phdprojectsorg@gmail.com
+91 98946 59122

NEXTGEN-6G-QoE: DESIGN A NOVEL FRAMEWORK FOR ENHANCING REAL-TIME VIDEO QUALITY OF EXPERIENCE IN NEXT-GENERATION WIRELESS NETWORKS

Introduction

The development of 360° Video streaming applications has created an increased demand for end-user experience or full "Quality of Experience (QoE)" in mobile wireless networks rather than simply a "Quality of Service (QoS)" framework based upon predetermined factors such as bandwidth, latency, etc… When used in conjunction with other forms such as Edge Computing and Artificial Intelligence, the synergy between these technologies will allow users to have an enhanced immersive viewing experience while simultaneously providing users with video quality greater than what is presently available through today's mobile and Wi-Fi networks. There has been much recent advancement in terms of new QoE-aware technologies for video streaming, including but not limited to tile-based adaptive video streaming systems, viewport prediction techniques, cross-layer optimization strategies, Deep-learning Based Video Quality Estimations, and Reinforcement Learning Based Rate Adaptation Methods. Although there have been improvements to performance characteristics associated with video delivery capabilities through these various products, they also share many limitations. For example, multi-tiered tile-based architectures frequently have degradation at the viewport level and most frame-level QoE optimization can't cope with rapidly changing networks. Most models for predicting QoE today, unfortunately, do not perform effectively in dynamic conditions such as those created when users do not remain in one location or when users experience complicated time-spatial patterns of distortion. Additionally, methods used today to adapt QoE prediction models using reinforcement learning algorithms generally exhibit considerable computational complexity and do not typically permit quick adaptations, making it more challenging to use in dense or mobile network contexts. The following text describes an adaptive video quality enhancement architecture that meets the requirements of future 6G networks using both reinforcement and deep-learning techniques. The architecture uses Swin-T as the backbone for spatiotemporal feature engineering. Hence, the Swin-T-based framework can accurately identify frame-level spatial and temporal distortions. This enhances the accuracy and robustness of QoE prediction under highly dynamic video and network conditions. Based on the predicted QoE, a novel DRL-Guided ARARAT framework is used to dynamically optimize bitrate selection, resource allocation, and network slicing in real time. To enhance the performance of the video quality, a hybrid VOT-CNN-QOE framework is used to enhance the QOE video prediction. To develop QoE-aware adaptation policies, the DRL agent will continuously monitor the state of the cross-layer network, including Video Quality Metrics as well as User Playback Conditions. With this information, the DRL Agent will jointly optimize all their "Quality of Experience Metrics (QoE)" of Video Quality, Latency, Buffering Behavior, and Bandwidth Utilization. This will improve Playback Stability and efficiently utilize Network Resources in the presence of Bandwidth Fluctuations and User Mobility. In general, the proposed methodology effectively addressed issues associated with low video QoE, predicting QoE, spatio-temporal complexity, and limited dynamic adaptability.

1.1 Research aim & scope

The purpose of this research is to design and develop a next generation based enhancing video quality of experiences to optimize the video quality by predicting the video performance in dynamic environments.

1.2 Research Objectives

- To design a model for scalable multiple layers of frames to encode each tile and enhance the quality of tile layers.
- To enhance the QoE for frame by frame basis of video transmission.
- To provide a high video QoE prediction and significantly improve the performance of both mobile devices and personal computers.
- For better capture of inter frames in video, and also compute the differences of temporal distortions.
- To adapt to dynamic environments, the quality of video and superior performance should be high.

2. Problem Statement

2.1 Specific Problem Statement

Reference 1

Title: "VRCT: A Viewport Reconstruction-Based 360◦ Video Caching Solution for Tile-Adaptive Streaming"

Concept

They propose a method for caching 360° video based on a "viewport-reconstruction approach (VRCT)". VRCT, as outlined in the research, utilizes a cache-management approach based on user QoE. The proposed work offers a QoE-driven trigger mechanism for determining if and when reconstruction should occur based upon the conditions of the cache and the network.

Problem Defined

However, it is difficult to construct the viewport with different layers. For this video, the quality of the video has decreased.

