Review of Computer Engineering Research
2026 Vol. 13, No. 3, pp. 212-226
ISSN(e): 2410-9142
ISSN(p): 2412-4281
DOI: 10.18488/76.v13i3.5149

QoE-driven adaptive bitrate selection for cloud–edge 5G video streaming using Bi-LSTM and XGBoost

Kumbha Ravi Kumar[1+]
Ayyagari Srinagesh[2]

[1]*Department of Computer Engineering, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, Andhra Pradesh, Pin-522510, India.*
[1]*Email: khumbaravikumar@gmail.com*
[2]*Department of Computer Science and Business Systems, RVR and JC College of Engineering, Chowdavaram, Guntur, Andhra Pradesh, 522019, India.*
[2]*Email: srinagesh@rvrjc.ac.in*

(+ Corresponding author)

Article History
Received: 6 February 2026
Revised: 25 May 2026
Accepted: 8 June 2026
Published: 4 September 2026

Keywords
5G networks
Adaptive bitrate selection
Bi-LSTM
Cloud–edge computing
Quality of experience
Video streaming
Whale optimization algorithm
XGBoost.

ABSTRACT

The video over 5G networks is accelerating its rapid growth. Adaptive bit-rate selection mechanism intelligence is introduced to achieve the optimal user Quality of Experience (QoE) in dynamic conditions of the network. This paper presents a quick and intelligent QoE-based adaptive bit rate selection system for 5G video streaming with the help of cloud-edge that incorporates a hybrid approach of Machine Learning and Deep Learning. The study uses a Whale Optimization algorithm (WOA) to optimally fine-tune both Bi-LSTM and XGBoost hyperparameters to improve the model's performance, and our augmented hybrid system demonstrated much better performance in predicting the accuracy of QoE, accuracy in selecting the required bitrate, decreased bitrate switching, and less computational latency than the traditional ML models and isolated deep learning models. Experimental findings prove that the proposed method is always better than the baseline methods in terms of balanced accuracy, macro-F1 score, and QoE-relevant Top-2 accuracy, with more than 89% accuracy in Top-2 bitrate selection and low inference latency, which can be deployed on the edge. The results reveal that the developed cloud edge hybrid system can be used in real-time, QoE-aware adaptive video streaming in 5G-based networks. Our experiment has demonstrated that the system described in this paper can be deployed on the cloud-edge-based 5 G video streaming systems in real-time.

Contribution/Originality: In this study, a novel adaptive bitrate framework based on cloud–edge QoE is proposed, which combines temporal prediction by Bi-LSTM and decision-making by XGBoost to guarantee the low latency of the system and further optimizes the two algorithms together based on the Whale Optimization Algorithm. It is different from the previous works because it integrates predictive learning, hybrid modeling, and meta-heuristic tuning of a real multi-operator 5G data set for real-time deployment.

1. INTRODUCTION

The dynamic expansion of video streaming has recently emerged as one of the primary sources of information traffic in the fifth-generation (5G) wireless networks, which puts a new burden on the intelligence of networks and the management of resources. High data rates are not enough, and the quality should be constant in highly dynamic network conditions brought about by user mobility, variable channel quality, and network congestion, which is needed by applications like ultra-high-definition video, live streaming, and immersive multimedia services. Adaptive bitrate (ABR) streaming has become an important method to tackle these issues by reacting dynamically to fluctuations in the network to dynamically change video quality; nevertheless, traditional ABR strategies tend to poorly capture the high-frequency network changes in 5G networks, causing frequent changes in bitrate, rebuffering, and poor Quality of Experience (QoE). The recent introduction of cloud-edge computing, as well as data-driven intelligence, leads to an increased desire to use fast, predictive, and QoE-aware bitrate selection frameworks with the ability to run in real time and utilize all capabilities of 5G networks to their full extent. Mobile video traffic is now an overwhelming part of data transmission in 5G and beyond-5G (B5G) networks as video streaming services have proliferated intensely. Mobility-intensive applications, high-resolution video, and low-latency live streaming are services with high requirements for the network throughput, latency, and reliability. Adaptive bitrate (ABR) streaming has been widely used in order to deal with the changing wireless conditions, allowing the dynamic adaptation of the quality of the video according to the available network resources. More recent works have underlined the importance of edge computing in the process of improving adaptive streaming through proximity to end-users and as such, minimizing latency and maximizing responsiveness. Additionally, cloud-edge architecture integration has become one of the promising models in addressing the computational complexity and scalability requirements of intelligent video streaming systems.

1.1. Problem Statement and Motivation

Nevertheless, these developments notwithstanding, the need to have a good Quality of Experience (QoE) among mobile users is a thorny issue in 5G settings. The presence of rapid channel changes, non-uniform mobility patterns, and non-uniform traffic profiles is likely to cause inaccurate throughput estimation, frequent switching of bitrates, rebuffering, and poor video quality. Current adaptive streaming techniques are often based on either lightweight edge-based heuristics or cloud-centric learning, which find it difficult to trade off the accuracy of prediction with latency in computation and real-time adaptive capabilities in highly dynamic network scenarios. As a result, there is an urgent demand for rapid, precise and QoE-oriented bitrate selection mechanisms capable of running effectively in cloud-edge-enabled 5G systems. Moreover, the hyperparameter selection is commonly done manually or by grid-based techniques, which result in poor performance and higher cost of computation. It is this drawback that drives the creation of a hybrid, predictive, and optimally tuned learning system that will provide fast and reliable decisions on bitrate selection with a better QoE.

The video traffic has become the largest consumer of data in the fifth-generation (5G) and beyond-5G wireless networks due to the rapid growth of mobile video streaming services. Video delivery in high-definition, as well as ultra-high-definition, and the mobility of the users, along with the heterogeneous network access conditions, pose a high demand in terms of network flexibility and service quality. The adaptive bitrate (ABR) streaming has thus emerged as one of the major mechanisms of ensuring an acceptable Quality of Experience (QoE) by being capable of changing the video quality in response to network conditions. It has been recently shown that edge-based adaptation can substantially improve responsiveness and decrease the latency in comparison with conventional cloud-only deployments; thus, edge intelligence becomes the key element of the video streaming system of the next generation [1]. Cloud-edge collaborative architectures have been proposed to enhance adaptive video streaming in 5G and B5G networks to further enhance scalability and the ability of learning. By spreading learning and decision-making workloads between the cloud and the edge, these structures can better take advantage of global wisdom at the same time as they allow rapid local adaptation. Cloud-edge learning models have been demonstrated to be effective in managing heterogeneous traffic and mobility-based dynamics in large-scale Internet of Things and mobile streaming applications [2]. Nevertheless, their performance depends on network state estimation and the time of bitrate decisions a lot. The prediction of accurate throughput and the network condition is still a primary issue in the edge-assisted ABR system. Radio condition differences, user velocity, and handover have the potential to cause severe prediction error, which causes common bitrate swings and reduced QoE. Strong throughput estimation schemes have also been suggested to address these challenges, though typically based on small feature sets or reactive adaptation schemes that do not work in highly dynamic 5G situations [3]. Consequently, there is a growing need to have predictive intelligence capable of tapping into temporal dependencies of network behavior. Although there have been huge achievements in the adaptive bitrate (ABR) streaming of 5G networks, there are still a number of gaps. To begin with, the current methods are mainly based on cloud-centric deep learning models or lightweight edge-based heuristics, where a hybrid approach based on predictive cloud-edge ABR, which is also capable of integrating temporal network-state prediction and low-latency edge decision-making, has not been studied in depth so far. Second, it is common that most ABR solutions based on learning are highly predictive but not well balanced between computational complexity and real-time deployability, which restricts their deployment in the dynamic 5G cases of mobility and fast channel dynamics. Third, hyperparameter choice in the majority of previous works is conducted either by manual tuning or grid-search with no systematic metaheuristic optimization methods that can enhance robustness and generalization when conditions of class imbalance and heterogeneous networks are involved. They restrict the design of a predictive, computationally efficient, and maximally tuned cloud-edge system to select bitrates based on QoE.

