

QoE Prediction for Gaming Video Streaming in O-RAN Using Convolutional Neural Networks

GEORGIOS KOUGIOUMTZIDIS¹, ATANAS VLAHOV² (Graduate Student Member, IEEE),
VLADIMIR K. POULKOV¹ (Senior Member, IEEE), PAVLOS I. LAZARIDIS³ (Senior Member, IEEE),
AND ZAHARIAS D. ZAHARIS⁴ (Senior Member, IEEE)

¹Department of Communication Networks, Technical University of Sofia, 1000 Sofia, Bulgaria

²Intelligent Communication Infrastructure Laboratory, Sofia Tech Park, 1784 Sofia, Bulgaria

³School of Computing and Engineering, University of Huddersfield, HD1 3DH Huddersfield, U.K.

⁴School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

CORRESPONDING AUTHOR: G. KOUGIOUMTZIDIS (e-mail: gkougoumtzidis@tu-sofia.bg)

This work was supported in part by the European Union through the Horizon 2020 Marie Skłodowska-Curie Innovative Training Networks Programme "Mobility and Training for Beyond 5G Ecosystems (MOTOR5G)" under Grant 861219, and in part by the "Intelligent Communication Infrastructure Laboratory" at Sofia Tech Park, Sofia, Bulgaria.

ABSTRACT The growing popularity of online and cloud gaming applications is reshaping the landscape of the entertainment industry and acting as a key driver of market growth. However, the dependency of these applications on network resources poses significant challenges to the communication infrastructure. This is particularly critical as network performance plays a key role in influencing user satisfaction during gameplay. Inevitably, these inherently interactive applications are also closely linked to the concept of quality of experience (QoE), which expresses the perceived quality of a service by end-users. In this paper, we leverage deep learning methodologies to develop an objective QoE prediction model. Specifically, the proposed prediction model investigates the effect of wireless network operation on the QoE of gaming video streaming. Employing a tailored multi-headed convolutional neural network (multi-headed CNN), the model can predict in real-time the transmission-related QoE value using measurable quality of service (QoS) parameters. To validate the effectiveness of the model, tests and evaluations were conducted in an open radio access network testbed environment equipped with O-RAN-compatible interfaces.

INDEX TERMS Convolutional neural network, deep learning, deep neural network, gaming video, open radio access network (Open RAN), quality of experience (QoE), QoE prediction.

I. INTRODUCTION

THE VIDEO game industry has grown at an astonishing rate in recent years, and gaming has now transformed from a niche hobby to one of the largest markets in the entertainment industry [1]. Gaming reached an all-time high in terms of revenue and user engagement during the COVID-19 global pandemic. In 2022, the global video game industry's revenues surpassed the USD 200 billion mark for the first time, while the number of players worldwide was estimated at 3 billion. Currently, the video game industry exceeds the combined revenues of the music and film industries and is only trailed by television [2]. Therefore, online and cloud gaming can be seen as an increasingly

important research area, attracting the interest of both academia and industry. Consequently, the development of wireless networks capable of supporting the demanding service requirements of gaming video streaming is a necessary condition for the further growth of gaming applications.

Thus, it has become vital for Internet service providers (ISPs) to provide the best possible gaming experience to end-users, making efficient use of the available physical network infrastructure. However, understanding the mechanisms that determine quality as experienced by end-users is a highly complex process. This is due to the fact that these mechanisms include technical parameters such as terminal equipment, network state and video compression,

the gaming environment, as well as the subjective perception of the players [3]. To this end, quality of experience (QoE) is the metric that indicates the perceived quality of a service by the end-user, taking into account all the factors involved. In the context of wireless communication networks, QoE has become a key tool for evaluating the quality of their operation. This is because it allows a more in-depth knowledge of how the quality of service (QoS) factors of a network affect the perceived quality of communication services by end-users.

Legacy wireless networks, however, are unable to meet the network traffic and radio resource requirements of today's gaming video streaming applications as they exhibit limited bandwidth, high latency, and inefficient network resource management. These challenges require new wireless network designs that offer customized communication solutions based on programmable, flexible, and scalable architectures. A significant opportunity for transforming the wireless and mobile networks is offered by open radio access network (Open RAN) technology, as suggested by the O-RAN Alliance [4]. Based on two key concepts, openness and intelligence, Open RAN can be seen as a disruptive technology in the wireless communications ecosystem, capable of enabling new markets and encouraging innovation.

In addition, the development of QoE prediction models can make a decisive contribution to optimizing the operation of wireless networks, complementing the innovative design of the Open RAN network architecture. The benefits of accurate QoE prediction in a wireless communication system with respect to system performance arise from the ability to predictively analyze the patterns and dynamic characteristics of QoE. This analysis can serve as the basis for the implementation of a network optimization policy. Consequently, accurate QoE prediction facilitates the effective planning for the allocation of the finite radio and computing resources, enhances network energy efficiency, mitigates network congestion, and leads to reduced capital expenditure (CapEx) and operational expenditure (OpEx), ultimately enhancing profitability. Real-time QoE prediction should be regarded as an important element in the design of next-generation wireless communication networks, particularly when transmitting services with intrinsic interactivity, such as gaming applications.

In this paper, we present a deep learning-based objective QoE prediction model for gaming video streaming applications. The prediction model is based on a customized multi-headed convolutional neural network (multi-headed CNN) applied to an O-RAN-compliant wireless communication network. The goal of the prediction model is to measure the influence of wireless network operation on gaming video quality as perceived by end-users in real-time. Since the prediction model can forecast the network/transmission-related QoE value using purely quantifiable QoS factors, the QoE evaluation can be categorized as objective. The QoS parameters of bandwidth, latency, packet loss, and jitter are captured and recorded by a QoS monitoring system

designed with open-source monitoring tools. The prediction model first converts the collected QoS parameters into mean opinion score (MOS) values, and then the multi-headed CNN network analyzes the interdependencies among them to derive the overall QoE value for the wireless gaming video streaming service. To the authors' knowledge, this is the first study to measure and predict the QoE of gaming video streaming over Open RAN. Furthermore, it gives for the first time a benchmarking of deep neural network (DNN) algorithms for this kind of application.

The paper is organized as follows: Section II provides a review of state-of-the-art models for gaming video QoE prediction. Section III analyzes the QoE assessment methodologies in connection to gaming video applications. Moreover, it explores the gaming video QoE influencing factors (IFs) and evaluates the network/transmission-related IFs. In addition, it examines the QoS/QoE mapping approaches. Section IV introduces the suggested QoE prediction model. First, the Open RAN testbed is described, then the QoS monitoring system and dataset creation process, the QoS/QoE mapping model, and finally the deep learning-based QoE prediction model. Section V presents a comparative study of deep learning methodologies and provides the performance assessment findings for the suggested prediction model. Section VI concludes with closing observations.

II. RELATED WORK

Early work on QoE prediction for wireless video streaming focuses on conventional video formats without considering gaming video applications, which due to their highly interactive nature exhibit particular features and network requirements.

In [5], a DNN method for QoE prediction in mobile video transmission is presented. This approach uses a mobile phone application to collect data related to user QoE. The process involves the collection of a significant amount of data that includes both subjective ratings and network parameters. A DNN is then constructed to discriminate the links between these network parameters and subjective QoE ratings.

In [6], a real-world dataset derived from a mobile operator is employed to create a correlation between network-side parameters and user QoE in video streaming applications. This involves the utilization of a deep learning model for the initial forecast of the channel path loss. Subsequently, this prediction is applied to forecast the MOS for mobile video streaming. It is important to note that the trained model cannot be directly applied to a different geographical area, which does not allow the generalization of this approach.

In [7], a model integrating mobile edge computing (MEC) and software-defined networking (SDN) is proposed to allocate resources and reduce latency for 3D high-definition video. The model uses an actor-critic based deep reinforcement learning algorithm for viewport prediction and QoE optimization and a long-short memory network (LSTM) for bandwidth and viewport prediction. The model

can adaptively assign the best transmission throughput based on observations to maximize QoE.

In [8], an end-to-end framework for video QoE prediction is presented. Initially, a mixture of deep learning methodologies, including word embedding and a 3D convolutional neural network (C3D), are used to extract generalized features. These features are then fused and fed into a neural network to learn representations. The acquired representation is then used as input for tasks related to classification and regression.

In [9], an explainable artificial intelligence (XAI) model is used to overcome the need to ensure high levels of explainability in AI models. The study contrasts fuzzy decision tree models with classical decision tree models and random forest (RF) classifiers on a QoE classification dataset. The findings of the comparison demonstrate that fuzzy decision trees are easier to interpret and exhibit comparable performance, particularly in detecting stalling events in video streaming applications.

In [10], a Bayesian network (BN) model aiming to predict the re-buffering ratio by concentrating on a selected set of QoE parameters is introduced. Specifically, a neural network-based approach is suggested to validate the BN's accurate representation of stalling data patterns. The study concludes by demonstrating that models incorporating hidden variables and contextual information exhibit enhanced performance on QoE-related metrics.

In [11], a tutorial focused on implementing QoE measurement and prediction in video streaming applications using supervised machine learning (ML) algorithms is presented. The tutorial is structured in three parts. Initially, an approach for applying video streaming QoE prediction models based on supervised learning is described. Next, the development of ML-based models for QoE prediction and measurement in 5G/6G networks is examined. Finally, a benchmark analysis is provided, which evaluates the performance of state-of-the-art supervised learning ML models.

In recent years, there has been increasing interest in developing QoE prediction models adapted to the particular characteristics of wireless gaming video streaming applications.

In [12], a no-reference frame-based measurement system for gaming video quality performance is created using the support vector regression (SVR) algorithm. The SVR training includes nine frame-level indices as input features and video multimethod assessment fusion (VMAF) scores as reference data. This prediction model is characterized by low complexity, as it is based on features that are available in real-time.