Solution

In this paper, we have proposed a scalable video encoding for multiple layers. The proposed framework has enhanced the quality of experiences and significantly improved the viewport quality.

Reference 2

Title: "Quality of Experience Oriented Cross-layer Optimization for Real-time XR Video Transmission"

Concept

The goal is to enhance the user's "quality of experience (QoE)" through the development of a cross-layer transmission framework for real-time XR video. The optimization of cross-layer transmission for near real-time XR video transfer from XR servers and BS is presented in this paper. To facilitate the exchange of resources between a wireless BS and XR servers, an integrated cross-layer transmission framework that aids in providing feedback to the wireless BS regarding their scheduling of available wireless resources and dynamically adjust bitrate to maintain acceptable quality.

Problem Defined

However, is that it is still very difficult to dynamically modify video frames and optimize the QoE under changing network conditions.

Solution

We have proposed LSTM base Deep-Q network for video adaptive streaming for video transmission in frame by frame basis, and the proposed model enhances the QoE.

Reference 3

Title: "Optimizing Mobile-Friendly Viewport Prediction for Live 360-Degree Video Streaming"

Concept

Due to the increased number of people using multimedia devices day-to-day as well as the advancements made in immersive media technologies, VR, and AR, 360-degree video content is becoming increasingly popular. With this new technology, users are able to experience their environment from a 360-degree perspective, which is quite different from traditional video formats. However, due to the increased amount of information contained within 360-degree video, they require much larger amounts of bandwidth to be streamed than traditional video formats.

Problem Defined

In video streaming, bitrate adaptive, they have achieved all and enhanced the quality. But still, it has a limitation in very poor prediction performance.

Solution

To tackle this issue, we have proposed a hybrid method, "A hybrid VOT-Driven CNN-Based Perceptual QoE Prediction for Real-Time Video Streaming" for next-generation 6G networks based on video prediction performance enhancement.

Reference 4

Title: "Blindly Assess Quality of In-the-Wild Videos via Quality-aware Pre-training and Motion Perception"

Concept

This paper proposes a DNN-Based BVQA Method in the in-the-Wild. It implemented both model transfer learning and both spatial appearance and temporal motion domains to take advantage of the information derived from the source models and to utilize it in our BVQA approach.

Problem Defined

However, there is a limitation in the spatio-temporal interaction approach, which is still a difficult undertaking due to the restriction of the current model.

Solution

This work proposes a new and effective VQA method utilizing transformers, allowing for the better differentiation of temporal variations to better capture temporal distortions through differentiating temporal variations in the VQA process by utilizing an approved Swin T backbone model and through finding temporal differences in extracted features and finding a regression model for those features to produce qualities of frames that are aware of temporal distortion.

Reference 5

Title: “Real-time rate control of WebRTC video streams in 5G networks: Improving quality of experience with Deep Reinforcement Learning”

Concept

The purpose of this article is to describe a Model-Free Deep Reinforcement Learning approach to optimize User Experience by managing the transmitted Media Stream Data Rate to the uplink.

Problem Defined

However, due to dynamic environments, it has a complex approach to reach the video-enhanced quality.

Solution

To tackle this issue, we have proposed novelty as “A Novel DRL-Guided ARARAT Approach for Adaptive Video Quality Enhancement” for dynamic adaptability approaches, and also enhance the video quality.

2.2 Overall Problem Statement

In this paper, the present issues were video frame distortion; spatio-temporal challenges, poor QoE performance, and decreasing quality are mentioned below.

- **Decreasing QoE**: Existing research suffers from reduced quality of experience when tiles with different quality versions or layered representations are combined within the same viewport, leading to visible inconsistencies and degraded perceptual video quality.
- **Insufficient QoE Optimization:** In previous research that relies on frame-level video adjustment often fails to achieve sufficient QoE optimization, as fine to maintain consistent video quality under dynamic network conditions.
- **Poor QoE prediction:** In existing QoE prediction approaches, frequent quality variations and unstable video characteristics significantly degrade prediction accuracy, resulting in poor QoE estimation accuracy, resulting in poor QoE estimation performance.
- **Challenging in strategy:** Existing research faces significant challenges in designing effective strategies due to the complex interaction between spatio-temporal processes, making it difficult to accurately model and optimize video QoE in real-time environments.
- **Complexity in dynamic adaptability:** In previous research, exhibit high complexity and insufficient dynamic adaptability, as they struggle to continuously adjust video quality in response to rapidly changing networks.

