Sparse ConvNet and Policy Driven Encoder for Enhancing Cloud-based QoE in 5G and Beyond Networks

Tejomurtula Himaja[1], Dr. S Neeraja[2], Dr. Suresh Chittineni[3]

[1]Department of Electrical, Electronics & Communications Engineering, Gitam Institute of Technology, Gitam Deemed to Be University.
[2]Associate Professor, Department of Electrical, Electronics & Communications Engineering, Gitam Institute of Technology, Gitam Deemed to Be University.
[3]Professor, Department of Computer Science and Engineering, Gitam Institute of Technology, Gitam Deemed to Be University.

KEYWORDS

Cloud computing, video streaming, Quality of Experience, Quality of Service, Video encoding, QoE Management

ABSTRACT

Optimizing video streaming quality for extended-duration videos in cloud-based Quality of Experience (QoE) presents significant challenges due to the dynamic nature of network conditions, user engagement patterns, and resource allocation demands. Hence, a novel Sparse ConvNet and Policy Driven Encoder for Enhancing Cloud-based QoE in 5G and Beyond Networks is proposed for optimizing video streaming quality and resource allocation in real-time. Existing neural networks struggle fail to solve this data traffic issue, which increases computational burden and degraded QoE for users, this leads to customer's dissatisfaction. Thus, Sparse Graph Attention ConvNet (SGA ConvNet) is introduced to analyze the data traffic. This reduces computational complexity while optimizing network resources in real-time, ensuring a balance between operational costs and QoE. Moreover, traditional video encoding algorithms struggle with rapid scene changes and high motion, which increase data volume and computational intensity, reducing overall efficiency. To address this, Policy-Driven Variational Encoder (PD-VE) is introduced, enhancing adaptability to network conditions and improving the viewer's streaming experience by alleviating congestion and latency issues. The suggested method outperforms existing techniques in optimizing video streaming quality and resource allocation, achieving higher accuracy, precision, while minimizing RMSE and encoding time, as demonstrated by experimental results.

1. INTRODUCTION

Cloud computing has revolutionized IT, providing scalable and flexible resources over the Internet on a pay-as-you-go basis. In the realm of 5G and next-gen networks, cloud-based Quality of Experience (QoE) methods are vital for improving service delivery. They enhance reliability, reduce latency, and adjust network performance dynamically, all of which elevate user satisfaction. This integration supports intelligent traffic management, efficient content distribution, and dynamic resource allocation, making the user experience smoother. Interaction rate, reflecting user engagement with content through likes, shares, and comments, is a key indicator of success [1-3].
Streaming popular content with interactive features like live chats, polls, and Q&A boosts engagement by fostering community involvement. However, these features also introduce

challenges such as higher development costs, the need for ongoing technical support, and increased data traffic, which raises operating expenses [4-6].

Algorithms such as Dynamic Resource Allocation (DRA) and Proportional Fairness effectively manage server resources and bandwidth but struggle with handling traffic surges and preventing resource underutilization. Content Delivery Network (CDN) algorithms, including those for mapping and edge computing, enhance content distribution and reduce latency but still face challenges in real-time load balancing and server optimization. Additionally, Natural Language Processing (NLP) techniques like sentiment analysis help moderate user interactions but often struggle with language nuances and slang. The key is to balance cost, complexity, and performance to ensure an optimal user experience [7-9].

Complex video scenes, with high motion and detail, generate large files that can strain bandwidth and lead to buffering or quality issues, particularly in areas with weak internet connections. Long video streams also put a strain on device hardware, leading to overheating and performance lags. Maintaining consistent video quality and synchronization during long streams, especially for live or interactive content, is difficult. Overcoming these issues requires enhanced hardware, better network infrastructure, and optimized encoding techniques [10-12].

To address these challenges, video compression methods like H.264 and H.265 help reduce file sizes, though they can be computationally intensive, leading to longer encoding times and potential quality loss. Adaptive bitrate streaming adjusts resolution and bitrate based on network conditions but struggles with predicting the best settings for complex scenes over long durations. While these algorithms are crucial, they still face hurdles with scalability, quality retention, and content security [13-15].

1.1 Main Contribution

SGA ConvNet: A method that incorporates the Stream Bayes Change Point Identifier (BCPI) in its hidden layer to optimize video quality while minimizing computational demands. The sparse attention mechanism ensures efficient resource usage in real time, maintaining a balance between cost and QoE.

Policy-Driven Variational Encoder (PD-VE): Combined with the Diamond Rood Search Algorithm (DRS) and Adaptive Rood Pattern Search (ARPS), PD-VE dynamically adjusts streaming parameters, enhancing video quality and extending streaming durations while cutting costs.

These methods tackle the dual challenge of improving video quality and reducing computational load in interactive streaming. The paper is organized as follows: Section 2 reviews existing research and identifies gaps; Section 3 introduces the proposed methods and their algorithms; Section 4 evaluates the effectiveness of the approach through comparisons; and Section 5 concludes with a summary.

2. SPARSE CONVNET AND POLICY DRIVEN ENCODER FOR ENHANCING CLOUD-BASED QOE IN 5G AND BEYOND NETWORKS

A novel **Sparse ConvNet and Policy-Driven Encoder** is proposed to optimize network efficiency and preserve Quality of Experience (QoE) in 5G and beyond networks for extended-duration video streaming. As data traffic increases due to live chat and real-time feedback, managing bandwidth and minimizing latency is crucial to avoid congestion and rising operational costs. Traditional neural networks fail to address this due to complex calculations, degrading QoE. The Sparse Graph Attention ConvNet (SGA ConvNet) is introduced to analyze traffic and reduce operational costs. It integrates the Stream Bayes Change Point Identifier (BCPI) within its hidden layer to adjust data traffic dynamically, optimizing efficiency while reducing computational overhead. The SGA ConvNet uses real-time monitoring to detect and adjust to traffic changes, ensuring efficient resource allocation and enhanced QoE.

