Developing a Quality Prediction Model for Wireless Video Streaming Using Machine Learning Techniques

Emtnan Alkhowaiter[1†], Ibrahim Alsukayti[2††], Mohammed Alreshoodi[3†††]
361200286@qu.edu.sa skiety@qu.edu.sa MO.ALRESHOODI@qu.edu.sa
[1†2††]Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia
[3†††] Department of Applied Science, Unizah Community College, Qassim University, Buraydah, Saudi Arabia

Abstract
The explosive growth of video-based services is considered as the dominant contributor to Internet traffic. Hence it is very important for video service providers to meet the quality expectations of end-users. In the past, the Quality of Service (QoS) was the key performance of networks but it considers only the network performances (e.g., bandwidth, delay, packet loss rate) which fail to give an indication of the satisfaction of users. Therefore, Quality of Experience (QoE) may allow content servers to be smarter and more efficient. This work is motivated by the inherent relationship between the QoE and the QoS. We present a no-reference (NR) prediction model based on Deep Neural Network (DNN) to predict video QoE. The DNN-based model shows a high correlation between the objective QoE measurement and QoE prediction. The performance of the proposed model was also evaluated and compared with other types of neural network architectures, and three known machine learning methodologies, the performance comparison shows that the proposed model appears as a promising way to solve the problems.

Keywords:
DNN, QoS, QoE, prediction, H.264.

.1. Introduction

Video services have been gaining a dominant place in Internet traffic, especially with the use of portable devices (tablets, smartphones, etc.) [1]. Video traffic requires a lot of resources because it is highly sensitive to any problem during the traffic delivery process. This is more evident where error-prone wireless networks are used to transmit videos to portable devices. As a result, the end-user experience would be adversely affected by different levels of degradation. In addition, nowadays devices come with highly sophisticated functionality and take the expectations of consumers regarding video quality to that of broadcast level, in a way that makes full use of device capabilities [2][3] . As a result, it becomes inevitable to move from an approach that relies on the quality of service (QoS) to that based on the quality of experience (QoE)[4].

QoS reflects only network performance that does not directly indicate user satisfaction. However, controlling the QoS parameters for a video transmission system is important to evaluate video quality from a user's perspective known as Quality of Experience (QoE). As a subjective metric, QoE involves human dimensions. It ties together the experience of application and network performance, user perception, and expectations. QoE-based video quality evaluation is difficult since the user experience is subjective and difficult to be quantified and measured due to a high cost in terms of time, manual effort, and money. Therefore, several recent research studies have been focusing on the idea of developing objective NR quality prediction models. In this regard, several machine learning-based techniques have been applied to realize no-reference models for QoE prediction [5], such as Fuzzy Logic Systems, Neural Networks, and Decision Tree. The ITU classification of objective quality assessment models [6] outlined the efficiency of the hybrid QoE prediction model in the development of an effective generic prediction model. This is the adopted approach in this research work.

In this paper, a hybrid no-reference prediction model based on the Deep Neural Networks (DNN) is proposed for predicting QoE of wireless video streaming. As the main characteristics of the prediction model are entered on the basis of DNN without relying on any human element. This model can provide QoE prediction for video streaming in a real-time manner, which can be located over a wireless network at an intermediate measuring point.

The structure of the paper is as follows; Section 2 discusses the related works. A description of the experimental setup for dataset collection is provided in Section 3. The QoE measurements and dataset collection are presented in Section 4. The proposed model is presented in Section 5. Section 6 demonstrates the performance evaluation. Finally, the conclusion is presented in section 7.

2. Related work

In recent years, several research studies focused on the prediction of the video QoE based on machine learning methods. The study presented [7] a no-reference model based on ANN for predicting HD video coded using the

Manuscript received March 5, 2021
Manuscript revised March 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.3.31

H.265 / HEVC and VP9 codecs in the mobile network environment. The results show that the prediction accuracy was slitly improved by the ANN method when compared to the regression-based approach. Work in [8] presented a hybrid model based on fuzzy reasoning systems (FIS) to predict video quality with a no-referenced prediction model in wireless networks. The proposed model is evaluated based on random neural networks (RNN), where the simulation result shows that the FIS-based model has better performance. In addition, the work in [9] studied the effect of QoS parameters on 4kUHD video transmission and built a NR predictive model based on FIS for the cognitive quality of codec H.265 4kUHD videos in wireless transmission environments. SSIM was used as an objective quality metric. The FIS-based model was also evaluated against RNN-based models and demonstrated significant performance.