In [13], two no-reference lightweight ML methods for predicting QoE in gaming video streaming are introduced. These models are constructed using the algorithms of SVR, Gaussian process regression (GPR), random forest (RF), and artificial neural network (ANN). Due to their simplicity, both models can serve as the initial phase of a real-time optimized online gaming QoE management framework.

In [14], a method is proposed to improve the quality of compressed gaming content by using generative super-resolution contradictory networks (SRGAN). This involves using a DNN in tandem with an adversarial network to produce higher-resolution images. The proposed approach incorporates adaptations to the generative network, including modifications to the skip connections and loss function. These modifications improve the information flow within the network, leading to increased perceived quality.

In [15], a no-reference lightweight module for assessing video quality in gaming content is created, utilizing the random forest regression (RFR) algorithm. This method forecasts video quality scores solely based on the recorded video, prioritizing features that are straightforward to calculate. Additionally, it incorporates minimal features in order to establish a model capable of making real-time QoE predictions.

In [16], a technique is put forward to develop a CNN-derived metric for assessing the quality of gaming videos. The CNN undergoes training with the VMAF objective quality model as a reference, and further refinement is achieved through subjective image quality evaluations. Additionally, a novel temporal pooling approach based on frame-level predictions is introduced for the prediction of gaming video QoE.

In [17], a methodology for real-time reduced reference evaluation of gaming video quality is introduced. This approach is built upon a psychometric curve-fitting method characterized by low complexity. The model employs ML techniques, including decision tree regression (DTR) and ANN. The proposed solution selects the most pertinent objective features while minimizing complexity. Subsequently, it models the relationship between these features and the reference quality by incorporating human visual system (HVS) psychometric perception.

In [18], a QoE estimation model is introduced, incorporating both gaming and non-gaming videos and relying on the algorithms of CNN and RF. The CNN is trained using an objective metric, enabling it to grasp video artifacts, and is fine-tuned using blockiness and blurriness scores derived from a small image quality dataset. The temporal and frame-level information of videos is employed by an RF model to estimate video QoE. The model's low complexity renders it well-suited for real-time applications.

In [19], a streamlined quality prediction model for gaming video streaming is suggested, utilizing CNN. This prediction model incorporates a hard pairwise ranking loss, enabling it to prioritize the discrimination of similar pairs. Additionally, an efficient adapted distillation model is integrated, resulting in minimal performance loss.

The first set of state-of-the-art QoE prediction models mentioned above focuses on the wireless streaming of conventional video formats and cannot be generalized for use in gaming video streaming applications, as these models do not take into account the specific aspects and characteristics of this type of video. The second set of QoE

prediction models mainly takes into account frame-level IFs of the gaming video format, such as blurriness, naturalness, blockiness and complexity, as well as the influence of spatiotemporal features and psychometric parameters. Our approach differs from these methods, concentrating on the influence of the transmission channel on gaming video QoE. Across various applications, the majority of existing literature on QoE assessment and prediction neglects the thorough examination of network/transmission-related IFs. In instances where the impact of the wireless transmission channel is considered, it is typically done using network simulators and emulators. Our study stands out as the first to utilize a real 4G/5G wireless network, thoroughly investigating the impact of network/transmission-related IFs on gaming video streaming QoE under real-world conditions. Moreover, our work is centered on developing an entirely objective QoE prediction model through the application of QoS/QoE mapping functions. Specifically, we consider real-world network design scenarios based on Open RAN technology, where the use of subjective QoE evaluation and prediction models is impractical due to their limitations, as discussed in Section III-A. Finally, this is the first paper that provides a comprehensive study on deep learning techniques for QoE prediction in wireless gaming video streaming applications. The choice to develop deep learning-based prediction models is due to the effectiveness of their operation when trained with huge amounts of data, such as those encountered in real mobile networks, which tends to completely replace the use of statistical analysis and conventional ML models.

III. GAMING VIDEO QOE ASSESSMENT

The idea behind the development of the QoE metric is to reflect the quality of a service as perceived by end-users. As a result, in a wireless network ecosystem, QoE indicates how the network operation affects the perceived quality of communications services. The QoE evaluation is conditional on the comparison between the anticipated quality characteristics that define the user's expectations and the perceived characteristics derived from the physical stimulus.

A contributing factor to the difficulty in comprehending gaming video QoE mechanisms is that gaming, as opposed to conventional media applications, can be viewed as a human-machine interaction. Therefore, typical approaches to assessing the influence of transmission on media distribution are not applicable. Furthermore, parameters such as the content, the backend platform on which the game was developed, the user interface, the wireless communication channel, and the user features can all significantly affect QoE [20].

A. PREDICTION METHODOLOGIES

The literature reports two ways of evaluating QoE, subjective and objective evaluation [21]. In subjective methods, participants evaluate the quality of a service after being

subjected to a set of tests or stimuli. These techniques use psychophysical and psychometric methodologies to quantify evaluators' perception of service quality, as well as qualitative methodologies to determine which IFs affect QoE and to what extent [22]. Evaluators, as a general rule, rate a variety of perceived quality attributes on a MOS scale that spans from 1 to 5 (i.e., poor to excellent), reflecting the degree of contentment with a service [23]. The main advantage of subjective evaluations lies in the fact that, due to the direct reception of data from end-users, they achieve precise results. These results can subsequently be used as a reference for training and validating QoE prediction models [24].

Because subjective methodologies have boundaries such as being laborious, costly, unable to perform real-time use, and non-reproducible, there has been a substantial interest in creating objective models that predict QoE using exclusively measurable qualitative characteristics of communication networks. The idea behind objective methods is to predict QoE values that are close to the evaluations of subjective techniques. The fundamental benefit of objective approaches is their ease of use and adjustment, because the assessment procedure requires quantifiable QoS parameters and mathematical models that connect these parameters to QoE values. The drawback of these methodologies is their inaccuracy, as they calculate the QoE value; hence, their results are an estimate and not an actual representation of the quality perceived by end-users [25].

B. GAMING VIDEO QOE NETWORK/TRANSMISSION INFLUENCING FACTORS

Any characteristic of a user, system, service, or application that can have an impact on the perceived quality of a service by the end-user can be referred to as a QoE IF [26]. The purpose of this paper is to investigate the effect of the operation of the wireless communication network on QoE. For this reason, our work focuses on the study of network/transmission-related IFs. These IFs are affected by transmission channel losses and refer to network QoS parameters, including latency, jitter, bandwidth and packet loss [27].

Wireless gaming video streaming and cloud gaming applications require extensive Internet resources when connecting between client and server. Due to the need to exchange a huge amount of multimedia data with a server, a user's wireless connection must be able to send this data with the lowest possible delay [28]. Therefore, these types of applications are susceptible to variations in network QoS parameters, which makes it necessary to ensure adequate values for latency, jitter, bandwidth, and packet loss to achieve optimal QoE:

- The *latency* perceived by an end-user is associated with the time interval from the execution of the user's commands to the occurrence of the subsequent game event on the screen. Therefore, the impact of delay on QoE is greatly affected by the game's characteristics [29].

- *Jitter* has a discernible influence on the QoE of online and cloud gaming applications [30]. The presence of jitter induces an unsmooth visual impression of the game, as due to its effect, the display of video frames occurs with a fluctuating latency [31].
- The impact of *bandwidth* limitation on QoE has proven to be particularly important for cloud gaming applications, as they rely on real-time gaming content streaming. This requires a consistent and high-rate data transfer to ensure smooth gameplay. Bandwidth constraints can lead to buffering, lower resolution and increased latency, negatively impacting QoE [32].
- *Packet loss* has a significant impact on QoE in gaming applications, with levels as low as 1% resulting in an important decline in the end-user experience. Excessive packet loss reduces visual quality, leading to lower frame rates and a poor gaming experience [33].

C. QoS/QoE MAPPING

The operating principle of QoS/QoE mapping is to calculate QoE using only countable values of the network QoS parameters. To specify the association between these parameters and QoE level, it is essential to create a correlation model. The purpose of this model is to compute QoE values using fitting mathematical models [34]. The prediction model we have developed is based on the logistic mapping function, IQX hypothesis, Weber-Fechner law, and Steven's power law, as follows:

1) To convert the countable QoS parameters into MOS scores, an appropriate mapping function is required. Mapping functions can be either linear or nonlinear [35]. However, since objective quality metrics are seldom uniform, the linear mapping functions tend to understate the outcome. Therefore, typically, nonlinear mapping functions are employed, as they obtain more accurate correlations. One of the most commonly used mapping functions is the *logistic*, as shown in the below formula [22]:

$$QoE = a / (1 + e^{-b(QoS-c)}). \quad (1)$$

where a , b and c coefficients are adjustable parameters.

2) The *IQX hypothesis* is an exponential method that describes QoE as a parameterized negative exponential function of a QoS impairment attribute. QoE can be defined as a function of n influence factors I_j , $1 \leq j \leq n$:

$$QoE = \Phi(I_1, I_2, \dots, I_n) \quad (2)$$

IQX hypothesis focuses on a single influence component, $I = QoS$, to obtain the primary correlation $QoE = f(QoS)$ [36]. We fit the differential equation below, considering a linear connection at the QoE level:

$$\frac{\partial QoE}{\partial QoS} = \sim -(QoE - \gamma). \quad (3)$$

where γ coefficient defines the sensitivity of the observed QoE to changes in QoS. This equation's outcome is specified as an exponential function reflecting the fundamental relationship of the IQX hypothesis:

$$QoE = \alpha e^{-\beta QoS} + \gamma \quad (4)$$

where a , b and c coefficients are adjustable parameters.

3) The *Weber-Fechner Law* (WFL) is a logarithmic approximation that relates the perceptual capacities of the human sensory system to the awareness of scarcely perceptible changes of a salient stimulus [37]. The following differential equation explains it:

$$\frac{\partial Perception}{\partial Stimulus} \sim \frac{1}{Stimulus}. \quad (5)$$

Therefore, the ensuing mathematical equation is logarithmic and can be utilized to express the coupling of stimulus and perception as follows:

$$QoE = a \log(QoS) + b. \quad (6)$$

where a , and b coefficients are adjustable parameters.