For high-bitrate streaming, the Policy-Driven Variational Encoder (PD-VE) is introduced, incorporating the Diamond Rood Search Algorithm (DRS) and Adaptive Root Pattern Search (ARPS) to optimize video encoding. This framework integrates with the Proximal Policy Optimization-based Variational Autoencoder (PPO-VAE), enabling continuous policy optimization to adjust bitrate distribution and visual quality. Real-time adjustments to network conditions mitigate issues like latency, congestion, and bandwidth throttling, improving the streaming experience.

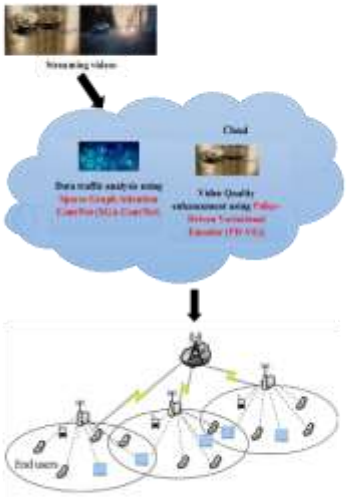


Figure 1. Overall block diagram of the proposed model

The architecture of the suggested model is shown in Figure 1. The process begins with cloud-based video streaming services catering to a group of end users. It integrates advanced data traffic analytics using SGA ConvNet to ensure real-time monitoring of data traffic, which enhances resource efficiency. Subsequently, the system employs the PD-VE to dynamically adapt video encoding techniques based on current network conditions, resulting in improved video quality. This continuous interaction between end users and cloud services guarantees the delivery of high-quality video streams. The flow illustrates how data analysis and enhancement techniques work together to optimize streaming for users, leading to better video quality and effective traffic management. Overall, this integrated approach significantly boosts user satisfaction and retention on the platform.

2.1.Sparse Graph Attention ConvNet

The SGA ConvNet addresses data traffic management and resource allocation in video streaming services by using real-time data analysis and probabilistic inference to detect and respond to changes quickly. By integrating BCPI, it identifies sudden shifts in traffic that indicate network congestion or spikes in user engagement, enhancing video quality while reducing computational strain.

The SGA ConvNet processes video streaming data, organizing it into a graph structure where each node represents data points and edges represent connections. The GNN analyzes the graph to identify traffic variations, as shown in equation (1):

$$T\left(S_i = n | Z_{i-1}, A_{current_{i-1}}, A_{current_i} = 1\right) = f_0(n) = sf.f(n) + \sum_{j \in t_{dp}} \emptyset_j . f_j(n) \quad (1)$$

Where S_i is the state at time i, $A_{current_i} = 1$ indicates the activation of the current context, Z_{i-1} is the previous state, and $\emptyset_j$ represents the weighted contributions from previous states. This analysis helps the system transition efficiently based on dynamic traffic patterns.

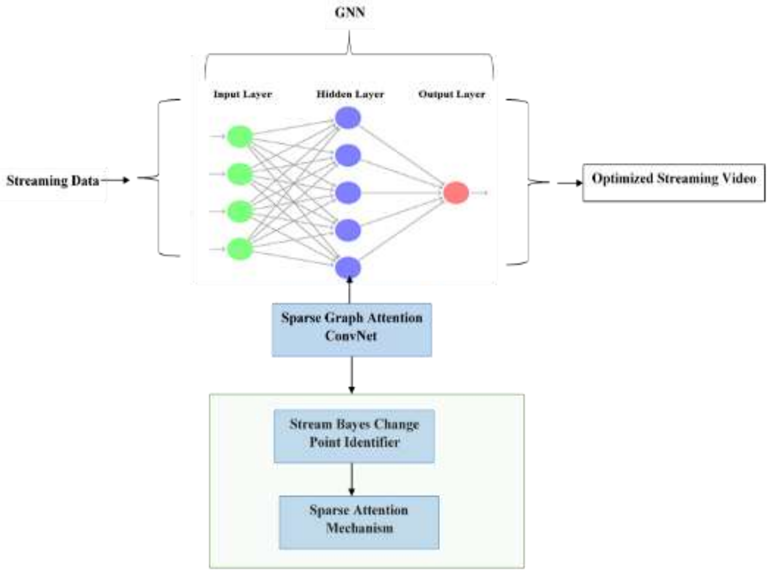


Figure 2: Structure of Sparse Graph Attention ConvNet

The architecture of the Sparse Graph Attention ConvNet (SGA ConvNet) begins by organizing streaming data through a Graph Neural Network (GNN), which captures relationships between data points. BCPI is integrated in the hidden layer to detect changes in data traffic dynamically, improving adaptability. The system uses Bayesian methods to detect and respond to shifts in network conditions and user demand in real-time, as shown in equation (2):

$$P(HT|S(t)) = \frac{\int_0^{t_{s_1}} \frac{e^{-\left(\log s - \mu_t^i(d_1)\right)^2}}{s \times \sigma_t^i(d) \times \sqrt{2\pi}} dx \times P(H)}{P\left(S(t)\right)} \quad (2)$$

Where $S(t)$ represents the observed data, $t.\ \mu_t^i$ and σ_t^i are the mean and standard deviation of the log-normal distribution, respectively. BCPI updates the probability of changes based on past and present data, ensuring optimized video streaming.

Next, BCPI detects critical changes in data traffic, represented by equation (3):

$$P(req_t|req_{t-1}) \begin{cases} HT(req_{t-1} + 1) & if\ req_t = 0 \\ 1 - HT(req_{t-1} + 1) & if\ \ req_t =\ req_{t-1} + 1 \\ 0 & Otherwise \end{cases} \quad (3)$$

This allows the system to detect abrupt shifts in traffic, improving responsiveness and resource allocation.