Simultaneously, other surveys and studies have stressed the need for QoE-based design in cloud-edge adaptive streaming systems. Although significant advances have already been achieved, a number of unresolved problems, such as achieving a balance between prediction accuracy and computational complexity, minimizing adaptation latency, and making them deployable in real-time, have been identified. These problems indicate the necessity of light but smart models, which can work effectively on the edge and are trained and optimized in the cloud [4]. Adaptive streaming solutions based on learning, especially with deep reinforcement learning, have been investigated as a solution to the limitations of heuristic ABR algorithms. Even though these methods are able to reach better QoE in the controlled setting, they are commonly characterized by high training complexity, sluggish convergence, and sensitivity with respect to hyperparameter values. All these disadvantages restrict their viability in mobile networks in real-time and large-scale applications [5]. Furthermore, most of the learning-oriented approaches are centralized optimization, and this approach might not fully leverage the advantages of edge computing. Recently, mobile edge computing (MEC)-based bitrate adaptation and bandwidth allocation schemes have been suggested to enhance resource utilization and minimize latency. Although the methods contribute to responsiveness in adapting, they are usually based on simplified decision models and lack predictive ability, which is susceptible to abrupt fluctuations in the network [6]. Also, multiple users and services coordination is a problematic area when it comes to dense network deployment. Adaptive streaming models that are decentralized and edge-cache-based have been explored to overcome the issue of scalability and multicast efficiency. These techniques show better network usage and fairness, though their optimization mechanisms tend to focus on network-level goals rather than end-user QoE and no longer necessarily reflect end user dynamics [7]. Moreover, the majority of the current decentralized solutions are not based on complex predictive learning models.

The first edge-assisted adaptive streaming models have provided the basis for the inclusion of edge computing into video delivery systems. Although these schemes can effectively decrease latency and increase adaptation speed, they mostly use rule-based or shallow learning methods, which restrain their capacity to model complex temporal relationships and maximize the QoE under 5G changing conditions [8]. Taken as a whole, these constraints drive the creation of a fast, predictive, and QoE-centered adaptive bitrate selection framework that takes advantage of hybrid deep learning, machine learning, and metaheuristic optimization on a cloud-edge architecture. The central purposes of the given research include creating a predictive, QoE-conscious bitrate-adjustment system that can be deployed in real-time at the edge of the cloud and testing its efficiency on the basis of real 5G network measurements. The most important contributions of the work are the following:

- Proposed hybrid Bi-LSTM and XGBoost adaptive bitrate selection framework of 5G video streaming.
- Combination of a metaheuristic WOA-based hyperparameter optimization strategy to enhance the performance of learning and computational efficiency.
- Thorough analysis of a real 5G multi-operator dataset in a variety of mobility and traffic conditions.

- Evidence of large enhancements in the accuracy of prediction of QoE, bitrate stability and latency relative to traditional machine learning methods and methods based on standalone deep learning.

In this paper, it is suggested to overcome the above challenges by proposing a fast QoE-based adaptive bitrate selection framework to support 5G video streaming over the cloud-edge. The suggested solution is to integrate the predictive learning of future network states with a bidirectional long short-term memory (Bi-LSTM) network with an Extreme Gradient Boosting (XGBoost) model to be deployed at the edge to effectively make a bitrate decision. In order to improve the accuracy of prediction and the training efficiency, a Whale Optimization Algorithm (WOA) is used to optimally adjust the hyperparameters of the two learning models. Decades of experiments performed on real-world 5G multi-operator datasets of various mobility and traffic conditions show that the proposed framework can greatly enhance the QoE, bitrate switching, and low computational latency relative to the current adaptive streaming schemes. The research paper is relevant to the literature, considering that it suggests a hybrid cloud-edge QoE-based adaptive bitrate system that combines temporal deep learning and gradient boosting. The new joint Bi-LSTM-XGBoost estimation procedure optimized through the Whale Optimization Algorithm is used in this study. The current study can be considered one of the few studies that have studied the predictive bitrate selection using real multi- operator 5G datasets. The article brings the pioneering logical examination of the integration of the temporal state prediction, QoE-based categorization, and metaheuristic hyperparameter enhancement within a single cloudedge platform.

2. LITERATURE SURVEY

The last several years have seen intensive research on the improvement of adaptive bitrate (ABR) video streaming over mobile and 5G networks with a specific focus on edge computing and data-driven intelligence. The initial research was devoted to adaptation responsivity improvement through shifting bitrate decision logic to the network edge, which minimized the latency and maximized perceived quality. As cloud-edge architectures have evolved, researchers have become increasingly interested in studies of collaborative learning models that can take advantage of global knowledge in the cloud as well as the awareness of context at the edge. Simultaneously, machine learning and deep learning methods have been extensively implemented to capture the dynamics of complex networks, forecast the changes in throughput, and maximize Quality of Experience (QoE) in conditions of heterogeneous mobility and traffic patterns. Although these methods have proved significant progress over the conventional heuristic-based ABR algorithms, they also expose some drawbacks associated with prediction accuracy, computation complexity, and real-time deployability, which further encourage research and development of more efficient and robust adaptive streaming algorithms.

Edge computing has been commonly perceived as a successful enabler in enhancing the adaptive video streaming performance through the minimization of latency and the increase of content availability. The initial steps toward this direction examined bitrate adaptation techniques in edge caching systems, in which video material is stored nearer to customers to address the issue of backhaul congestion. Tran et al. [9] designed a bitrate adaptation scheme that is optimized to edge caching conditions, and this scheme showed better stability of streaming and less rebuffering. Although useful in exploiting cached material, these solutions are based mainly on response adaptation using current network conditions and fail to include predictive insight into upcoming network dynamics. The QoE-centric design has become more and more popular as scholars aim to emphasize the concept of throughput-free adaptation to user-perceived quality optimization. Strategies to do cross-layer bitrate allocation have been studied to explicitly consider application-level video parameters and the lower-layer network measures. In Yeznabad et al. [10], a cross-layer bitrate distribution framework using QoE was proposed in the context of adaptive streaming using MEC, and the user satisfaction gains were significant. Nonetheless, the scheme proposed has close coordination between protocol layers and might come at an extra signaling cost, which restricts scale in a dense 5G implementation. Equally, Yu et al. [11] introduced a QoE-based adaptive streaming scheme to edge-assisted cellular networks, making it a priority on fast adaptation and the enhancement of QoE, but the model is reliant primarily on short-term network measurements and offers no long-horizon prediction. As the field of machine learning methods has developed, a number of papers have examined the issue of learning-based adaptive bitrate selection to better represent complex network behavior. The recent development is the emergence of meta-learning as an exciting paradigm of enhancing generalization in heterogeneous environments. MetaABR was a meta-learning-based framework suggested by Li et al. [12] to adapt fast to new conditions of the network using little training data. Even though it is an effective method to enhance adaptability, meta-learning methods typically add more training complexity and can be limited to real-time deployment because of the overhead associated with model initialisation and adaptation.