4) The *Stevens' power law* (SPL) is a psychophysics law that describes how the intensity of a physical stimulus affects human perception [38]. The following equation can be used to describe SPL:

$$P(S) = KS^b. \quad (7)$$

where P stands for human perception as a product of stimulus strength S , K is a constant that varies with the measurement setting, and the exponent b denotes the kind of stimulus and defines the curvature of the function power.

IV. QoE PREDICTION MODEL

The suggested QoE prediction model is based on a multi-headed CNN and applied to a small-scale experimental wireless network created using Open RAN technology. The user-side QoS monitoring system has been built with open-source monitoring tools. The collected QoS attributes of bandwidth, delay, jitter and packet loss are provided to the QoS/QoE mapping model, which maps them to MOS values. Finally, to calculate the overall QoE value, the multi-headed CNN prediction model analyzes the interdependencies between the mapped input values.

A. OPEN RAN TESTBED

Open RAN technology has the potential to revolutionize mobile RAN and is expected to be a key element in the creation of next-generation wireless networks. In line with the O-RAN Alliance specifications, the Open RAN architecture will be able to provide customized communication services to meet the demanding and diverse service demands of innovative usage scenarios, including online and cloud gaming applications [39].

The Open RAN implementation operates in a small-scale testbed environment, supporting both 4G and non-standalone

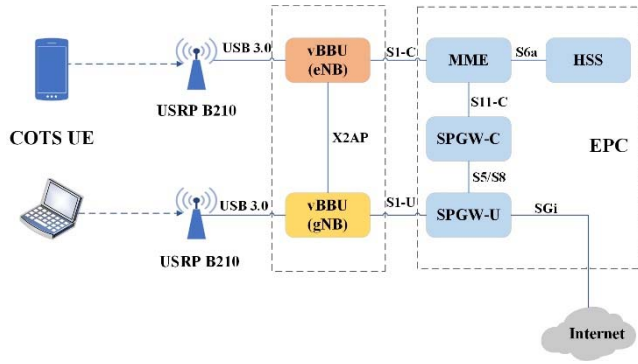


FIGURE 1. Monolithic gNB (5G NSA) architecture [42].

5G (NSA) architectures. It is based on the open-source software package of OpenAirInterface (OAI) [40]. The testbed supports three different node base station (eNB) architectures (monolithic, split option 2 and split option 7.2), a next generation node base station (gNB) architecture (monolithic) and an evolved packet core (EPC) architecture [41]. All network components are installed on the same physical hardware, a personal computer (PC) with 64 GB of RAM and an Intel i9 3.6 GHz CPU. Transmission in eNB is configured in frequency division duplex transmission (FDD) mode with a bandwidth of 20 MHz in band 7, while transmission in gNB is configured in time division duplex transmission (TDD) mode with a bandwidth of 40 MHz in band n78. The following are the main characteristics of the testbed, as illustrated in Fig. 1 [42]:

- Remote radio head (RRH) is implemented for all RAN architectures via a universal software radio peripheral (USRP) B210 in single-input single-output (SISO) mode, connected to the PC with a universal serial bus (USB) 3.0 cable.
- For split option 7.2, the distributed unit (DU) implements higher-layer physical layer (high-PHY) functions, including medium access control (MAC) and radio link control (RLC) layer, while the central unit (CU) deploys radio resource control (RRC), packet protocol data convergence process (PDCP) and service data adaptation process (SDAP). The two units are connected to each other via an F1 interface.
- The virtual baseband unit (vBBU) is present in the cases of 4G and 5G monolithic architectures, performing the functions of the CU and DU units.
- The EPC consists of the 3GPP modules based on the home subscriber server (HSS), the mobility management entity (MME), the service gateway (SGW) and the packet gateway (PGW). However, the service and packet gateways (SPGWs) have separate user (SPGW-U) and control (SPGW-C) layers.

B. QOS MONITORING AND DATASET GENERATION

The QoS monitoring system uses the open-source monitoring tools Prometheus [43], Telegraf [44] and Grafana [45] for the

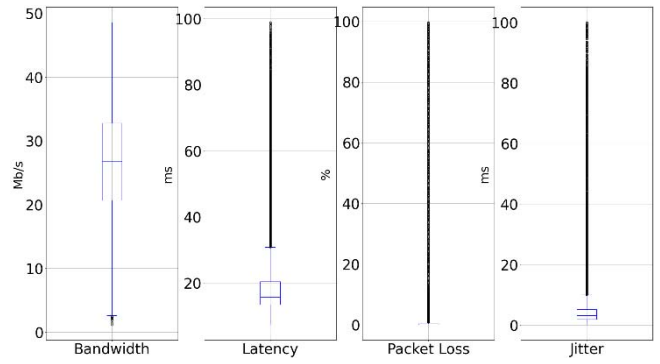


FIGURE 2. Five number summary of QoS parameters.

collection of the network/transmission metrics of bandwidth, jitter, latency and packet loss. It can be classified as a user-centric QoS monitoring system since network/transmission parameters are collected from the user side using end-user device probes to provide application-level metrics [46].

The dataset for training the neural network was generated by monitoring the QoS parameters during the wireless transmission of YouTube gaming videos. Data was collected using a variety of video streaming loads from November 2022 to January 2023. The monitoring period was 8 weeks in total, and data was collected with a measurement interval of 1 min, resulting in a total of 80640 data samples for each QoS parameter. Videos with a resolution of 2560x1440 pixels (2K) and frame rate of 60 FPS were used. The five-number summary of the dataset (i.e., minimum, first quartile, median, third quartile and maximum) is shown in Fig. 2 and the collected samples of QoS parameters are shown in Fig. 3.

C. QOS/QOE MAPPING MODEL

In order to construct the QoS/QoE mapping model, we used the logistic, IQX, WFL and SPL mapping methods by applying formulas (1), (4), (6) and (7) to the measurements of the network QoS parameters, respectively. To obtain an approximate curve $y = f(x)$ that best fits the discrete set of measurement points (x_i, y_i) , where $i = 1, 2, 3, \dots, n$, we used the curve fitting method. In particular, we adopted the method of least squares, which is one of the most commonly used methods to find the curve that best fits a given dataset [47]. In the formula $f(x)$, the coefficients a , b , and c of the mapping functions are adjustable parameters. The aim of the least squares method is to determine these parameters to minimize the fitting error, i.e., the variance between the data values y_i and the y -values $f(x_i)$ on the fitted curve. The residuals are defined as the variances between the observed y -values and those given by the fitted curve at the x -values where the data were originally collected.

Let the data points be $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ where x is the independent variable and y is the dependent variable. The deviation error e_i of the fitted curve $f(x)$ from each data

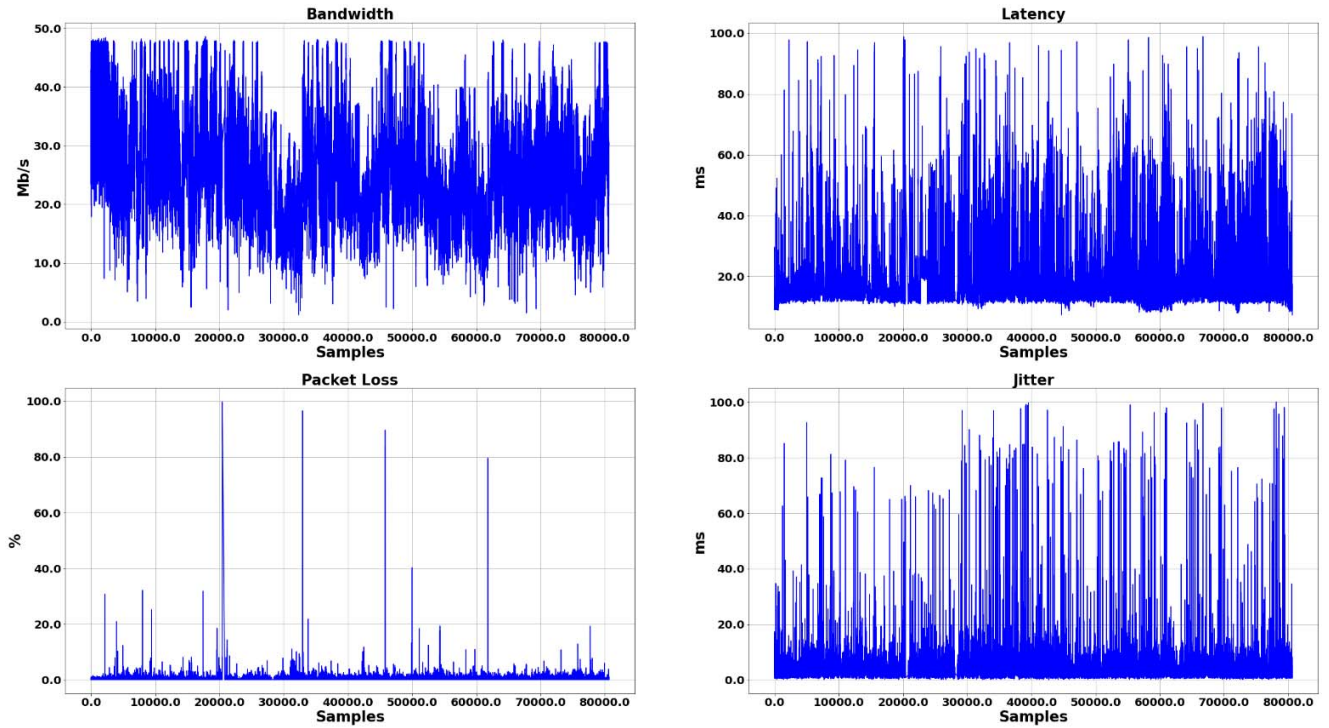


FIGURE 3. Samples of QoS parameters.

point is determined as follows:

$$\begin{aligned} e_1 &= y_1 - f(x_1), \\ e_2 &= y_2 - f(x_2), \\ &\vdots \\ e_n &= y_n - f(x_n). \end{aligned} \quad (8)$$

According to the principle of the least squares, the best fitting curve has the property that the sum of the squares of errors in formula (8) is minimum, and hence, the calculated value of the parameters a , b , and c minimizes the error e_i .

$$\sum_1^n e_i^2 = \sum_1^n [y_i - f(x_i)]^2. \quad (9)$$

To implement the least squares curve fitting we used Python’s SciPy, NumPy, and Pandas open-source libraries [48]. To evaluate the curve fitting accuracy, we used the metrics of R^2 and mean squared error (MSE) [34].

