Computer Systems
Science and Engineering

Tech Science Press

DOI: 10.32604/csse.2023.036864
Article

Adaptive Learning Video Streaming with QoE in Multi-Home Heterogeneous Networks

S. Vijayashaarathi[1,*] and S. NithyaKalyani[2]

[1]Department of Electronics and Communication, Sona College of Technology, Salem, 636005, Tamil Nadu, India
[2]Department of Information Technology, K. S. R. College of Engineering, Tiruchengode, 637215, Tamil Nadu, India
*Corresponding Author: S. Vijayashaarathi. Email: vijayashaarathis@gmail.com
Received: 14 October 2022; Accepted: 21 December 2022

Abstract: In recent years, real-time video streaming has grown in popularity. The growing popularity of the Internet of Things (IoT) and other wireless heterogeneous networks mandates that network resources be carefully apportioned among versatile users in order to achieve the best Quality of Experience (QoE) and performance objectives. Most researchers focused on Forward Error Correction (FEC) techniques when attempting to strike a balance between QoE and performance. However, as network capacity increases, the performance degrades, impacting the live visual experience. Recently, Deep Learning (DL) algorithms have been successfully integrated with FEC to stream videos across multiple heterogeneous networks. But these algorithms need to be changed to make the experience better without sacrificing packet loss and delay time. To address the previous challenge, this paper proposes a novel intelligent algorithm that streams video in multi-home heterogeneous networks based on network-centric characteristics. The proposed framework contains modules such as Intelligent Content Extraction Module (ICEM), Channel Status Monitor (CSM), and Adaptive FEC (AFEC). This framework adopts the Cognitive Learning-based Scheduling (CLS) Module, which works on the deep Reinforced Gated Recurrent Networks (RGRN) principle and embeds them along with the FEC to achieve better performances. The complete framework was developed using the Objective Modular Network Testbed in C++ (OMNET++), Internet networking (INET), and Python 3.10, with Keras as the front end and Tensorflow 2.10 as the back end. With extensive experimentation, the proposed model outperforms the other existing intelligent models in terms of improving the QoE, minimizing the End-to-End Delay (EED), and maintaining the highest accuracy (98%) and a lower Root Mean Square Error (RMSE) value of 0.001.

Keywords: Real-time video streaming; IoT; multi-home heterogeneous networks; forward error coding; deep reinforced gated recurrent networks; QoE; prediction accuracy; RMSE

1 Introduction

In recent years, the increasing popularity of IoT and wireless communication networks has enabled users to access their networks and stream videos anywhere. The proliferating wireless infrastructure offers a wide range of broadcasts, including Wireless Fidelity (Wi-Fi), Wireless Local-Area Networks (WLAN), Worldwide Interoperability for Microwave Access (Wi-MAX), Institute of Electrical and Electronics Engineers 802.11 (IEEE 802.11), and even Mobile Cellular Communication [1,2].

According to a report [3], due to the exponential growth of these wireless technologies, real-time video will most likely account for 90% of the increase in network traffic by 2023. Single wireless networks cannot deliver good video-sharing quality compared to their controlled capacity, fragility, and irregular coverage. While cellular networks like Universal Mobile Telecommunications Service (UMTS) and Global System for Mobile (GSM) communication might offer a stable connection, they fall short when it comes to providing the highest Quality of Service (QoS). Despite having excellent coverage and faster data rates, Long Term Evolution (LTE) and Wi-MAX are not widely used [4]. According to the talks, users must provide their devices with several connectors to connect to several networks simultaneously and have multi-home access. To deliver a better QoE in heterogeneous networks with multiple-homed clients, effective coding schemes and Adaptive Forward Error Coding (AFEC) methods are used in existing work. [5–7]. Nonetheless, all existing results only use static network patterns to predict future designs, ignoring the possible relationship between static and future that influences the failure of these algorithms as the network changes dynamically [8–10].

To develop a better QoE and better performance, recent studies have explored the advantages of the machine and DL architectures in AFEC for real-time video transmission in multi-path environments. Long Short-Term Memory (LSTM) [11] has recently attracted many researchers to embed these networks with AFECs. Despite implementing Deep Learning (DL) methods that improve performance, video streaming applications in multi-homed clients continue to suffer from packet losses, distortion, low PDR, and latency [12–15]. Motivated by the above drawbacks, this paper proposes a novel intelligent network and content-aware framework to achieve a better QoE with high performance. The Adaptive Reinforced Gated Recurrent Neural Networks (ARGRN) are introduced for scheduling packets according to different network conditions. This is the first type to incorporate a Gated Recurrent Neural Network (GRNN) with AFEC to improve QoE and performance. It may open a new gateway for video streaming in multi-path research.

The main contributions of this research are listed as follows:

1. Embed the Reinforced Gated Recurrent Networks (RGRN) with FEC to solve the problem of streaming video packets in a network that changes over time.
2. Network Content-Aware Transmission (NCAT) is adopted to save bandwidth, thus reducing the transmission delay and increasing performance.
3. Extensive experiments and novel evaluation measures have been adopted to prove the excellence of the proposed model when compared with other existing algorithms.

The paper's organization is as follows: The related works by different writers are included in Section 2. The specific operating system of the recommended model is discussed in Section 3. The experimental setup, performance analysis, and comparisons with current systems are described in Section 4. Finally, after the future improvements, Section 5 presents the conclusion.

2 Related Works

By assessing the continued results of Adaptive Video Streaming (AVS) for heterogeneous video transmission, the author [16] investigates the crucial concept of QoE designs. Finally, by incorporating a coordinated telecom framework, this study brings the hypothetical models closer to plausible execution. However, the community genuinely requires modern approaches. In areas like video coding, multiuser communication, and broadcasting networks, both academia and business will need to work together to come up with effective techniques.

