TELEMÁTICA
REVISTA DIGITAL de las Tecnologías
de la Información y las Telecomunicaciones

Revista Telemática, Vol. 21, No. 1, Enero-Marzo, 2022, p.56- 70 ISSN 1729-3804

OBJECTIVE METHODS FOR VIDEO QUALITY ASSESSMENT USING DEEP LEARNING

Omar Antonio Hernández Duany[1], Caridad Anías Calderón[2], Jose Alberto Padilla Morani[3]

[1 2 3] Universidad Tecnológica de la Habana "José Antonio Echeverría", CUJAE,
Ave. 114 #11901 / Ciclovía y Rotonda. Marianao.
[1]e-mail:omar.hd@tele.cujae.edu.cu
[2]e-mail:cacha@tesla.cujae.edu.cu
[3]e-mail:ftjose.pm@tele.cujae.edu.cu

ABSTRACT

The objective methods of video quality evaluation without references make possible the identification of degradations and noises contained in videos, which are produced by several external factors present in the context of different telecommunication processes: acquisition, compression, processing, or transmission. Identifying degradations in the video is essential to improve its quality. Other processes need to be carried out later, such as detection, recognition of objects, or simply reproducing videos with higher quality. The methods of video quality assessment have evolved from the stage in which it was performed only from the perceptual or subjective evaluation of the quality based on the experience of experts. At the current stage, objective methods predominate using artificial intelligence techniques that perform the analysis from machine learning and are designed using convolutional neural networks. These processes are bio-inspired since they emulate how the human brain identifies noises and distortions in video sequences. This paper summarizes the study of the state of scientific knowledge in the field of video quality assessment methods based on deep learning, which has been developed in the last five years by the international scientific community. We aim to expand the potential for executing the process when large volumes of audiovisual information need to be managed in real-time.

KEYWORDS: noises, distortions, video quality assessment, convolutional neural networks, deep learning

MÉTODOS OJETIVOS PARA LA EVALUACIÓN DE LA CALIDAD DE VIDEO EMPLEANDO DEEP LEARNING

RESUMEN

Los métodos objetivos de la evaluación de la calidad de video sin referencias posibilitan la identificación de degradaciones y ruidos contenidos en videos, que son producidos por diversos factores externos presentes en el contexto de diferentes procesos de telecomunicaciones: adquisición, compresión, procesamiento o transmisión. El proceso de identificación de degradaciones en video resulta indispensable para lograr el mejoramiento de su calidad y tiene importancia para otros procesos que deben ejecutarse ulteriormente, tales como: detección y reconocimiento de objetos, o simplemente la reproducción de videos con una calidad superior. Los métodos de evaluación de la calidad de videos han evolucionado desde la etapa en que se realizaba únicamente a partir de la evaluación perceptual o subjetiva de la calidad basado en la experiencia de expertos, hasta la etapa actual en la que predominan los métodos objetivos que emplean técnicas de inteligencia artificial que realizan el análisis a partir del aprendizaje automático y son diseñados empleando las redes neuronales convolucionales. Estos procesos son bio-inspirados dado que emulan el modo en el que el cerebro humano identifica ruidos y distorsiones en las secuencias de video. Este artículo resume el estudio del estado del conocimiento científico en el campo de los métodos de evaluación de la calidad de video basados en el aprendizaje profundo, que han sido desarrollados en el lustro más reciente por la comunidad científica internacional, pretendiendo ampliar las potencialidades para la ejecución del proceso en circunstancias en las que se requiere gestionar grandes volúmenes de información audiovisual en tiempo real.

PALABRAS CLAVES: ruidos, distorsiones, evaluación de la calidad de video, redes neuronales convolucionales, aprendizaje profundo.

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

1. INTRODUCTION

In the most recent decade, there has been a significant increase in the number and diversity of applications in the field of Telecommunications that require processing huge volumes of audiovisual information to manage a wide range of organizational processes of various application domains, among which are: telematics services, content management platforms, streaming or video traffic over multiple communications networks, among others [1]. On the other hand, it is essential to consider the evolution of mobile devices and their remarkable impact on the daily lives of many people and organizations, which have contributed to developing a wide range of online services over various dynamic network environments.

Video management demands as one of its fundamental processes the video quality assessment (VQA), which aims to estimate the quality of video frame sequences or video files previously stored, depending on the type of application. An analysis of the emerging VQA methods is necessary for developing applications with increasing effectiveness using hybrid models that work automatically [2]. Making the VQA process more effective requires designing and implementing applications that efficiently use computational infrastructures to achieve real-time execution.

Many academic and business organizations in the international telecommunications industry are working on research to develop new VQA methods for use in a wide range of computer vision applications. VQA methods are of great importance for the processes of improvement and analysis of content in various fields of scientific knowledge, making them indispensable for heterogeneous decision-making processes in application domains such as: in medicine: microscopy, radiography, ultrasonography, tomography, magnetic resonance imaging; in surveillance and security; image analysis, object recognition, identification of biometric features, counting and inspection of objects; in autonomous driving: aerial, maritime and terrestrial uncrewed vehicles; precision agriculture: herd control, livestock growth control, fruit ripening control and optimal time for harvesting [3].