In [10], the researchers presented a NR model to predict user satisfaction in terms of QoE using the RNN approach. They studied the impact of five QoS parameters on the video QoE. The QoS parameters are represented as inputs on the neural networks. The authors argue that the presented results of the proposed model are very useful for network operators to achieve user satisfaction based on the network performance viewpoint. Work in [11] used six different machine learning methods to build a prediction model, including Logistic Regression, Linear Discriminant Analysis, Decision Tree (DT), K-Nearest-Neighbors, Support Vector Machine, and Gaussian Naive Bayes. The DT-based model provided an exciting result by extracting only the two most important features - the bitrate and user profile that gives a good result of about 86% accuracy in predicting the QoE. Furthermore, authors in [12] proposed a quality prediction model based on RNN for video streaming over long-term evolution (LTE) networks. In this study, PSNR was used as an objective quality metric. The QoS parameters that affect the video quality over LTE networks were chosen from the application and network layers, where the H.264 codec is used. The results indicate that the proposed prediction model was able to achieve an accuracy of about 93% when compared to objective values. Table 1 presents a comparative summary of a number of related works.

It can be seen that the majority of the reviewed quality prediction models for video streaming relied on RNN or FIS as a machine learning methodology. Also, the RNN was used in a simple way with few hidden layers. This paper presents the development of a non-intrusive QoE prediction model which incorporates a deep neural network (DNN) method. It takes into consideration a set of QoS parameters critical to wireless video streaming. The DNN is a neural network but comes with an additional level of complexity, specifically with more than two layers. Sophisticated mathematical modeling is applied in DNNs which allows complex, yet effective, data processing [13].

Table 1. comparative summary of a number of related works

Aspect / Model	A QoS	No QoS parameter	Subjective Metrics	Objective Metrics	Application	Simulation/ test bed	Technology	Reference	Video coding parameters	Artificial adaptive technique	Codec
Pal et al. [7]	√	√	√	√	VS	Simu./ test-bed	wireless	NR	√	ANN	H.265 and VP9
Danish et al.[8]	√	√	√	√	VS	Simu.	wireless	NR	√	FIS	H.264
ALreshoodi et al. [9]	√	√	√	√	VS	Simu.	wireless	NR	√	FIS	H.265
Danish et al.[10]	√	√		√	VS	Simu.	wireless	NR	√	RNN	H.264
Shalala et al.[11]	√		√		VS	lab	wireless	NR	√	DT, LR, LDA, KNN, NB and SVM	H.264
Ghalut et al. [12]	√	√		√	VS	Simu./ test-bed	wireless	NR	√	RNN	H.264

3. Dataset Collection Setup

3.1 Video Sequences

The encoded video files in YUV format are taken from the publicly available video database and published work in[14][15]. The video sequences were classified into three types (low motion, moderate motion, and high motion), while the resolution is classified into (HD and SD). The selected video files were encoded with different Quantization parameter (QP) values (16, 24, 32, 40, 48). The H.264/Advanced Video Coding (AVC) Joint Model (JM) reference codec [16][17] was used in the encoding and decoding process. RTP packetization and encapsulation in IP packets were applied to the network abstraction layer (NAL) units. Fig.1. shows how the video transmissions.

Fig. 1. video transmissions

3.2 QoS Parameters and Simulation Scene

Table 2 presents the simulated QoS parameters and their values. The selected QoS parameters at the encoding level are the resolution (R) and content type (CT). In addition, two QoS parameters were selected from the network QoS level, packet loss rate (PLR) and the mean burst loss (MBL).

Table 1: QoS parameters.

Parameters	Values
Video Resolution (R)	SD, HD
Content type (CT)	Low, Moderate, High motion
Packet loss ratio (PLR)	1, 2.5, 5, 7.5, 10
Mean burst length (MBL)	1, 2.5, 5, 7.5

The simulation scenario is shown in Fig.2. The PLR and MBL are generated by the wireless transmission error that is based on the Gilbert-Elliot model [18][19]. The PLR is distributed uniformly and the MBL is bursts distribution along the packet loss trace. The error model was used for describing packet loss traces in transmission channels and simulating communications links with error performance. In order to increase data confidence, each tested condition was simulated 10 different times. An objective video quality metric was then utilized to assess the distorted video sequences, as explained in the following subsection.

