation and service behaviors of a certain 5G network area $\mathcal{D}$, which covers a set of network cells $\mathcal{C} = \{C_1, \ldots, C_n\}$, i.e., $\bigcup_{i=1}^{n} C_i = \mathcal{D}, C_x \bigcap C_y = \emptyset$ and $x \neq y$. And the evaluation score of a specific cell C_i is calculated based on a series of network performance criteria, which can be written as follows:

$$R_i^c = F\left(\mathbf{N}_i, \mathbf{E}_i, \mathbf{A}_i\right), \tag{1}$$

where R_i^c is the corresponding score of C_i. $\mathbf{N}_i$ denotes the cellular network indicator data of transmission including criteria such as latency, bandwidth, packet loss rate, etc. $\mathbf{E}_i$ and $\mathbf{A}_i$ denote the energy efficiency data and coverage capability data of C_i, respectively. Although different conventional evaluation methods involve different network performance indicator systems of criteria, the overall performance of a cellular network area can be adequately evaluated based on the above three perspectives.

However, conventional network performance evaluation can not generate evaluation results with fine spacial granularity to meet the growing demands of geographically fine-grained 5G network optimization and management.

Therefore, fine-grained network performance evaluation is also studied in this section. Consider that the performance of a 5G network $\mathcal{N}^{cee}$ is evaluated with the serving group of mobile users, i.e., $\mathcal{M}_u = \{Mu_1, \ldots, Mu_m\}$, and the coverage spatial area of it is $\mathcal{D}$. In particular, the spatial domain of the network area $\mathcal{D}$ is partitioned into a collection of equidistant sub-regions that exhibit no overlap, i.e.,

$$\mathcal{D}_{sub} = \{D_1, \ldots, D_k | \bigcup_{i=1}^{n} D_i = \mathcal{D}, D_x \bigcap D_y = \emptyset, x \neq y\}. \tag{2}$$

It should be noted that the coverage area of each sub-region is typically smaller than a network cell, i.e., $\mathcal{D}(D_i) < \mathcal{D}(C_i)$, in order to facilitate the network performance evaluation of 5G networks with finer granularity.

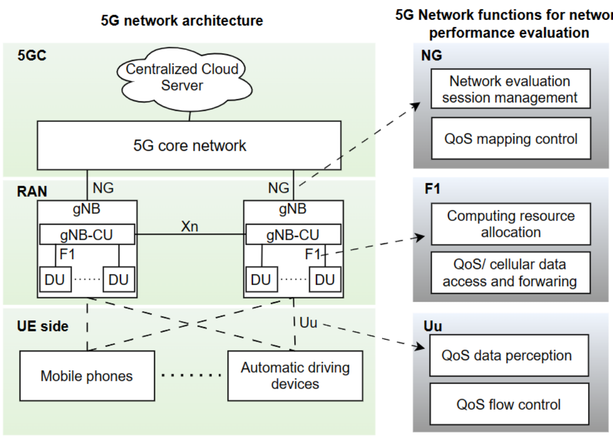


FIGURE 1. Deployment of proposed network performance evaluation scheme in practical 5G networks.

Since conventional network performance evaluation as presented in (1) generates evaluation results on the basis of network-level data, the fine-grained network performance evaluation should be implemented based on smaller-scale network performance data. In the current network operation systems, the user QoS data can be obtained directly from network management devices and provide network performance reference from a more geographically granular perspective. Without loss of universality, we consider a specific sub-region D_i, the performance evaluation score of it can be obtained based on user QoS data as:

$$\begin{aligned} R_i^s &= F(\mathbf{Q_{i_1}}, \ldots, \mathbf{Q_{i_l}}), \\ \mathcal{M}_u^i &= \{Mu_{i_1}, \ldots, Mu_{i_l} | Mu_{i_1}, \ldots, Mu_{i_l} \in \mathcal{D}(D_i)\}, \end{aligned} \tag{3}$$

where $\mathbf{Q_{i_j}}$ represents the QoS data of mobile user Mu_{i_j}, $Mu_{i_j} \in \mathcal{M}_u^i$, $i_j = i_1, \ldots, i_l$, and $\mathcal{M}_u^i$ denote a group of mobile users which distribute in the coverage domain $\mathcal{D}(D_i)$ of fine-grained network area D_i.

B. DEPLOYMENT IN PRACTICAL 5G NETWORKS

In order to achieve high-performance network performance evaluation, the deployment of proposed network performance evaluation scheme in practical 5G network architecture is applied and elaborated in this section. Owing that different parts of 5G networks have markedly different functions and characteristics, deploying different functional modules of network performance evaluation in corresponding 5G network interfaces can facilitate the integration of powerful and complete centralized network management mechanism, disperse flexible network edge computation capability, and massive user terminal data. Since the 5G network consists of 5G core network (5GC) and radio access network (RAN), the deployment of network performance evaluation in practical 5G architecture can be introduced mainly from the above two perspectives, with different network functional interfaces achieving relevant specific operation and management functions for 5G network performance evaluation.

As illustrated in Figure 1, the 5GC is connected with the RAN through the NG interface, which is responsible for the establishment and management of network performance evaluation session. Moreover, the NG interface can also achieve the QoS mapping which maps user-uploaded multimodal network data such as numerical text, video and so on to specific QoS data which can be further evaluated. The 5G RAN mainly consists of 5G base stations (gNodeBs), inside which a centralized unit (CU) is connected to multiple distributed units (DUs) through the F1 interface. In the procedure of 5G network performance evaluation, the F1 interface is accountable for allocation of computing resources to facilitate the training of intelligent evaluation models, and the access management of cellular-scale network performance data is also achieved by the F1 interface. What's more, user experience data collected by DUs is also forwarded by this interface. The Uu interface of RAN takes control of the interaction between RAN and user side, as vast user experience data such as QoS flow is controlled and managed by the Uu interface, which is crucial for user data collection and perception so as to further facilitate fine-grained network performance evaluation. Moreover, please note that both the cellular data and user experience data involved in the network performance evaluation are collected from different scenario-based network services of practical private 5G network platform.

III. AI-ENABLED 5G NETWORK PERFORMANCE EVALUATION ALGORITHM WITH FINE GRANULARITY AND HIGH ACCURACY

In order to achieve high-performance 5G network performance evaluation, an artificial intelligence-enabled model is proposed in this section. In particular, the proposed intelligent model aims to simultaneously generate both the conventional cellular-scale and under-explored fine-grained network performance evaluation results, based on different types of input data. The concrete modules of relevant data collection and processing, the specific architecture design and the corresponding training strategies are introduced in detail in the following sections.

A. DATA COLLECTION

Since both the cellular-scale and fine-grained network performance evaluations require numerous network multimodal data, it is of great significance to efficiently collect network performance data such as the network-level data and user QoS data from totally different network devices. The specific data collecting procedure in each part of the network operation systems is elaborated as follows:

- **Mobile user devices:** Amount network terminal devices in 5G networks are utilized to collect user QoS data, which can reflect users' complete perception of network status and service quality and is the main data of fine-grained evaluation as introduced in (3). In particular, the collected user QoS data can be transmitted into the 5G gNodeB for further storage and utilization.
- **5G base stations (gNodeBs):** While the CU in the gNodeB controls several DUs to achieve edge services deployment and relevant physical-layer functions, each DU takes charge of a certain cellular network area. Therefore, with the help of these functional units, network-level performance indicator data of cellular-scale network performance evaluation can be collected and further uploaded to the 5GC for massive data storage and processing.
- **Central cloud servers:** In the central cloud, massive global network performance data including user QoS data and network-level performance indicator data are stored, which can be utilized for global computing with high efficiency such as model training and data analysis based on powerful computing resources of 5G.