3. Proposed Methodology

The significance of this research is to enhance the QoE of video in dynamic networks and provide secure video transmission and quality in next generation networks system. The proposed methodology for the wireless network communication system is given below. Fig. 1 depicts the overall proposed architecture.

- Network Environment Configuration
- Data Collection & Preprocessing
- Feature Engineering
- Deep learning based QOE Prediction
- Adaptive Resource and Network Slice Allocation for QoE Optimization

Network Environment Configuration

In this work, a **QoE-aware AI-driven 6G network environment configuration method** is proposed to support real-time video streaming under highly dynamic conditions. The network is modeled as a heterogeneous 6G architecture consisting of sub-THz and mmWave access points integrated with edge computing nodes. Unlike conventional 4G/5G systems that rely on static QoS-based configurations, the proposed method dynamically configures access selection, network slicing, and resource availability based on real-time QoE predictions. Edge intelligence is employed to execute QoE estimation and deep reinforcement learning-based control, enabling ultra-low latency decision-making. A dedicated video network slice is adaptively adjusted in response to predicted QoE degradation, while realistic network impairments such as bandwidth fluctuation, mobility-induced handovers, and packet loss are continuously introduced and monitored. This closed-loop QoE-driven configuration enables autonomous, adaptive, and user-centric network behavior, making it suitable for next-generation 6G video services.

Data Collection & Preprocessing

A consistent collection and preparation approach for data collection has been created in order to facilitate the gathering of both the dynamics and user's perception of the quality of video over the next generation of wireless systems, through use of the "NextGen-6G-QoE" framework. The data is gathered from **"6G Global Virtual Multilingual Education Dataset"** including duration, protocols, and service. Real-time streams with Adaptative bitrate profiles, diverse resolutions, and encoding formats are transmitted over emulated 6G-oriented network environments characterized by ultra-low latency, high mobility, and fluctuating bandwidth. After collecting data from the sources, the data has been preprocessed, Data preprocessing will include extracting and resizing the frames, normalizing all visual and network features, minimizing noise and artifacts added by compression or during recording, handling missing data, and creating labeled datasets that can be utilized to increase the precision and efficacy of QoE prediction algorithms to prepare video-based systems for video processing.

Feature Engineering

Once the data collection and preprocessing have been completed, it has been sent to the feature extraction process to process the data to deal with missing values, to remove irrelevant data, and to choose the precise data for the subsequent phase of examining the data. The approach proposed in this paper is the use of the "**Swin T-backbone method-based feature selection"** for differentiating the spatio-temporal distortions in video capturing in the next generation of 6G. In the feature engineering plays a critical role in accurately modeling real-time video quality of Experience under highly dynamic next-generation wireless conditions. Multi-domain features are extracted and transformed from the network, video content, user device, and application layers to capture the complex interactions affecting perceived video quality. This engineered feature set enables robust learning of QoE patterns and supports adaptive, real-time video optimization in next-Generation 6G wireless networks.

Deep learning based QOE Prediction

A Hybrid method is proposed, namely **"A 6G-Enabled Hybrid VOT-Driven CNN Framework for Human-Centric Perceptual QoE Prediction in Real-Time Video Streaming"** for next-generation 6G networks based video prediction performance enhancement. The hybrid method is the combination of "**CNN-QOE** (Convolutional neural network)", and "**VOT** (Video Object Tracking)". The framework is built utilizing low-latency processing technology that makes it possible to deploy using Edge assistance in effectively creating video streaming systems that are enabled by the use of 6G technologies to provide a better experience from the user and application perspectives by utilizing the VOT scores to help with QoE-based adaptation. Therefore, an accurate method of maintaining reliable dynamic transmission parameter optimization as the networks change will be enhanced as a result of this unique VOT-based approach for accurately predicting QoE performance, thus providing a scalable and user-centric architecture that dramatically improves video prediction results as compared to the legacy frame-based or standalone CNN-QoE method in future wireless networks.