Manuscript received March 5, 2021
Manuscript revised March 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.3.31

Fig. 2. Conceptual illustration of the simulation scene.

4. QoE ASSESSMENT

The objective quality assessment method is used as a cost-effective alternative to produce results that are comparable with those of subjective testing and can be used for real-time measurement of video quality. In this paper, the video quality metric (VQM)[20] was used for video quality measurement. This was introduced by the National Telecommunications and Information Administration (NTIA), and assessed by the video quality experts group (VQEG). Then, ANSI [21] and ITU [22] have standardized it. The VQM scale is from 0 to 1, where 0 represents complete loss and 1 represents original quality. In this work, the VQM also normalized to the subjective Mean Opinion Scores (MOS) through the equation:

$$MOS = 5 - 4VQM \qquad (1)$$

MOS is a numerical measure ranging from five grades (1:Bad, 2:Poor, 3:Fair, 4:Good, 5:Excellent) [3]. After measuring the VQM of all degraded video sequences, the collected dataset was then divided into training (60%) and testing (40%) datasets.

5. QoE PREDICTION MODEL BASED ON DNN

We used in this work the DNN method to build the proposed prediction model. Deep learning is a sub-field of machine learning that basically applies neural networks with multiple hidden layers[23]. Deep Learning is based on algorithms inspired by the biological structure and functioning of a brain in order to support machines with artificial intelligence [24]. The DNN consists of an input and output layer in addition to a set of hidden layers in between. Each layer implements specific types of sorting and sorting in a process that some refer to as a "feature hierarchy". A common use of this complex model of the neural network is to deal with unclassified or disorganized data.

The proposed DNN-based model was implemented using the python programming language. We initialized our model in Keras using a sequential model. A dense layer is a fully connected layer. The model requires the data input to the first layer where five are used as input variables to the model (R, CT, PLR, MBL). We used four hidden layers, as illustrated in Fig.3. The efficient Adam optimizer is used with a default learning rate of 0.001. Tangent hyperbolic (Tanh) function is used to perform the computation of activation for the Dense layer. Each layer has 64 neurons. After building and compiling the model, the next step is to fit the DNN model into our dataset.

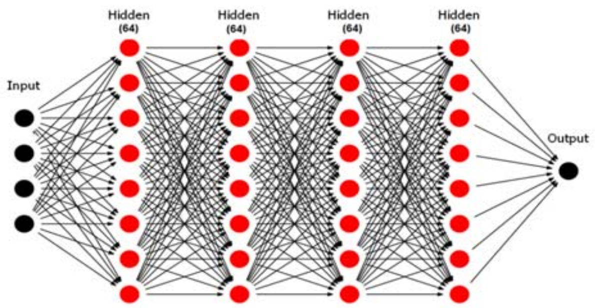

Fig. 3. Conceptual illustration of the simulation scene.

The DNN model learns and maps the correlation between measured QoE and QoS parameters. As part of the learning phase, the DNN is trained on the dataset of measured QoE. After that, the system becomes able to perform QoE prediction based on a combination of QoS parameters as shown in Fig.4. A comparison is then established between the predicted QoE output (MOS) and the objectively measured QoE. This is done based on the calculation of two main factors: the Root Mean Squared Error (RMSE) and the correlation coefficient.

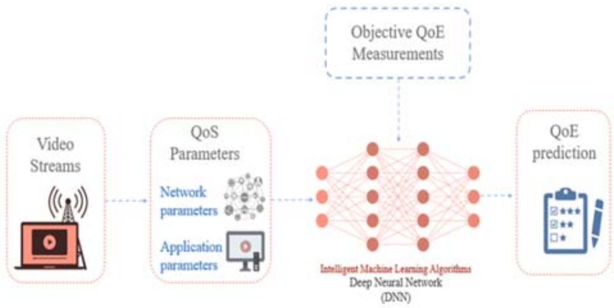

Fig. 4. DNN mapping the QoS parameters and measured QoE.

6. Results and discussion

6.1 Validation using The Testing Dataset

As mentioned in section 4, the collected dataset was divided into training and testing datasets. The testing dataset was used towards practical validation of the proposed DNN-based model. As aforementioned, the correlation coefficient and the Root Mean Squared Error (RMSE) are calculated to effectively compare the predicted QoE output (MOS) against the objectively measured QoE (testing dataset), as shown in Fig 5. The R2 scored 95% and RMSE was 0.22 which shows that the predicted QoE is highly correlated with measured QoE. This means that the proposed DNN-based model effectively succeeded in the prediction of end-users perception.