Bandwidth prediction and adaptive streaming have also been done using recurrent neural networks. In Huu et al. [13], a GRU-based bandwidth forecast algorithm was integrated with QoE-based bitrate control of heterogeneous mobile networks, which obtained better prediction and smoother playout. Although such advantages have been made, single deep learning models can be prone to overfitting and sensitivity to hyperparameter choice, particularly in the context of high-mobility 5G conditions. Moreover, deep networks may be training-intensive at scale, and edge deployment of lightweights is difficult. There are also several studies that are dedicated to the optimization of bitrate selection based on explicit QoE metrics and user bandwidth properties. Naik and Madhavi [14] suggested a QoE-based scheme of video delivery that optimizes the video bitrate with the help of network and user bandwidth information. Although useful in educational and content delivery systems, these techniques normally make certain that there is relative stability in the user behavior, and do not explicitly consider the high-frequency wireless channel fluctuations that may occur in mobile situations. Recently, federated learning has been studied to achieve collaborative model training without a centralized collection of data. Wong et al. [15] explored effective client selection mechanisms in asynchronous federated learning in adaptive bitrate streaming with emphasis on scalability and privacy conservancy. Federated approaches can, however, bring communication overhead and convergence delays, especially when dealing with highly dynamic networks. In addition to the classical video streaming, other applications in which the concept of QoE optimization is examined include extended reality (XR) video transmission. Pan et al. [16] investigated the QoE optimization of real-time XR video with energy constraints, proving the significance of the combination of QoE and system-level constraints. As insightful as they are, these works are application-specific, and it is unclear how they can be directly applied to general adaptive bitrate streaming frameworks within cloud-edge 5G systems. To conclude, current studies prove that there has been great advancement in edge-assisted, QoE-based, and learning-based adaptive video streaming. However, the majority of methods are based on the use of either single learning paradigms or reactive adaptation mechanics, and little effort has been put into hybrid predictive models with systematic hyperparameter optimization. These remarks inspire the creation of a cohesive cloud-edge system that integrates temporal deep learning, effective machine learning classifier, and metaheuristic optimization to attain quick, precise, and QoE-based bitrate alteration in 5G practical circumstances.

3. PROPOSED METHODOLOGY

3.1. System Architecture and Workflow Overview

The suggested framework uses a cloud-edge collaborative structure to facilitate quick and QoE-oriented adaptive bitrate choice to 5G video streaming. The general flow of work is shown in Figure 1. Centralized learning, historical data aggregation, and model training are done by the cloud layer, and real-time inference and bitrate decision-making are done by the edge layer. At the cloud interface scale, extensive historical data of 5G networks, user mobility, and video traffic are gathered and preprocessed. A Bi-LSTM network is conducted using a Bidirectional Long Short-Term Memory network that the network learns temporal relationship in the historical observations to forecast future network conditions. The model parameters are trained and updated periodically and relayed to edge nodes. On the edge, instantiated streaming context and predicted network states output by the Bi-LSTM model are added and input to an Extreme Gradient Boosting (XGBoost) classifier at the edge. The XGBoost model is used to choose the best video bitrate level that optimizes user QoE and reduces bitrate switching and rebuffering. A Whale Optimization Algorithm (WOA) is used to optimize the hyperparameters of the Bi-LSTM and XGBoost models to ensure that they perform better and require less time to train.

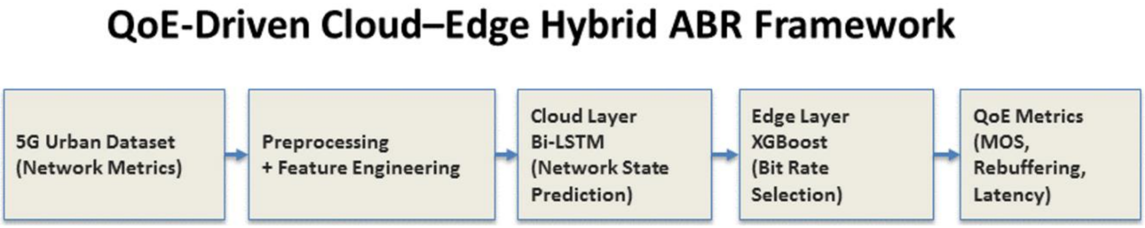

Figure 1. Workflow of the proposed QoE-driven cloud–edge adaptive bitrate selection framework.

In this research, the experiments are done on the Urban Multi-Operator QoE-Aware Dataset, which is available on Mendeley Data (DOI: 10.17632/dx5xyyfz2y.1). The dataset comprises real-world cellular network measurements of various mobile network operators in a dense urban area, and as such is representative of real-life conditions of 4G/5G video streaming. This information was collected using massive field data collection during walking and driving mobility conditions, which represented realistic temporal local variations in network performance. Readings are captured as time-series traces when the cells are active, and they also incorporate network and radio indicators that have a direct bearing on video streaming Quality of Experience (QoE). These indicators are categorized as throughput-related (e.g., downlink and uplink bitrate), radio signal quality indicators (e.g., signal strength and quality), and mobility-related. As the dataset does not give explicit application-layer bitrate decisions or subjective QOE scores, a label of the bitrate classes is received based on downlink throughput characteristics. The learning task is defined as the next-step prediction of bitrate class based on a sliding time window, which is a simulated decision-making process of adaptive bitrate (ABR) streaming systems where network dynamics are characterized by changing conditions. This proxy labeling method is commonly used in QoE-aware streaming studies and allows the significant assessment of the bitrate adaptation methods. The dataset is inherently class imbalanced at the levels of bitrates, which is an accurate cellular usage pattern. Instead of artificially balancing the data, the original distribution is maintained, and the imbalance-sensitive evaluation measures, such as balanced accuracy and F1-score, measured macro-averaged are used to provide an equal and realistic evaluation of the performance [17]. Table 1 shows the dataset details.

Table 1. Dataset details.

Attribute	Value
Dataset source	Mendeley data
Network technologies	LTE / 5G
Number of operators	Multiple (≥2)
Total raw samples	≈ 18,000
Total samples used (after preprocessing)	≈ 18,000
Mobility scenarios	Walking, driving
Temporal nature	Time-series
Sliding window length	10 samples
Prediction target	Next-step bitrate class
Number of bitrate classes	3 (Low / Medium / High)
Input feature categories	Throughput, radio quality, mobility
Total input features (after encoding)	≈ 95
Train / Validation / Test Split	70% / 15% / 15% (session-aware)

3.2. Mathematical Analysis

To formally characterize the behavior of the proposed cloud--edge QoE-driven adaptive bitrate framework, this subsection presents a mathematical analysis of the bitrate decision process and the associated Quality of Experience

(QoE) optimization objective. Let $\mathcal{B} = \{b^{(1)}, b^{(2)}, \ldots, b^{(L)}\}$ denote the discrete set of available bitrate levels, and let $\mathrm{x}_t \in \mathbb{R}^d$ represent the observed multivariate 5G feature vector at time index t, constructed from radio signal measurements, transport-layer statistics, user mobility information, and streaming context parameters. Using a sliding observation window $\mathrm{X}_t = \{\mathrm{x}_{t-T+1}, \mathrm{x}_{t-T+2}, \ldots, \mathrm{x}_t\}$ of length T, the Bi-LSTM predictor estimates the future network state $\hat{\mathrm{s}}_{t+1}$, which is subsequently utilized by the edge-side decision model to select the streaming bitrate $b_t \in \mathcal{B}$. The following analysis formulates the QoE objective and examines the trade-offs among delivered video quality, rebuffering penalties, and bitrate switching costs, providing theoretical insight into the stability and latency characteristics of the proposed framework for real-time 5G video streaming [18].