Adaptive Resource and Network Slice Allocation for QoE Optimization

A Novel method is proposed, namely **"QoE-Driven DRL-Guided ARARAT for Adaptive video quality Enhancement in Next- Generation 6G Networks"** for Next-Generation 6G networks based Video quality enhancement. The novel method is the combination of "**DRL** (Deep Reinforcement Learning)" and "**ARARAT** (A Collaborative Edge-Assisted Framework for HTTP Adaptive Video Streaming)". An RL agent continuously observes cross-layer networks and video states and learns QoE-aware policies that adapt bandwidth, scheduling priority, and slice allocation resources in response to fluctuating channel conditions and traffic demands. By optimizing QoE-centric that accounts for video quality, latency, and buffering, the framework ensures efficient slice allocation utilization while maintaining stability and enhancing real-time video quality of Experience in next-generation wireless environments.

Diagram

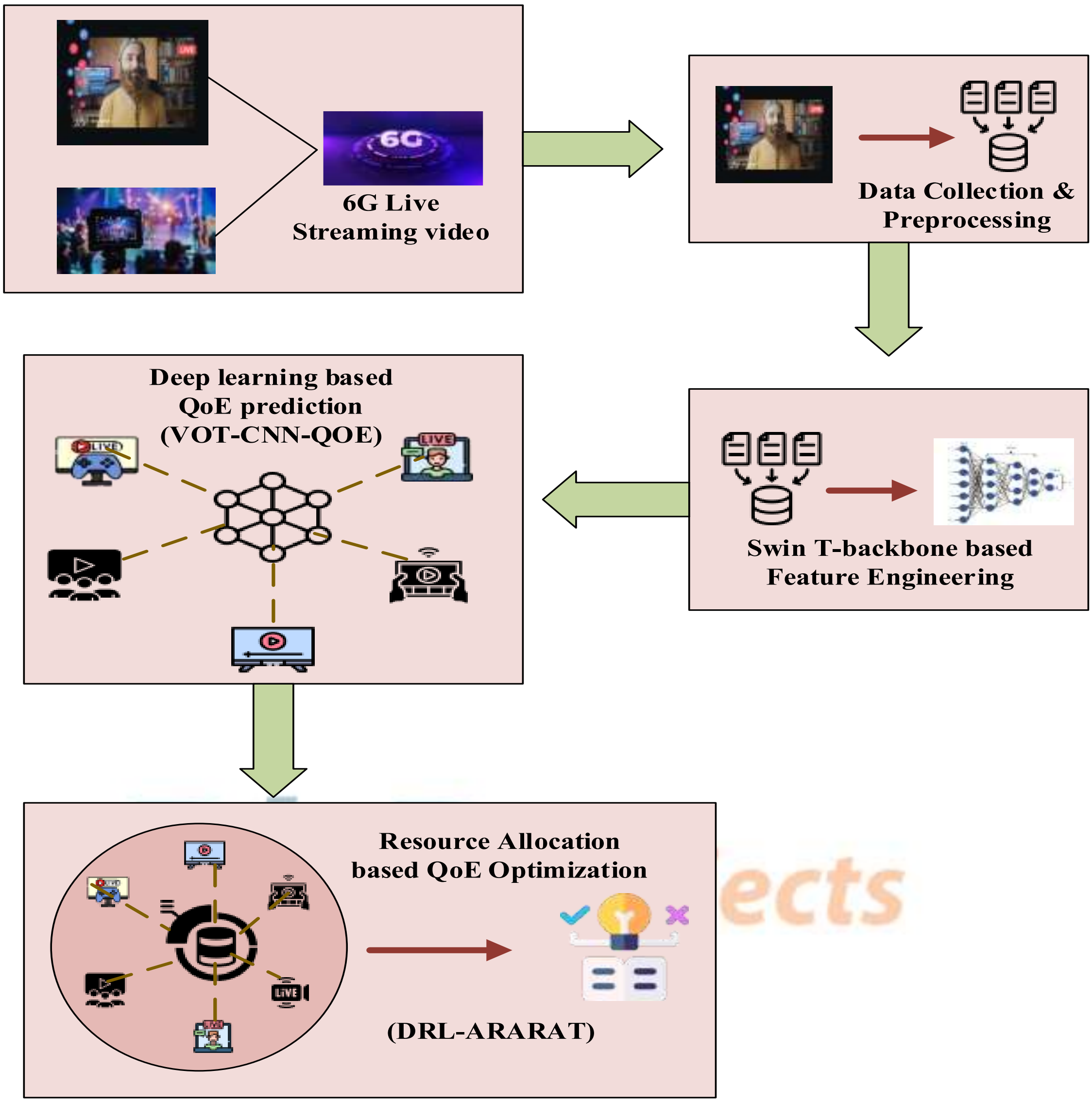

Fig 1: Overall proposed architecture

Research Highlights

- When the data is collected from sources, the collected data is used to start preprocessing.
- By using **"Swin T-backbone"** in the feature extraction method and differentiates the variations between spatio-temporal distortion.
- In order to improve the experience regarding the viewing of online videos, the article proposes a hybrid approach for its "Quality of Experience (QoE)" assessment and

performance enhancement called **“VOT-CNN QoE”**. For the purpose of predicting and improving video quality experience.

- To enhance the optimization and allocation for dynamic adaptability, we have proposed a novel method named **“DRL-ARARAT”**. This method is used to adapt to dynamic environments and optimize the video quality of the experience.

Reference 6

Title: “Improving quality of experience in adaptive low-latency live Streaming”

Concept

The latency stems from the fact that the client is receiving the content in segments rather than a continuous stream. The amount of data that must be buffered by the client will need to be large enough to withstand fluctuations in network throughput and provide for continuous playback of the content without interruption. To optimize the “Video Quality (VQ)” and minimize potential stalls during playback, the client utilizes an “Adaptive Bitrate (ABR)” algorithm to determine what quality to request for each segment of the video to achieve a balance between VQ and stall avoidance to enhance the “Quality of Experience (QoE)” of the user.

Limitation