B. PREPROCESSING FOR MULTIMODAL NETWORK DATA

In order to facilitate the model training of proposed intelligent model and further generation of evaluation results, the original data of cellular-scale and fine-grained evaluation should be processed before inputting them into the intelligent model. As shown in Figure 2, there are several steps to process the original sampled data collected directly from the network operation devices, which are as follows:

- **Positive conversion:** To unify the evaluation trend of different input data and better facilitate the generation of accurate performance evaluation results, the preprocessing layer of positive conversion is set firstly. And the positive conversion aims to convert the input different types of criteria to benefit criteria, the larger values of which represent better indicator performance. In particular, the positive conversion is conducted on specific network indicators, i.e,

$$x_p = Pos(x_c), \quad x_c \in \{\mathbf{N}_o, \mathbf{E}_o, \mathbf{A}_o, \mathbf{Q}_o\}, \tag{4}$$

  where x_c denotes the specific performance indicator of sampled network transmission indicator data set $\mathbf{N}_o$, energy efficiency data set $\mathbf{E}_o$, coverage capability data set $\mathbf{A}_o$, or QoS data set $\mathbf{Q}_o$. And $Pos(\cdot)$ denotes the positive conversion function.
- **Normalization:** In order to equalize the scales of different types of input data and prevent the measurement bias, which significantly effect the performance of model training, the normalization of indicator data is employed to process the original data, i.e.,

$$x_n = Norm(x_p), \quad x_p \in \{\mathbf{N}_p, \mathbf{E}_p, \mathbf{A}_p, \mathbf{Q}_p\}, \tag{5}$$

  where x_n denotes the indicator after the normalization layer, and $\mathbf{N}_p, \mathbf{E}_p, \mathbf{A}_p, \mathbf{Q}_p$ denote data sets of performance evaluation indicator after the preprocessing of positive conversion corresponding to $\mathbf{N}_o, \mathbf{E}_o, \mathbf{A}_o, \mathbf{Q}_o$, respectively.
- **Vectorization:** Due to the demand that the input data of the deep learning model should be in the form of vector, vectorization is conducted on the data after the normalization and positive conversion, i.e.,

$$\begin{aligned} \mathbf{x}_{in} &= vec(\mathcal{T}), \\ \mathcal{T} &= \{\mathbf{N}_n, \mathbf{E}_n, \mathbf{A}_n, \mathbf{Q}_n\}, \end{aligned} \tag{6}$$

  where $\mathbf{x}_{in}$ denotes the input vector of the intelligent model and vec$(\cdot)$ denotes the vectorization function which should be conducted on both the training data and the validation data. $\mathcal{T}$ denotes the indicator set after the normalization and $\mathbf{N}_n, \mathbf{E}_n, \mathbf{A}_n, \mathbf{Q}_n$ are corresponding to $\mathbf{N}_p, \mathbf{E}_p, \mathbf{A}_p, \mathbf{Q}_p$.

Due to the rapid development of data feature engineering, it is feasible to utilize the existing simulation tools to fast achieve indicator data preprocessing illustrated in (4), (5) and (6). The specific procedures of positive conversion, normalization and vectorization of the cellular-scale and fine-grained input data are elaborated in the following section of simulation.

C. THE NETWORK ARCHITECTURE OF PROPOSED SCHEME

As illustrated in Figure 2, in the proposed CNN-based network performance evaluation scheme, the network-level data including network transmission data $\mathbf{N}$, energy efficiency data $\mathbf{E}$ and coverage capability data $\mathbf{A}$ after data process are input into the intelligent model to generate the cellular-scale evaluation results set $\mathbf{r}^c$ according to (1), while the user QoS data $\mathbf{Q}$ after the data process are fed into the model to obtain fine-grained evaluation results $\mathbf{r}^s$. Please note that the intelligent models of cellular-scale and fine-grained evaluation are trained independently with totally two kinds of different weights but an identical shared attention mechanism-based CNN structure illustrated in Figure 2, which mainly includes the following parts:

- **Convolution layer for feature extraction:** Since the amounts of both the cellular-scale and fine-grained input data are huge, leading to extremely complex local patterns and correlations between input data and final evaluation results. Therefore, the convolution layer is set to initially extract the features of input data and reduce the input data dimension for further learning of data

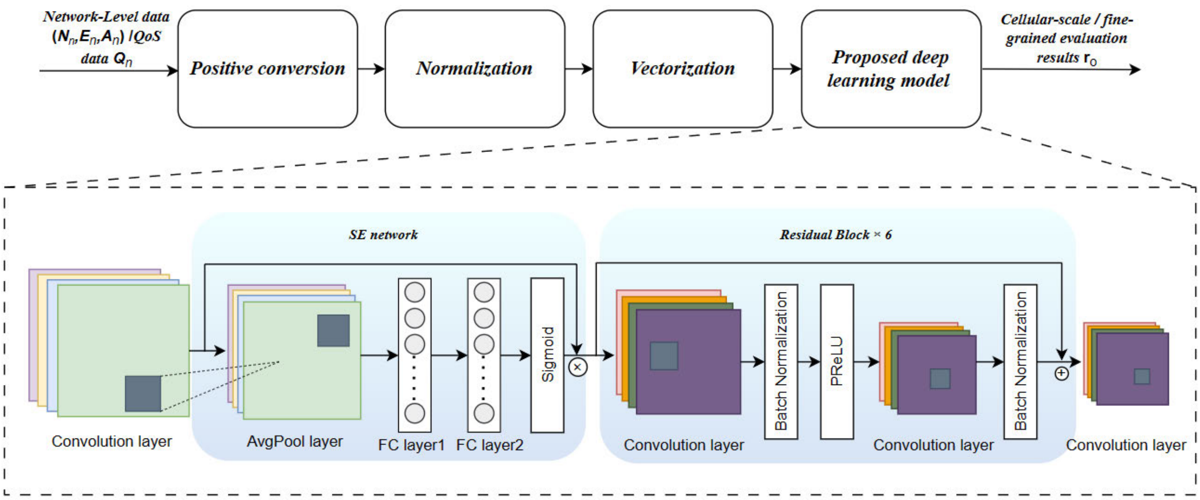


FIGURE 2. Proposed AI-enabled network performance evaluation scheme.

distribution by convolution operation, i.e.,

$$\mathbf{y}_c = \mathbf{x}_{in} * \mathbf{k}, \tag{7}$$

where $\mathbf{y}_c$ denotes the output of the convolution layer with the input vector $\mathbf{x}_{in}$. And $\mathbf{k}$ is the learning kernel of convolution operation. In particular, the shape of $\mathbf{y}_c$ can be controlled to the desired value by adjusting the size of convolution kernel, and typically the size of $\mathbf{y}_c$ is smaller than $\mathbf{x}_{in}$ by convolution operation, which achieves the dimension reduction of input data features.

- **SE network for filtrating better feature channels:** Owing that the convolution operation mainly integrates different features spatially, the differences between data features of different feature channels are ignored, which weakens the feature learning ability of the model. Therefore, to filtrate most informative channel-level features and suppress those that are not important, SE network are employed in this layer with the following two main functional operations: Squeeze operation is first achieved by the average pooling (AvgPool) layer, as shown in Figure 2, to shrink the feature maps $\in \mathbb{R}^{w \times h \times c}$ through spatial dimensions ($w \times h$) and obtain channel-level global features of the input, i.e.,

$$\mathbf{z}_s = F_{sq}(\mathbf{y_c}) = \frac{1}{h \times w} \sum_{i=1}^{h} \sum_{j=1}^{w} \mathbf{y}_c(i,j), \mathbf{z}_s \in \mathbb{R}^c, \tag{8}$$

where $\mathbf{z}_s$ is the output of the squeeze operation with the input convoluted $\mathbf{y}_c$. Two full connection layers with the sigmoid activation function are employed following the AvgPool layer to capture the complex non-linear correlations between different feature channels of $\mathbf{z}_s$, which is the core of the excitation operation and can be expressed as

$$\mathbf{z}_e = F_{ex}(\mathbf{z}_s) = \text{Sigmoid}(W_2 W_1 \mathbf{z}_s), \tag{9}$$

where $W_1 \in \mathbb{R}^{\frac{c}{r} \times c}$ and $W_2 \in \mathbb{R}^{c \times \frac{c}{r}}$ denote the weights of the first the second full connection layers, respectively. $\mathbf{z}_e$ denotes the corresponding output. In particular, the first full connection layer aims to reduce the dimensions with the dimension-reducing coefficient r, while the second full connection layer recover vector to its original dimensions. Then, features of different channels are given different weights, which are multiplied over the original input $\mathbf{y}_c$ and can make the model acquire more powerful capacity of distinguishing between different channel-level features. And the output of the SE network can be written as

$$\mathbf{x}_r = F_{scale}(\mathbf{y}_c, \mathbf{z}_e) = \mathbf{y}_c \cdot \mathbf{z}_e, \tag{10}$$

where $\mathbf{x}_r$ denotes the output vector of the SE network which is also the input vector of the next residual blocks.