3.2.1. Problem Formulation and Notation

Let $\mathcal{B} = \{b^{(1)}, b^{(2)}, \ldots, b^{(L)}\}$ denote the discrete set of available bitrate levels (Classes).

At each segment time index t, the objective is to choose $b_t \in \mathcal{B}$ that maximizes user QoE while minimizing stalling and excessive bitrate switching.

A multivariate feature vector is observed at time t.

$$\mathrm{x}_t = [\mathrm{RSRP}_t, \mathrm{RSRQ}_t, \mathrm{SINR}_t, \hat{\tau}_t, \ell_t, p_t, v_t, B_t, \ldots]^\top \in \mathbb{R}^d \quad (1)$$

Where RSRP_t, RSRQ_t, and SINR_t represent 5G radio signal metrics, $\hat{\tau}_t$ is measured throughput, ℓ_t is end-to-end latency, p_t is packet loss, v_t is user mobility speed (walking/driving), and B_t is buffer occupancy. The dimension d depends on the selected monitoring features.

A sliding window of length T constructs the model input sequence:

$$\mathrm{X}_t = \{\mathrm{x}_{t-T+1}, \mathrm{x}_{t-T+2}, \ldots, \mathrm{x}_t\} \quad (2)$$

Using X_t captures temporal dependencies and mobility-induced variations that cannot be represented by a single instantaneous measurement.

3.2.2. Bi-LSTM Network for Predictive Network State Estimation

The cloud-side Bi-LSTM predictor learns temporal dynamics in 5G conditions to estimate the future network state. Given $\mathbf{X}_t$, the model predicts $\hat{\mathbf{s}}_{t+1}$, where $\mathbf{s}_{t+1}$ may represent one-step-ahead throughput category, congestion risk, or a compact state embedding used for bitrate decision [19].

For each time step, the LSTM gates are defined as:

$$\mathrm{f}_t = \sigma(\mathrm{W}_f \mathrm{x}_t + \mathrm{U}_f \mathrm{h}_{t-1} + \mathrm{b}_f) \quad (3)$$

$$\mathrm{i}_t = \sigma(\mathrm{W}_i \mathrm{x}_t + \mathrm{U}_i \mathrm{h}_{t-1} + \mathrm{b}_i) \quad (4)$$

$$\mathrm{o}_t = \sigma(\mathrm{W}_o \mathrm{x}_t + \mathrm{U}_o \mathrm{h}_{t-1} + \mathrm{b}_o) \quad (5)$$

$$\tilde{\mathrm{c}}_t = \tanh(\mathrm{W}_c \mathrm{x}_t + \mathrm{U}_c \mathrm{h}_{t-1} + \mathrm{b}_c) \quad (6)$$

$$\mathrm{c}_t = \mathrm{f}_t \odot \mathrm{c}_{t-1} + \mathrm{i}_t \odot \tilde{\mathrm{c}}_t \quad (7)$$

$$\mathrm{h}_t = \mathrm{o}_t \odot \tanh(\mathrm{c}_t) \quad (8)$$

Where $\sigma(\cdot)$ is the sigmoid function, $\odot$ denotes element-wise multiplication, $\mathrm{W}_{(\cdot)}$ and $\mathrm{U}_{(\cdot)}$ are learnable weight matrices, and $\mathrm{b}_{(\cdot)}$ are biases.

A Bi-LSTM uses both forward and backward hidden sequences.

$$\mathrm{h}_t^{\mathrm{Bi}} = [\mathrm{h}_t^{\rightarrow}; \mathrm{h}_t^{\leftarrow}] \quad (9)$$

and predicts the next network state as:

$$\hat{\mathbf{s}}_{t+1} = \phi(\mathrm{W}_p \mathrm{h}_t^{\mathrm{Bi}} + \mathbf{b}_p) \quad (10)$$

Where $\phi(\cdot)$ is a linear or nonlinear mapping depending on whether $\mathbf{s}$ is continuous or categorical.

Parameter explanation:

$\mathrm{h}_t \in \mathbb{R}^H$ is the hidden state with H units, and $\mathrm{c}_t \in \mathbb{R}^H$ is the cell state.

T is the sequence length, and d is the feature dimension.

Bi-LSTM is selected because it captures richer temporal context than a unidirectional model, improving robustness under mobility and abrupt fluctuations typical of 5G.

3.2.3. Edge-Side XGBoost for QoE-Driven Bitrate Decision

At the edge, the bitrate selection is formulated as a multi-class classification over $\mathcal{B}$ using predicted state $\hat{\mathrm{s}}_{t+1}$ (Optionally concatenated with instantaneous context).

The selected bitrate is:

$$b_t = \arg\max_{b \in \mathcal{B}} P(b \mid \hat{\mathbf{s}}_{t+1}) \qquad (11)$$

XGBoost learns an additive ensemble of decision trees:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(\mathbf{z}_i), f_k \in \mathcal{F} \qquad (12)$$

Where $\mathbf{z}_i$ is the model input (e.g., $\hat{\mathbf{s}}$ plus context), K is the number of trees, and $\mathcal{F}$ is the space of regression trees.

The objective is:

$$\mathcal{L} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \qquad (13)$$

With regularization:

$$\Omega(f_k) = \gamma T_k + \frac{1}{2}\lambda \|\mathbf{w}_k\|^2 \qquad (14)$$

Where T_k is the number of leaves in tree k, $\mathbf{w}_k$ are leaf weights, and γ, λ control model complexity.

A QoE metric is used to guide learning and evaluation.

$$\mathrm{QoE}_t = \alpha\, q(b_t) - \beta\, R_t - \delta\, |b_t - b_{t-1}| \qquad (15)$$

Where $q(b_t)$ is a quality score associated with bitrate b_t, R_t is rebuffering duration (or stall penalty), and $|b_t - b_{t-1}|$ penalizes switching.

$\alpha, \beta, \delta > 0$ weight video quality, rebuffering penalty, and switching penalty, respectively.

N is the number of training samples for XGBoost, and K is the number of trees.

XGBoost is selected for edge deployment due to low inference latency, strong performance on tabular features, and robustness with limited compute, enabling fast bitrate decisions.

3.2.4. WOA-Based Hyperparameter Optimization

To avoid suboptimal manual tuning and to improve training efficiency, WOA optimizes the hyperparameters of both Bi-LSTM and XGBoost.

Let the combined hyperparameter vector be:

$$\Theta = [H, \eta_{\mathrm{LSTM}}, \mathrm{Dropout}, \mathrm{Batch}, K, \mathrm{Depth}, \eta_{\mathrm{XGB}}, \ldots] \qquad (16)$$

Where H is the number of Bi-LSTM hidden units, η_{LSTM} and η_{XGB} are learning rates, and the remaining terms are model-specific hyperparameters.