One critical limitation of this proposed approach is that while it can model how users are moving between different sets of networks at higher speeds, it does not have the necessary capability to accurately represent very rapid fluctuations in bandwidth, these types of unpredictable or intermittent connections are especially common for mobile/wireless users, and when this happens, it can negatively impact the quality of experience for the end-user.

Reference 7

Title: “Improving Video Quality by Predicting Inter-Frame Residuals Based on an Additive 3D-CNN Model”

Concept

This paper describes a new way to improve the quality of videos with a 3D “Convolutional Neural Network (CNN)”. Basically, this new type of CNN receives five frames at a very low bitrate. Based on the first and fifth frames, this new type of CNN then predicts whether the frames were improperly compressed. Finally, this new CNN can reconstruct the first frame into an improved version of high-quality video. The additive CNN 3D model can predict the amount of inter-frame redundancy eliminated by compression.

Limitation

A significant drawback of the suggested additive 3D-CNN architecture is its extremely high computational requirements. Utilizing spatiotemporal 3D convolutional kernels adds considerable time for processing as well as a large amount of memory usage, which further diminishes applicability to mobile and edge devices that have very limited resources and thus are unable to use the networks in real-time operations.

Reference 8

Title: "Multi-neural network based tiled 360◦ video caching with Mobile Edge Computing"

Concept

The research described in this paper proposed a method for storing parts of a 360 degree video in multiple local servers based on the most commonly viewed areas of content rather than all areas of content being Cache Location Best Testing Tool cached. Previous work had relied on data from the viewing habits of users to predict which areas of video content they were most likely to look at in the future. In contrast to previous studies, the proposed method utilized deep learning algorithms to classify tiles based solely upon user characteristics rather than viewing statistics of prior videos.

Limitation

Using multiple neural networks to predict tile popularity, provide user viewport estimates, and manage cache results in a high level of computational and memory consumption at the edge.

Reference 9

Title: "QoE-Aware Analysis and Management of Multimedia Services in 5G and Beyond Heterogeneous Networks"

Concept

An initial analysis of Quality of Experience in multimedia services over 5G will be presented followed by a discussion of the weaknesses in Mobile Networks. A framework will be developed to improve Quality of Experience based on the findings collected from the first two analyses to facilitate the seamless delivery of personalized immersive video experiences and the ability to interact with these videos from anywhere, at any time and using any device.