A sparse attention mechanism is then applied to optimize resource allocation by prioritizing key features, reducing computational overhead. The attention score is calculated as:

$$\alpha_j(n) = \frac{exp\left(f_j^T \times f_j(n)\right)}{\sum_{n \in \sqcup p(j)} exp\left(f_j^T \times f_j(n)\right)} \tag{4}$$

Where $\alpha_j(n)$ represents the attention score for feature $f_j(n)$, with $\sqcup\, p(j)$being the set of features related to j. The sparse attention mechanism is applied in equation (5):

$$H = \alpha_j(n) = Attention\left(Q_c, K_{rf}, V_a\right) = softmax \frac{Q_c K_{rf}^T}{\sqrt{d_n}} V_a \tag{5}$$

Here, Q_c is the context of the current request, K_{rf} represents historical features, and V_a represents the values that provide context to the attention mechanism. This optimization ensures computational efficiency while improving network performance.

The SGA ConvNet system, through Bayesian inference and sparse attention, efficiently adapts to dynamic traffic conditions, optimizing video streaming quality and user experience.

2.2.Policy-Driven Variational Encoder

The Policy-Driven Variational Encoder adjusts encoding strategies in real-time to optimize video quality based on network fluctuations and user requirements. It utilizes a variational autoencoder (VAE) to compress and encode video efficiently, dynamically allocating bitrate to maintain quality even with limited bandwidth. The encoder processes video frames using a DRS method, combining DSA for motion vector estimation with ARPS to refine the search for optimal encoding.

The DSA method estimates motion by calculating the Mean Absolute Difference (MAD) between reference and current frames using equation (6):

$$MAD = \frac{1}{BM^2} \sum_{i=0}^{BM-1} \sum_{j=0}^{BM-1} \left| Ref_f\left(x_{p_{mv}+i}, Y_{p_{mv}+j}\right) - Curr_f\left(x_{0+i}, Y_{0+j}\right) \right| \tag{6}$$

The motion vector cost is computed to evaluate the best candidate position, as defined in equation (7):

$$Cost_{total}(p_{mv}) = \sum_{i=0}^{BM-1} \sum_{j=0}^{BM-1} \left| Ref_f\left(x_{p_{mv}+i}, Y_{p_{mv}+j}\right) - Curr_f\left(x_{0+i}, Y_{0+j}\right) \right|^2 \tag{7}$$

After minimizing the cost, the optimal motion vector is identified using equation (8):

$$MV_{Ini} = \arg min_{p_{mv}} (Cost_{total}(p_{mv})) \tag{8}$$

ARPS refines the motion vector by adjusting the search pattern based on initial estimates, focusing on areas with significant motion, enhancing accuracy, and reducing encoding artifacts. The final motion vector is used by the encoder to predict current block content, improving compression efficiency.

The optimized motion vectors are incorporated into the VAE framework, which minimizes a combined loss function with reconstruction error and a regularization term. The VAE loss is expressed as:

$$Loss_{VAE} = |Residual|_2^2 + D_{KL}(q(z|x) \| p(z)) + \beta \left\| l - Ex(x) \right\|_2^2 \tag{9}$$

This approach ensures efficient video compression and enhances streaming quality while reducing bandwidth usage.

Figure 3 illustrates the Policy-Driven Variational Encoder (PD-VE) process. Video data is input and motion vectors are initially estimated, refined using DSA, and enhanced with ARPS. After ensuring vector accuracy, they are processed through the VAE framework. The VAE loss function is calculated, and the Proximal Policy Optimization (PPO) algorithm adjusts the encoding strategy in real-time to maintain high-quality video streaming even during bandwidth fluctuations, as shown in equation (10):

$$Loss_{PPO}(\theta) = Ex_t min\left[ra_t(\theta)\widehat{AE}_t, clip(ra_t(\theta), 1-\epsilon, 1+\epsilon)\widehat{AE}_t\right] \quad (10)$$

Where:
$Loss_{PPO}(\theta)$ denotes the PPO loss function.
Ex_t represents adaptability across time steps.
$ra_t(\theta)$ is the action efficacy under the new policy compared to the previous one.
$\widehat{AE}_t$ is the advantage estimate, guiding policy improvement.
The clipping method $clip(ra_t(\theta), 1-\epsilon, 1+\epsilon)$ stabilizes training by restricting policy updates.
This dynamic adaptation allows for accurate motion vector representation, minimizing encoding artifacts. The system continuously refines encoding strategies, improving compression and streaming quality.

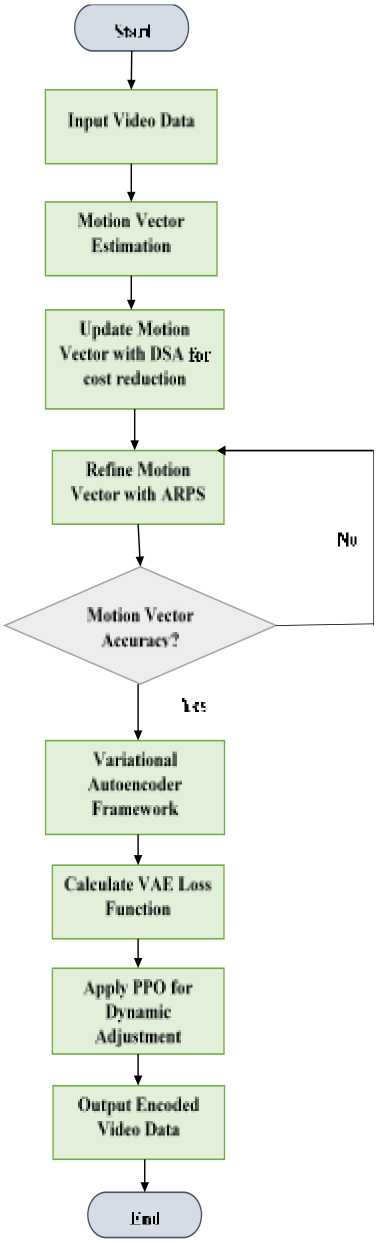


Figure 3: Flowchart of Policy-Driven Variational Encoder

3. RESULTS AND DISCUSSION

This section evaluates the performance of the recommended method, highlighting its effectiveness in optimizing resource allocation and improving video quality for cloud-based QoE in streaming applications. A comparison analysis shows how the integrated strategy ensures better performance and a seamless user experience by addressing network congestion and bandwidth throttling during long-form video streaming.