WOA updates candidate solutions using encircling behaviour.

$$\mathrm{D} = |\mathrm{C} \cdot \mathrm{X}^*(t) - \mathrm{X}(t)| \qquad (17)$$

$$\mathrm{X}(t+1) = \mathrm{X}^*(t) - \mathrm{A} \cdot \mathrm{D} \qquad (18)$$

Where $\mathbf{X}^*(t)$ is the current best solution and $\mathbf{X}(t)$ is the current whale position.

The coefficient vectors are:

$$\mathrm{A} = 2a\mathrm{r} - a \qquad (19)$$

$$\mathrm{C} = 2\mathrm{r} \qquad (20)$$

With $\mathbf{r} \sim U(0,1)$ and a decreasing linearly from 2 to 0 over iterations.

The fitness function used for optimization is:

$$\mathcal{F}(\Theta) = \eta_1 \overline{\mathrm{QoE}} - \eta_2 \overline{\mathrm{Latency}} - \eta_3 \overline{\mathrm{Switch}} \qquad (21)$$

and the best hyperparameters are selected as:

$$\Theta^* = \arg\max_{\Theta} \mathcal{F}(\Theta) \quad (22)$$

$\eta_1, \eta_2, \eta_3 > 0$ weight QoE gain against latency and switching. Overlines denote averages over the validation set.

The reason behind the use of WOA is that it is a gradient-free algorithm, can use mixed discrete/continuous hyperparameters, requires less effort to use than manual tuning and is more likely to converge to a high-performing configuration. The mathematical formulation used in the proposed research will give a unified analytical perspective of the proposed cloud-edge adaptive bitrate framework [20-23]. The analysis expressly identifies the trade-offs between the quality of delivered video, rebuffering avoidance and the stability of the bitrate through integrating predictive network state estimation, alongside QoE-guided bitrate selection. Metaheuristic hyperparameter optimization is another method that can be used to increase the robustness of models and has a low computational cost. The provided analytical understandings create a theoretical foundation to the performance improvement exhibited in the experimental outcomes and prove the practicability of the suggested solution to be implemented in real-time in the dynamical 5G video streaming setup [24, 25]. The general training, optimization, and edge-side bitrate choice process of the proposed framework is described in Algorithm 1.

Algorithm 1: Proposed QoE-Driven ABR Selection Using Cloud–Edge Bi-LSTM + XGBoost with WOA
Input:
Historical 5G dataset $\boldsymbol{\mathcal{D}}$ (radio, transport, mobility, and streaming features),
Output:
Selected bitrate b_t for each video segment at the edge.

1. Collect and preprocess the dataset $\boldsymbol{\mathcal{D}}$ by removing missing values and outliers and normalizing all features.
2. Construct time-series input sequences X_t = {x_{t−T+1}, x_{t−T+2}, …, x_t} for all training samples.
3. Initialize the Bi-LSTM model f_θ for future network-state prediction and the XGBoost model g_ψ for bitrate classification.
4. Initialize the Whale Optimization Algorithm (WOA) population with M candidate hyperparameter vectors Θ^(m) = (θ^(m), ψ^(m)).
5. Evaluate the initial population on the validation dataset and select the best candidate Θ*.
6. For iteration i = 1 to I:
   a. Train Bi-LSTM model f_{θ^(m)}.
   b. Predict future network states ŝ_{t+1}.
   c. Train XGBoost model g_{ψ^(m)} using predicted states.
   d. Compute validation fitness:
      Fitness = η_1 × Avg(QoE) − η_2 × Avg(Latency) − η_3 × Avg(BitrateSwitch).
   e. Update Θ* if fitness improves.
   f. Update whale positions using WOA update rules.
7. Retrain final Bi-LSTM f_{θ*} and XGBoost g_{ψ*} models.
8. Deploy trained models to edge nodes.
9. For each video segment time t:
   a. Collect recent feature sequence X_t.
   b. Predict next network state ŝ_{t+1} = f_{θ*}(X_t).
   c. Compute bitrate class probabilities p = g_{ψ*}(ŝ_{t+1}).
   d. Select bitrate b_t = argmax_{b∈$\boldsymbol{\mathcal{B}}$} p(b).
   e. Stream video segment at bitrate b_t.
   f. Log playback KPIs for periodic cloud updates.
10. End For

4. EXPERIMENTAL RESULTS AND DISCUSSION

All experiments are conducted on Urban Multi-Operator QoE-Aware dataset that offers real-world time-series of cellular measurements of many operators in a dense urban area. The features of the input will be throughput-related indicators (e.g., downlink/uplink bitrate) and radio/network quality indicators (e.g., signal strength/quality and mobility-related variables), which make the proposed model learn dynamic network behavior applicable to adaptive video streaming. In order to have a fair and leakage-free assessment, data are split based on a session-aware approach such that samples of a session are not used in both training and test sets. A sliding window made up of temporary samples is used to create the temporal samples and the learning goal is to predict the following bitrate class which is representative of realistic ABR decision-making in time-varying network conditions. The hybrid cloud-edge architecture presented uses a Bi-LSTM network at the cloud to take into account all temporal dependencies and produce predicted network-state representations. Such estimated representations are then depicted to an XGBoost classifier that is edge-resident to choose low-latency bitrate representations. They consist of standard machine learning tools (Random Forest and XGBoost on tabular data), and end-to-end neural networks (GRU and Bi-LSTM classifiers). As the real-life bitrate classes are skewed, the evaluation is reported by the overall accuracy and the balanced accuracy, macro-F1, and weighted F1. Also, Top-2 bitrate selection accuracy is cited as being a QoE-oriented measure as choosing an adjacent bitrate level generally maintains apparent quality with less rebuffering danger.

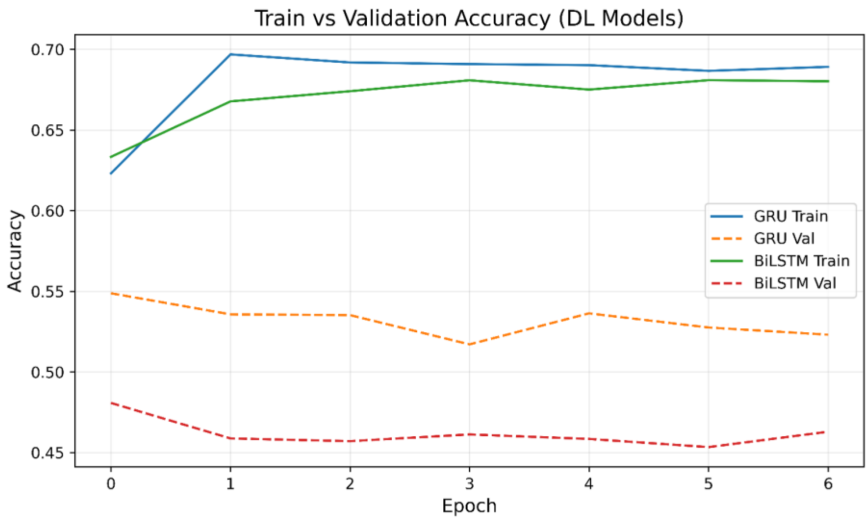

Figure 2. Training accuracy curves of all models.

Figure 2 and Figure 3 depict the curves of training and validation accuracy and loss of the GRU and the Bi-LSTM models. Both models are characterized by a stable convergence behavior with no serious overfitting, which means that temporal network patterns are learned effectively. A slight smooth convergence is evident with the Bi-LSTM model over the GRU since the former has a two-way structure, thus its ability to capture the future and past temporal aspects. These findings show that recurrent neural networks are appropriate to modeling time-varying conditions of cellular networks in adaptive video streaming.