Author [17] presents LSTM-QoE, an intermittent brain network-based QoE expectation model using an LSTM organization. The LSTM-QoE is a set of streaming LSTM units created to represent the intricate nonlinear effects and transitory situations connected with time-varying QoE. Based on an analysis of different persistent QoE datasets that are available to the public, this method shows that it has the potential to represent the QoE characteristics well. The comparison between the proposed model and the top-performing QoE forecast models in this structure demonstrates the proposed model's remarkable performance across these datasets. Additionally, by showing approaches to the QoE expectation, this system illustrates the practicality of the state space point of view for the LSTM-QoE. But despite all this, this framework's main flaw is that it has trouble handling heterogeneous information over time.

Author [18] showed Video Quality Aware Resource Allocation (ViQARA), a perceptual QoE-based resource allocation (RA) method for video web-based in cell organizations. ViQARA employs the most recent consistent QoE models and combines the summed-up and reasonable RA methodologies. This structure illustrates how ViQARA, compared to conventional throughput-based RA methods, may offer a notable increase in the users' perceived QoE and a surprise drop in rebufferings. The planned computation is also there to allow better QoE streamlining of the usually available resources when the mobile network doesn't have enough resources or has long delays in sending packets. However, the prices of network ventures and Content Delivery Network (CDN) use fees go up with this technique.

Another quality-conscious multi-source video web-based conspiracy for Content-Centric Networking (CCN) is suggested by [19]. First, various storage methods are considered for the material conveyance of video recordings between CCN hubs. Second, a versatile video real-time system with adequate storage. Adaptive Video Streaming with Distributed Caching (AVSDC) calculation is designed to maintain QoE while sharing data between good sources. The conveyance of AVS is taken into account in the AVSDC computation. It consequently alters the layers in video transmission when there is source switching in light of a QoE model that illustrates the effect of slowing down. In terms of the QoE determined by human emotional tests, experimental results show that AVSDC computation works better than dynamic versatile spilling over HyperText Transfer Protocol HTTP (DASH) in the CCN stage. The fundamental attraction of this arrangement is that it promotes computational complexity.

A novel cooperative QoE-based versatile mindful video real-time conspiracy sent to Mobile Edge Computing (MEC) servers is proposed [20]. The author's suggested plan can be implemented to maintain a suitable QoE level for each client throughout the entire video conferencing. Broad reproductions have investigated how the implied conspiracy will be carried out. In contrast to prior methods, the results demonstrate that high productivity is obtained by a coordinated effort across MEC servers, leveraging unambiguous window size transformation, cooperative prefetching, and handover among the edges. However, this system's drawback is the requirement for increased transmission capacity for streaming.

The author [21] proposes the CNN-QoE as a more potent TCN-based model for constantly predicting the QoE, which exhibits consecutive information features. By learning about upgrades to further increase the accuracy of QoE expectations, the proposed model uses TCN's advantages to get over this framework's computational complexity limitations. A thorough investigation of this structure shows that the presented framework may be able to provide high QoE forecast execution on both PCs and mobile devices, which would be better than existing methods. In either scenario, the arrangement results in handover delays.

Flex-Steward (FS) is a programme made by [22] that improves the joint QoE of flexible video for multiple clients in real time while sharing bottleneck data transfer. The term *'Joint QoE Improvement'* refers to improving QoE consistency among users of different video devices using separate services and having multiple demands. FS facilitates learning at the network edge and transmits a flexible bitrate computation based on Neural Networks (NN). A developed NN model is needed to make recommendations for the right bitrate for video bits to be sent by clients with the same bottleneck transmission capacity. In terms of joint QoE enhancement, the results proved how FS reduces shamefulness by between 10.9% and 41.7%. This system's primary constraint is that it necessitates time complexity and additional resources throughout the cycle.

The author [23] examines and develops HTTP adaptive streaming's processing and data-transmission capabilities for live streaming web-based video with different edge rates and objectives. In order to evaluate the viewer experience of live video channels, this system also presents an asset-aware QoE model. The structure then provides a QoE-driven HAS to support a new channel paradigm to enhance the usual client QoE. Using a heuristic solution, this framework transforms the boost issue into a Multidimensional Knapsack Problem (MKP). The results of the experimental analysis projected the suggested method's viability compared to benchmark setups. However, this system required greater computation energy.

Author [24] offers a novel Adaptive Bitrate Algorithm (ABR) calculation that can reduce traffic volume while keeping QoE higher than the goal. Customers' desires (or) CDN budgets might be considered when streaming service providers evaluate the intended QoE. Each meal chooses an acceptable bitrate by assessing QoE and traffic volume so that all bitrate designs support the approaching few bits based on future throughput and a cushion change forecast. The QoE is better than existing computations while reducing network traffic by an average of 18.3% to 51.2% in the adaptable environment and by 1.2% to 38.3% in the broadband location, following the flow-based reproduction. In any case, this structure has handover problems when it uses long-range communication. Based on the related works, the research gap is that AVS applications in multi-homed clients that still have problems like packet loss, PDR, EED, and many others (Table 1) need an intelligent system.

- Maximized the average user's QoE.
- Long-term video quality is improved without reducing the delayed performance.
- Reduce the amount of time that services are unavailable.
- The designed approach handles the streaming of different video bitrates under congested network conditions to provide sufficient video quality.
- A designed framework ensured near-optimal satisfaction and efficiency.
- Minimized resource consumption, better throughput, and reduced delay.

Table 1: Summary of related works

Authors	Methodology	Merits	Demerits
Liu et al.	QoE-driven HTTP adaptive streaming (HAS)	Maximized average user QoE	High energy consumption
Kimura et al.	ABR	Minimized network traffic	High bandwidth
Feng et al.	Long-term rate control scheme	Less delay	High energy consumption
Haotong et al.	Virtual network embedding (VNE)	Minimize the time of service interruption	Increased resource utilization
Samira et al.	Content-aware and path-aware (CAPA)	High video quality	High energy consumption
Eksert et al.	Intra and inter-cluster link scheduling	Near-optimal satisfaction and efficiency	Increased delay and bandwidth
Guanyu et al.	Video transcoding in ABR streaming	Less resource consumption	But increased time complexity
Xiongli et al.	Blind DL-driven method	High throughput	High resource complexity
Ghadiyaram et al.	QoE-live mobile stall video database-II	Less delay	High computational complexity

3 Proposed Methodology

Fig. 1 shows the proposed model for intelligent and AFEC-based video transmission over multiple-communication interfaces. The proposed system consists of (a) a module for Intelligent Content Extraction, (b) a cognitive Learning-based Scheduling and FEC Module (CLS-FEC) (c) a module for Channel Monitoring.