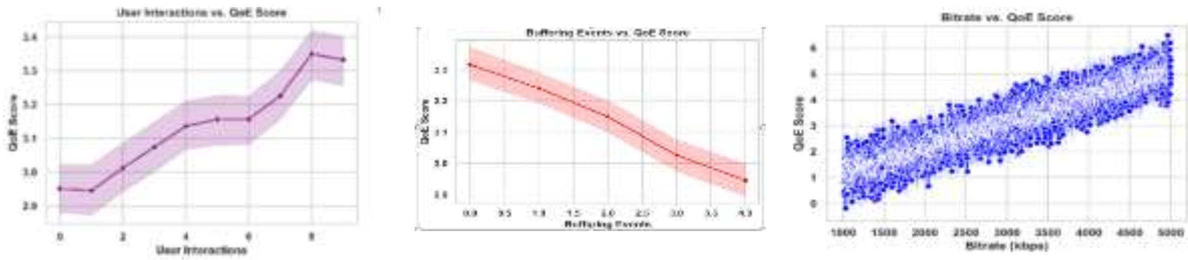
3.1. System configuration

The suggested approach has been modelled using Python, and tests have been done to fully evaluate its performance and efficacy in improving resource allocation and video streaming quality by adjusting user demand fluctuations and system factors.

Software	: Python
OS	: Windows 10 (64-bit)
Processor	: Intel i5
RAM	: 8GB RAM

3.2. Simulated output of the suggested model

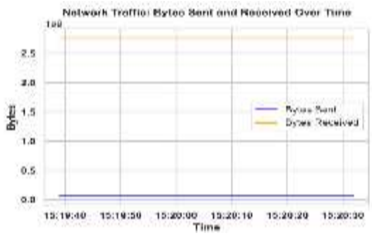
This section explains the simulated output of the model using Sparse ConvNet and Policy-Driven Variational Encoder to optimize resource allocation and video streaming quality in cloud-based QoE systems.

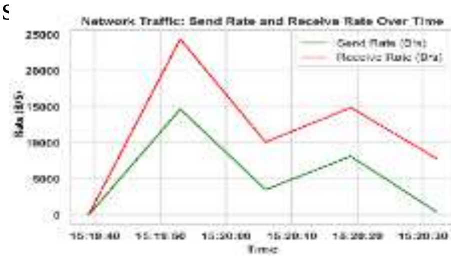


(a) (b) (c)

Figure 4: (a) QoE Score Vs User interface, (b) Buffering events vs QoE score, (c) Bitrate vs QoE score of suggested models

Figure 4 shows how user interactions boost QoE scores, with sparse attention improving engagement and resource allocation (a). It also emphasizes the importance of minimizing buffering, as more buffering reduces QoE scores (b). Lastly, it highlights the correlation between bitrate optimization and improved QoE, s itrates (c).

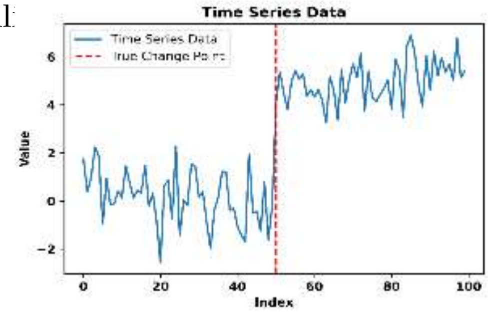




(a) (b)

Figure 5: Network Traffic (a) Bytes sent and received over time, (b) Sent Rate and receive rate over time of suggested model

Figure 5 shows network traffic dynamics: (a) stable bytes sent (~1e9) and received (~2e9) over time, indicating efficient data transmission; (b) varying send and receive rates, with the send rate peaking at 25,000 B/s and the receive rate showing inverse fluctuations. The results highlight the need for monitoring to maintain stabil nt in streaming.

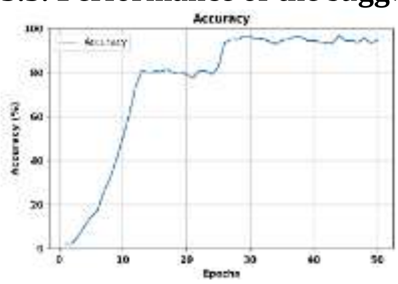


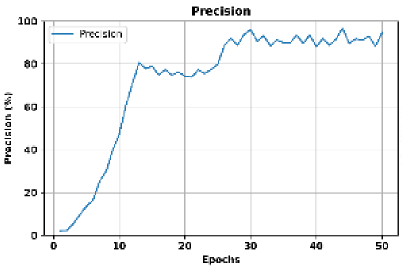
(a) (b)

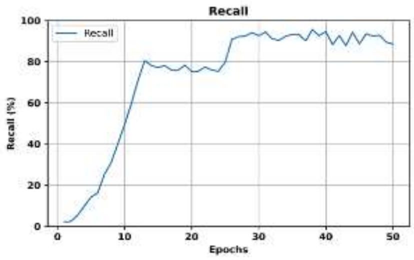
Figure 6: (a) Change Point Scores (b) time series data of the proposed model

In Figure 6, (a) shows the Change Point Scores, which fluctuate at lower levels before spiking above the 0.8 threshold around index 40, signaling significant data shifts and improving responsiveness in applications like network optimization. (b) depicts the time series data, with scores ranging from -2 to 6. The model detects sharp changes, reaching a peak of 6.5 at index 90, demonstrating its ability to quickly identify shifts and enhance adaptive streaming and network performance.