Limitation

The proposed analysis does not adequately account for users' mobility, continuous network changes, and the frequent transferring of users between access points as well as variations in traffic loads, all of which are constant throughout daily operations. Due to the inability to fully capture the continually changing nature of QoE variables, this will effect the accuracy of all QoE modelling efforts and subsequently all QoE-related management decisions, thereby decreasing the ability to operate successfully in real time.

Reference 10

Title: "PPO-ABR: Proximal Policy Optimization based Deep Reinforcement Learning for Adaptive Bitrate streaming"

Concept

They propose a new type of method for creating "Adaptive Bit Rate (ABR)" systems, namely, a Deep Reinforcement Learning method focused on "Proximal Policy Optimization (PPO)", known as PPO-ABR. The method is designed to improve the "Quality of Experience (QoE)" for a streamer's viewing experience by increasing both the sampling efficiency of the algorithm and the number of epochs of minibatch updates while implementing a clipped probability ratio between both the new and the old policies.

Limitation

The major downside to using PPO-based ABR Models as a training strategy is that they are very complex and require a lot of attention to select appropriate hyperparameters. Additionally, these models will take significantly longer to converge reliably when being trained on video data, due in part to the necessity of hyperparameters with PPO policy updates.

Reference 11

Title: "Progressive Frame Patching for FoV-based Point Cloud Video Streaming"

Concept

The purpose of this study was to make an update to create a point cloud video for use on a video streaming website. Created our point cloud video with multiple iterations that improved over time. Rather than streaming individual frames, they used features of scalability in point cloud coding when using an octree structure to stream all frames that were within the same timeframe at once.

Limitation

The accuracy of the prediction of the "Field of View (FoV)" is a major limiting factor for this progressive frame patching method; if the FoV estimation is not accurate or is delayed, the patches that are sent are likely to be irrelevant and therefore will waste the available bandwidth, causing a resultant decrease in the overall user QoE.

Reference 12

Title: "Improving UE Energy Efficiency through Network-aware Video Streaming over 5G**"**

Concept

EnDASH-5G, which synchronizes data downloads through a network-aware system, is an evolution of the MPEG - DASH "Adaptable Bit Rate (ABR)" Streaming Protocol that is widely used by today's internet video streaming services. Examples include Amazon Prime and Netflix.

Limitation

A major drawback of the suggested method is that using frequent communication of network condition state information brings added complexity to both control and signal transmission operations. This added complexity could somewhat offset any expected energy savings, especially in cases of highly dynamic network states where information will need to be sent more often, thus reducing the potential efficiencies gained from the proposed solution.

Reference 13

Title: “QoE-Driven Optimization in 5G O-RAN Enabled HetNets for Enhanced Video Service Quality”

Concept

This article will propose “QoE Enhancement Function (QoE2F)”, a new app that enhances its capabilities by efficiently allocating resources to end customers who want high-resolution video services.

Limitation

The QoE optimization models may not efficiently scale in Dense HetNet deployments that have a high number of users, as well as many different cells and ways of varying traffic.

Reference 14

Title: “Lavie: high-quality video generation with cascaded latent diffusion models”

Concept

The authors propose the development of a unified platform for the generation of video through a cascaded framework of video latent diffusion models. These will consist of a T2V base video generation model, the original model, a temporal interpolation model, and a video resolution enhancement model for super-resolution.

Limitation

Although extensive cascading increases the quality of video sequences, long video sequences remain difficult to maintain in terms of temporal coherence over time.

Reference 15

Title: “Resource Allocation Algorithm for Sensing Video Transmission over Cell-Free Radio Access Networks”

Concept

This research presents a way of identifying the best MCS selection procedure for QoE in large-scale, real-time Media Services, for example, 3D Point Cloud Videos. Because of the high amount of transmitted data in large-scale 3D point cloud video, packet losses are difficult to recover from and have a major effect on QoE.