Table 1 shows the curve fitting accuracy results for the IQX, WFL, SPL and logistic mapping functions, whereas Fig. 4 shows the associated curves. The QoS/QoE mapping mechanism, as shown in Fig. 4, transforms the measured values of bandwidth, jitter, delay and packet loss into MOS values. It is worth mentioning that in the case of latency, jitter and packet loss, their impact on perceived service quality develops in reverse proportion to the degradation, i.e., the higher the QoS value, the lower the objective quality. In the case of bandwidth, however, the higher the QoS score, the higher the objective quality. For this reason, the algebraic

TABLE 1. Curve fitting accuracy.

| | R^2 | MSE |
|----------------------|-------|------------------------|
| IQX Bandwidth | 0.99 | $7.33 \cdot 10^{-15}$ |
| IQX Latency | 0.99 | $7.85 \cdot 10^{-18}$ |
| IQX Packet Loss | 0.99 | $5.46 \cdot 10^{-19}$ |
| IQX Jitter | 1 | 0 |
| WFL Bandwidth | 0.99 | $1.12 \cdot 10^{-14}$ |
| WFL Latency | 1 | 0 |
| WFL Packet Loss | 1 | 0 |
| WFL Jitter | 1 | 0 |
| SPL Bandwidth | 0.99 | $3.18 \cdot 10^{-14}$ |
| SPL Latency | 1 | 0 |
| SPL Packet Loss | 1 | 0 |
| SPL Jitter | 1 | 0 |
| Logistic Bandwidth | 0.99 | $3.43 \cdot 10^{-15}$ |
| Logistic Latency | 0.99 | $14.06 \cdot 10^{-14}$ |
| Logistic Packet Loss | 0.99 | $1.06 \cdot 10^{-13}$ |
| Logistic Jitter | 0.99 | $1.44 \cdot 10^{-14}$ |

signs in equations (1), (4), (6), and (7) have been adjusted accordingly.

D. DEEP LEARNING PREDICTION MODEL

The development of the QoE prediction model is based on deep learning techniques, namely CNN. Originally designed for two-dimensional image data, CNN can also be used to model time series prediction problems, as they can automatically learn features from sequence data, handle multivariate data, and instantly generate a vector for multistep prediction. One of the key advantages of CNN is that it can generate representations for fixed-size frames, permitting

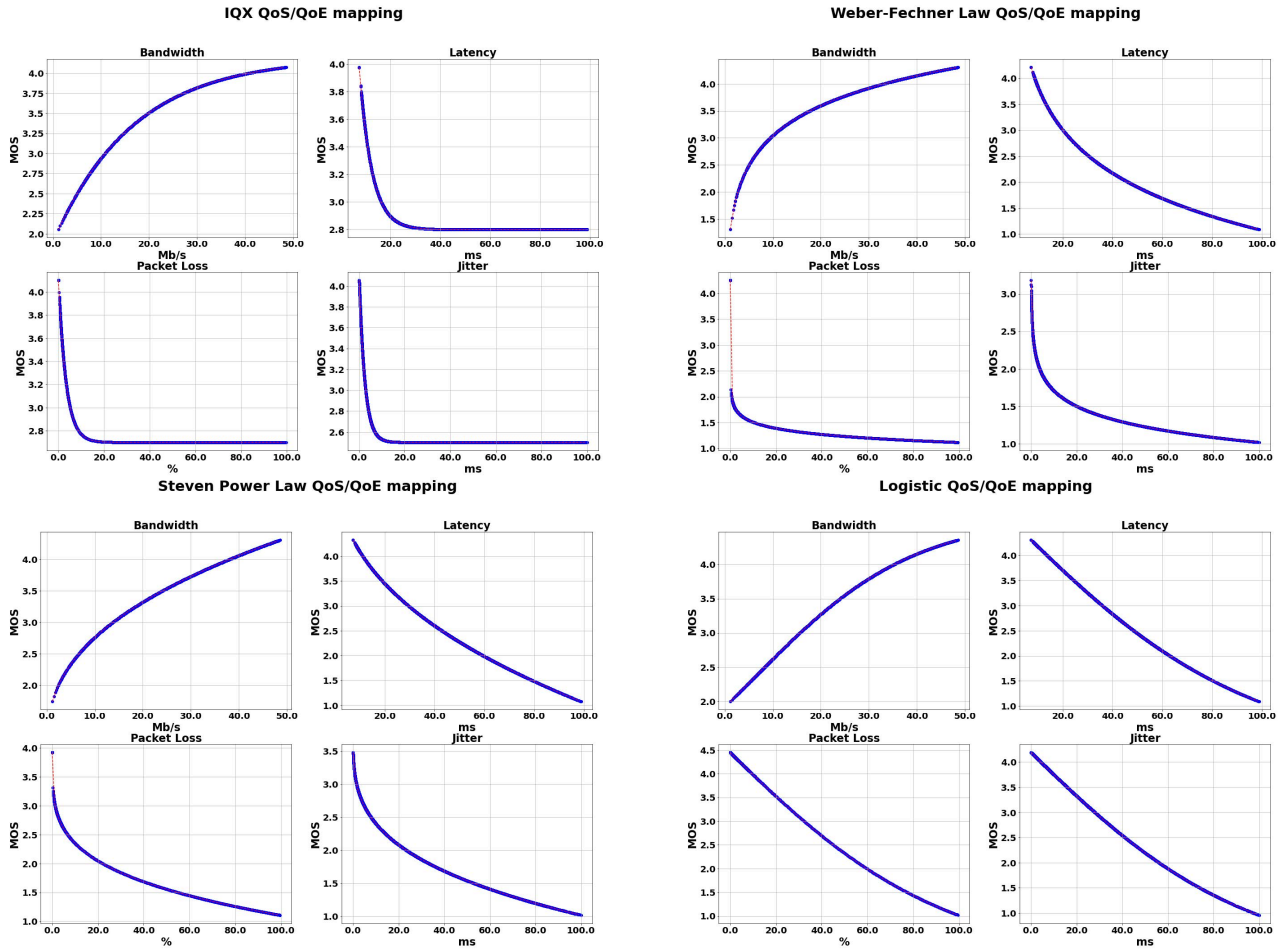


FIGURE 4. Curve fitting for QoS/QoE mapping.

the expansion of the network's frame size through multi-level stacking. This ensures exact control over the maximum length of the dependencies to be modeled. Furthermore, the convolutions can be applied in parallel, allowing an entire input sequence to be processed in its whole during both training and evaluation, which results in faster training and testing procedures [49]. As a rule, the architectural design of CNN consists of an input

layer, an output layer and multiple hidden layers. These layers are referred to as convolutional layers, aggregation layers and fully connected layers, respectively. More specifically, the convolution layers are responsible for the convolution or cross-correlation process, which maps multidimensional parameters using kernels, i.e., locally connected filters. The convolution performs for each position p of the output y , the following operation [50]:

$$yp_y = \sum_{p_G \in G} w_{p_G} x(p_y + p_G). \quad (10)$$

where p_G stands for the positions in the receptive field G of the convolutional filter W , which corresponds to the receptive field of the neurons in the inputs at a convolutional layer. W denotes the weights which are distributed in the

input representation. Typically, when the inputs and outputs contain M and N filters respectively, the convolutional layer would need $M \times N$ filters to perform the convolution process. The CNN architecture is based on three key principles [51]: 1) sparse interactions, 2) parameter sharing, and 3) equivalent representations, which has the effect of reducing the volume of model features, and the number of necessary factors, minimizing the possibility of overfitting during model training.

A subclass of the CNN architecture is the CNN model with multiple heads, in which a separate CNN sub-model or a separate head is used for each input variable. In multi-headed architecture, adaptation is required in the design of the model and in the preprocessing of the training and test datasets. In terms of model design, a separate CNN model must be defined for each of the input variables, with appropriate modifications to the number of layers and hyperparameters tuning. More specifically, a sub-model must be developed for each input variable, which accepts a one-dimensional sequence of input data and outputs a planar vector containing a summary of the feature learning. The set of these vectors can be merged by concatenation in order to develop a large vector that, prior

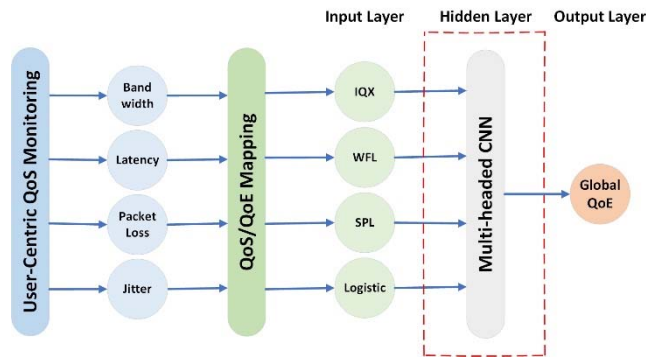


FIGURE 5. Deep learning QoE prediction model architecture.

to forecasting, is interpreted by a number of fully connected layers.