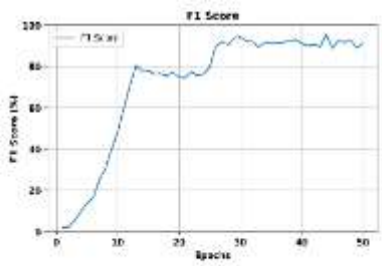
3.3. Performance of the suggested model







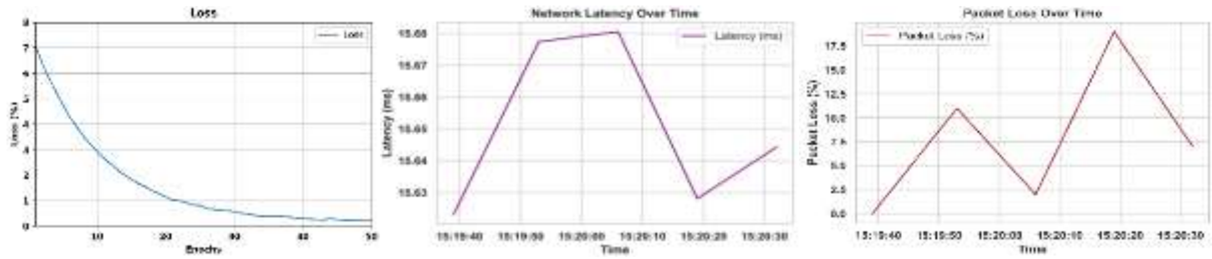
(a) (b) (c)



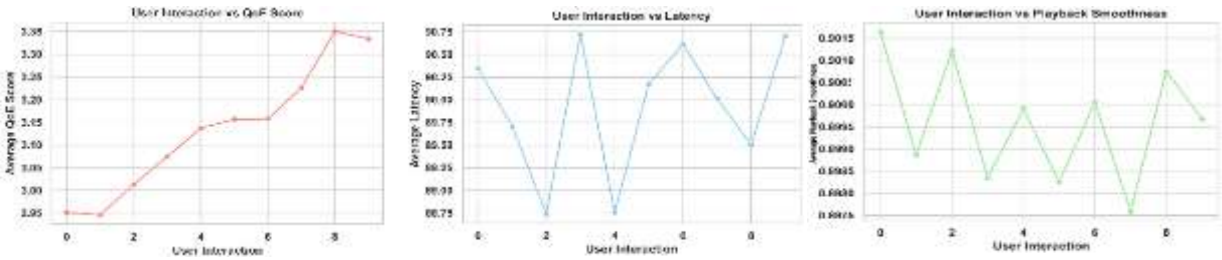
(d)

Figure 7: (a) Accuracy (b) Precision (c) Recall (d) F1 Score of the suggested model

Figure 7 demonstrate the effectiveness of the integrated technique (SGA ConvNet, BCPI, PD-VE) in managing network fluctuations and bitrate allocation, ensuring steady and efficient video streaming for cloud-based QoE in 5G and beyond networks. (a) The model achieves 98% accuracy by the 50th epoch, highlighting rapid performance improvement during training. (b) The model reaches a peak precision of 97.7%, indicating effective training and optimization, especially with PD-VE's adaptive motion estimation. (c) The recall improves from 45% at 10 epochs to 97% at 50 epochs, thanks to SGA ConvNet's effective data analysis. (d) The F1-score rises from 43% to 97.7%, reflecting the model's robust performance improvements due to the SGA ConvNet's data flow pattern analysis.



(a) (b) (c)



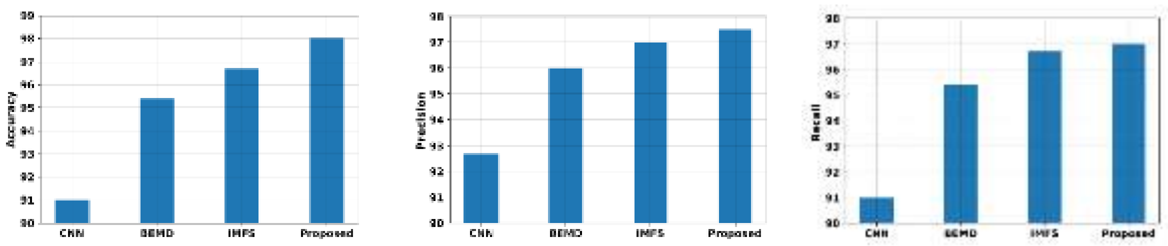
(d) (e) (f)

Figure 8: (a) Loss (b) Latency (c) Packet Loss (d) Average QoE Score (e) Average Latency (f) Average Playback Smoothness of the Suggested Model