Limitation

That said, future work will focus on extending models to have better communication efficiency while also improving visual QoE for users. In addition, semantic communication is looking to increase integration through more semantic-aware resource mapping, cross-layer optimization, and distributed decision making about industrial Metaverse and AR/VR applications.

Reference 16

Title: "PreCNet: Next-Frame Video Prediction Based on Predictive Coding"

Concept

PreCNet was developed upon a well-supported and widely used framework, and optimized to perform extremely well on an urban robotic driving dataset, which had a strong base established in the time before the introduction of the Urban Summer Project.

Limitation

From our comparison of PreCNet's quantitative performance against other architectures, it appears that PreCNet performs worse than the other architectures in predicting the next 15 frames. They believe that the architecture that uses a recurrent approach, where the outputs of a previous prediction are used as inputs for the next time step, will not necessarily be the best architecture for predicting the next frame, and also for predicting multiple frame predictions at one time.

Reference 17

Title: "Real-time video anomaly detection for smart surveillance"

Concept

This paper describes a fully automated surveillance system developed by combining background subtraction, convolutional autoencoders, and object detection. In BS, the pixels of an image are modeled as a mixture of Gaussians. The results of BS were used as input to convolutional autoencoders for the purpose of filtering out abnormal events from the normal events in real time. The results of this process are that the system can automatically detect and classify signs of violence and other threats in the images captured by the surveillance camera.

Limitation

Detecting anomalies while they happen in near real-time requires endlessly processing digital video, which will require many resources with much higher computational complexity than when processing still images, and this increases the hardware required to deploy for edge and low-power computing applications.

Reference 18

Title: “Optimizing Prediction of YouTube Video Popularity Using XGBoost”

Concept

The XGBoost model was used in this research to predict how popular videos would be based on several features, including selecting appropriate features to use in the model, fusing multiple sources of information, using min-max normalization, and adjusting certain hyperparameters like gamma, eta, learning rate, etc. The model produced an overall accuracy rate, additionally enhanced the Tuned Xgboost model's performance by increasing its overall accuracy and Precision rates, and also performed an analysis comparing our findings with other literature regarding the classification of unpopular videos.

Limitation

XGBoost does not include in itself a model that can represent how things have changed through use over time. As a result, this is not sufficient to capture how an asset will gain or lose value over time.

Reference 19

Title: “QoE-based Mobility-aware Collaborative Video Streaming on the Edge of 5G”

Concept

An innovative collaborative QoE-based video streaming solution for end-user mobility has been proposed and implemented on multiple "MEC (Mobile Edge Computing)" servers. Our implementation of this solution allows for maintaining the quality of experience for all end users throughout their entire video session.

Limitation

The extent to which a mobility-aware optimization system can improve service quality depends critically on the availability of accurate information regarding the predicted movement of users. Failure to correctly predict users' future movements results in suboptimal use of available resources and ultimately leads to a decrease in service quality from a user's perspective.

Reference 20

Title: "Reinforcement Learning-Based Rate Adaptation in Dynamic Video Streaming"

Concept

This research presents DQNReg, a Q-Network deep reinforcement-learning designed to improve deep Q-learning, which is the basis for all existing RL algorithms. By integrating QoE into the design of the reward function based on segments, provide an adaptive learning process that allows the algorithm to converge on maximizing the QoE outcome. To illustrate this method of adaptation using RL-based adaptation, have used a case study to conduct a trace-based simulation on both a standard wireless local area network and a fifth-generation 5G mobile network.

Limitation

It takes a lot of training to develop or train a better reinforcement learning model. Training is time-intensive and requires significant amounts of computing resources.

References

1. Ye, Z., Li, Q., Ma, X., Zhao, D., Jiang, Y., Ma, L., ... & Muntean, G. M. (2023). VRCT: A viewport reconstruction-based 360 video caching solution for tile-adaptive streaming. *IEEE Transactions on Broadcasting*, *69*(3), 691-703.