To develop the deep learning prediction model, we used TensorFlow [52] and Keras [53]. The suggested QoE prediction model is built using a multistep, multi-headed CNN with multivariate input. The multivariate input, as illustrated in Fig. 5, comprises the mapped values of bandwidth, jitter, latency, and packet loss. The hidden layer is based on a multi-headed CNN designed to resolve sequence problems. Since the number of elements in input and output sequences might differ and prediction entails providing the next value in an actual sequence, the task of sequence prediction is made particularly challenging. These kinds of applications are typically framed as prediction problems involving one or more input time steps and one output time step. In our case, we use the QoS measurements of the previous 2 days to forecast the QoE of the next day.

In data preparation we resampled the dataset into non-overlapping 60-minute sequences. Consequently, the 80640 minute input data were resampled into hourly samples, yielding a total of 1344 samples. The multi-headed CNN is designed as a multistep multivariate time series prediction model that predicts the next day's QoE values based on the previous two days' data. In particular, to predict the QoE of 24 forward time steps (1 day x 24 hours), it uses 48 hourly samples (2 days x 24 hours) as backward time steps. The prediction model has 16 input sequences corresponding to the mapped values of the QoS parameters and 1 output that yields the overall QoE value.

The structure of the multi-headed CNN model is shown in Fig. 6. The model does not consider the data to have time steps; instead, it treats it as a sequence on which convolutional layers may apply filters similar to a 1D image. During model development, the input and flatten layers are kept in lists in order to determine the model inputs and the function of the max pooling layer. The model expects 16 arrays as input, one for each of the sub models. This is required at all stages of the process, including training, evaluation, and prediction with a final model. For this reason, a list of 16 3D matrices [*samples, time steps, features*] has been created, where the column of features is

equal to one, as it represents the one-dimensional input sequences.

V. PREDICTION MODEL EVALUATION

The aim of this paper is to introduce a model designed to predict the QoE of wireless gaming video streaming in real-world conditions. The development of the proposed prediction model is based on deep learning techniques, while its effectiveness is evaluated through comparison with state-of-the-art deep learning models. This choice is based on the superior performance demonstrated by deep learning models when trained with huge amount of data, such as that encountered in real-world mobile networks. Therefore, statistical analysis and traditional ML models are not considered, and instead, the performance of the proposed model is evaluated against deep learning models belonging to the CNN and RNN classes.

A. DEEP LEARNING MODELS FOR COMPARISON

The multi-headed CNN is compared with the most widely used models of the CNN class, including CNN, multi-channel CNN, temporal convolutional network (TCN), and residual CNN-LSTM (ResCNN-LSTM), as well as with the well-established models of the recurrent neural network (RNN), LSTM, and gated recurrent unit (GRU):

- *CNN* is suitable for multistep time series forecasting problems. It can be used in two ways: either in a recursive forecasting approach, in which the model produces one-step forecasts and the outputs are provided as inputs for subsequent forecasts, or in a direct forecasting approach, in which a model is constructed for each time step to be forecast. A major advantage of using CNN is that it can provide forecasts using numerous 1D inputs, which is advantageous in problems where the output is a function of a multivariate input sequence.
- *Multi-channel CNN* is an enhanced version of CNN designed to predict the value of the next time step using each of the variables in the time series input sequence. This is achieved by feeding each unidimensional time series into the model as a separate input channel. The model then reads each input sequence into a separate set of filter maps, thus learning features from each input time series variable. This method is useful for problems in which the output sequence is a function of observations at previous time steps of various independent features.
- *TCN* is a variant of CNN for sequence modelling problems that incorporates elements of RNN and CNN design. The key difference between TCN and other CNNs is that it uses causal and dilated convolutions. Causal convolutions prompt the model to learn the relationship between time steps while maintaining the natural time sequence. This differs from other CNNs, whose algorithms use all available data in a sequence. The dilation approach allows TCN to process more steps

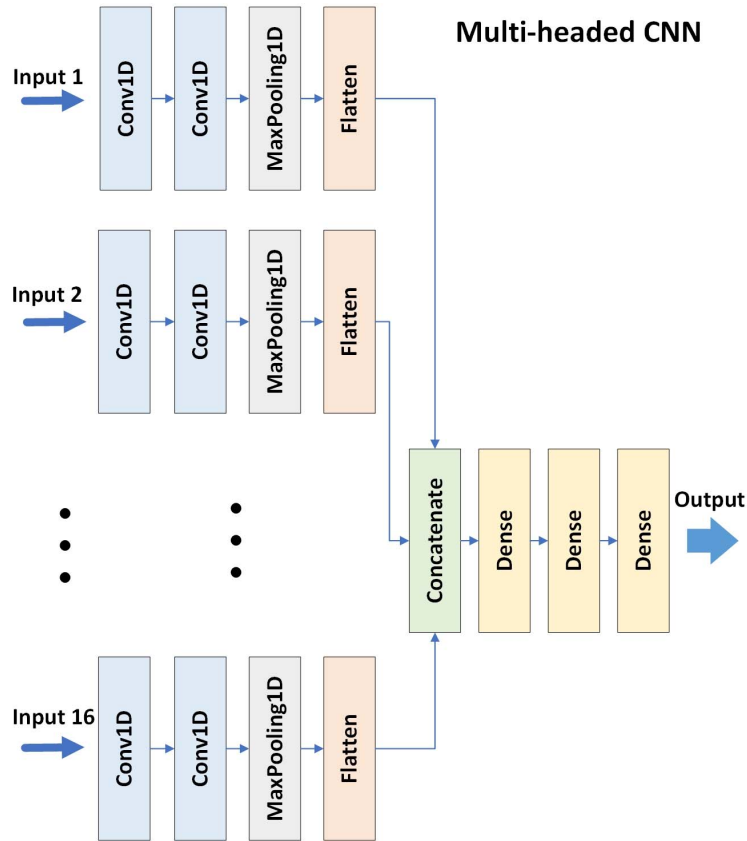


FIGURE 6. Structure of the multi-headed CNN.

in the time series as it progresses deeper into the layers. Typically, these networks are trained faster than RNN.

- *LSTM*, a dedicated type of RNN, is specifically developed for applications where the use of a common RNN proves to be ineffective. It excels at capturing long-term dependencies, having an architectural design specifically tailored to address the primary shortcomings of RNN. These shortcomings include the challenge of preserving past state information and addressing the problems of exploding and vanishing gradients commonly associated with conventional RNN.
- *ResCNN-LSTM* is a CNN framework that addresses the vanishing gradient problem, allowing for the construction of networks with a huge number of convolutional layers. This is accomplished by the use of skip connections, in which some layers are skipped and the output of the preceding layer is passed on to the current position, enhancing the model's operation. Another purpose of skip connections is to allow for better gradient flow and to guarantee that critical attributes are conveyed down to the network's final layers without increasing computational load.
- *RNN* is a modified version of feedforward neural networks (FNN), specifically designed for processing sequential or time series data. Unlike conventional FNN, which is suitable for handling unrelated data, RNN excel in scenarios where the values in a sequence are

dependent on each other. It can be trained to retain historical information from previous inputs, allowing it to predict future values within the sequence. This ability to capture the dependencies between values makes RNN an effective choice for modeling data with sequential relationships.

- *GRU* stands out as an improved variant in the RNN category, addressing the challenge of vanishing gradients faced by traditional RNN. Similar to LSTM, GRU addresses this issue by incorporating two gates - the update gate and the reset gate - to regulate the data stream. What makes GRU salient is its simpler architecture compared to LSTM; it does not have a separate cell state and has only one hidden state. This streamlined design simplifies the model and also improves the training efficiency, making GRU a more efficient and advanced option.

B. COMPARATIVE EVALUATION OF PREDICTION MODELS

The prediction models use each of the 16 time series input variables to predict the total value of QoE for the next 24 time steps. This is achieved by including each one-dimensional time series as a separate input sequence, to which an internal representation is then interpreted from the output. This is because, instead of predicting a single feature, our approach requires multivariate inputs, as the output sequence is a