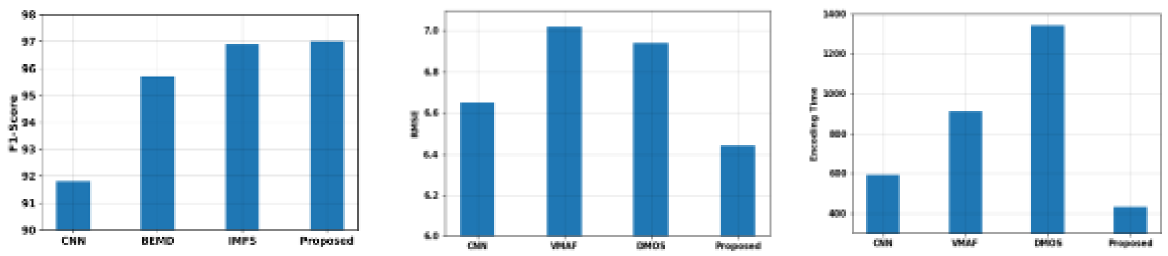
Figure 8 highlight how the suggested model works by monitoring and optimizing various performance metrics over time. Here's a summary of their functions: (a) This figure tracks the model's loss, showing how quickly it learns and improves over time. A decreasing loss indicates better performance, demonstrating the model's efficiency in learning and adapting. The model's loss starts at 7% but rapidly decreases to 0.2% by the 50th epoch, indicating optimal performance through SGA ConvNet integration. (b) This figure shows how the model responds to changes in network latency, reflecting the model's ability to adapt quickly to fluctuations in network performance, ensuring stable streaming. Latency starts at 15.67 ms, fluctuates slightly, and stabilizes at 15.64 ms, showing the model's responsiveness to network performance changes. (c) This tracks packet loss over time, identifying periods of instability. The model adjusts to minimize packet loss, improving overall network reliability during video streaming. Packet loss rises to 17.5% but stabilizes at 7%, reflecting transient network instability, with SGA ConvNet improving resilience. (d) This shows how user experience (Quality of Experience, or QoE) evolves based on interaction. A higher QoE score reflects improved user engagement, which is optimized by real-time data analysis using the model. The average QoE score starts at 2.95, rises to 6, and ultimately achieves 3.32, demonstrating enhanced user experience through real-time data analysis. (e) This figure illustrates how increased user interaction affects latency, with the model using adaptive strategies to minimize latency during high engagement periods. (f) This shows how user interaction influences playback smoothness, with the model optimizing video playback quality by analyzing and adjusting to user behavior.

3.4. Comparative analysis of the suggested method

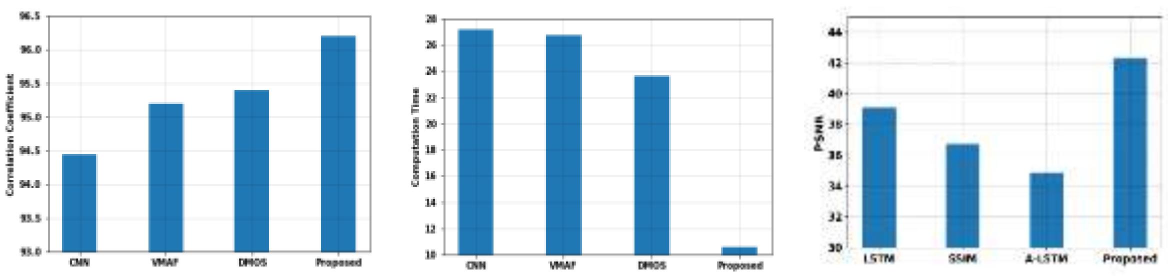
This section reviews the performance and effectiveness of the suggested method using key indicators like accuracy, precision, recall, F1-Score, Encoding Time, Computation Time, RMSE, Correlation Coefficient, PSNR, and average bitrate. These metrics demonstrate how the integrated strategy enhances resource allocation and video streaming quality through detailed performance comparisons.



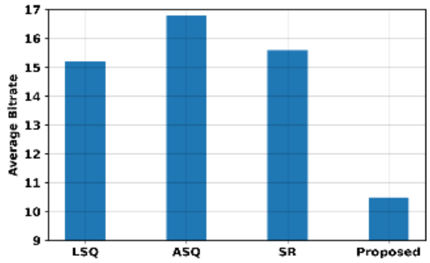
(a) (b) (c)



(d) (e) (f)



(g) (h) (i)



(j)

Figure 9: Comparision of (a)Accuracy (b)Precision (c)Recall (d)F1-Score (e)RMSE (f) Encoding Time (g)Correlation Coefficient (h)Computation Time (i)PSNR (j)Bitrate of the Suggested Model

Figure 9 highlights the significant advancements of the suggested method, which outperforms existing approaches in several metrics. It achieves a Peak Signal-to-Noise Ratio (PSNR) of 42.3 dB, surpassing LSTM, SSIM, and DMOS. The average bitrate of 10.5 is lower than LSQ, ASQ, and SR. The method also excels with 98% accuracy, 97.7% precision, 97% recall, and 97.7% F1-Score. It minimizes prediction errors (RMSE of 6.44) and encoding time (430 seconds), while maintaining a high correlation coefficient (96.2), demonstrating its superiority in optimizing resource allocation and video streaming quality.

4. CONCLUSION

In conclusion, the suggested method effectively addresses the challenges associated with video streaming services, particularly for extended-duration videos. Traditional approaches often struggle with high operational costs and inadequate responsiveness to dynamic data traffic patterns. By integrating SGA ConvNet with BCPI, the suggested method offers real-time analysis of data traffic, ensuring efficient resource allocation and improved QoE. Furthermore, the introduction of PD-VE, enhanced by DRS and ARPS, significantly optimizes motion estimation, resulting in reduced computational complexity and enhanced visual fidelity. The outcomes demonstrate the effectiveness of the suggested method, achieving an impressive accuracy of 98%, alongside precision and F1-Score of 97%, and recall rates of 97%. Additionally, the method minimizes prediction errors with an RMSE of 6.44 and exhibits efficient processing with an encoding time of just 430 seconds and an average bitrate of 10.5. A high correlation coefficient of 96.2 further underscores the model's capability to accurately capture data relationships, reflecting its superiority in optimizing resource allocation and enhancing video streaming quality. These advancements demonstrate the efficacy of the suggested method compared to traditional techniques, paving the way for more reliable and efficient cloud-based video streaming services.