- **Residual blocks for down-sampling:** To adequately learn the feature distribution of input data and generate accurate and precise evaluation results, multiple residual blocks in a tandem way are employed for down-sampling and deep feature learning while avoiding the problem of gradient vanishing by utilizing residual learning mechanism, i.e,

$$\mathbf{y}_r = R(\mathbf{x}_r) + \mathbf{x}_r, \tag{11}$$

where $\mathbf{x}_r$ and $\mathbf{y}_r$ denote the input and output vectors respectively. And $R(\mathbf{x}_r)$ denotes the residual mapping correlation between the true distributions of the input and output, which is learned by the residual blocks layer. As shown in Figure 2, there are two convolution layers utilized in each residual block, which achieves feasible outputs of down-sampling while considering feature distribution correlations between input data and 5G network performance. Moreover, to guarantee the convergence of proposed intelligent model during the training phase, batch normalization layer is added following each convolution layer in the residual blocks. To enhance the expressive ability of the model and feasibility of convoluted outputs, the activation function of parametric rectified linear unit (PReLU) is exploited following the first batch normalization layer.

- **Evaluation results output layer based on convolution:** With the down-sampling results provided by the multiple residual blocks, the final output performance evaluation results for evaluated 5G networks can be generated in the form of evaluation numerical scores, which can be written as

$$\mathbf{r}_o = [S_1, \ldots, S_l] = \mathrm{Conv}(\mathbf{y}_{lr}), \tag{12}$$

where $\mathbf{r}_o$ denotes the final output performance evaluation results set of intelligent model and $S_1, \ldots, S_l$ denote the evaluation scores of l cellular network areas or fine-grained network areas, which depends on the type of input data. $\mathbf{y}_{lr}$ denotes the output vector of the last residual block and $\mathrm{Conv}(\cdot)$ denotes the convolution operation in the last convolution layer.

D. THE TRAINING STRATEGIES OF PROPOSED SCHEME

In the training phase of the proposed intelligent model, the training strategy is of great significance to generate accurate cellular-scale evaluation results and precise fine-grained evaluation results. However, the following two challenges of training the proposed intelligent model are existed as:

- First, during the design of specific objective training task of the model, the labels of input training data, i.e., the measured cellular or fine-grained evaluation results are uncertain and lacking, which makes it difficult for the intelligent model to learn the correlation between the input original vectors and true performance evaluation scores.
- Second, due to the differences between the cellular-scale and fine-grained input training data, which are cellular-scale performance data collected from the cloud-edge devices and QoS data collected from the edge-end devices respectively, the training strategy including the label calculation as aforementioned and loss function design should be totally different.

To overcome the above issues, the data labelling approaches for cellular-scale and fine-grained performance evaluation are proposed in this section. Besides, the corresponding loss function designs of each of them are studied to guarantee the accuracy of generating evaluation scores.

1) WSA-BASED PERFORMANCE EVALUATION APPROACH FOR DATA LABELING

Although conventional network performance evaluation methods can not achieve high precise evaluation results, it is feasible for them to calculate the labels for training data of cellular-scale network areas. In particular, owing that the input training data of cellular-scale and fine-grained evaluation are quite different, the specific calculation of labels should be different and designed separately.

In this section, a weighted sum approach (WSA)-based method is exploited to calculate the data labels, whose core idea is to generate the performance evaluation scores by averaging the input data. And the corresponding data labelling procedures of cellular-scale and fine-grained evaluation are as follows:

- **Data labelling for cellular-scale evaluation:** As for the data labelling of cellular-scale evaluation, without loss of generality, we focus on a specific 5G network area C_n with the cellular scale, the label of it can be expressed as

$$S_n^c = \frac{1}{k_1}\sum_{i=1}^{k_1}\alpha_n n_i + \frac{1}{k_2}\sum_{j=1}^{k_2}\alpha_e e_j + \frac{1}{k_3}\sum_{m=1}^{k_3}\alpha_c a_m, \tag{13}$$

Algorithm 1 Paradigm of the Proposed Scheme for Cellular-Scale evaluation

Training phase:
Step 1. Process network transmission data, energy efficiency data and coverage capacity data by (4),(5) and (6) to obtain input vector $\mathbf{x}_{in}$.
Step 2. Calculating data labels of the input data by (13) to obtain $\mathbf{s}_c$.
Step 3. Initialize the parameters θ of the intelligent model with initial training settings.
Step 4. Iteratively train the intelligent model and by minimizing (15).
Step 5. Return trained parameters θ of the proposed intelligent model.
Inference phase:
Step 1. Repeat the operations in *step 1* for testing data.
Step 2. Input vectorial testing data into trained intelligent model.
Step 3. Return cellular-scale network performance evaluation results.

where $n_i \in \mathbf{N}_n = [n_1, \ldots, n_{k_1}]$ denotes the network-level transmission data, $e_i \in \mathbf{E}_n = [e_1, \ldots, e_{k_2}]$ denotes the energy efficiency data and $a_i \in \mathbf{A}_n = [a_1, \ldots, a_{k_3}]$ denotes the network-level coverage capability data of C_n. And $\alpha_n, \alpha_e, \alpha_c$ are corresponding weights. Please note that $\mathbf{N}_n$, $\mathbf{E}_n$ and $\mathbf{A}_n$ are standardized numerically and converted positively to guarantee $S_n^c \in (0, 100)$ according to (4) and (5).

- **Data labelling for fine-grained evaluation:** Similarly, as for the data labelling for fine-grained network performance evaluation, we still consider a coarse cellular network area C_m, which can be constructed via area aggregation among several fine-grained areas $D_1, \ldots, D_s$, and the label of it can be calculated based on WSA method and QoS data as

$$S_m^s = \frac{1}{m_l}\sum_{i=1}^{m_l}\alpha_q q_i,$$

$$\mathcal{M}_u = \{Mu_{m_1}, \ldots, Mu_{m_l} | Mu_{m_1}, \ldots, Mu_{m_l} \in \mathcal{D}(C_m)\}, \quad (14)$$

where S_m^s is the evaluation score, i.e., the data label of the network area C_m. And q_i denotes the QoS score of the mobile user Mu_{m_i} which locates in the coverage region $\mathcal{D}(C_m)$ of C_m.

In the current 5G network operation systems, WSA is commonly utilized to evaluate the network performance and quality for cellular-scale network coverage areas with high accuracy, especially with adequate input data. Therefore, WSA is selected to calculate the performance evaluation scores of cellular-scale network area so as to overcome the challenge of lacking data labels.