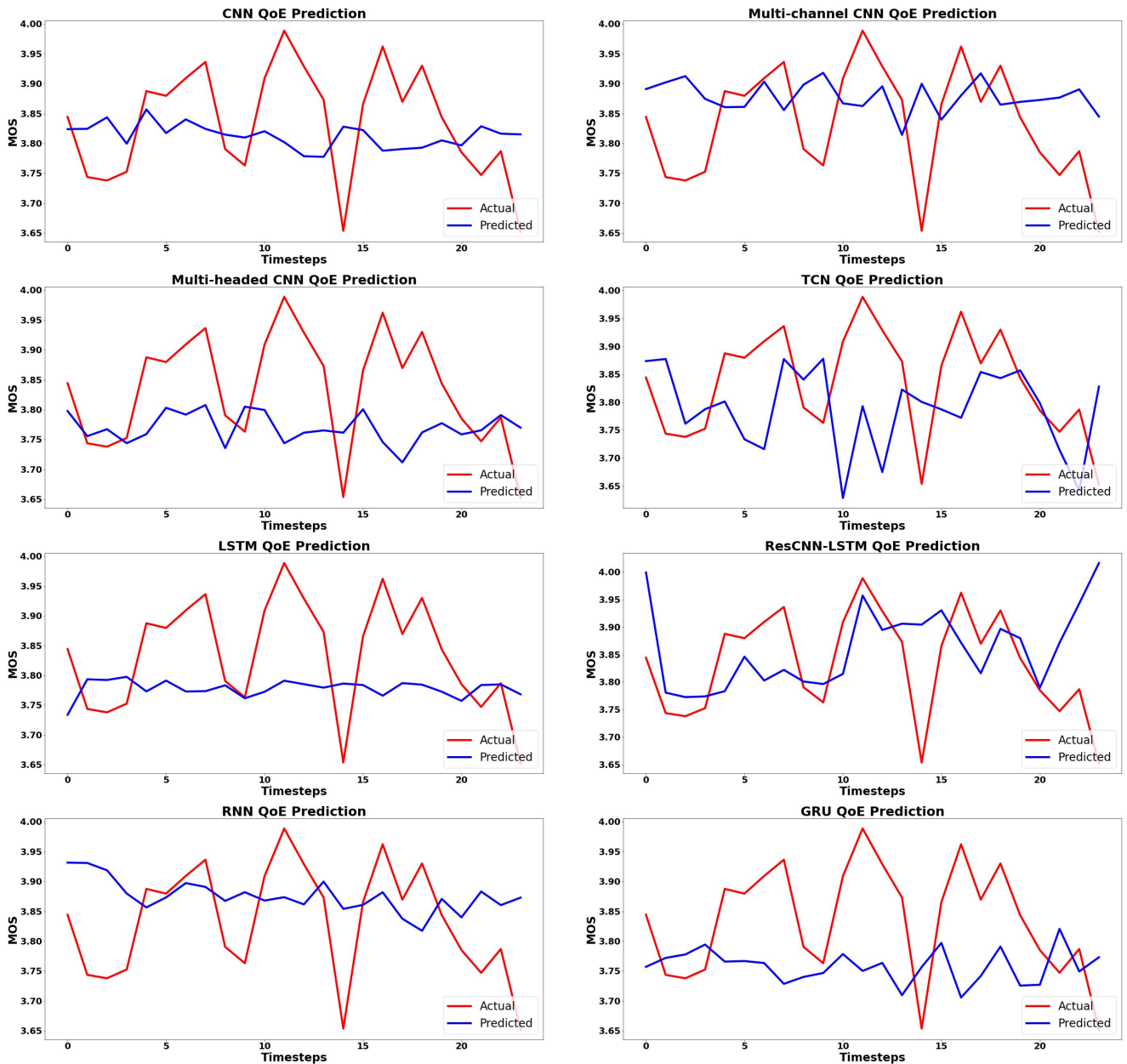


FIGURE 7. Representation of actual versus predicted QoE value.

product of observations of previous time steps affected by multiple independent features.

The 8 weeks of QoS monitoring resulted in 56 days of data collection. These 56 days were distributed as follows: the training data set consists of 42 days, the test data set consists of 7 days, and the validation data set consists of 7 days. We employed Keras callbacks to increase the efficiency of training when assembling the models. In particular, we utilized ModelCheckpoint to store model weights at key time steps, EarlyStopping to terminate the process when the observed evaluation metric no longer improves, and ReduceLROnPlateau to minimize the learning rate when the observed metric no longer improves.

We utilized the MSE, root-mean-square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and median absolute error (MedAE) metrics

to assess the accuracy of the prediction models [34]. Table 2 presents the findings of the prediction accuracy metrics. We can see that the suggested multi-headed CNN model prevails over the other models by outperforming them in all

the accuracy scores. It is important to emphasize that all the prediction models show moderate discrepancies between them, which validates the suitability of the CNN and RNN class models in addressing time series data modelling problems.

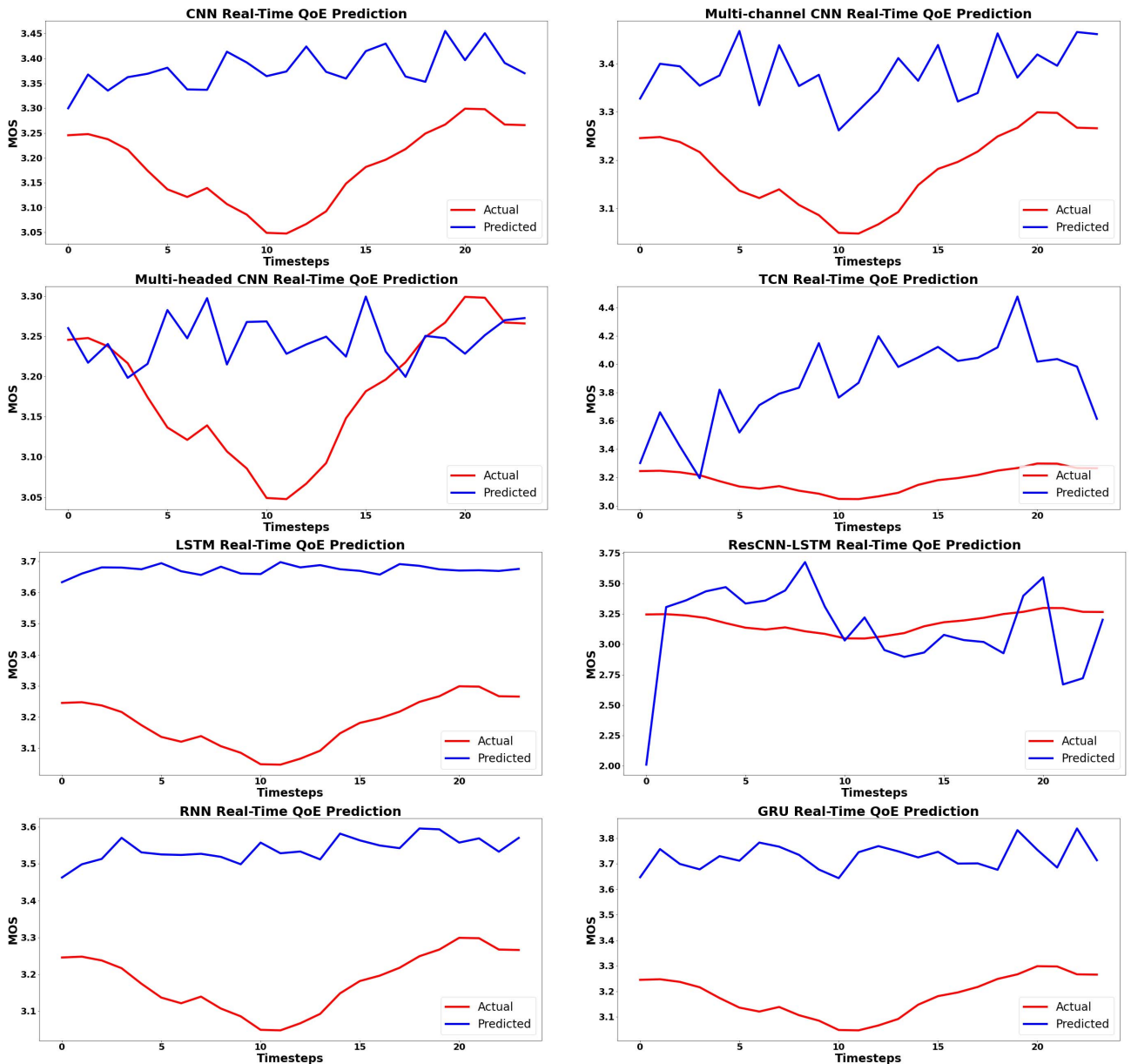


FIGURE 8. Representation of real-time QoE prediction versus actual values.

The results of the QoE prediction of the next day over the validation dataset values of the previous 2 days are depicted in Fig. 7. The next day's prediction corresponds to 24 forward time steps, and the validation basis of the previous 2 days corresponds to 48 backward time steps. Table 3 includes the mean QoE values as derived from the considered prediction models. We can observe that the proposed model approximates the actual QoE value with greater accuracy compared to the rest of the prediction models.

The reason for using mean QoE values is that this aggregate metric summarizes the performance of deep learning models across the entire dataset and observation time. This facilitates comparison between models and provides a strong indication of the expected level of QoE predicted by

the models. In addition, the deep learning models exhibit stochastic behavior during the training phase due to the use of mini-batch sampling and the dropout regularization technique. Utilizing average QoE values helps mitigate the effects of this stochasticity by providing a summary statistic that reflects overall performance by smoothing out the noise introduced during training.

C. REAL-TIME QOE PREDICTION

The trained models are used to predict QoE in real-time, as shown in Fig. 8, where QoE values are predicted for 24 time steps forward. Predicted versus actual QoE values are presented, based on real-time QoS monitoring and QoS/QoE mapping. The experimental data collection was performed over a 4-hour period and was based on QoS monitoring

TABLE 2. QoE prediction model evaluation.

| | MSE | RMSE | MAE | MAPE (%) | MedAE |
|-------------------|----------------|----------------|----------------|----------------|----------------|
| CNN | 0.03241 | 0.18005 | 0.14492 | 3.92163 | 0.13820 |
| Multi-channel CNN | 0.03432 | 0.18527 | 0.14835 | 4.03715 | 0.13218 |
| Multi-headed CNN | 0.01342 | 0.11588 | 0.09786 | 2.52839 | 0.09664 |
| TCN | 0.03581 | 0.18924 | 0.15303 | 4.14064 | 0.13052 |
| LSTM | 0.03250 | 0.18029 | 0.14196 | 3.85048 | 0.13774 |
| ResCNN-LSTM | 0.03245 | 0.18016 | 0.14317 | 3.87785 | 0.13492 |
| RNN | 0.03450 | 0.18576 | 0.14212 | 3.88562 | 0.11938 |
| GRU | 0.03232 | 0.17978 | 0.14060 | 3.82902 | 0.12912 |

TABLE 3. QoE predicted mean value

| | Mean QoE (MOS) |
|-------------------|----------------|
| Actual | 3.76110 |
| CNN | 3.75376 |
| Multi-channel CNN | 3.78851 |
| Multi-headed CNN | 3.76470 |
| TCN | 3.74875 |
| LSTM | 3.75908 |
| ResCNN-LSTM | 3.75194 |
| RNN | 3.79294 |
| GRU | 3.77089 |

TABLE 4. Real-time QoE prediction model evaluation.

| | MSE | RMSE | MAE | MAPE (%) | MedAE |
|-------------------|----------------|----------------|----------------|----------------|----------------|
| CNN | 0.23826 | 0.48811 | 0.48011 | 15.1477 | 0.47310 |
| Multi-channel CNN | 0.01761 | 0.13272 | 0.10765 | 3.43980 | 0.09638 |
| Multi-headed CNN | 0.01362 | 0.11674 | 0.09386 | 3.00307 | 0.07228 |
| TCN | 0.56402 | 0.75101 | 0.72555 | 22.8735 | 0.72082 |
| LSTM | 0.23745 | 0.48729 | 0.48058 | 15.1615 | 0.47333 |
| ResCNN-LSTM | 0.31307 | 0.55952 | 0.54857 | 17.2956 | 0.54257 |
| RNN | 0.11620 | 0.34089 | 0.32962 | 10.4182 | 0.32056 |
| GRU | 0.30079 | 0.54844 | 0.54065 | 17.0485 | 0.53445 |

of the gaming video streaming via the Open RAN testbed, providing a total of 240 minute samples. These samples are used to evaluate the accuracy of the real-time QoE prediction in backward time steps and are the dataset against which the ability of the model to generalize well to new data is validated.