REFERENCES

[1] [1]Palit, B., Sen, A., Mondal, A., Zunaid, A., Jayatheerthan, J. and Chakraborty, S., 2023. Improving ue energy efficiency through network-aware video streaming over 5g. IEEE Transactions on Network and Service Management.

[2] Huang, X., He, L., Wang, L. and Li, F., Towards 5G: Joint optimization of video segment caching, transcoding and resource allocation for adaptive video streaming in a multi-access edge computing network. IEEE Transactions on Vehicular Technology, Vol. 70, 2021, No. 10, pp. 10909-10924.

[3] Pérez, P., Corregidor, D., Garrido, E., Benito, I., González-Sosa, E., Cabrera, J., Berjón, D., Díaz, C., Morán, F., García, N. and Igual, J., Live free-viewpoint video in immersive media production over 5G networks. IEEE Transactions on Broadcasting, Vol. 68, 2022, No. 2, pp. 439-450.
[4] Tanjung, D., Kim, J.D., Kim, D.H., Lee, J., Kim, S. and Jung, J.Y., QoE Optimization in DASH-Based Multiview Video Streaming. IEEE Access, 2023.
[5] Dinaki, H.E., Shirmohammadi, S., Janulewicz, E. and Côté, D., Forecasting video QoE with deep learning from multivariate time-series. IEEE Open Journal of Signal Processing, Vol. 2, 2021, pp. 512-521.
[6] Bampis, C.G., Li, Z., Katsavounidis, I., Huang, T.Y., Ekanadham, C. and Bovik, A.C., Towards perceptually optimized adaptive video streaming-a realistic quality of experience database. IEEE Transactions on Image Processing, Vol. 30, 2021, pp. 5182-5197.
[7] binti Burhanuddin, L.A., Liu, X., Deng, Y., Challita, U. and Zahemszky, A., QoE optimization for live video streaming in UAV-to-UAV communications via deep reinforcement learning. IEEE Transactions on Vehicular Technology, Vol. 71, 2022, No. 5, pp. 5358-5370.
[8] Taha, M., Ali, A., Lloret, J., Gondim, P.R. and Canovas, A., An automated model for the assessment of QoE of adaptive video streaming over wireless networks. Multimedia Tools and Applications, Vol. 80, 2021, No. 17, pp. 26833-26854.
[9] Ma, X., Li, Q., Zou, L., Peng, J., Zhou, J., Chai, J., Jiang, Y. and Muntean, G.M., QAVA: QoE-aware adaptive video bitrate aggregation for HTTP live streaming based on smart edge computing. IEEE Transactions on Broadcasting, Vol. 68, 2022, No. 3, pp. 661-676.
[10] Taha, M., Canovas, A., Lloret, J. and Ali, A., A QoE adaptive management system for high definition video streaming over wireless networks. Telecommunication Systems, Vol. 77, 2021, No. 1, pp. 63-81.
[11] Haung, Y.R., A qoe-aware strategy for supporting service continuity in an mcc environment. Wireless Personal Communications, Vol. 116, 2021, No. 1, pp. 629-654.
[12] Ky, J.R., Mathieu, B., Lahmadi, A. and Boutaba, R., Ml models for detecting qoe degradation in low-latency applications: A cloud-gaming case study. IEEE Transactions on Network and Service Management, 2023.
[13] Najjar, A., Mualla, Y., Singh, K.D., Picard, G., Calvaresi, D., Malhi, A., Galland, S. and Främling, K., One-to-many negotiation qoe management mechanism for end-user satisfaction. IEEE Access, Vol. 9, 2021, pp. 59231-59243.
[14] Zhang, K., Chen, L., An, Y. and Cui, P., A QoE test system for vehicular voice cloud services. Mobile Networks and Applications, Vol. 26, 2021, pp. 700-715.
[15] Kesavan, S., Saravana Kumar, E., Kumar, A. and Vengatesan, K., An investigation on adaptive HTTP media streaming Quality-of-Experience (QoE) and agility using cloud media services. International Journal of Computers and Applications, Vol. 43, 2021, No. 5, pp. 431-444.
[16] Hosseinzadeh, M., Shankar, K., Apostolaki, M., Ramachandran, J., Adams, S.E., Sekar, V. and Sinopoli, B., CANE: A Cascade Control Approach for Network-Assisted Video QoE Management. IEEE Transactions on Control Systems Technology, 2023.
[17] Sultan, M.T. and El Sayed, H., QoE-aware analysis and management of multimedia services in 5G and beyond heterogeneous networks. IEEE Access, 2023.
[18] Agarwal, B., Togou, M.A., Ruffini, M. and Muntean, G.M., Qoe-driven optimization in 5g o-ran-enabled hetnets for enhanced video service quality. IEEE Communications Magazine, Vol. 61, 2022, No. 1, pp. 56-62.