2) LOSS FUNCTION DESIGN BASED ON CALCULATED LABELS

As aforementioned, the goal of proposed network performance evaluation scheme is to generate accurate and precise evaluation results based on the input cellular-scale network data or user QoS data. And owing that in the previous section, we obtain data labels of cellular coverage areas with different types of network data, the specific loss function designs of training corresponding cellular-scale and fine-grained attention mechanism-based intelligent models are different and are elaborated as:

Algorithm 2 Paradigm of the Proposed Scheme for Fine-Grained evaluation

Training phase:
Step 1. Process QoS data by (4), (5) and (6) to obtain input vector $\mathbf{x}_{in}$.
Step 2. Calculating data labels of the input data by (14) to obtain $\mathbf{s}_c$.
Step 3. Initialize the parameters θ of the intelligent model with initial training settings.
Step 4. Iteratively train the intelligent model and by minimizing (16).
Step 5. Return trained parameters θ of the proposed intelligent model.
Inference phase:
Step 1. Repeat the operations in *step 1* for testing data.
Step 2. Input vectorial testing QoS data into trained intelligent model.
Step 3. Return fine-grained network performance evaluation results.

- **Cellular-scale model loss function design:** During the loss function design of cellular-scale evaluation, owing that the scales of the output results, i.e, the generated evaluation scores of the intelligent model based on input network-level data, are the same as original data labels, the distortion between generated output results $\mathbf{r}_o^c$ and the real cellular data labels $\mathbf{s}_c$ should be minimized and the euclidean distance is employed in the loss function which can be expressed as
$$L_c = ||\mathbf{s}_c - \mathbf{r}_o^c||_2^2, \quad (15)$$
where $\mathbf{r}_o^c = [S_1, \ldots, S_l]$ denotes the generated evaluation scores of l cellular-scale network areas. And $\mathbf{s}_c = [S_1^c, \ldots, S_l^c]$ denotes the data labels of the corresponding l cellular-scale network areas, which are calculated by (13)
- **Fine-grained model loss function design:** Distinct from the cellular-scale evaluation, the scales of the generated evaluation results by the intelligent model trained for fine-grained network performance evaluation are much smaller than the origin cellular-scale data label acquired by (14), which makes it difficult to directly utilize the generated evaluation results to design loss function. Therefore, we exploit the results generated by the fine-grained intelligent model to first estimate the real data labels $\mathbf{s}_s$ calculated based on conventional methods according to (14), and then minimize the distance between the estimated $\hat{\mathbf{s}_s}$ and real label set $\mathbf{s}_s$ in the euclidean space. The loss function of training the fine-grained intelligent model can be written as
$$L_s = ||\hat{\mathbf{s}_s} - \mathbf{s}_s||_2^2, \quad (16)$$
where $\hat{\mathbf{s}_s} = [\hat{S_1^s}, \ldots, \hat{S_N^s}]$ denotes the estimated data score set of N cellular-scale network areas constructed via area aggregation among M fine-grained network areas where $M > N$, while $\mathbf{s}_s = [S_1^s, \ldots, S_N^s]$ denotes the data label set calculated by (14). And as for utilizing the fine-grained generated results $\mathbf{r}_o^s = [S_1, \ldots, S_M]$ to estimate the real data labels, we consider a specific cellular-scale network area C_m which covers s fine-grained areas $D_1, \ldots, D_s$, the specific estimated label of it can be calculated as
$$\hat{S_m^s} = \frac{1}{s}\sum_{i=1}^{s} S_j, \quad (17)$$
where $\hat{S_m^s}$ is the corresponding estimated data label of original data label S_m^s acquired by (14). And S_j is the generated evaluation result of the intelligent model, presenting the evaluation score of fine-grained area D_j.

In particular, stochastic gradient descent (SGD) algorithm is utilized to update the proposed intelligent model by minimizing the loss functions presented in (15) and (16) for cellular-scale and fine-grained model training, and the paradigms of the proposed evaluation schemes applying two kinds of intelligent models with different evaluation granularity are illustrated in algorithm 1 and 2, respectively.

IV. EXPERIMENTAL RESULTS BASED ON PRACTICAL 5G NETWORK OPERATION DATA

In this section, experimental results are supplied to present the network performance evaluation of multiple 5G scenarios, i.e., generating the comprehensive evaluation scores of

TABLE 1. The platform of 5G vertical applications.

Category	Central cloud server	Edge server in gNodeB	Mobile user terminal
Deployment location	CNs	RAN	User side
Number	>1000	>40000	>500000
Sampled data type	Network-level data	Network-level data/QoS data	QoS data
Data format	XDR	XDR	XDR
Involved evaluation level	Cellular	Cellular and fine-grained	fine-grained
Specific equipment	Centralized computing center	Base station	Mobile phones

various network areas with different precision. First, the performance of model convergence is observed to measure the robustness of the proposed model. Then, to facilitate the distinct and intuitive observations of generated scores of fine granularity, we also transfer them into visible graphs. Finally, to verify the high performance of our proposed intelligent model, performance assessment based on correlation analyses of both the cellular-scale and fine-grained intelligent models are conducted.

In particular, in order to enhance the applicability and effectiveness of the proposed model in the real 5G networks, we sampled data from the practical enterprise 5G network platform. As illustrated in Table.1, the sampled data applied for network performance evaluation from cellular scale to fine-grained scale is mainly sampled from the central cloud servers, edge servers in gNodeBs and mobile user terminals. The number of the central cloud servers of sampling is more than 1000, while the number of edge servers is more than 40000. Moreover, the number of mobile user terminals is more than 500000. Besides, both the cellular-scale network data and user QoS data after sampling are preprocessed following Section III-B, and the detailed settings are provided in Table 3, where $max(\cdot)$ and $min(\cdot)$ denote functions of obtaining maximum and minimum values of input data respectively while $np2tensor(\cdot)$ denotes the data transformation function which turns numerical matrix to tensor.

A. DATASETS AND EXPERIMENTAL SETTINGS

As shown in Table.2, we employ sampled cellular network transmission data, energy efficiency data, coverage capacity data and user QoS data of 5G networks in three representative 5G empowering scenarios, including smart treatment, internet of industry and intelligent transportation. Please note that the involved network data is collected from practical private 5G network platform, and the specific data collection devices in practical 5G network architecture has been aforementioned in III-A. The detailed information of the datasets are elaborated as follows:

- **Internet of Industry:** As for the cellular-scale performance evaluation in the scenario of internet of industry, we employ totally 21000 sampled network-level data including transmission data, energy efficiency data and coverage capacity data as training data to train the proposed intelligent model, while additional 2000 sampled data of identical specification are utilized for validation. And the sampled data are collected every hour from 5G network operation devices continuously for two weeks. And for training the fine-grained intelligent model, 80000 QoS data sampled every 40 minutes for consecutive two weeks are selected as training data, while the rest 5000 QoS data are used for validation.
- **Intelligent Transportation:** In the scenario of intelligent transportation, there are 40000 sampled network-level data collected every hour for continuous two weeks used for training the proposed intelligent model to achieve cellular-scale evaluation, and 2000 additional network-level data are set for validation. And as for the fine-grained evaluation, 80000 QoS data sampled every 40 minutes for two weeks are used for training and other 5000 QoS data are set as validation data.
- **Smart Treatment:** To train the cellular-scale intelligent model of the smart treatment, 35000 network-level data sampled every hour for two weeks are input in the training procedure and additional 2000 testing network-level data of the same specification are employed. And when it turns to the training of the fine-grained intelligent model, 80000 QoS data sampled every 40 minutes for two weeks are exploited, while other 5000 QoS data are used for validation.

In particular, as shown in Table.2, during the iterative training process of our proposed intelligent model, the number of training epochs is set as 200, with the initial learning rate set as 0.01, which is reduced to 0.001 after 100 training rounds. Moreover, the batch size is set as 128.

B. MODEL CONVERGENCE PERFORMANCE ASSESSMENT

Since the scales of input cellular network data and user QoS data are extremely large, it is of great significant to guarantee the model convergence. In particular, as the loss functions presented in (15) and (16) are customized and self-designed as the distance measurement between the real network performance labels and generated evaluation results, the lower convergence values of loss functions denote higher performance evaluation accuracy of the proposed intelligent model. Therefore, the model convergence is first observed in this paper when training cellular-scale and fine-grained intelligent models. In particular, in order to facilitate observation of the loss function curves, the loss values in each round of training the proposed intelligent model in three simulated 5G scenarios are averaged equally.