The development of VQA methods at present is mainly aimed at increasing the effectiveness of this process in an automated mode to achieve compatibility with the quality criteria of the human visual system (HVS), having the peculiarity that it is required to process increasing volumes of audiovisual information. According to the CISCO Annual Report 2018-2023 it was predicted by 2021 that 80% of Internet traffic would be video and an estimated 3 trillion minutes of video per month would be viewed in real-time [4]. It is likely that under the conditions of the COVID-19 pandemic, online interchange would have considerably exceeded these estimates. In any case, it is required to work on the development of tools that perform the VQA process more and more efficiently, based on the design of new algorithms that should contribute to the efficient identification of noises and distortions in video sequences, without using reference videos.

In this paper, an approach to studying the distinctive features of a group of reference-free video quality assessment methods is carried out. The authors of this article have considered that these methods are the most relevant among those using convolutional neural networks (CNN). We have selected from the analysis of more than 40 methods developed in the last five years by various research groups at the international level, which are an expression of the evolution of research in this field of science. This study is the basis for designing and implementing new VQA solutions with potentialities for real-time processing of multiple video streams, which require maximizing the performance of hardware infrastructures to reduce the associated execution times significantly.

2. Digital video: noises and distortions

Digital video is the representation of a sequence of moving images resulting from the capture of a real signal through a video camera. Each image of the video sequence is called a video frame [5]. Each frame represents a two-dimensional matrix containing several pixels that determine the resolution and is the smallest homogeneous unit in the color that is part of the digital image. Digital video encodes an electrical signal stored in standard formats, constituted by the size of the image, the number of frames transmitted per second (frame rate) and the pixel aspect ratio (PAR-Pixel Aspect Ratio). This is a mathematical ratio describing how the width of a pixel compares to its height in a digital image. The default aspect ratio of (16:9) encodes video in the following resolutions: 2160p: 3840x2160, 1440p:2560x1440, 1080p:1920x1080, 720p:1280x720, 480p:854x480, 360p:640x360, 240p:426x240 [6].

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

Video acquisition and processing can introduce distortions or artifacts in the video signal, which represent the alteration of the shape of a signal as it passes through a given system by changing its amplitude, phase or frequency in an unequal proportion. It can occur during the acquisition, compression, processing or transmission processes [7]. Artifacts related to the acquisition process may include visible effects due to interlaced scanning, aliasing (temporal and spatial) or distortions due to perspective. The sampling area on the sensor limits the resolution and may cause contrast inversions. The video quality perceived in the frames of a digital video sequence may be degraded by the noise that mixes with the useful video signal and causes the appearance of undesirable effects [8].

In a multimedia environment, an image can be described as a two-dimensional representation of a scene and in the particular case of a video, it can be treated as a sequence of images or frames containing additional information from the temporal point of view. During the capture of a digital video additional noises and distortions can also be added [9]. In natural scenes, noises can occur due to rainy, predominantly dark, or very light environments, also with the appearance of different atmospheric particles or haze. Movements and vibrations can also cause them; by microwave background radiation or interference of electronic devices in a close location, among others [10].

Nowadays, video coding standards use motion compensation methods and block-based coding schemes. In that case the decoded video can be affected by one or more compression effects such as blurring, false contours, block effect, flickering, and motion compensation errors [11]. In any case, the quality of digital video processing systems varies according to the dynamic characteristics of the input video signal concerning the amount of motion or spatial detail. The video can also be degraded by signal weakness, fragmentation, noise, freeze frames or corrupted frames, or desynchronization. It is necessary to consider that video transmissions are information carriers that enable the prediction of data loss that can cause errors at the receiver, which can be rectified with frame processing from the analysis of preceding frames [12].

In the last two decades, there has been a wide diffusion of digital video that has almost completely replaced analog video, although in many cases a large number of videos that were initially recorded with analog technology and are the result of analog-digital conversion are still available in digital format. Video quality is a characteristic associated with transmissions or video processing. It represents a measure of the degradation of its features [13]. The human visual system (HVS) is endowed with capabilities for effectively identifying distortions, discarding various aspects related to the characteristics of the applications in which they are being used and the regions of the frame where noise or distortions occur. The perceived video quality is also highly dependent on the type and technical characteristics of the device used for playback [14].

The exponential growth of audiovisual information volumes and the transfer speed of video streams to be processed for the execution of various decision-making processes in organizations or simply the exhibition of better quality videos requires a priori quality assessment. This process can identify degradations in frames of the video sequence to reduce noise and eliminate distortions in the preprocessing stage [15].

3. Video quality Assessment methods

The VQA process involves identifying degradations in the frames that compose a video, so that it contributes to the detection of inherent distortions and noises [16]. It is an essential process in the preprocessing stage. It is performed from the analysis by executing various methods of detecting characteristic features of the neighborhood of each of the pixels that compose each frame of a video sequence. The scientific community has developed various methods that implement VQA methods based on multiple theoretical criteria that contribute to the realization of quality estimates with a high degree of accuracy [17]. These methods are indispensable tools as a starting point to improve the quality of video and image sequences so that noise and distortions are reduced or eliminated.