[19] Hayyolalam, V., Otoum, S. and Özkasap, Ö., Dynamic QoS/QoE-aware reliable service composition framework for edge intelligence. Cluster Computing, Vol. 25, 2022, No. 3, pp. 1695-1713.
[20] Dinaki, H.E., Shirmohammadi, S., Janulewicz, E. and Côté, D., Forecasting video QoE with deep learning from multivariate time-series. IEEE Open Journal of Signal Processing, Vol. 2, 2021, pp. 512-521.
[21] Taraghi, B., Live Low-Latency Cloud-based Adaptive Video Streaming Evaluation Framework. In Proceedings of the 2nd Mile-High Video Conference, 2023, May, pp. 158-158.
[22] Zong, T., Mao, Y., Li, C., Liu, Y. and Wang, Y., Progressive Frame Patching for FoV-based Point Cloud Video Streaming. arXiv preprint arXiv:2303.08336, 2023.
[23] Yaseen, M.U., Anjum, A., Fortino, G., Liotta, A. and Hussain, A., Cloud based scalable object recognition from video streams using orientation fusion and convolutional neural networks. Pattern Recognition, Vol. 121, 2022, pp. 108207.
[24] Weishan, Zhang., Haoyun, Sun., Dehai, Zhao., Liang, Xu., Xin, Liu., Huansheng, Ning., Jiehan, Zhou., Yi, Guo., Su, Yang. A Streaming Cloud Platform for Real-Time Video Processing on Embedded Devices. IEEE Transactions on Cloud Computing, Available from: 10.1109/TCC.2019.2894621, 2021.
[25] Taraghi, B., Hellwagner, H. and Timmerer, C., LLL-CAdViSE: live low-latency cloud-based adaptive video streaming evaluation framework. IEEE Access, Vol. 11, 2023, 25723-25734.
[26] He, Q., Wang, Z., Zeng, H., Zeng, Y. and Liu, Y., Svga-net: Sparse voxel-graph attention network for 3d object detection from point clouds. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 36, 2022, June, No. 1, pp. 870-878.
[27] Manjunatha, Peddareddygari, Bayya, Reddy., Sheshappa, Shagathur, Narayanappa. Bayesian K-means clustering based quality of experience aware multimedia video streaming. Indonesian Journal of Electrical Engineering and Computer Science, Available from: 10.11591/ijeecs. v33.i1. pp. 612-621, 2024
[28] Alami, R., Maillard, O. and Féraud, R., Restarted Bayesian online change-point detector achieves optimal detection delay. In International conference on machine learning, PMLR, 2020, November, pp. 211-221.
[29] Zhang, M., Wu, S., Gao, M., Jiang, X., Xu, K. and Wang, L., Personalized graph neural networks with attention mechanism for session-aware recommendation. IEEE Transactions on Knowledge and Data Engineering, Vol. 34, No. 8, 2020, pp.3946-3957.
[30] Fani, M., Ghofrani, F. and Yazdi, M., Modifying ARPS method by use of global motion information for fast and precise motion estimation. In 2015 9th Iranian Conference on Machine Vision and Image Processing (MVIP), IEEE, 2015, November, pp. 115-119.
[31] Rani, M.S. and Valantina, G.M., A Novel Hybrid ARPS Motion Estimation Algorithm for Video Coding in VLSI Architecture. In 2020 International Conference on Communication and Signal Processing (ICCSP), IEEE, 2020, July, pp. 0685-0690.
[32] Xu, H., Cai, Z., Takabi, D. and Li, W., Audio-visual autoencoding for privacy-preserving video streaming. IEEE Internet of Things Journal, Vol. 9, 2021, No. 3, pp. 1749-1761.
[33] Wang, Y., He, H. and Tan, X., Truly proximal policy optimization. In Uncertainty in artificial intelligence, PMLR, 2020, August, pp. 113-122.
[34] Yaseen, M.U., Anjum, A., Fortino, G., Liotta, A. and Hussain, A., Cloud based scalable object recognition from video streams using orientation fusion and convolutional neural networks. Pattern Recognition, Vol. 121, 2022. pp. 108207.

[35] Wang, X., Khan, A., Wang, J., Gangopadhyay, A., Busart, C. and Freeman, J., An edge–cloud integrated framework for flexible and dynamic stream analytics. Future Generation Computer Systems, Vol. 137, 2022, pp. 323-335.
[36] Utke, M., Zadtootaghaj, S., Schmidt, S., Bosse, S. and Möller, S., NDNetGaming-development of a no-reference deep CNN for gaming video quality prediction. Multimedia Tools and Applications, 2022, pp.1-23.
[37] Ali, M., Anjum, A., Rana, O., Zamani, A.R., Balouek-Thomert, D. and Parashar, M., RES: Real-time video stream analytics using edge enhanced clouds. IEEE Transactions on Cloud Computing, Vol. 10, 2020, No. 2, pp. 792-804.

BIOGRAPHIES OF AUTHORS

	**Himaja Tejomurtula** is currently serving as a Senior Program Manager at Amazon Development Center, Bangalore, India. She holds a M. Tech in Wireless & Mobile Communications from Jawaharlal Nehru Technological University. Himaja has over 10 years of experience in program management and has successfully delivered cross-functional projects, specializing in strategic planning, process improvement, and stakeholder engagement. She has worked with Oracle India Pvt. Ltd. and Amazon Development Center in various roles, including Senior Program Manager and Compliance Operations. Her key areas of expertise include Generative AI, Data Visualization, Lean Six Sigma, and Agile Methodologies. She is passionate about leveraging innovative solutions to drive program success and optimize resource utilization.
	**Dr. S. Neeraja** is presently working as an Associate Professor in the Department of Electrical, Electronics and Communications Engineering, GITAM (Deemed to be University). She has received Ph.D. degree from Andhra University in the year 2013. She has completed a major research project DST-SERB, under the Early Career Research Award (ECRA) Scheme and CARS Project. She has over 17 years of teaching experience. She has published more than 40 research papers in various reputed International/National journals/Conferences. Her areas of interest are Wireless and Mobile Communications, Wireless Sensor Networks, and MIMO/OFDMA.
	**Dr. Suresh Chittineni** is currently serving as a Professor at GITAM (Deemed to be University), Vishakhapatnam, Andhra Pradesh. He holds a Ph.D. in Computer Science and Systems Engineering from Andhra University, and an M. Tech in Networking and Internet Engineering from SJCE, Mysore. With over 17 years of experience in the education management sector, he has also served in various leadership roles, including Director of Skill Development and Associate Director Administration (Vice-Principal) at GITAM. Dr. Chittineni has been instrumental in managing and developing departments, teaching, and research. His research interests lie in areas like E-Learning, SQL, Analytical Skills, and Computer Science. He has published several research papers on topics like encrypted cloud data, data classification, and cryptography.