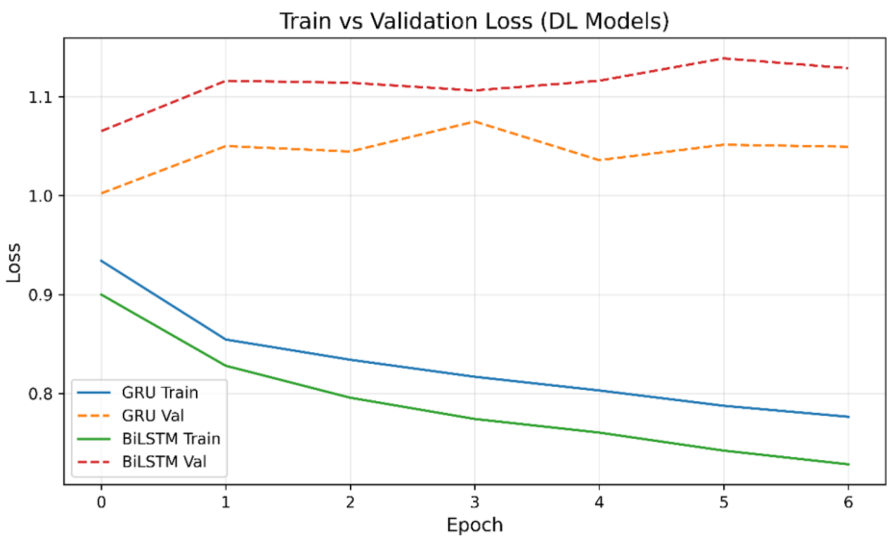

Figure 3. Training loss curves of all models.

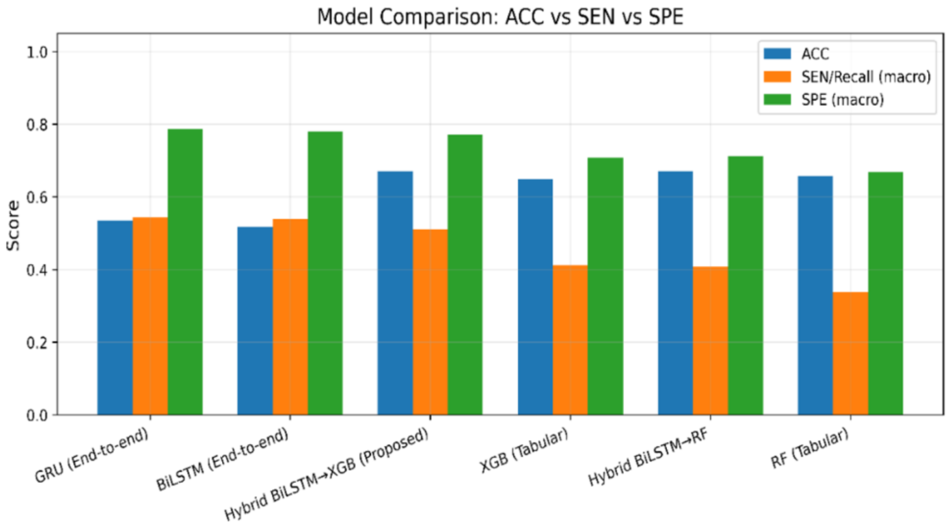

Figure 4. Model comparison curve for accuracy, sensitivity, and specificity.

The comparison of the accuracy, sensitivity (recall), and specificity as groups is displayed in Figure 4. Most models are highly specific in that they provide good avoidance of infeasible bitrate selections, but the hybrid model suggested is more sensitive and specific. This consistent acting means a better response to changing network conditions, which means the need to avoid the under-utilization of video bitrates and over-selection aggression.

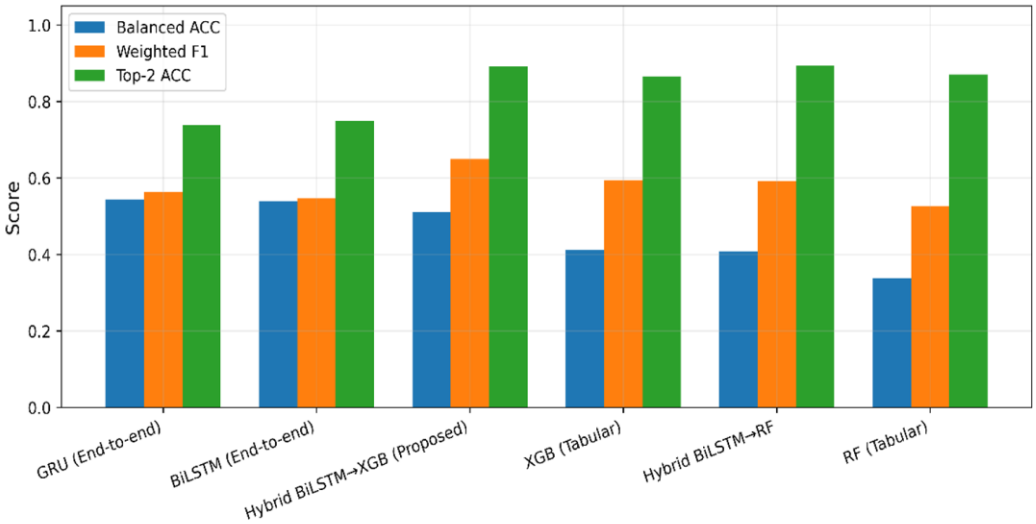

Figure 5. Grouped comparison curve for balanced accuracy, weighted F1, Top-2 ACC.

Figure 5 emphasizes the accuracy of the Top-2 bitrate selection, a QoE-based measure of the real acceptability of the adjacent bitrate selection. The hybrid framework proposed has a Top-2 accuracy of more than 89% which is better than standalone deep learning and classical machine learning baselines. This finding is especially important to adaptive video streaming because obtaining a near-optimal bitrate is likely to deliver similar perceptual QoE with less risk of rebuffering.

Table 2. Performance comparison of proposed and baseline models on the test dataset.

Model	ACC	Balanced ACC	Macro-F1	Weighted F1	Top-2 ACC
GRU (End-to-end)	0.53	0.54	0.48	0.56	0.74
Bi-LSTM (End-to-end)	0.52	0.54	0.47	0.55	0.75
RF (Tabular)	0.66	0.34	0.28	0.53	0.87
XGB (Tabular)	0.65	0.41	0.42	0.59	0.86
Hybrid Bi-LSTM→RF	0.67	0.41	0.39	0.59	0.89
Hybrid Bi-LSTM→XGB (Proposed)	0.67	0.51	0.51	0.65	0.89

Table 2 provides a summary of the performance of proposed and baseline models on the test dataset. Although both hybrid and tabular models have similar overall accuracy, accuracy is not enough because of the fact that the bitrate classes are skewed. Thus, the focusing on balanced accuracy and macro-averaged F1-score are used. The Hybrid Bi-LSTM-XGBoost model is the best in terms of balanced accuracy score and macro-F1 showing better robustness among minority classes of bitrates. Moreover, the suggested model achieves higher than 89% Top-2 accuracy which proves its capability to pick the nearly best bitrate rates that maintain the perceptual QoE when the network is dynamically changing.