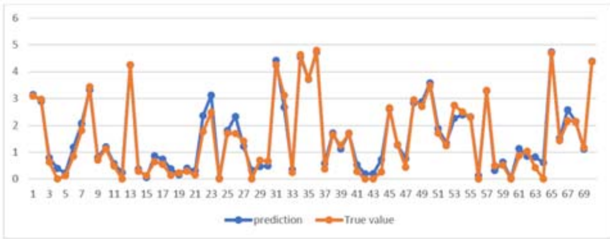

Fig. 5. Measured QoE vs. predicted QoE.

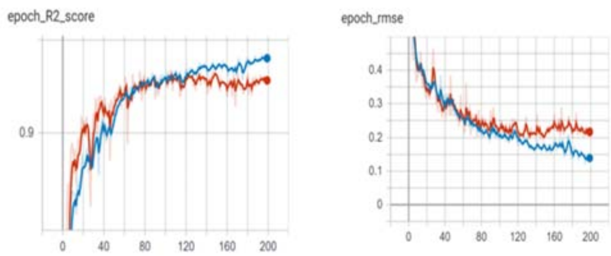

Fig.6. R2 and RMSE agents' number of epochs.

6.2 Evaluation by Other ML Methods

In this work, the performance of the proposed DNN-based model was also compared with three other machine learning models. Three ML methods were used for the performance comparison step, namely: Linear Regression, KNN, and Decision Tree. To implement these methods, we used the Sklearn package and its classes: Linear Regression, K neighbors Repressor with 3 neighbors, and Decision Tree Repressor. As we can see from table 3, we found that the proposed DNN-based model shows a better performance.

Table 3. Performance comparison

	Linear Regression	KNN	Decision Tree	DNN
RMSE	0.33	0.35	0.23	0.22
R^2	0.88	0.86	0.94	0.95

6.3 Evaluation by Other Neural Networks Architectures

We also compared the performance of DNN-based model with other types of neural networks architectures, namely Convolutional neural networks (CNN) [25] and Artificial Neural Networks (ANN)[26]. The results of the performance comparison in terms of R2 and RMSE are shown in Fig.7. and Fig.8. The three models used the same optimizers 'adam'. The DNN-based model is significantly better than the two other neural network architectures considered in this testing phase. Fig.9. and Fig 10 show the R2 and RMSE agents' number of epochs.

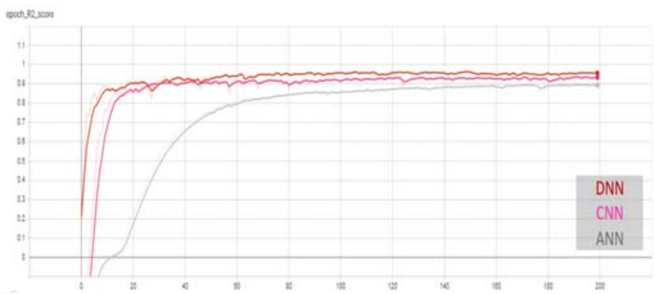

Fig.7.comparing DNN, CNN, and ANN-based on R2 against epochs.

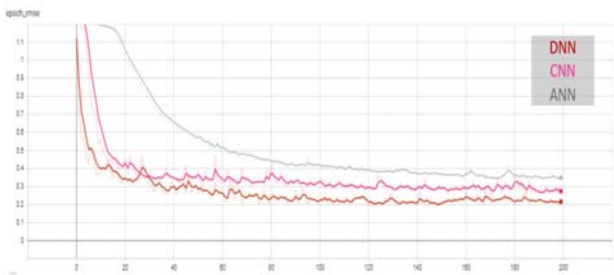

Fig.8.comparing DNN, CNN, and RNN based on RMSE against epochs.

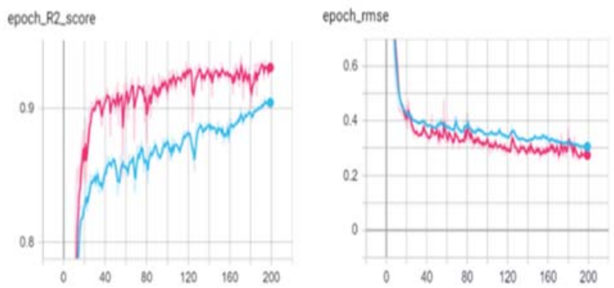

Fig.9. CNN-based model: R2 and RMSE agents' number of epochs.