TABLE 2. Datasets and simulation settings.

Category		Internet of Industry cellular	Internet of Industry fine-grained	Intelligent Transportation cellular	Intelligent Transportation fine-grained	Smart Treatment cellular	Smart Treatment fine-grained
Train number		21000	80000	40000	80000	35000	80000
Test number		2000	5000	2000	5000	2000	5000
Sampled frequency		1 hour	40 minutes	1 hour	40 minutes	1 hour	40 minutes
Lasting sampled time		2 weeks	2 weeks	2 weeks	2 weeks	2 weeks	2 weeks
Training	epochs	200					
	learning rate	0.01 at the begining and 0.001 after 100 rounds					
	optimizer	SGD					
	batchsize	128					

TABLE 3. Specific preprocessing of postive conversion, normalization and vectorization.

Category	Indicator $\mathbf{x}$	Pos conversion	Normalization	Vectorization
Cellular network-level data	Air latency	$max(\mathbf{x}) - x_i$	$\frac{100\times(x_i^p - min(\mathbf{x}^\mathbf{P}))}{max(\mathbf{x}^\mathbf{P})-min(\mathbf{x}^\mathbf{P})}$	$np2tensor(\mathbf{x}^n)$ by python
	Air port flow	-		
	Wireless resource utilization	-		
	Packet loss ratio	$max(\mathbf{x}) - x_i$		
	CQI good ratio	-		
	Successful access ratio	-		
	Cell handover ratio	-		
	NR cell available ratio	-		
	Power consumption	$max(\mathbf{x}) - x_i$		
	Coverage area	-		
QoS data	Bandwidth	-	$\frac{100\times(x_i^p - min(\mathbf{x}^\mathbf{P}))}{max(\mathbf{x}^\mathbf{P})-min(\mathbf{x}^\mathbf{P})}$	$np2tensor(\mathbf{x}^n)$ by python
	User latency	$max(\mathbf{x}) - x_i$		
	User latency jitter	$max(\mathbf{x}) - x_i$		
	User packet loss ratio	$max(\mathbf{x}) - x_i$		

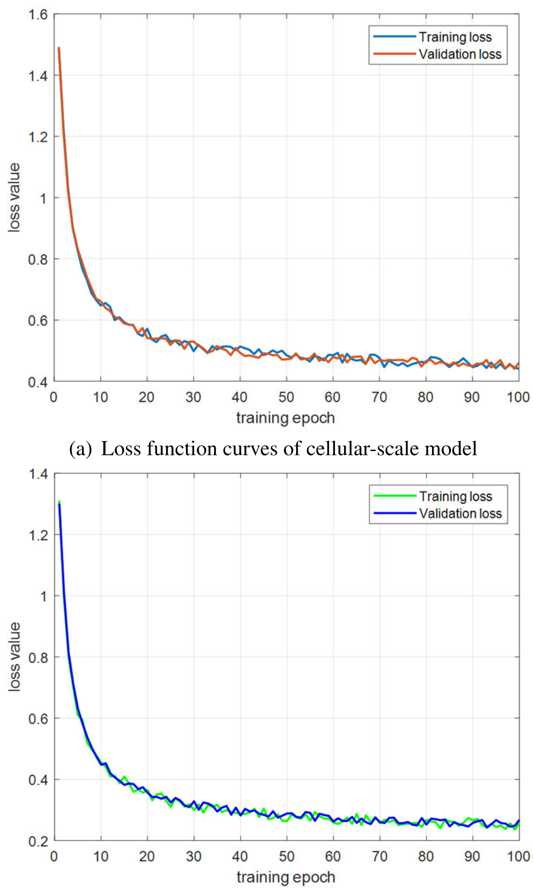
As shown in Figure.3, Figure.3 (a) denotes loss function curves of cellular-scale evaluation, where blue solid line represents training loss curve and orange solid line represents validation loss curve. And Figure.3 (b) denotes loss function curves of fine-grained intelligent model, where green solid line represents training loss curve and blue solid line represents validation loss curve. In particular, although the number training epochs is set as 200 in the previous section, the curves of two kinds of intelligent models converge to a certain range after approximately 60 rounds. Therefore. in this section, the loss curves within 100 rounds are illustrated in Figure.3.

And it is anticipated that the proposed model has a great convergence performance for both the cellular-scale and fine-grained evaluation, which denotes the training strategies are efficient to guarantee the overall model performance and high accuracy of generating the evaluation results. Moreover, great convergence performance of the proposed intelligent model also denotes the model's robustness and splendid learning capacity after training. And the validation loss curves fitting well with the training loss curves as illustrated in Figure.3 also denotes that the proposed intelligent model has a great generalization performance.

C. VISILABLE EVALUATION RESULTS OF OUR PROPOSED AI-ENABLED SCHEME WITH HIGH GRANULARITY

Since the fine-grained intelligent model can generate more geographically precise performance evaluation results, presenting more details of network performance and user experience that can not be achieved by conventional evaluation methods. In order to observe the geographical high precision of generating evaluation results by the intelligent model more distinctly, visible illustration of the output fine-grained evaluation scores matched to corresponding fine-grained network areas is provided. In particular, we transfer the generated network performance evaluation results of the selected three representative 5G scenarios into visible heat maps, which use color gradients to highlight patterns, concentrations, and variations within the data, making it easier to present generated results in an intuitive and visually appealing way. In particular, we present visible performance evaluation results of fine-grained network areas that are geographically close and correlated.

Figure 4, Figure 5 and Figure 6 illustrate the visible fine-grained evaluation results of the simulated 5G scenarios of internet of industry, intelligent transportation and smart treatment, respectively. In each depicted figure, we present visible evaluation results of three groups of fine-grained network areas, where the deeper color denotes the higher performance evaluation score. And it can be observed that the performance evaluation results of sampled fine-grained network areas in internet of industry and intelligent transportation are higher than those in smart treatment. And with the proposed fine-grained intelligent model generating geographically precise evaluation results, more details of network performance can be observed and further analysed, which can not be achieved by conventional evaluation methods.



(a) Loss function curves of cellular-scale model

(b) Loss function curves of fine-grained model

FIGURE 3. Loss function curves of training cellular-scale and fine-grained intelligent models.

Taking the presented fine-grained evaluation results of intelligent transportation as an example, which is shown in Figure 5, the network performance of the selected group $\mathcal{D}_4$ is relatively better than $\mathcal{D}_5$ and $\mathcal{D}_6$. Moreover, the fine-grained network areas in $\mathcal{D}_4$ with higher scores distribute more intensively than those in $\mathcal{D}_5$ and $\mathcal{D}_6$, which denotes the central-cloud and edge devices deployed in $\mathcal{D}_4$ have better coverage capacity than those in $\mathcal{D}_5$ and $\mathcal{D}_6$. Similar analysis is also applied to other 5G network scenarios, which denotes that the proposed fine-grained intelligent model is beneficial to the precise network performance analysis and further network optimization.

D. CORRELATION ANALYSIS RESULTS OF OUR PROPOSED SCHEME

Owing that the proposed intelligent model generates network evaluation results based on cellular network data or user QoS data, the correlation between the output scores and input indicator data should be strong so as to guarantee the high accuracy of the proposed intelligent model, i.e., for cellular-scale evaluation, the generated results are ought to be strongly correlated with network-level data including network transmission data, energy efficiency data and coverage capacity data, while for fine-grained evaluation, the correlation between output results and input QoS data should be high. Therefore, we carry out correlation analysis for both the cellular-scale and fine-grained intelligent models, which are elaborated in the following sections.