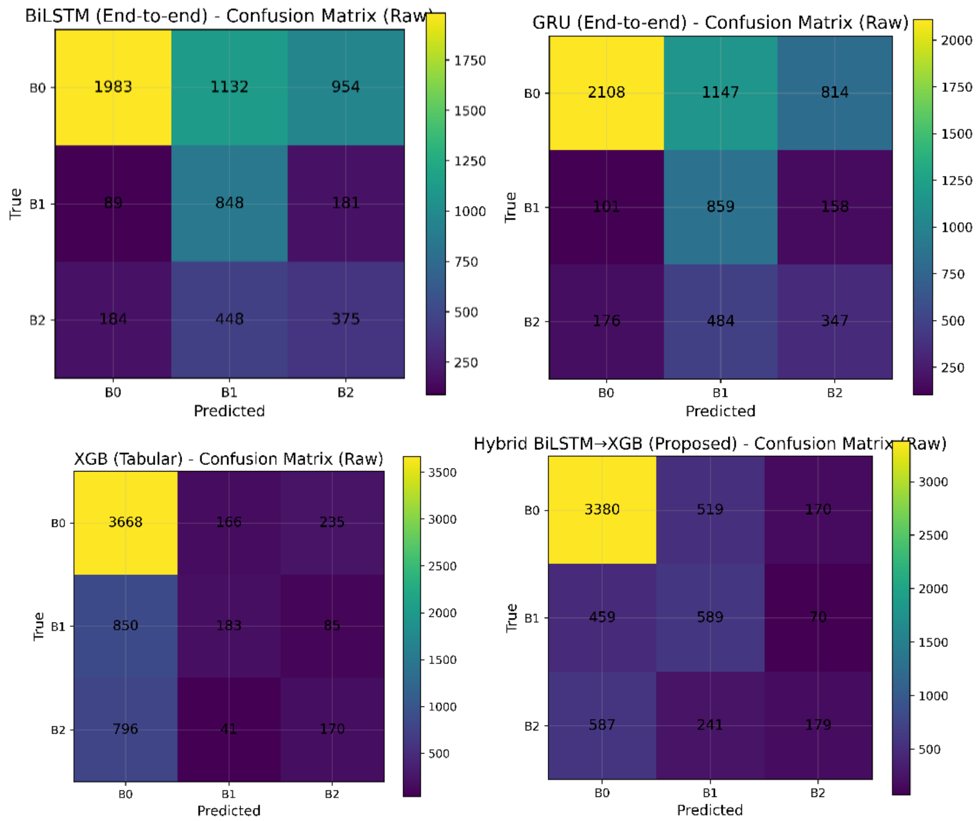

Figure 6. Confusion matrix of all models.

Figure 6 demonstrates the confusion matrix of the current and suggested models. The confusion matrices indicate that misclassification between the minority bitrate classes is reduced by the hybrid model than the baseline strategies. ROC and precision-recall curves also show a better discrimination ability, as the macro-averaged AUC values are high in comparison to tabular and standalone deep learning models. The findings validate the strength of the proposed framework in a variety of network conditions. The experimental findings prove that the discussed Hybrid Bi-LSTM-XGBoost framework is always the most successful in relation to end-to-end deep learning and machine learning-only baselines in terms of various evaluation perspectives. Although the general classification performance of the models is similar because of implying inherent imbalance in the classes of the bitrates, the best-balanced performance, weighted F1-score, and Top-2 performance demonstrated by the proposed method suggests greater robustness to the majority and minority bitrate classes.

The metric comparisons of the groups show that deep learning models by themselves can be strongly trained but experience significant generalization deterioration in case the models are tested on unknown test data. By comparison, the hybrid architecture is able to effectively combine learning of temporal features using Bi-LSTM and discriminative decision boundaries using XGBoost, leading to better test-time stability and by-class discrimination. This trend is also supported by the confusion matrix analysis, in which the proposed model greatly minimizes the misclassification between neighboring bitrate classes in comparison to those of baseline approaches. Notably, the accuracy of the proposed model reaches Top-2 of 89, indicating that in cases where the precise prediction of the bitrate is not made, the choice of a bitrate is still within a perceptually acceptable range. This is especially true of adaptive video streaming systems, where adjacent bitrate choices usually maintain user Quality of Experience (QoE) and ensure the reduction of rebuffering and bitrate gating. The loss curves of training and validation further affirm that the hybrid methodology reduces overfitting in the pure deep learning models and the convergence is not affected by epochs.

All these findings confirm the efficiency of the suggested architecture to adaptive bitrate classification in a realistic network environment and support its appropriateness to the QoE-sensitive streaming applications. This research is based on proxy bitrate labeling of class based on throughput measurements instead of end-user subjective perceptual quality based on QoE scores which maybe not sufficiently accurate in perceiving changes in quality. The analysis is done on one publicly available multi-operator dataset and real-time implementation in a live 5G network setting has not been substantiated. Also, a multi-user fairness and the large-scale edge coordination were not explicitly modeled.

5. CONCLUSION

The present paper proposed a hybrid deep learning and machine learning-based adaptive-bitrate selection framework of cloud-edge enabled 5G video streaming. The suggested architecture will combine a Bi-LSTM network in the cloud layer to predict the temporal network state and an XGBoost classifier in the edge with it's lightweight to select the bitrate within the allotted low latency.

The proposed design allows achieving good prediction accuracy and deployment efficiency by separating time models with real-time decision-making. Well-scaled experiments were based on real-life multi-operator cellular value on a large dataset acquired in an urban environment of dense density under different mobility conditions. Compared to standalone deep learning and classical machine learning baselines, the suggested hybrid Bi-LSTM-XGBoost model showed better results in terms of the balanced accuracy and macro-averaged F1-score which proves its resilience in the conditions of the bitrate distribution of classes. Furthermore, the framework delivered very large Top-2 bitrate selection accuracy, meaning that near-optimal bitrate selections are made even when the network conditions are changing rapidly, a very important factor in ensuring perceptual Quality of Experience is maintained.

The experimental findings also showed that as compared to the traditional accuracy measures that can conceal the decline in performance in minority bitrate classes, the suggested structure achieves a desirable sensitivity-specificity ratio, which minimizes under-utilization and over-selection of aggressive bitrate. The confusion matrix and ROC study revealed that the majority of the misclassification is between the consecutive bitrate levels that have minimal effects on the perceived video quality. Moreover, the measurements of the inference latency ensured that the edge-deployed XGBoost decision module meets the real-time requirements, which makes the offered approach appropriate to real-life cloud-edge 5G streaming systems. In the future, such a framework will be generalized to multi-objective optimization where energy efficiency, network fairness, and user mobility prediction are looked at. Moreover, the policy adaptation strategies and cross-layer optimization strategies based on reinforcement learning will be considered in order to increase the long-term stability of QoE in ultra-dense and heterogeneous 5G and beyond-5G networks.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] S. Yau, S. Awiphan, J. Bootkrajang, and J. Chaijaruwanich, "Edge-centric quality adaptation scheme for mobile video streaming," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 6953-6965, 2025. https://doi.org/10.1109/OJCOMS.2025.3601666

[2] H. Zhan, L. Fan, C. Li, X. Lei, and F. Li, "Cloud–edge learning for adaptive video streaming in B5G Internet of Things systems," *IEEE Internet of Things Journal*, vol. 11, no. 24, pp. 40140-40148, 2024. https://doi.org/10.1109/JIOT.2024.3450477