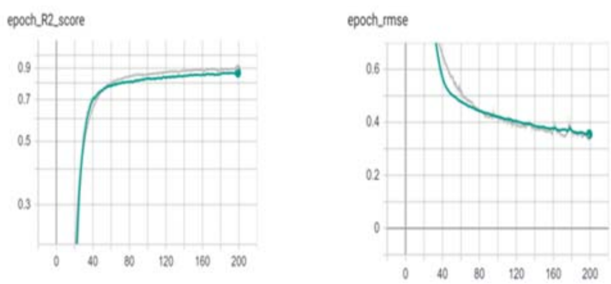

Fig.10. RNN-based model: R2 and RMSE agents' number of epochs

7. Conclusion

This research presents the concept of QoS/QoE correlation for video QoE prediction of video streaming in wireless networks. The main aim of this work is to develop a no-reference video quality prediction model based on DNN methodology. effective mapping of a set of critical QoS parameters to the QoE (represented by MOS scores) of video streaming traffic to conduct the main aim of this research. Furthermore, the wireless transmission errors model has been used in the coded video packets to practically simulate communications links with error performance. The collected QoE video dataset was constructed to help in the development of a robust objective QoE prediction model based on DNN. The assessment of the quality of the decoded video frames was carried out using the VQM to provide an arithmetic mean over the sequence's video frames.

The proposed DNN-based model was validated by the testing dataset, other types of neural network architectures, and three other machine learning models as well. From the results, it is clear that the DNN-based model shows better performance on all validation methods. A high correlation between the predicted QoE and measured QoE was achieved by the proposed DNN-based model. The results of this work also confirmed that the QoS parameters selection is important to achieve an efficient prediction accuracy.

In our future work, the proposed DNN-based model can be applied to other potential applications, and other QoS parameters can be incorporated for improving quality prediction to a higher level in an end-to-end manner.

References

[1] E. Summary, "Cisco Visual Networking Index : Forecast and Methodology , 2018 – 2023," updated March 9 2020.

[2] G. Dimopoulos, I. Leontiadis, P. Barlet-Ros, and K. Papagiannaki, "Measuring video QoE from encrypted traffic," *Proc. ACM SIGCOMM Internet Meas. Conf. IMC*, vol. 14-16-Nove, no. July 2018, pp. 513–526, 2016, doi: 10.1145/2987443.2987459.

[3] J. Joskowicz, R. Sotelo, and J. C. L. Ardao, "Towards a general parametric model for perceptual video quality estimation," *IEEE Trans. Broadcast.*, vol. 59, no. 4, pp. 569–579, 2013, doi: 10.1109/TBC.2013.2277951.

[4] M. Alreshoodi and J. Woods, "Survey on Qoe\Qos Correlation Models Formultimedia Services," *Int. J. Distrib. Parallel Syst.*, vol. 4, no. 3, pp. 53–72, 2013, doi: 10.5121/ijdps.2013.4305.

[5] S. Aroussi and A. Mellouk, "Survey on machine learning-based QoE-QoS correlation models," *2014 Int. Conf. Comput. Manag. Telecommun. ComManTel 2014*, pp.

200–204, 2014, doi: 10.1109/ComManTel.2014.6825604.

[6] A. Schwind, M. Seufert, Ö. Alay, P. Casas, P. Tran-Gia, and F. Wamser, "Concept and implementation of video QoE measurements in a mobile broadband testbed," *TMA 2017 - Proceedings of the 1st Network Traffic Measurement and Analysis Conference*. 2017, doi: 10.23919/TMA.2017.8002921.

[7] L. Skorin-Kapov, M. Varela, T. Hoßfeld, and K.-T. Chen, "A Survey of Emerging Concepts and Challenges for QoE Management of Multimedia Services," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 14, no. 2s, pp. 1–29, 2018, doi: 10.1145/3176648.

[8] E. Danish, A. Fernando, M. Alreshoodi, and J. Woods, "A hybrid prediction model for video quality by QoS/QoE mapping in wireless streaming," *2015 IEEE Int. Conf. Commun. Work. ICCW 2015*, pp. 1723–1728, 2015, doi: 10.1109/ICCW.2015.7247429.