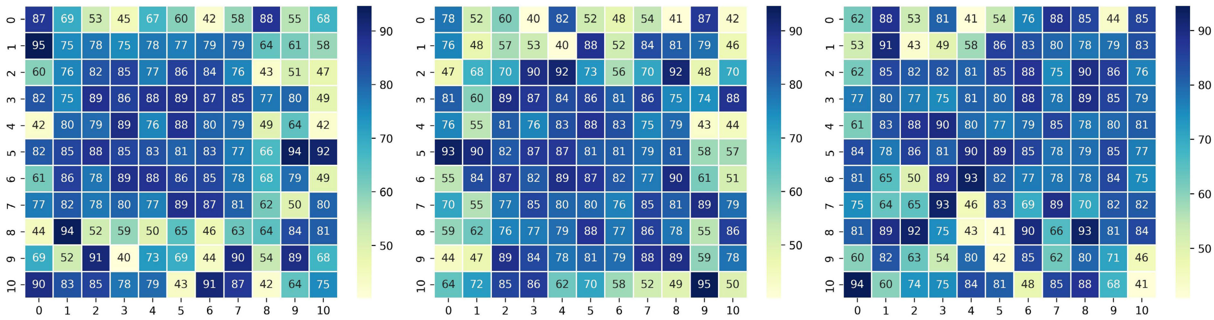
TABLE 4. Correlation analysis results of our proposed scheme (cellular level).

5G scenarios		Pearson	Spearman	Kendall
Internet of industry	$\mathcal{N}_1 - \mathcal{C}_1$	0.801	0.826	0.815
	$\mathcal{N}_2 - \mathcal{C}_2$	0.764	0.758	0.779
	$\mathcal{N}_3 - \mathcal{C}_3$	0.833	0.842	0.859
	$\mathcal{N}_4 - \mathcal{C}_4$	0.813	0.798	0.812
Intelligent transportation	$\mathcal{N}_5 - \mathcal{C}_5$	0.778	0.742	0.790
	$\mathcal{N}_6 - \mathcal{C}_6$	0.825	0.827	0.836
	$\mathcal{N}_7 - \mathcal{C}_7$	0.878	0.844	0.868
	$\mathcal{N}_8 - \mathcal{C}_8$	0.860	0.855	0.832
Smart treatment	$\mathcal{N}_9 - \mathcal{C}_9$	0.808	0.813	0.829
	$\mathcal{N}_{10} - \mathcal{C}_{10}$	0.871	0.855	0.869
	$\mathcal{N}_{11} - \mathcal{C}_{11}$	0.880	0.864	0.847
	$\mathcal{N}_{12} - \mathcal{C}_{12}$	0.812	0.806	0.825

1) CORRELATION ANALYSIS FOR CELLULAR-SCALE EVALUATION

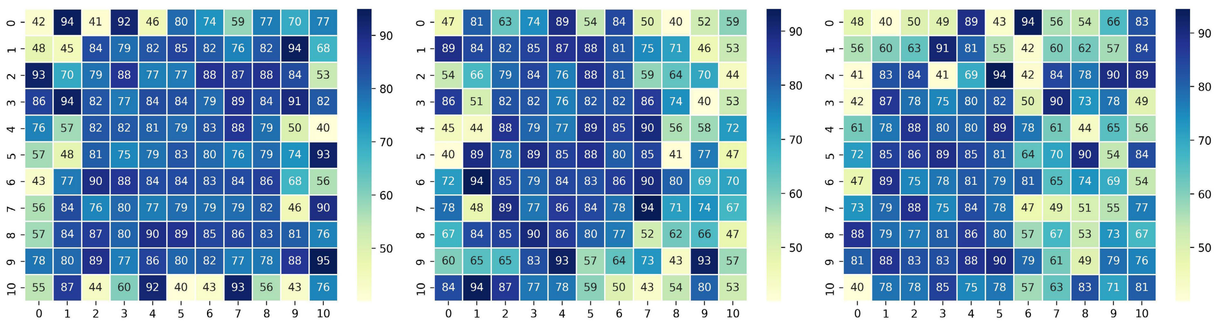
Since the cellular-scale network performance evaluation results are generated by the intelligent model based on network-level indicator data, the output scores should be strongly correlated to them. In this section, three kinds of widely applied correlation coefficients between the input network-level data and output scores are calculated, including Pearson, Spearman and Kendall correlation coefficients. In particular, we select four groups of cellular network areas in each 5G scenario, and calculate the above three kinds of correlation coefficients between the cellular output results in each group and the corresponding network-level data. As shown in Table.4, $\mathcal{N}_i = \{\mathbf{N}_i, \mathbf{E}_i, \mathbf{A}_i\}$ $(i = 1, \ldots, 12)$ denotes the network-level data including network transmission data $\mathbf{N}_i$, energy efficiency data, $\mathbf{E}_i$ and coverage capacity data $\mathbf{A}_i$ corresponding to cellular network group $\mathcal{C}_i$ (i $= 1, \ldots, 12$), which consists of multiple cellular 5G network areas.

It is plain to see that in Table.4, the correlation coefficients between sampled network-level data and corresponding output evaluation results are relatively high. In the internet of industry, the maximum correlation coefficient is 0.859 of Kendall correlation coefficient between $\mathcal{N}_3$ and $\mathcal{C}_3$, while the minimum correlation coefficient is 0.758 of Spearman correlation coefficient between $\mathcal{N}_2$ and $\mathcal{C}_2$, which denotes the correlation in this scenario is strong while the correlation between $\mathcal{N}_3$ and $\mathcal{C}_3$ is relatively higher than that between $\mathcal{N}_2$ and $\mathcal{C}_2$ and the generating accuracy of the intelligent model applied in this scenario is also proved to be high. Similar analysis are likewise suitable for the other two scenarios, where the maximum value of correlation is 0.880 of smart treatment while the minimum value is 0.742 of intelligent transportation and it can be concluded that the



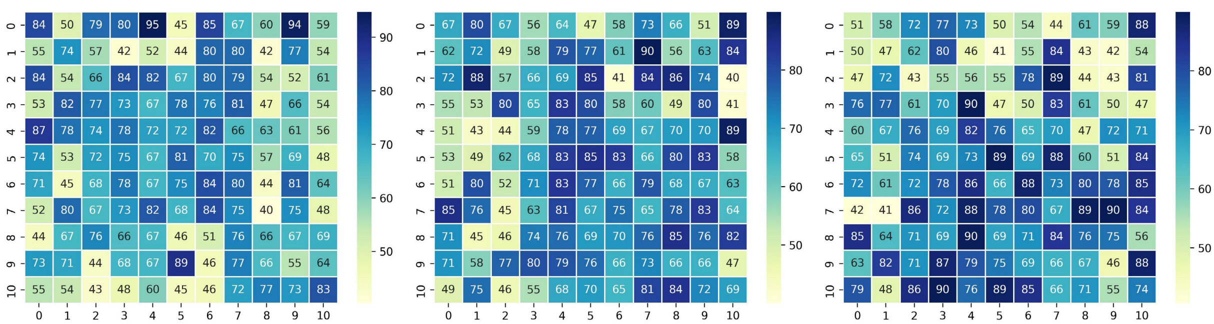
(a) fine-grained network areas $\mathcal{D}_1$ (b) fine-grained network areas $\mathcal{D}_2$ (c) fine-grained network areas $\mathcal{D}_3$

FIGURE 4. **Visible fine-grained evaluation results of internet of industry.**



(a) fine-grained network areas $\mathcal{D}_4$ (b) fine-grained network areas $\mathcal{D}_5$ (c) fine-grained network areas $\mathcal{D}_6$

FIGURE 5. **Visible fine-grained evaluation results of intelligent transportation.**



(a) fine-grained network areas $\mathcal{D}_7$ (b) fine-grained network areas $\mathcal{D}_8$ (c) fine-grained network areas $\mathcal{D}_9$

FIGURE 6. **Visible fine-grained evaluation results of smart treatment.**

overall correlation coefficients in the selected 5G scenarios are high, which indicates that the proposed cellular-scale intelligent model has a high accuracy and great model performance.