In the first module, the transmitted video is converted into corresponding frames by a frame splitter, in which the principle of saliency mapping extracts the actual contents. Also, the channel properties of multi-paths are monitored and collected by the channel monitor in a pipelined fashion. Finally, the combined parameters from the above modules are then passed to the CLS-FEC module, which consists of AFEC for forwarding error correction and DL architecture for scheduling methods. The video transmitter then transmits the packets to the different users. The decoding of the video stream is done on the receiver. When a packet reaches the receiver side, the decoder puts it together correctly and shows it to the user's video player.

3.1 Network Model

Considering two transmission pathways, a heterogeneous wireless network with P routing paths. The Gilbert simulator is used to model the losses on each path. The path state $x(t)$ is taken to be either 1 (good) or 0 (bad),'t' at the time. If $x(t) = 1$, the packet will be executed, but the results will be lost; if $x(t) = 0$, the packet will be destroyed. Let's suppose that M_t is the maximum transmission unit and that 'O' represents the frame's output bits. The following is the expression for 'N', the number of video frames: Eq. (1)

$$N = O/Mt \tag{1}$$

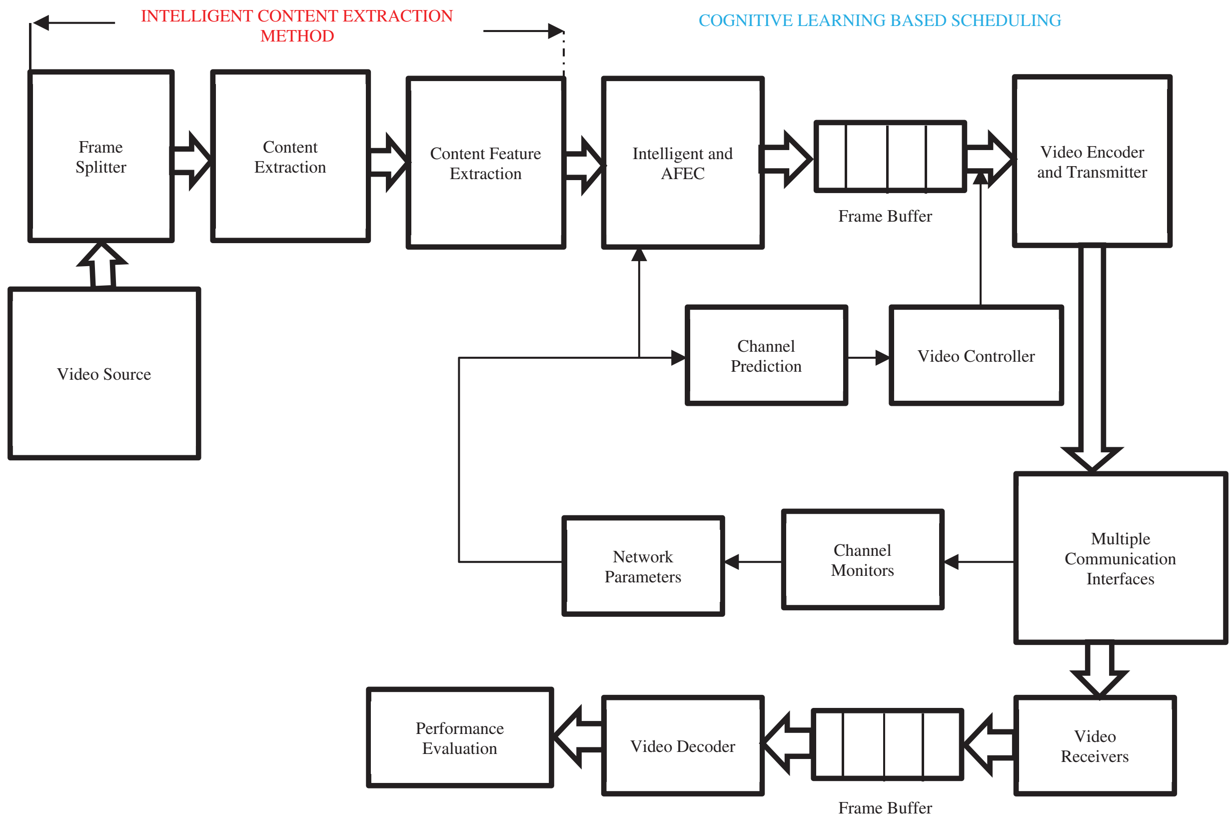


Figure 1: Proposed model for intelligent and AFEC for video transmission

The transmission loss rate on the path *"p"* for each *"g"* content frame is expressed by the tuple of pathways *"p"* with the size of *"N_p"* as Eq. (2)

$$F(g,p) = 1/\mathrm{N}p \sum_{i=0}^{Np} \beta(C(p)^i == 1) \tag{2}$$

where, C(p) = n{1,0,1,0,1,0,1,0,1 . . . 1}. From Eq. (3), $\boldsymbol{\beta}$ → indicator function value. In ability networks, the exponential distribution described in [25] can be used to figure out how likely it is that packets will be lost along path "p" after the deadline "T."

$$F''(g,p) = P(D(g,p) > \tag{3}$$

where D (g, p) is the EED for the content frame. *"g"* includes delivery processing and propagation latency. The total delay is calculated for each path that carries the video streaming flow and is expressed as Eq. (4)

$$X(T) = X(p) \cdot T * F(g,p) \tag{4}$$

For calculating the video streams, X(T) → size of the cumulative sub-streaming flow [0,t] over the path p, and X(p) → long-term average video streaming rate, the packet path is used. As mentioned in [26], the model employs a work-conserving queueing system, which is then used to calculate the overall delay. As a result, the overall EED for the frame *"g"* full path is the sum of all queuing delays associated with the path *"p"*. Mathematically, the overall EED is expressed as follows, Eq. (5):

$$D\left(g,p\right)=\frac{F\left(g,p\right)}{X\left(p\right)-}+F\left(d\right) \tag{5}$$

where F (d) is the fixed EED on each path, the proposed system uses the distortion model from [24], where d is the fixed delay on each path.