[3] S. Yau, S. Awiphan, J. Bootkrajang, and J. Katto, "A robust throughput estimation in edge-assisted adaptive bitrate streaming networks," *IEEE Access*, vol. 13, pp. 152598-152607, 2025. https://doi.org/10.1109/ACCESS.2025.3602651

[4] W. Wang, X. Wei, W. Tao, M. Zhou, and C. Ji, "Quality of experience-oriented cloud-edge dynamic adaptive streaming: Recent advances, challenges, and opportunities," *Symmetry*, vol. 17, no. 2, p. 194, 2025. https://doi.org/10.3390/sym17020194

[5] J. Luo, F. R. Yu, Q. Chen, and L. Tang, "Adaptive video streaming with edge caching and video transcoding over software-defined mobile networks: A deep reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 1577-1592, 2020. https://doi.org/10.1109/TWC.2019.2955129

[6] W. Zhou, Y. Lu, W. Pan, Z. Chen, and J. Qin, "Dynamic bitrate adaptation and bandwidth allocation for MEC-enabled video streaming," *IEEE Communications Letters*, vol. 28, no. 9, pp. 2121-2125, 2024. https://doi.org/10.1109/LCOMM.2024.3438941

[7] L. Zhong, M. Wang, C. Xu, S. Yang, and G.-M. Muntean, "Decentralized optimization for multicast adaptive video streaming in edge cache-assisted networks," *IEEE Transactions on Broadcasting*, vol. 69, no. 3, pp. 812-822, 2023. https://doi.org/10.1109/TBC.2023.3254165

[8] M. Kim and K. Chung, "Edge computing assisted adaptive streaming scheme for mobile networks," *IEEE Access*, vol. 9, pp. 2142-2152, 2021. https://doi.org/10.1109/ACCESS.2020.3047373

[9] A.-T. Tran, N.-N. Dao, and S. Cho, "Bitrate adaptation for video streaming services in edge caching systems," *IEEE Access*, vol. 8, pp. 135844-135852, 2020. https://doi.org/10.1109/ACCESS.2020.3011517

[10] Y. F. Yeznabad, M. Helfert, and G.-M. Muntean, "QoE-driven cross-layer bitrate allocation approach for MEC-supported adaptive video streaming," *IEEE Transactions on Network and Service Management*, vol. 21, no. 6, pp. 6857-6874, 2024. https://doi.org/10.1109/TNSM.2024.3453992

[11] J. Yu, H. Wen, G. Pan, S. Zhang, X. Chen, and S. Xu, "Quality of experience oriented adaptive video streaming for edge assisted cellular networks," *IEEE Wireless Communications Letters*, vol. 11, no. 11, pp. 2305-2309, 2022. https://doi.org/10.1109/LWC.2022.3200830

[12] W. Li, X. Li, Y. Xu, Y. Yang, and S. Lu, "MetaABR: A meta-learning approach on adaptative bitrate selection for video streaming," *IEEE Transactions on Mobile Computing*, vol. 23, no. 3, pp. 2422-2437, 2024. https://doi.org/10.1109/TMC.2023.3260086

[13] T. V. Huu, S. Van Pham, T. N. T. Huong, and H.-C. Le, "QoE aware video streaming scheme utilizing GRU-based bandwidth prediction and adaptive bitrate selection for heterogeneous mobile networks," *IEEE Access*, vol. 12, pp. 45785-45795, 2024. https://doi.org/10.1109/ACCESS.2024.3382155

[14] N. V. Naik and K. Madhavi, "Adaptive QoE-based video delivery by optimizing the bitrate of the video course using network bandwidth and user bandwidth," *Concurrency and Computation: Practice and Experience*, vol. 36, no. 16, p. e8114, 2024. https://doi.org/10.1002/cpe.8114

[15] Y. J. Wong, M.-L. Tham, B.-H. Kwan, Y. C. Chang, A. Mokraoui, and F. Ke, "Efficient client selection for asynchronous federated learning for adaptive bitrate streaming," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 21, no. 12, pp. 1-22, 2025. https://doi.org/10.1145/3765759

[16] G. Pan, S. Xu, S. Zhang, X. Chen, and Y. Sun, "Quality of experience optimization for real-time XR video transmission with energy constraints," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 10, pp. 15883-15888, 2024. https://doi.org/10.1109/TVT.2024.3404266

[17] M. Kabeer, R. Nordin, M. Behjati, and F. Y. binti Mohd Shaharuddin, "An Urban multi-operator QoE-aware dataset for cellular networks in dense environments," *arXiv preprint arXiv:2506.22484*, 2025. https://doi.org/10.48550/arXiv.2506.22484

[18] J. Zhang, H. Wu, X. Tao, and X. Zhang, "Adaptive bitrate video streaming in non-orthogonal multiple access networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3980-3993, 2020. https://doi.org/10.1109/TVT.2020.2972363

[19] I. Triki, R. El-Azouzi, and M. Haddad, "NEWCAST: Joint resource management and QoE-driven optimization for mobile video streaming," *IEEE Transactions on Network and Service Management*, vol. 17, no. 2, pp. 1054-1067, 2020. https://doi.org/10.1109/TNSM.2019.2952498

[20] L. Du, L. Zhuo, J. Li, J. Zhang, X. Li, and H. Zhang, "Video quality of experience metric for dynamic adaptive streaming services using DASH standard and deep spatial-temporal representation of video," *Applied Sciences*, vol. 10, no. 5, p. 1793, 2020. https://doi.org/10.3390/app10051793

[21] R. U. Mustafa, M. T. Islam, C. Rothenberg, S. Ferlin, D. Raca, and J. J. Quinlan, "DASH QoE performance evaluation framework with 5G datasets," in *2020 16th International Conference on Network and Service Management (CNSM)*, Izmir, Turkey, 2020, pp. 1–6. https://doi.org/10.23919/CNSM50824.2020.9269111

[22] V. Ravikumar, G. U. Srikanth, S. K. Manigandan, and A. Vishnukumar, "Res-K2G-MAN: A residual knowledge graph attention framework for intrusion detection in online music education over public cloud networks," *International Journal of Information Technology*, vol. 17, no. 9, pp. 5681-5686, 2025. https://doi.org/10.1007/s41870-025-02797-2

[23] N. Madyavanhu and V. Kumar, "Utilizing multi-population ant colony system and exponential grey prediction model for multi-objective virtual machine consolidation in cloud data centers," *International Journal of Information Technology*, vol. 17, no. 9, pp. 5441-5453, 2025. https://doi.org/10.1007/s41870-024-01949-0

[24] S. A. Alzakari *et al.*, "Enhanced heart disease prediction in remote healthcare monitoring using IoT-enabled cloud-based XGBoost and Bi-LSTM," *Alexandria Engineering Journal*, vol. 105, pp. 280-291, 2024. https://doi.org/10.1016/j.aej.2024.06.036

[25] A. U. R. Butt, S. Perveen, M. Asif, R. M. Tawafak, S. Shafiq, and I. Khan, "Optimizing intelligent traffic management for smart cities using ensemble learning algorithms," in *Computational Intelligent Systems: Applications of AI and Machine Learning*, M. I. Uddin, F. Aziz, and H. Jafari, Eds., Singapore: Springer Singapore, 2026, pp. 57–72. https://doi.org/10.1007/978-981-95-2212-5_4