The suggested multi-headed CNN model outperforms the CNN, multi-channel CNN, TCN, LSTM, ResCNN-LSTM, RNN, and GRU prediction models in all accuracy measures, as shown in Table 4. We can also see that the real-time QoE prediction accuracy scores for all models are lower than the values obtained when evaluating the models. This is due to the fact that the experimental real-time dataset is limited in size, which affects the intrinsic variance of the input data, leading to a larger intrinsic variance of the training dataset compared to the validation dataset. Due to its greater ability to handle small datasets, the multi-headed CNN model achieves higher accuracy than the rest of the

TABLE 5. Real-time QoE predicted mean value.

| | Mean QoE (MOS) |
|-------------------|----------------|
| Actual | 3.18467 |
| CNN | 3.66479 |
| Multi-channel CNN | 3.28782 |
| Multi-headed CNN | 3.26618 |
| TCN | 3.91023 |
| LSTM | 3.66526 |
| ResCNN-LSTM | 3.73323 |
| RNN | 3.51430 |
| GRU | 3.72533 |

prediction models. The prediction accuracy improves when the volume of the dataset is increased through long hours of data collection, but we are interested in verifying the usefulness of the prediction model in real-world situations, where the model should be able to provide predictions in a limited amount of time.

Table 5 presents the real-time average QoE value obtained from the considered forecasting models. This value indicates the overall performance of the Open RAN testbed as determined by the composition of the dependencies among the 16 input variables. As mentioned previously, the multi-headed CNN model approximates more accurately the actual QoE value compared to state-of-the-art deep learning prediction models. This demonstrates that the multi-headed CNN is highly effective in handling small-sequence problems, making it suitable for real-time applications, where, due to time constraints, wireless network performance evaluation relies on feeding a limited amount of new data into the QoE prediction model.

VI. CONCLUSION

Real-time QoE prediction should be thought of as a key component in the design of next-generation wireless communication networks, especially with regard to the transmission of inherently interactive and increasingly demanding gaming video streaming applications. In this work, we provide a QoE prediction model based on deep learning techniques. More specifically, we present a multi-headed CNN-based objective QoE prediction model for gaming video streaming applications. The prediction model is evaluated on an Open RAN testbed and is capable of quantifying the impact of wireless network operation on gaming video quality using the QoS parameters of bandwidth, latency, packet loss, and jitter.

The suggested prediction model outperforms state-of-the-art deep learning models, including CNN, multi-channel CNN, TCN, LSTM, ResCNN-LSTM, RNN, and GRU. The multi-headed CNN model has greater prediction accuracy and can generalize to new data more successfully. Furthermore, it is suitable for time series forecasting problems with multivariate inputs since a distinct CNN sub-model is employed for each input variable and an output planar vector is created that comprises the feature learning summary. In addition, our findings show that the

multi-headed CNN model is ideal for real-time forecasting applications since it operates effectively with small data sequences.

This is the first study to present a comparative analysis of deep learning algorithms from the CNN and RNN class for QoE prediction in wireless gaming video streaming applications. Furthermore, this is the first time a QoE prediction model for gaming video streaming has been developed and evaluated on an Open RAN testbed.

REFERENCES

- [1] *2019 Year in Review: Digital Games and Interactive Media: SuperData*, Nielsen Holdings Media co., New York, NY, USA: 2020.
- [2] "Statista." Accessed: Jun. 2023. [Online]. Available: <https://www.statista.com>
- [3] F. Metzger et al., "An introduction to online video game QoS and QoE influencing factors," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1894–1925, 3rd Quart., 2022.
- [4] *O-RAN: Towards an Open and Smart RAN*, O-RAN Alliance, Alfter, Germany, Oct. 2018.
- [5] X. Tao, Y. Duan, M. Xu, Z. Meng, and J. Lu, "Learning QoE of mobile video transmission with deep neural network: A data-driven approach," *IEEE J. Select. Areas Commun.*, vol. 37, no. 6, pp. 1337–1348, Jun. 2019.
- [6] G. Huang, O. Ercetin, H. Gokcesu, and G. Kalem, "Deep learning-based QoE prediction for streaming services in mobile networks," in *Proc. 18th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Thessaloniki, Greece, 2022, pp. 327–332.
- [7] P. Zhou et al., "QoE-aware 3D video streaming via deep reinforcement learning in software defined networking enabled mobile edge computing," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 1, pp. 419–433, Jan.–Mar. 2021.
- [8] H. Zhang, L. Dong, G. Gao, H. Hu, Y. Wen, and K. Guan, "DeepQoE: A multimodal learning framework for video quality of experience (QoE) prediction," *IEEE Trans. Multimedia*, vol. 22, no. 12, pp. 3210–3223, Dec. 2020, doi: [10.1109/TMM.2020.2973828](https://doi.org/10.1109/TMM.2020.2973828).
- [9] A. Renda, P. Ducange, G. Gallo, and F. Marcelloni, "XAI models for quality of experience prediction in wireless networks," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, 2021, pp. 1–6.
- [10] V. Vasilev, J. Leguay, S. Paris, L. Maggi, and M. Debbah, "Predicting QoE factors with machine learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2018, pp. 1–6.
- [11] A. Ahmad, A. B. Mansoor, A. A. Barakabitze, A. Hines, L. Atzori, and R. Walshe, "Supervised-learning-based QoE prediction of video streaming in future networks: A tutorial with comparative study," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 88–94, Nov. 2021.
- [12] S. Zadtootaghaj, N. Barman, S. Schmidt, M. G. Martini, and S. Möller, "NR-GVQM: A no reference gaming video quality metric," in *Proc. IEEE Int. Symp. Multimedia (ISM)*, 2018, pp. 131–134.
- [13] N. Barman, E. Jammeh, S. A. Ghorashi, and M. G. Martini, "No-reference video quality estimation based on machine learning for passive gaming video streaming applications," *IEEE Access*, vol. 7, pp. 74511–74527, 2019.
- [14] N. J. Avnaki, S. Zadtootaghaj, N. Barman, S. Schmidt, M. G. Martini, and S. Möller, "Quality enhancement of gaming content using generative adversarial networks," in *Proc. Twelfth Int. Conf. Quality Multimedia Exp. (QoMEX)*, 2020, pp. 1–6.
- [15] S. Göring, R. R. Rao, and A. Raake, "nofu—A lightweight no-reference pixel based video quality model for gaming content," in *Proc. 11th Int. Conf. Qual. Multimedia Exp. (QoMEX)*, 2019, pp. 1–6.
- [16] M. Utke, S. Zadtootaghaj, S. Schmidt, S. Bosse, and S. Möller, "NDNetGaming-development of a no-reference deep CNN for gaming video quality prediction," *Multimed Tools Appl.*, vol. 81, no. 3, pp. 3181–3203, Jan. 2022.
- [17] S. Van Damme, M. T. Vega, J. Heyse, F. De Backere, and F. De Turck, "A low-complexity psychometric curve-fitting approach for the objective quality assessment of streamed game videos," *Signal Process. Image Commun.*, vol. 88, Oct. 2020, Art. no. 115954.
- [18] S. Zadtootaghaj et al., "DEMI: Deep video quality estimation model using perceptual video quality dimensions," in *Proc. IEEE 22nd Int. Workshop Multimedia Signal Process. (MMSP)*, 2020, pp. 1–6.
- [19] S. Wen, S. Ling, J. Wang, X. Chen, Y. Jing, and P. L. Callet, "Subjective and objective quality assessment of mobile gaming video," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2022, pp. 1810–1814.
- [20] S. Möller, S. Schmidt, and J. Beyer, "Gaming taxonomy: An overview of concepts and evaluation methods for computer gaming QoE," in *Proc. 5th Int. Workshop Qual. Multimedia Exp. (QoMEX)*, Klagenfurt am Wörthersee, Austria, 2013, pp. 236–241.
- [21] G. Kougioumtzidis, V. Poulkov, Z. Zaharis, and P. Lazaridis, "QoE assessment aspects for virtual reality and holographic telepresence applications," in *Proc. 6th EAI Int. Conf. Future Access Enablers Ubiquitous Intell. Infrastruct. (FABULOUS)*, 2022, pp. 171–180.
- [22] R. Schatz, T. Hößfeld, L. Janowski, and S. Egger, "From packets to people: Quality of experience as a new measurement challenge," in *Data Traffic Monitoring and Analysis (Lecture Notes in Computer Science, 7754)*, E. Biersack, C. Callegari, and M. Matijasevic, Eds. Berlin, Germany: Springer, 2013, pp. 219–263.
- [23] *Mean Opinion Score (MOS) Terminology*, ITU-T. P.800.1, Int. Telecommun. Union, Geneva, Switzerland, 2006.
- [24] M. A. Alreshoodi and J. C. Woods, "Survey on QoE/QoS correlation models for multimedia," *Int. J. Distrib. Parallel Syst.*, vol. 4, no. 3, pp. 53–72, 2013.
- [25] Y. Wang, W. Zhou, and P. Zhang, *QoE Management in Wireless Networks* (Springer Briefs in Electrical and Computer Engineering). Cham, Switzerland: Springer, 2017.
- [26] K. Brunnström et al., "Qualinet white paper on definitions of quality of experience," in *Proc. Output 5th Qualinet Meet.*, 2013, pp. 1–25.
- [27] *Influence Factors on Gaming Quality of Experience*, ITU-T G.1032, Int. Telecommun. Union, Geneva, Switzerland, 2017.
- [28] O. S. Peñaherrera-Pulla, C. Baena, S. Fortes, E. Baena, and R. Barco, "Measuring key quality indicators in cloud gaming: Framework and assessment over wireless networks," *Sensors*, vol. 21, no. 4, p. 1387, Feb. 2021.
- [29] U. Reiter et al., "Factors influencing quality of experience," in *Quality of Experience: Advanced Concepts, Applications and Methods* (T-Labs Series in Telecommunication Services), S. Möller and A. Raake, Eds. Cham, Switzerland: Springer, 2014, pp. 55–72.
- [30] P. Quax, A. Beznosyk, W. Vanmontfort, R. Marx, and W. Lamotte, "An evaluation of the impact of game genre on user experience in cloud gaming," in *Proc. IEEE Int. Games Innov. Conf. (IGIC)*, 2013, pp. 216–221.
- [31] M. Ries, P. Svoboda, and M. Rupp, "Empirical study of subjective quality for massive multiplayer games," in *Proc. 15th Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Bratislava, Slovakia, 2008, pp. 181–184.
- [32] Z.-Y. Wen and H.-F. Hsiao, "QoE-driven performance analysis of cloud gaming services," in *Proc. IEEE 16th Int. Workshop Multimedia Signal Process. (MMSP)*, 2014, pp. 1–6.
- [33] K.-T. Chen, Y.-C. Chang, H.-J. Hsu, D.-Y. Chen, C.-Y. Huang, and C.-H. Hsu, "On the quality of service of cloud gaming systems," *IEEE Trans. Multimedia*, vol. 16, no. 2, pp. 480–495, Feb. 2014.
- [34] G. Kougioumtzidis, V. Poulkov, Z. D. Zaharis, and P. I. Lazaridis, "A survey on multimedia services QoE assessment and machine learning based prediction," *IEEE Access*, vol. 10, pp. 19507–19538, 2022.
- [35] J. Korhonen, N. Burini, J. You, and E. Nadernejad, "How to evaluate objective video quality metrics reliably," in *Proc. 4th Int. Workshop Qual. Multimedia Exp. (QoMEX)*, Melbourne, VIC, Australia, 2012, pp. 57–62.
- [36] T. Hößfeld, P. Tran-Gia, and M. Fiedler, "Quantification of quality of experience for edge-based applications," in *Proc. 1st Int. Teletraffic Congr. (ITC)*, 2007, pp. 361–373.
- [37] P. Reichl, S. Egger, R. Schatz, and A. D'Alconzo, "The logarithmic nature of QoE and the role of the weber-fechner law in QoE assessment," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Cape Town, South Africa, 2010, pp. 1–5.
- [38] S. Khorsandroo, R. M. Noor, and S. Khorsandroo, "The role of psychophysics laws in quality of experience assessment: A video streaming case study," in *Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICACCI)*, 2012, pp. 446–452.