3.2 Intelligent Content Extraction Module

The content extraction module employs the visual saliency technique for extracting the visual content from the video frames. In order to determine the salient objects from the movies, a saliency model incorporating the cues of object intensity, color, and motion has been developed [27,28]. A simple statistical model that works on both videos and images. This method combines conditional random fields with local information like color and motion signals to create saliency maps. Several bottom-up techniques have been presented in order to identify the prominent items in films. A multiscale method for video saliency map computation combines motion cues to extract features from movies. This article employs one-dimensional Gated Recurrent Units (GRU) to reduce pixel processing. The full explanation of GRU is discussed in Section 3.4. It is found that each frame of the video consists of static content (background) and specific content (foreground). In this GRU-based process, saliency maps represent each frame's different contents based on the pixels' color, intensity, and luminance values. The pixels with saliency content are labeled as the "higher-pixels", whereas those without saliency content are labeled as the "lower-pixels". Both categories of pixels are stored in the same buffer, which is used for encoding and transmission according to the network characteristics. Table 2 shows the GRU specifications used for saliency extractions.

Table 2: Specification of the GRU used for the extraction of saliency maps

Parameters used for GRU training	Specification
Number of cells	10
Learning rate	0.001
Dropout ratio	0.2
Number of the hidden layers	100

3.3 Network Channel Status Monitor

The network channel monitor is responsible for collecting the path status information from the multiple heterogeneous paths and directing it to the cognitive learner, where the properties are used to train the proposed DL method. Based on the network parameters, the proposed model predicts the best QoS path among the multiple paths and schedules the frame packets according to the network parameters. The network characteristics, such as Available Bandwidth (B_a), Received Signal Strength (RSS), End-to-End Delay (EED), and noise and distortion level (Signal to Noise Ratio (SNR)) are measured and then used for training the proposed model.

3.4 Cognitive-Based Scheduling and Adaptive Forward Error Correction

The proposed FEC ensembles predict paths based on network and content properties by combining Q-Learning and GRU networks. The descriptions of Q-learning and GRU are evident in the preceding section.

3.4.1 Q-Learning Concepts

A saliency model, including the signals of object intensity, color, and motion, has been constructed in order to identify the salient items from the videos [29]. A straightforward statistical method applies to both images and videos. This method combines local information like color and motion signals with conditionally random fields to produce the saliency maps. The most significant elements in movies have been identified using different bottom-up approaches. A multiscale video saliency map computation method derives features from movies by combining motion cues. This article uses a one-dimensional GRU to reduce pixel processing. The aim of RL as feedback to the learning model is to maximize reward. According to Q-learning, significant reinforcement learning success It is the procedural version of the off-policy model-free method, often known as the Q-learning algorithm. The standard algorithm for resolving related problems is Q-learning. Using samples collected during interactions with the environment, the Q-function can approximate the state of action pairs [30,31]. Eq. (6) represents the discrete-time Q-function. A Markov Decision Process (MDP) generates Q-learning as a reinforcement learning method. This MDP establishes the criteria for the state action, the reward, and the likelihood that it will occur as (S, A, P, R) and $P^{a}_{zz'}$. Let z current state and z' is the next state with action value 'a'

$$P^{a}_{zz'} = Prob\{z_{i+1} = z' | z_t = z, a_t = a\} \tag{6}$$

The state reward function is given as [z and z'] is given as $[R^{a}_{z_t z_{t+1}} \cdot t]$. For the current state, the overall reward function is Eq. (7)

$$R_t = \sum\nolimits_{z_{t+1} \in Z} P^{a_t}_{s_t s_{t+1}} R^{a_t}_{z_t z_{t+1} | z_t = z, a_t = a} \tag{7}$$

3.4.2 Gated Recurrent Units

GRU is the most appealing LSTM variant [32,33]. This idea was proposed in [34,35], which combines the forget gate and input vector as a single vector. This network supports long-term sequences and also has long memories. The complexity is greatly reduced when compared with the LSTM network. The following Eqs. (8)–(11) are coined by the author to represent the characteristics of GRU:

$$h_t = (1 - x_t) \odot h_{t-1} + x_t \odot h_t \tag{8}$$

$$\widetilde{h}_t = g(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h \tag{9}$$

$$z_t = \sigma(W_h x_t + U_z h_{t-1} + b_z) \tag{10}$$

$$r_t = \sigma(W_h x_t + U_r h_{t-1} + b_r) \tag{11}$$

The overall GRU characteristic Eqs. (11) and (12) is represented by

$$P = GRU\left(\sum\nolimits_{t=1}^{n} [x_t, h_t, z_t, r_t \,(W(t), B(t), \eta\,(tannh))]\right. \tag{12}$$

where "$W_t \rightarrow$weights and $B_t \rightarrow$bias weights at the current instant, $Z_t, r_t \rightarrow$update and reset gates, $x_t \rightarrow$input feature at the current state, $y_t \rightarrow$output state, and $h_t \rightarrow$the module's output at the current instant".