[9] A. O. Adeyemi-Ejeye, S. D. Walker, M. Alreshoodi, and J. Woods, "Fuzzy logic inference system-based hybrid quality prediction model for wireless 4kUHD H.265-coded video streaming," *IET Networks*, vol. 4, no. 6, pp. 296–303, 2015, doi: 10.1049/iet-net.2015.0018.

[10] P. Anchuen, P. Uthansakul, and M. Uthansakul, "QOE model in cellular networks based on QOS measurements using Neural Network approach," *2016 13th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol. ECTI-CON 2016*, 2016, doi: 10.1109/ECTICon.2016.7561318.

[11] R. Shalala, R. Dubin, O. Hadar, and A. Dvir, "Video QoE Prediction Based on User Profile," *2018 Int. Conf. Comput. Netw. Commun. ICNC 2018*, pp. 588–592, 2018, doi: 10.1109/ICCNC.2018.8390347.

[12] T. Ghalut, H. Larijani, and A. Shahrabi, "Content-based video quality prediction using random neural networks for video streaming over LTE networks," *Proc. - 15th IEEE Int. Conf. Comput. Inf. Technol. CIT 2015, 14th IEEE Int. Conf. Ubiquitous Comput. Commun. IUCC 2015, 13th IEEE Int. Conf. Dependable, Auton. Se*, pp. 1626–1631, 2015, doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.245.

[13] H. Yingge, I. Ali, and K. Y. Lee, "Deep neural networks on chip - A survey," *Proc. - 2020 IEEE Int. Conf. Big Data Smart Comput. BigComp 2020*, pp. 589–592, 2020, doi: 10.1109/BigComp48618.2020.00016.

[14] I. Alsukayti and M. Alreshoodi, "Hybrid non-reference QoE prediction model for 3D video streaming over wireless networks," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 11, pp. 697–703, 2018, doi: 10.14569/ijacsa.2018.0911100.

[15] I. S. Alsukayti, "An Adaptive Neuro-Fuzzy Model for Quality Estimation in Wireless 2D/3D Video Streaming Systems," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 2, pp. 2291–2297, 2019, doi: 10.35940/ijitee.b6454.129219.

[16] R. T. Sataloff, M. M. Johns, and K. M. Kost, "TU-T, \Advanced video coding for generic audiovisual services, recommendation itu-t h.264," [Online]. Available: TU-T, %5CAdvanced video coding for generic audiovisual services, recommendation itu-t h.264.

[17] "J. V. T. J. of ISO/IEC MPEG & ITU-T VCEG, \H.264/avc jm reference software.," 2011.," [Online]. Available: http://avc.hhi.fraunhofer.de/.

[18] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channels," *Bell Syst. Tech. J.*, vol. 42, no. 5, pp. 1977–1997, 1963, doi: 10.1002/j.1538-7305.1963.tb00955.x.

[19] M. Zorzi, R. Racd, and L. B. Milstein, \A markov model for block errors on fading channels," in Personal, Indoor and Mobile Radio Communications, 1996. PIMRC'96., Seventh IEEE International Symposium on, vol. 3, pp. 1074{1078, IEEE, 1996.

[20] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312–322, 2004, doi: 10.1109/TBC.2004.834028.

[21] A. T1.801-2003, "Digital Transport of One-Way Video Signals - Parameters for Objective Performance Assessment," vol. 2003, 2003.

[22] Itu-T, "J.144 : Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference." 2004, [Online]. Available: http://www.itu.int/rec/T-REC-J.144-200403-I/en.

[23] N. Eswara *et al.*, "Streaming video QoE modeling and prediction: A long short-term memory approach," *arXiv*. 2018.

[24] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A Survey of Deep Learning and Its Applications : A New Paradigm to Machine A Survey of Deep Learning and Its Applications : A New Paradigm to Machine Learning," *Arch. Comput. Methods Eng.*, no. July, 2019, doi: 10.1007/s11831-019-09344-w.

[25] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1733–1740, 2014, doi: 10.1109/CVPR.2014.224.

[26] A. S. Ajrash, R. F. Ghani, and L. Al-Jobouri, "ANN based measurement for no-reference video quality of experience metric," *2019 11th Comput. Sci. Electron. Eng. Conf. CEEC 2019 - Proc.*, pp. 128–133, 2019, doi: 10.1109/CEEC47804.2019.8974336.