- [39] G. Kougioumtzidis, V. Poulkov, Z. D. Zaharis, and P. I. Lazaridis, "Intelligent and QoE-aware open radio access networks," in *Proc. 3rd URSI Atlantic Radio Sci. Meet. (AT-AP-RASC)*, Gran Canaria, Spain, 2022, pp. 1–4.
- [40] F. Kaltenberger, A. P. Silva, A. Gosain, L. Wang, and T.-T. Nguyen, "OpenAirInterface: Democratizing innovation in the 5G Era," *Comput. Netw.*, vol. 176, Jul. 2020, Art. no. 107284.
- [41] *O-RAN Use Cases and Deployment Scenarios*, O-RAN Alliance, Alfter, Germany, Feb. 2020.
- [42] A. Vlahov, G. Kougioumtzidis, A. Mihovska, and V. Poulkov, "Performance analysis of evolved RAN architectures with open interfaces," *J. Mobile Multimedia*, vol. 19, no. 1, pp. 239–262, Sep. 2022.
- [43] "Prometheus." Accessed: Dec. 2022. [Online]. Available: <https://prometheus.io/>
- [44] "InfluxData." Accessed: Dec. 2022. [Online]. Available: <https://www.influxdata.com/time-series-platform/telegraf/>
- [45] "Grafana." Accessed: Dec. 2022. [Online]. Available: <https://grafana.com/>
- [46] G. Kougioumtzidis, V. Poulkov, Z. Zaharis, and P. Lazaridis, "Machine learning for QoE management in future wireless networks," in *Proc. 34th Gen. Assembly Sci. Symp. Int. Union Radio Sci. (URSI GASS)*, Rome, Italy, 2021, pp. 1–4.
- [47] R. Levie, "Curve fitting with least squares," *Critical Rev. Anal. Chem.*, vol. 30, no. 1, pp. 59–74, Jun. 2010.
- [48] "Python." Accessed: Dec. 2022. [Online]. Available: <https://www.python.org/>
- [49] K. Wang et al., "Multiple convolutional neural networks for multivariate time series prediction," *Neurocomputing*, vol. 360, pp. 107–119, Sep. 2019.
- [50] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [51] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.
- [52] "Tensorflow." Accessed: Dec. 2022. [Online]. Available: <https://www.tensorflow.org/learn>
- [53] "Keras." Accessed: Dec. 2022. [Online]. Available: <https://keras.io/>



GEORGIOS KOUGIOUMTZIDIS received the B.Sc. degree in electronics engineering from the Alexander Technological Educational Institute of Thessaloniki, Thessaloniki, Greece, in 2007, the M.Sc. degree in wireless communication systems from the Open University of Cyprus, Nicosia, Cyprus, in 2017, and the M.A. degree in acoustic design and multimedia from Hellenic Open University, Patras, Greece, in 2018. He is currently pursuing the Ph.D. degree with the Faculty of Telecommunications, Technical University of Sofia, Sofia, Bulgaria. From 2013 to 2020, he was with Hellenic Telecommunications Organization S.A (OTE Group). He also holds an Early Stage Researcher position in European Union's Horizon 2020 MOTOR5G Project. His research interests include quality of experience enhancement in mobile and wireless networks, open radio access networks, machine learning, and extended reality and holographic telepresence communications.



ATANAS VLAHOV (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunications from the Technical University of Sofia, Bulgaria, in 2019 and 2021, respectively, where he is currently pursuing the Ph.D. degree with the Faculty of Telecommunication. His current research interests in the field of telecommunications are related to open radio access network and software defined networks.



VLADIMIR K. POULKOV (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees from the Technical University of Sofia (TUS), Sofia, Bulgaria. He has more than 30 years of teaching, research, and industrial experience, managing numerous educational, research and development and engineering projects in the field of telecommunications. He has been the Dean of the Faculty of Telecommunications, TUS and the Vice Chairman of the General Assembly of the European Telecommunications Standards Institute. He is currently the Head of the "Teleinfrastructure" Research and Development Laboratory of TUS; the Head of the "Intelligent Communication Infrastructures" Laboratory at the Research and Development and Innovation Consortium, Sofia Tech Park; and the Chairman of the "Cluster for Digital Transformation and Innovation," Bulgaria. He is the author of many scientific publications, tutoring B.Sc., M.Sc., and Ph.D. courses in the field of information transmission theory and wireless access networks.



PAVLOS I. LAZARIDIS (Senior Member, IEEE) received the M.Eng. degree in electrical engineering from the Aristotle University of Thessaloniki, Greece, in 1990, the M.Sc. degree in electronics from Université Pierre and Marie Curie (Paris 6), Paris, France, in 1992, and the Ph.D. degree from the École Nationale Supérieure des Télécommunications (ENST) Paris and Université Paris 6 in 1996. From 1991 to 1996, he was involved in research with France Télécom, and teaching with ENST Paris. In 1997, he became the Head of the Antennas and Propagation Laboratory, Télédiffusion de France/the France Télécom Research Center (TDF– C2R Metz). From 1998 to 2002, he was a Senior Examiner with the European Patent Office, The Hague, The Netherlands. From 2002 to 2014, he was involved in teaching and research with the ATEI of Thessaloniki, Greece, and Brunel University, London, U.K. He is currently a Professor of Electronics and Telecommunications with the University of Huddersfield, U.K. He has been involved in several international research projects, such as EU Horizon 2020 MOTOR5G and RECOMBINE and NATO-SfP ORCA, and he has published over 200 research articles and several national and European patents. He is serving as an Associate Editor for IEEE ACCESS. He is a member of the IET, a Senior Member of URSI, and a Fellow of the Higher Education Academy.



ZAHARIAS D. ZAHARIS (Senior Member, IEEE) received the B.Sc. degree in physics, the M.Sc. degree in electronics, the Ph.D. degree in antennas and propagation modeling for mobile communications, and the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1987, 1994, 2000, and 2011, respectively. From 2002 to 2013, he was with the Administration of the Telecommunications Network, Aristotle University of Thessaloniki, where he has been with the School of Electrical and Computer Engineering since 2013. He has been involved in several European research projects, such as Horizon 2020 MOTOR5G, Horizon 2020 RECOMBINE, and Horizon Europe 6G-ICARUS. He is the author of 82 scientific journal articles, 109 international conference papers, one national patent, five book chapters, and one textbook. His current research interests include design and optimization of antennas and microwave circuits, signal processing on smart antennas, development of evolutionary optimization algorithms, and neural networks. He is currently serving as an Associate Editor for IEEE ACCESS. He is a member of the Technical Chamber of Greece. Recently, he was elected the Chair of the Electron Devices/Microwave Theory and Techniques/Antennas and Propagation Joint Chapter of the IEEE Greece Section.