3.4.3 Adaptive Learning and Path Prediction

The proposed learning model ensembles the Q-GRU learning for scheduling method and the Markov Decision Process (MDP) and is used to select an proper path and schedule it in line with the available bandwidth, EED, SNR, and signal intensity. To avoid random study at the initial phase, the algorithm has been initialized with a partially pre-computed policy applied to the different values using Eqs. (13) and (14). The proposed Q-based GRU network receives input from the channel status monitor, which evaluations the best QoS path, and compares it to pre-calculated rules consisting of several heterogeneous network reward functions. Based on the computation, Q-learning ranks the different paths, schedules the content, and sends it to the FEC coder. Algorithm 1 represents the Q-learning-based QoS aware paths. Specifically, MDP for the QoS path environment defines a set of *"s"* of nodes and a group of *"A"* actions that allow an agent to move to different states. Other states in this case represent QoS-aware paths. The Reward Policy "R" defines the reward given by an action that selects the best path to transmit the video content without distortion. Finally, the main goal of the MDP is to find the optimal paths. More specifically, MDP is made up of a series of "n" discrete steps t = 0, 1, 2,..., l, in which an agent looks at each part of the network and chooses the best way to get from one network to another. An agent gets an immediate reward once the best path is selected. Rewards are modeled based on Eq. (15). Mathematically, the reward function for this decision on the QoS-aware path is modified.

$$R\left(P\right) = \sum\nolimits_{z_{t+1} \in Z} P^{A_t}_{s_t s_{t+1}} R^{A_t}_{z_t z_{t+1} | z_t = z, A_t = A} \tag{13}$$

where,

$$s_t = E\left(s\right) = \left\{Max\left(B, RSSI, SNR\right), Min\left(D\right)\right\} \tag{14}$$

$$s_{t+1} = E\left(s\right) = \left\{Max\left(B, RSSI, SNR\right), Min\left(D\right)\right\} \tag{15}$$

The mathematical models of how each state reaches the best rewards ensure that the best path is chosen. Here the bandwidth, Received Signal Strength Indication (RSSI), SNR, and EED play a major role in selecting the best path. The proposed DL model is embedded along with FEC coders that encode the data and correct the bit frames according to the nature of the path selected by the DL model. Finally, it sends each content frame to the video transmitter. The video transmitter transmits the video frames according to the chosen path.

3.4.4 Algorithm-1 for Proposed Model

Step 1. Initiate the streaming rate X(p), Fix the EED, channel

Step 2. Compute the bandwidth, RSSI, D, and SNR

Step 3. Compute the paths and arrange them in descending order

Step 4. If (R(p) == D(t, p) < T, where T = threshold reward function;

Select the best path from the stored data

Step 5. Allocate the Saliency Contents in the path selected

Step 6. Encode and transmit to the video controller

Step 7. Else

Step 8. Go to Step 3

Step 9. End

4 Result and Discussion

The simulation experimentation was conducted using OMNET++5.6 interfaced with the INET framework. The INET framework is a software plug-in for OMENT++ 5.6, which supports most wireless communication network interfaces such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Internet Protocol version 4 (IPv4), IEEE 802.11, and even IPv6. Hence, these are used for emulating heterogeneous wireless networks. A Python 3.10-based video codec has been developed and explored for various input applications. All the network properties are downloaded offline and used for training the proposed model. The DL model has been developed using the Tensorflow and Keras libraries. The server has one connection, and the client has three connection interfaces. An end-to-end connection is set up between the client and server by binding a pair of Internet Protocol (IP) addresses from the server to the clients. Table 3 lists the experiment's parameters, which can be found in the previous work [37].

Table 3: Parameters used for the experimentation

Experimental parameters	Specification
Bandwidth	350 kbps
Number of route	4
Capacity	50-350 kbps
SNR ranges	15-30 db
Number wireless networks	4
Loss rate	3 to 5%
Bit rates	600 kbps
Channel power	30 to 45 dBm

To prove the excellence of the proposed algorithm, existing algorithms such as Earliest Delivery Path First (EDPF), the Round-Robin (RR), Local Balancing Algorithm (LBA), and Recurrent Neural Network (RNN)-based Region of Interest (ROI) detectors were compared. The evaluation of performance is done in three parts: Video Quality Analysis (VQA), Exit Event Detection (EED), and Path Predictive Analysis (PPA).

4.1 Video Quality Analysis

For the analysis of video quality at the receiver side, Peak Signal-to-Noise Ratio (PSNR) and Mean Opinion Score (MOS) are used for evaluating the proposed model by using mathematical expressions as mentioned in Tables 4–6 show the PSNR of the different models for the different video streaming rates. The result proved the other models' PSNR at a streaming rate of 650 kbps. As the streaming rates increase, algorithms such as LB, RR, and EDPF have drastically reduced their performance, whereas LSTM-ROI and the proposed model have produced very good PSNR ranges of 33–40. But still, the usage of Q-GRU in the proposed network has outperformed the LSTM-ROI in creating a better PSNR for the increased streaming rates. The results show that PSNR decreases as the bit rate increases, which affects the video quality received on the user's side. Hence, the experimentation involves measuring MOS values at the receiver's end.

Table 4: Compares the various models at a video streaming rate of 600 kbps

Algorithms	PSNR (db)
EDPF	26
RR	27
LB	24
RNN-ROI	34
Proposed model	37

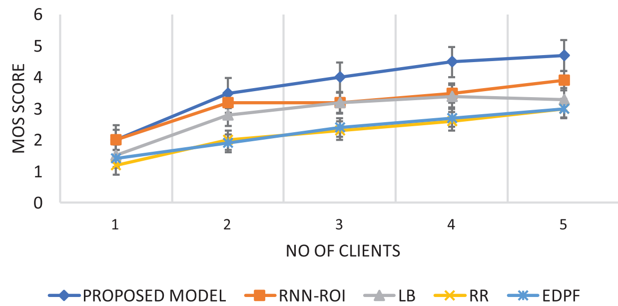
Table 5: A comparison of the various models at 850 kbps video streaming rate

Algorithms	PSNR (db)
EDPF	25.78
RR	25.4
LB	24
RNN-ROI	33.5
Proposed model	36.5

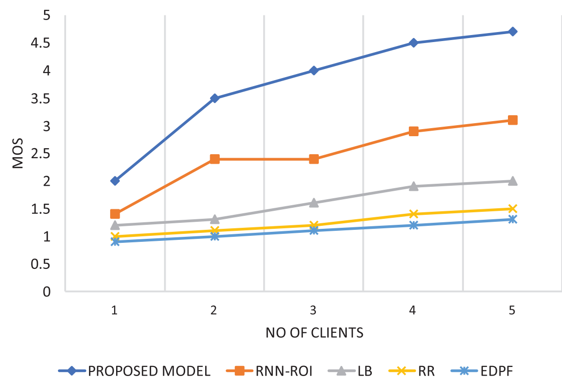
Table 6: Comparative analysis of the different models at video streaming rates of 1 Mbps