2. Pan, G., Xu, S., Zhang, S., Chen, X., & Sun, Y. (2024). Quality of experience-oriented cross-layer optimization for real-time XR video transmission. *IEEE Transactions on Circuits and Systems for Video Technology*, *34*(8), 7742-7755.
3. Zhang, L., Chen, P., Zhang, C., Pan, C., Long, T., Xu, W., ... & Liu, J. (2025). Optimizing Mobile-Friendly Viewport Prediction for Live 360-Degree Video Streaming. *IEEE Transactions on Mobile Computing*.
4. Li, B., Zhang, W., Tian, M., Zhai, G., & Wang, X. (2022). Blindly assess the quality of in-the-wild videos via quality-aware pre-training and motion perception. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(9), 5944-5958.
5. Smirnov, N., & Tomforde, S. (2024). Real-time rate control of WebRTC video streams in 5 G networks: Improving quality of experience with deep reinforcement learning. *Journal of Systems Architecture*, *148*, 103066.
6. Lyko, T., Broadbent, M., Race, N., Nilsson, M., Farrow, P., & Appleby, S. (2024). Improving the quality of experience in adaptive low-latency live streaming. *Multimedia Tools and Applications*, *83*(6), 15957-15983.
7. Azadegan, H., & Shirazi, A. A. B. (2023). Improving video quality by predicting inter-frame residuals based on an additive 3D-CNN model. *Journal of Visual Communication and Image Representation*, *90*, 103734.
8. Kumar, S., Bhagat, L., & Jin, J. (2022). Multi-neural network-based tiled 360° video caching with mobile edge computing. *Journal of Network and Computer Applications*, *201*, 103342.
9. Sultan, M. T., & El Sayed, H. (2023). QoE-aware analysis and management of multimedia services in 5G and beyond heterogeneous networks. *IEEE Access*, *11*, 77679-77688.
10. Naresh, M., Saxena, P., & Gupta, M. (2023, June). Ppo-abr: Proximal policy optimization-based deep reinforcement learning for adaptive bitrate streaming. In *2023, International Wireless Communications and Mobile Computing (IWCMC)* (pp. 199-204). IEEE.
11. Zong, T., Mao, Y., Li, C., Liu, Y., & Wang, Y. (2025). Progressive frame patching for FoV-based point cloud video streaming. *IEEE Transactions on Multimedia*.
12. Palit, B., Sen, A., Mondal, A., Zunaid, A., Jayatheerthan, J., & Chakraborty, S. (2023). Improving UE energy efficiency through network-aware video streaming over 5g. *IEEE Transactions on Network and Service Management*, *20*(3), 3487-3500.

13. Agarwal, B., Togou, M. A., Ruffini, M., & Muntean, G. M. (2022). Qoe-driven optimization in 5 G o-RAN-enabled HetNets for enhanced video service quality. *IEEE Communications Magazine*, *61*(1), 56-62.
14. Wang, Y., Chen, X., Ma, X., Zhou, S., Huang, Z., Wang, Y., ... & Liu, Z. (2025). Lavie: High-quality video generation with cascaded latent diffusion models. *International Journal of Computer Vision*, *133*(5), 3059-3078.
15. Liu, C., Yu, L., Sha, L., Zhang, C., & Wang, D. (2025). Resource Allocation Algorithm for Sensing Video Transmission Over Cell-Free Radio Access Networks. *IEEE Open Journal of the Communications Society*.
16. Straka, Z., Svoboda, T., & Hoffmann, M. (2023). PreCNet: Next-frame video prediction based on predictive coding. *IEEE Transactions on Neural Networks and Learning Systems*, *35*(8), 10353-10367.
17. Ali, M. M. (2023). Real-time video anomaly detection for smart surveillance. *IET Image Processing*, *17*(5), 1375-1388.
18. Nisa, M. U., Mahmood, D., Ahmed, G., Khan, S., Mohammed, M. A., & Damaševičius, R. (2021). Optimizing prediction of YouTube video popularity using XGBoost. *Electronics*, *10*(23), 2962.
19. Tuysuz, M. F., & Aydin, M. E. (2020). QoE-based mobility-aware collaborative video streaming on the edge of 5G. *IEEE Transactions on Industrial Informatics*, *16*(11), 7115-7125.
20. Hafez, N. A., Hassan, M. S., & Landolsi, T. (2023). Reinforcement learning-based rate adaptation in dynamic video streaming. *Telecommunication Systems*, *83*(4), 395-407.