2) CORRELATION ANALYSIS FOR FINE-GRAINED EVALUATION

During the procedure of correlation analysis for network performance evaluation with fine granularity, three kinds of commonly employed correlation coefficients between generated results of fine-grained network areas and corresponding QoS data sets are calculated, which are Pearson, Spearman and Kendall correlation coefficients. In particular, we select four non-overlapping groups of fine-grained5G networks to generate the performance evaluation results in each 5G scenario, as depicted in Table.5, where $\mathcal{D}_{sub_i}$ $(i = 1, \ldots, 12)$ denotes a set of fine-grained network performance evaluation results corresponding to a group of fine-grained areas which satisfies $\mathcal{D}_{sub_i} \bigcap \mathcal{D}_{sub_j} = \emptyset, i \neq j$. And $\mathbf{Q_i}$ denotes the QoS data set corresponding to $\mathcal{D}_{sub_i}$.

As illustrated in Table.5, the overall correlation coefficients in the selected three 5G scenarios are high, especially in the scenario of smart treatment. Taking the fine-grained areas $\mathcal{D}_{sub_{10}}$ as an example, the Pearson, Spearman and Kendall

TABLE 5. Correlation analysis results of our proposed scheme (sub-cellular level with fine granularity).

5G scenarios		Pearson	Spearman	Kendall
Internet of industry	$\mathbf{Q_1} - \mathcal{D}_{sub_1}$	0.794	0.785	0.805
	$\mathbf{Q_2} - \mathcal{D}_{sub_2}$	0.816	0.824	0.819
	$\mathbf{Q_3} - \mathcal{D}_{sub_3}$	0.786	0.798	0.806
	$\mathbf{Q_4} - \mathcal{D}_{sub_4}$	0.817	0.824	0.821
Intelligent transportation	$\mathbf{Q_5} - \mathcal{D}_{sub_5}$	0.835	0.851	0.849
	$\mathbf{Q_6} - \mathcal{D}_{sub_6}$	0.812	0.807	0.827
	$\mathbf{Q_7} - \mathcal{D}_{sub_7}$	0.852	0.847	0.843
	$\mathbf{Q_8} - \mathcal{D}_{sub_8}$	0.802	0.815	0.806
Smart treatment	$\mathbf{Q_9} - \mathcal{D}_{sub_9}$	0.851	0.843	0.844
	$\mathbf{Q_{10}} - \mathcal{D}_{sub_{10}}$	0.873	0.866	0.860
	$\mathbf{Q_{11}} - \mathcal{D}_{sub_{11}}$	0.885	0.879	0.862
	$\mathbf{Q_{12}} - \mathcal{D}_{sub_{12}}$	0.848	0.836	0.865

correlation coefficients between $\mathbf{Q_{10}}$ and $\mathcal{D}_{sub_{10}}$ are 0.873, 0.866 and 0.860 respectively, which denotes the performance evaluation results of $\mathcal{D}_{sub_{10}}$ are strongly correlated with the corresponding QoS data set $\mathbf{Q_{10}}$. The same analysis is also applied to the other two selected 5G scenarios, where the maximum correlation coefficient is 0.852 of Pearson correlation coefficient between $\mathbf{Q_7}$ and $\mathcal{D}_{sub_7}$ in intelligent transportation and the minimum correlation coefficient is 0.785 of Spearman correlation coefficient between $\mathbf{Q_1}$ and $\mathcal{D}_{sub_1}$ in internet of industry, which denotes the trained intelligent model has great performance and generating accuracy when applied to the fine-grained network performance evaluation.

V. CONCLUSION

In this paper, in order to evaluate the performance of networks in various 5G empowering scenarios with fine granularity and accuracy, the network performance evaluation scheme of 5G network is studied. Inspired by the research field of computer vision, we propose an intelligent network performance evaluation scheme, which applies attention mechanism to give different weights to different channel features and thus better filtrating task-concentrated features, thereby enhancing the robustness and learning capacity of the proposed AI model. Moreover, corresponding training strategies are provided. Finally, experimental results based on data collected from practical 5G network operation system are provided, which illustrates that our proposed model can generate evaluation results with high accuracy and fine granularity of various 5G network scenarios.

REFERENCES

[1] Z. Zhao, C. Feng, W. Hong, J. Jiang, C. Jia, T. Q. S. Quek, and M. Peng, ''Federated learning with non-IID data in wireless networks,'' *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1927–1942, Mar. 2022.

[2] O. O. Erunkulu, A. M. Zungeru, C. K. Lebekwe, M. Mosalaosi, and J. M. Chuma, ''5G mobile communication applications: A survey and comparison of use cases,'' *IEEE Access*, vol. 9, pp. 97251–97295, 2021.

[3] X. Zhu and Y. Chi, ''5G release 16 key technologies for vertical industries application,'' in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2021, pp. 1269–1274.

[4] A. Asadi, Q. Wang, and V. Mancuso, ''A survey on device-to-device communication in cellular networks,'' *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, 4th Quart., 2014.

[5] N. Kumar, S. Misra, J. J. P. C. Rodrigues, and M. S. Obaidat, ''Coalition games for spatio-temporal big data in Internet of Vehicles environment: A comparative analysis,'' *IEEE Internet Things J.*, vol. 2, no. 4, pp. 310–320, Aug. 2015.

[6] J. J. Rutherford, ''Wearable technology,'' *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 3, pp. 19–24, May/Jun. 2010, doi: 10.1109/MEMB.2010.936550.

[7] P. Binkley, ''Predicting the potential of wearable technology,'' *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 3, pp. 23–27, May/Jun. 2003.

[8] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, ''A survey on mobile edge computing: The communication perspective,'' *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart., 2017.

[9] Y. Peng, Q. Bingyi, Y. Shiji, and S. Huihui, ''Research on performance evaluation method of communication radio network based on qos performance,'' in *Proc. 2nd IEEE Int. Conf. Comput. Intell. Appl. (ICCIA)*, Sep. 2017, pp. 454–457.

[10] X. Shang, L. Chen, and Y. Nan, ''A network performance evaluation method based on objective weights determination for data center network,'' in *Proc. 13th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD)*, Jul. 2017, pp. 2843–2849.

[11] N. Hossein Motlagh, S. Kapoor, R. Alhalaseh, S. Tarkoma, and K. Hätönen, ''Quality of monitoring for cellular networks,'' *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 1, pp. 381–391, Mar. 2022.

[12] R.-G. Lazar, A.-V. Militaru, C.-F. Caruntu, C. Pascal, and C. Patachia-Sultanoiu, ''Real-time data measurement methodology to evaluate the 5G network performance indicators,'' *IEEE Access*, vol. 11, pp. 43909–43924, 2023.

[13] Q. Xu, S. Li, T. V. Do, K. Jia, and N. Yang, ''Performance analysis of cognitive radio networks with burst dynamics,'' *IEEE Access*, vol. 9, pp. 110627–110638, 2021.

[14] S. Zroug, I. Remadna, L. Kahloul, S. Benharzallah, and S. L. Terrissa, ''Leveraging the power of machine learning for performance evaluation prediction in wireless sensor networks,'' in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Jul. 2021, pp. 864–869.

[15] C. Wang, N. Yoshikane, and T. Tsuritani, ''Usage of a graph neural network for large-scale network performance evaluation,'' in *Proc. Int. Conf. Opt. Netw. Design Model. (ONDM)*, Jun. 2021, pp. 1–5.