Algorithms	PSNR (db)
EDPF	20
RR	20.4
LB	23
RNN-ROI	32.5
Proposed model	36.5

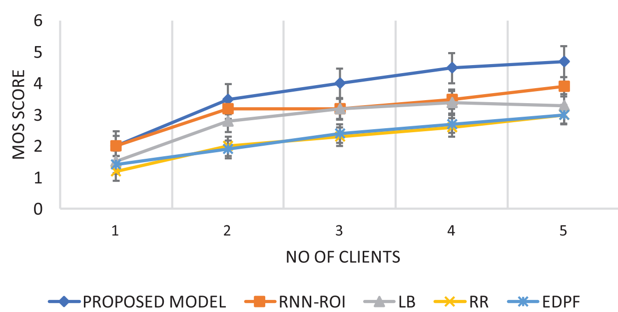
Figs. 2a–2c show the MOS performance of algorithms at different streaming rates. The output indicates that the LB, RR, and EDPF have degraded MOS scores ranging from 1.2 to 3 as the number of clients and streaming rate increase, while the LSTM-ROI and the proposed model have produced many suitable performances for the increased clients and streaming rate. But the proposed model has produced good video quality (MOS = 4.7) and has edged out the LSTM-ROI (MOS = 4.2) method for the increased streaming rates.



(a)



(b)

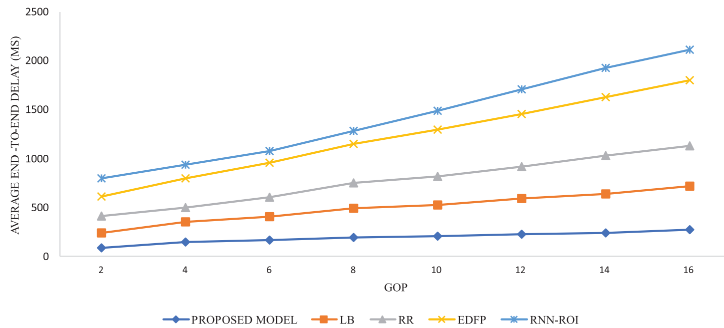


(c)

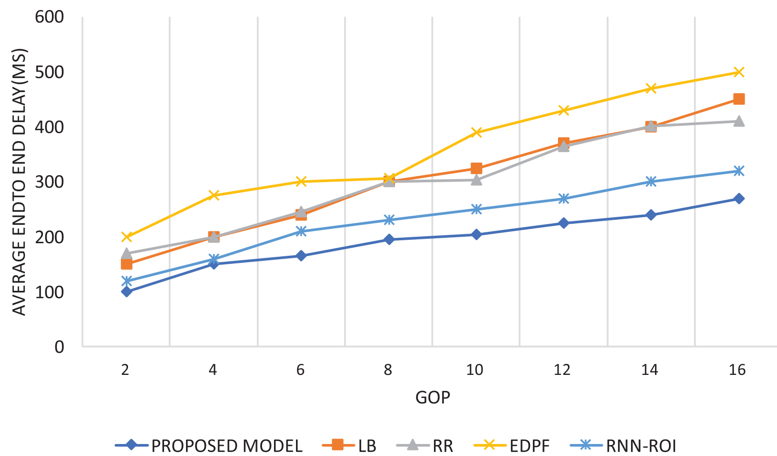
Figure 2: MOS performance with algorithms (a) Streaming rate = 650 kbps; (b) Streaming rate = 1 Mbps; (c) Streaming rate = 2 Mbps

4.2 End-to-End Delay Analysis

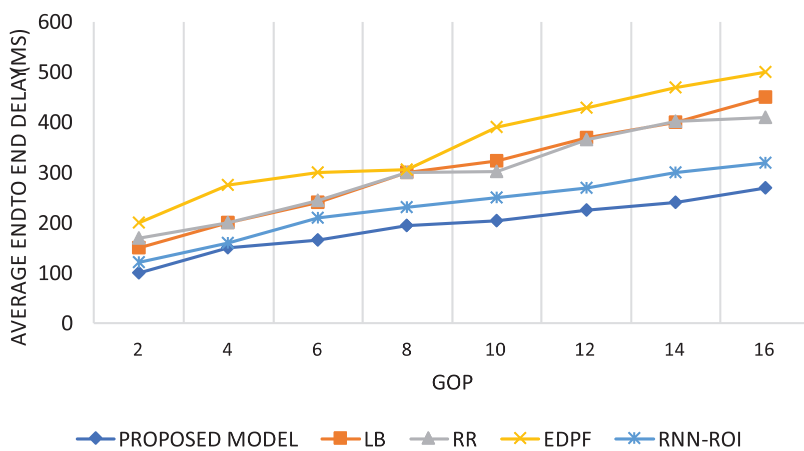
Figs. 3a–3c plot the average EED in Group Of Picture (GOP) units of various algorithms with varying streaming rates. Fig. 3a shows the EED analysis for lower bitrates. Fig. 3b shows the EED for the high bit rates. From Fig. 3c, it is clear that the EED of all five algorithms ranges from 100 to 300 ms. As the bit rates increase, the EED of the LB, RR, and EDPF drastically increases, even to 700 ms. The EED of the proposed model is less than 500 ms, even though the streaming rate increases. As a result, the EED goes up, but the quality of what is received goes up, as shown by the MOS results.