[16] M. Li, Y. Wang, Z. Wang, and H. Zheng, ''A deep learning method based on an attention mechanism for wireless network traffic prediction,'' *Ad Hoc Netw.*, vol. 107, Oct. 2020, Art. no. 102258. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870519310923

[17] V. S. Dommaraju, K. Nathani, U. Tariq, F. Al-Turjman, S. Kallam, P. K. Reddy, and R. Patan, ''ECMCRR-MPDNL for cellular network traffic prediction with big data,'' *IEEE Access*, vol. 8, pp. 113419–113428, 2020.

[18] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, ''Image segmentation using deep learning: A survey,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022.

[19] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, ''A survey of convolutional neural networks: Analysis, applications, and prospects,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022.

[20] U. Erdenebayar, H. Kim, J. U. Park, D. Kang, and K. J. Lee, ''Automatic prediction of atrial fibrillation based on convolutional neural network using a short-term normal electrocardiogram signal,'' *J. Korean Med. Sci.*, vol. 34, no. 7, 2019, Art. no. 1116574.

[21] Q. Zhang, D. Zhou, and X. Zeng, ''HeartID: A multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications,'' *IEEE Access*, vol. 5, pp. 11805–11816, 2017.

[22] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, ''Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks,'' *J. Sound Vib.*, vol. 388, pp. 154–170, Feb. 2017.

[23] K. He, X. Zhang, S. Ren, and J. Sun, ''Spatial pyramid pooling in deep convolutional networks for visual recognition,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.

[24] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, ''Dual path networks,'' in *Advances in Neural Information Processing Systems*, vol. 30, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2017.

[25] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, ''3D ShapeNets: A deep representation for volumetric shapes,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1912–1920.

[26] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, "Intelligent 5G: When cellular networks meet artificial intelligence," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 175–183, Oct. 2017.

[27] Y. Lee, "Classification of node degree based on deep learning and routing method applied for virtual route assignment," *Ad Hoc Netw.*, vol. 58, pp. 70–85, Apr. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870516303092

[28] V. Radu, N. D. Lane, S. Bhattacharya, C. Mascolo, M. K. Marina, and F. Kawsar, "Towards multimodal deep learning for activity recognition on mobile devices," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput., Adjunct (UbiComp)*, New York, NY, USA, 2016, pp. 185–188, doi: 10.1145/2968219.2971461.

[29] B. Mao, Z. Md. Fadlullah, F. Tang, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, "A tensor based deep learning technique for intelligent packet routing," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[30] Q. Zhang, M. Lin, L. T. Yang, Z. Chen, and P. Li, "Energy-efficient scheduling for real-time systems based on deep Q-learning model," *IEEE Trans. Sustain. Comput.*, vol. 4, no. 1, pp. 132–141, Jan. 2019.

[31] R. Atallah, C. Assi, and M. Khabbaz, "Deep reinforcement learning-based scheduling for roadside communication networks," in *Proc. 15th Int. Symp. Model. Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, May 2017, pp. 1–8.

[32] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[33] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," in *Proc. IEEE Int. Conf. Intell. Secur. Informat. (ISI)*, Jul. 2017, pp. 43–48.

[34] C. Zhang and P. Patras, "Long-term mobile traffic forecasting using deep spatio-temporal neural networks," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, New York, NY, USA, Jun. 2018, pp. 231–240, doi: 10.1145/3209582.3209606.

[35] T. Quisel, L. Foschini, A. Signorini, and D. C. Kale, "Collecting and analyzing millions of mHealth data streams," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*. New York, NY, USA, Aug. 2017, pp. 1971–1980, doi: 10.1145/3097983.3098201.

[36] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities and challenges," *Briefings Bioinf.*, vol. 19, no. 6, pp. 1236–1246, May 2017, doi: 10.1093/bib/bbx044.

[37] S. Garg, K. Kaur, N. Kumar, G. Kaddoum, A. Y. Zomaya, and R. Ranjan, "A hybrid deep learning-based model for anomaly detection in cloud datacenter networks," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 3, pp. 924–935, Sep. 2019.

[38] A. Iqbal, M.-L. Tham, and Y. C. Chang, "Convolutional neural network-based deep Q-network (CNN-DQN) resource management in cloud radio access network," *China Commun.*, vol. 19, no. 10, pp. 129–142, Oct. 2022.

[39] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092523122100477X

[40] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. 32nd Int. Conf. Mach. Learn.* (Proceedings of Machine Learning Research), vol. 37, F. Bach and D. Blei, Eds. Lille, France: PMLR, Jul. 2015, pp. 2048–2057.

[41] J. Lu, C. Xiong, D. Parikh, and R. Socher, "Knowing when to look: Adaptive attention via a visual sentinel for image captioning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3242–3250.

[42] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems*, vol. 27, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2014.

[43] D. Britz, A. Goldie, M.-T. Luong, and Q. Le, "Massive exploration of neural machine translation architectures," 2017, *arXiv:1703.03906*.

[44] J. Chorowski, D. Bahdanau, K. Cho, and Y. Bengio, "End-to-end continuous speech recognition using attention-based recurrent NN: First results," 2014, *arXiv:1412.1602*.

[45] W. Chan, N. Jaitly, Q. Le, and O. Vinyals, "Listen, attend and spell: A neural network for large vocabulary conversational speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 4960–4964.

[46] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," 2017, *arXiv:1709.01507*.

[47] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," 2017, *arXiv:1704.06904*.

QING ZHANG received the master's degree in information and communication systems from Beijing University of Posts and Telecommunications, in 2012. She is currently a Senior Engineer with China Unicom Intelligent Network Innovation Center, with a focus on developing innovative 5G network products, researching new mobile communication technologies, and technological development trends in related fields.

TAOYE ZHANG received the master's degree in business Administration from Zhejiang University, in 2002. He is currently a Senior Engineer with China Unicom Intelligent Network Innovation Center, with a focus on researching development management of network innovation products, researching network innovation technologies, and technological development trends in related fields.

BIN CHEN received the master's degree in communication and electronic systems from Beijing University of Posts and Telecommunications, in 1997. He is currently a Senior Engineer with China Unicom Intelligent Network Innovation Center, with a focus on researching development management of network innovation products, researching network innovation technologies, and technological development trends in related fields.

JI YAN received the B.S. degree in network engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2022. He is currently pursuing the Ph.D. degree in information and communication engineering with Beijing University of Posts and Telecommunications (BUPT). His research interests include artificial intelligence and 6G communications.

ZHONGYUAN ZHAO (Member, IEEE) received the B.S. degree in applied mathematics and the Ph.D. degree in communication and information systems from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2009 and 2014, respectively.

He is currently a Full Professor with BUPT. His research interests include fog computing/edge computing, content caching, and edge intelligence in wireless networks. He was a recipient of an Exemplary Editor Award twice, in 2017 and 2018, and the Best Paper Award from the IEEE CIT 2014 and WASA 2015. He served as an Editor for IEEE Communications Letters, from 2016 to 2020, and a Guest Editor for IEEE Access. He also serves as an Editor of IEEE Open Journal of the Communications Society.

XIAOFEI QIN received the bachelor's degree in electronic information technology and instruments from Zhejiang University, in 2010. He is currently with China Unicom Intelligent Network Innovation Center, with a focus on designing and managing innovative 5G private network products, planning private network product systems, and technological development trends in related fields.

CHAO CAI received the master's degree in computer science from Xi'an University of Electronic Science and Technology, in 2008. He is currently the Director of the 5G Innovation Center of China Unicom Intelligent Network Innovation Center. He has 16 years of experience in network planning, construction, innovative product research and development, and industry application promotion. He is mainly responsible for the research and development of China Unicom's innovative 5G network products, including MEC platform, 5G private network platform, and slicing acceleration/speed limit products. He has hosted and participated in more than ten national projects, published 20 papers, and won ten provincial and ministerial-level awards.

XIANKUI LUO received the B.S. degree in network engineering from Nanjing Agricultural University, Nanjing, China, in 2011. He is currently with China Unicom Intelligent Network Innovation Center. His current research interests include 5G-private networks, network architecture, and software system architecture.