(a)



(b)

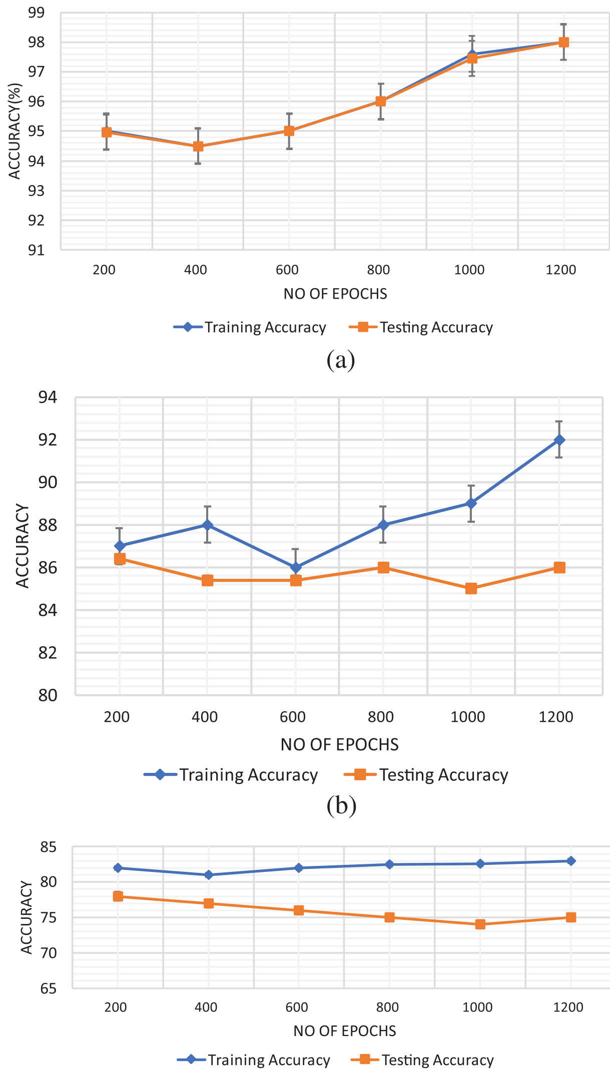


(c)

Figure 3: Average EED analysis for different algorithms (a) streaming rate of 650 kbps; (b) streaming rate of 1 Mbps; (c) streaming rate of 2 Mbps

4.3 Path Prediction Analysis

Figs. 4a–4c depict the PPA for learning models such as the proposed Q-GRU, RNN, and Hidden Markov Models (HMM) at different epochs, where prediction accuracy is used to assess the algorithm's strength. The proposed model has produced the highest prediction accuracy of 98%, with a lower RMSE error of 0.001. On the other hand, the accuracy of RNN is 92% with an RMSE of 0.2678, and HMM's accuracy is 84% with an RMSE of 0.458, respectively. The proposed model has outperformed the other existing algorithms due to its adaptive learning with less complex GRU networks. RNN has a problem with vanishing gradients, making it fail to learn the data features. This drawback of RNN significantly affects the performance, as evident from the output.



(a) (b) (c)

Figure 4: Path prediction analysis for the different algorithms: (a) proposed model; (b) RNN (without reinforcement); (c) Markov hidden models

5 Conclusion and Future Work

This research article presents a new framework for content-aware and network-aware video transmission in heterogeneous networks with multi-homed terminals. Even at lower bandwidths, the framework provides a higher QoE. The different network and channel properties were collected and calculated using the wireless network model. The system integrates Q-adaptive learning to enhance the QoE and schedules the packets based on the network characteristics. Additionally, saliency contents are extracted by 1-D GRU networks from the video pixels. These pixels are transmitted to the network according to the path decided by the intelligent learning algorithms. Also, adaptive FEC codes are used for error correction and make decisions based on the bit frames and networks. The novel experimental environment was created using OMNET++, INET, and Python 3.10 to support the different wireless terminals and to deploy the DL systems. Experiments show that the proposed method has worked better to improve QoE than other models that are already out there.

This framework can be enhanced in future work by paying more attention to video encoders and AFECs, which are trade-offs between error bits and redundant bits.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] B. Mariem, E. A. Adnen, J. Abderrazak and D. François, "Learning-based metaheuristic approach for home healthcare optimization problem," *Computer Systems Science and Engineering*, vol. 45, no. 1, pp. 1–19, 2023.

[2] R. Prabhu and S. Rajesh, "An advanced dynamic scheduling for achieving optimal resource allocation," *Computer Systems Science and Engineering*, vol. 44, no. 1, pp. 281–295, 2023.

[3] J. Noorul Ameen and S. Jabeen Begum, "Evolutionary algorithm based adaptive load balancing (EA-ALB) in cloud computing framework," *Intelligent Automation & Soft Computing*, vol. 34, no. 2, pp. 1281–1294, 2022.

[4] S. Dinesh and K. Neetesh, "Machine learning techniques in emerging cloud computing integrated paradigms: A survey and taxonomy," *Journal of Network and Computer Applications*, vol. 205, no. 103419, pp. 1–39, 2022.

[5] Q. Li, Y. Jiang, G. M. Muntean and L. Zou, "Learning-based joint QoE optimization for adaptive video streaming based on smart edge," *IEEE Transactions on Network and Service Management*, vol. 19, no. 2, pp. 1789–1806, 2022.

[6] L. Zhong, X. Ji, Z. Wang, J. Qin and G. M. Muntean, "A Q-learning driven energy-aware multipath transmission solution for 5G media services," *IEEE Transactions on Broadcasting*, vol. 68, no. 2, pp. 559–571, 2022.

[7] L. Jingfu and L. Jiuling, "Two-phase sample average approximation for video distribution strategy of edge computing in heterogeneous network," *Computer Communication*, vol. 182, no. 11, pp. 255–267, 2022.

[8] C. R. Debanjan, N. Sukumar and G. Diganta, "Video streaming over IoV using IP multicast," *Journal of Network and Computer Applications*, vol. 197, pp. 1–15, 2022.

[9] A. Tasnim, B. L. Asma and A. E. Sadok, "User behavior-ensemble learning based improving QoE fairness in HTTP adaptive streaming over SDN approach," *Advances in Computers*, vol. 123, no. 1, pp. 245–269, 2021.

[10] A. Mahmoud, S. Amin and T. Amir, "Deep learning for network traffic monitoring and analysis (NTMA): A survey," *Computer Communications*, vol. 170, no. 3, pp. 19–41, 2021.

[11] R. Wang, L. Si and B. He, "Sliding-window forward error correction based on reference order for real-time video streaming," *IEEE Access*, vol. 10, pp. 34288–34295, 2022.

[12] D. S. Rajput, R. Somula and R. K. Poluru, "A novel architectural model for dynamic updating and verification of data storage in cloud environment," *International Journal of Grid and High-Performance Computing*, vol. 13, no. 4, pp. 75–83, 2021.

[13] M. Vashishtha, P. Chouksey, D. S. Rajput, S. R. Reddy, M. P. K. Reddy *et al.,* "Security and detection mechanism in IoT-based cloud computing using hybrid approach," *International Journal of Internet Technology and Secured Transactions*, vol. 11, no. 5, pp. 436–451, 2021.

[14] J. Liu, W. Zhang, S. Huang, H. Du and Q. Zheng, "QoE-driven has live video channel placement in the media cloud," *IEEE Transactions on Multimedia*, vol. 23, pp. 1530–1541, 2021.

[15] T. Kimura, T. Kimura, A. Matsumoto and K. Yamagishi, "Balancing quality of experience and traffic volume in adaptive bitrate streaming," *IEEE Access*, vol. 9, pp. 15530–15547, 2021.

[16] C. Feng, Z. Jie, Z. Mingkui, W. Jiyan and L. Nam, "Long-term rate control for concurrent multipath real-time video transmission in heterogeneous wireless networks," *Journal of Visual Communication and Image Representation*, vol. 77, no. 102999, pp. 1–13, 2021.

[17] C. Haotong, H. Yue and Y. Xiang, "Towards intelligent virtual resource allocation in UAVs-assisted 5G networks," *Computer Networks*, vol. 185, no. 107660, pp. 1–18, 2021.

[18] A. Samira, E. R. Christian, T. Vanessa, K. Prakash and B. Imed, "Multipath MMT-based approach for streaming high-quality video over multiple wireless access network," *Computer Networks*, vol. 185, no. 107638, pp. 1–18, 2021.

[19] M. L. Eksert, Y. Hamdullah and O. Ertan, "Intra- and inter-cluster link scheduling in CUPS-based ad hoc networks," *Computer Networks*, vol. 185, no. 10765, pp. 1–18, 2021.

[20] G. Guanyu and W. Yonggang, "Video transcoding for adaptive bitrate streaming over edge-cloud continuum," *Digital Communications and Networks*, vol. 7, no. 4, pp. 598–604, 2021.

[21] C. Xiongli and S. Feng, "Blind quality assessment of omnidirectional videos using Spatio-temporal convolutional neural networks," *Optik*, vol. 226, no. 1, pp. 1–11, 2021.

[22] D. Ghadiyaram, J. Pan and A. C. Bovik, "A subjective and objective study of stalling events in mobile streaming videos," *IEEE Transactions on Circuits System Video Technology*, vol. 29, no. 1, pp. 183–197, 2019.

[23] N. Eswara, S. Chakraborty, H. P. Sethuram, K. Kuchi, A. Kumar *et al.,* "Perceptual QoE-optimal resource allocation for adaptive video streaming," *IEEE Transactions on Broadcasting*, vol. 66, no. 2, pp. 346–358, 2020.

[24] M. F. Tuysuz and M. E. Aydin, "QoE-based mobility-aware collaborative video streaming on the edge of 5G," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 11, pp. 7115–7125, 2020.

[25] T. N. Duc, C. T. Minh, T. P. Xuan and E. Kamioka, "Convolutional neural networks for continuous QoE prediction in video streaming services," *IEEE Access*, vol. 8, pp. 116268–116278, 2020.

[26] S. Yang, X. Yu and Y. Zhou, "Lstm and GRU neural network performance comparison study: Taking yelp review dataset as an example," in *Proc. of Int. Workshop on Electronic Communication and Artificial Intelligence*, Shanghai, China, pp. 98–101, 2020.

[27] A. E. Omer, M. S. Hassan and M. E. Tarhuni, "An integrated scheme for streaming scalable encoded video-on-demand over CR networks," *Physical Communication*, vol. 35, no. 100701, pp. 1–11, 2019.

[28] L. Hao, L. Weimin, Z. Wei and G. Yunchong, "A joint optimization method of coding and transmission for conversational HD video service," *Computer Communications*, vol. 145, no. 7, pp. 243–262, 2019.

[29] M. A. Samsuden, N. M. Diah and N. A. Rahman, "A review paper on implementing reinforcement learning technique in optimising games performance," in *Proc. of IEEE Int. Conf. on System Engineering and Technology*, Shah Alam, Malaysia, pp. 258–263, 2019.

[30] G. Chen, Y. Zhan, G. Sheng, L. Xiao and Y. Wang, "Reinforcement learning-based sensor access control for WBANs," *IEEE Access*, vol. 7, pp. 8483–8494, 2019.

[31] M. Ghermezcheshmeh, V. S. Mansouri and M. Ghanbari, "Analysis and performance evaluation of scalable video coding over heterogeneous cellular networks," *Computer Networks*, vol. 148, no. 9, pp. 151–163, 2019.

[32] A. E. Omer, M. S. Hassan and M. E. Tarhuni, "An integrated scheme for streaming scalable encoded video-on-demand over cr networks," *Physical Communication*, vol. 35, no. 2, pp. 1–11, 2019.

[33] M. Lahby, A. Essouiri and A. Sekkaki, "A novel modeling approach for vertical handover based on dynamic k-partite graph in heterogeneous networks," *Digital Communications and Networks*, vol. 5, no. 4, pp. 297–307, 2019.

[34] M. M. Hassan, I. K. T. Tan, B. Selvaretnam and K. H. Poo, "SINR-based conversion and prediction approach for handover performance evaluation of video communication in proxy mobile ipv6," *Computers and Electrical Engineering*, vol. 74, no. 8, pp. 164–183, 2019.

[35] M. A. Samsuden, N. M. Diah and N. A. Rahman, "A review paper on implementing reinforcement learning technique in optimising games performance," in *Proc. of IEEE Int. Conf. on System Engineering and Technology*, Shah Alam, Malaysia, pp. 258–263, 2019.