It is necessary to consider that during the processes of capture, storage, video transmission, noise and distortions (artifacts) can be added that produce undesirable effects in the video signal. So it is essential to perform VQA processes to identify the degradations present in each video sequence frame and outline the improvement work at the stage prior to its presentation [18]. The evaluation and prediction of video quality is required in a wide range of applications that are part of other heterogeneous organizational processes or content distribution services over various data networks; its realization becomes a critical task for the telecommunications sector [19].

Each of the video quality evaluation methods has its own design principles related to how they identify the impairments that affect the video signal and their contribution to the further improvement of video quality. These determine to a large extent, the effectiveness with which the removal of these degradations is performed in each of the frames during the stage before the execution of other essential processes, such as object detection, optical character recognition, multibiometric recognitions, among others [20]. VQA methods are classified into two main groups: subjective methods and objective methods. Subjective methods are associated with ratings that can be

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

performed by means of human competition, evaluation constitutes a natural way of measuring quality, which consists of the observation and classification by experts capable of providing high accuracy and reliability. They are very costly, and time-consuming, because they are performed based on human analysis known as quality of experience (QoE-Quality of Experience) by experts or groups of experts.

Subjective evaluations of video quality have been considered a reliable measurement system because they directly involve the users' opinion, which is averaged from a successive number of observations. It should be noted that subjective methods are not reproducible because they are based on the observation and classification of individual frames, so that a qualitative assessment can be formed. Under the current circumstances, subjective VQA methods are not the most viable option in automated quality monitoring and prediction applications in online systems or when multiple simultaneous video streams need to be processed in real time. These methods are more suitable for offline environments.

Subjective methods have been devised based on observing a set of standardized guidelines proposed by the ITU. Typically it uses a measurement scale based on the Mean Opinion Score (MOS), a measure used in the quality of experience and engineering in telecommunications, both in video and audio, although it is not restricted only to these modalities. ITU recommendation P.800.1 mentions several modes of reference to a MOS, such as VMOS, which provides a connection to the operation and maintenance systems for monitoring video service indicators. VQA methods have a high computational cost associated with them, so it is essential to minimize execution times to generate capabilities for processing increasing volumes of audiovisual information in the shortest possible time [21].

The objective methods of VQA are essentially mathematical models based on the analysis of the behavior of specific metrics in the compressed/degraded video sequence, it can be performed in two ways: by comparison with the original sequence, when available, or without the need to compare it with the original video sequence [22]. In objective methods, the evaluation is performed automatically by computer applications without human intervention. The fundamental purpose of objective VQA methods is to achieve results that are highly correlated with subjective quality measures. So, they are indispensable tools for the application domain of computer vision in applications where processing under time-restricted conditions is necessary. They make it possible to obtain objective assessments that reflect an approximation to the quality criteria of the human visual system in an automated way [23].

The new methods of video quality assessment constitute an area of research in continuous development to achieve increasingly effective mechanisms of degradation identification. They have the peculiarity that in the current circumstances it is necessary to inevitably manage huge volumes of audiovisual information that are usually disseminated through communication networks with a tendency for exponential growth. In these circumstances, it is required to handle video streams without human intervention, based on the design of real-time processing methods that achieve higher performance from the efficient use of the installed computational capacity [24].

3.1. Classification of objective methods of video quality assessment

Objective VQA methods are classified taking into consideration the original signal's full, partial or no availability. When comparing the degraded/compressed signal with the original video signal, they can be divided into three categories: complete reference methods, identified as FR-VQA (Full Reference-VQA), which perform the comparison concerning a reference video assumed to be of high quality. Reduced reference methods, identified as RR-VQA (Reduce Reference-VQA), in contrast to FR-VQA, only need parts of the reference video to perform the evaluation. The NR-VQA (Non-Reference-VQA) VQA methods perform the quality assessment of a distorted video without using any reference video, only the degraded video is worked with and is usually used in cases where the encoding method of the original video is unknown [25].

NR-VQA methods are classified according to the extraction of indicators at three fundamental levels. NR-P methods are based on the analysis of indicators such as blur, noise, block effects, or temporal deterioration. NR-B methods are based on coding information derived from the bitstream and packet header or can also be the result of a combination of both methods. In these methods, the indicators considered are bit rate, video motion, packet loss, and codec parameters, which are more closely related to the information transmission process. In addition, there are hybrid methods based on the combination of the two previous evaluation forms known as NR-P-B metrics.

Many of the objective quality assessment algorithms were originally developed for the context of images. Still, they have been extrapolated and improved in video processing by applying them to each frame of the sequence, averaging the quality levels measured throughout the video [26]. Methods have also been developed specifically for video, with the same purpose of identifying the degradations that can occur not only in the spatial dimensions of the sequence but

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

also in the temporal ones. The development of these methods can derive benefits from new high-performance hardware architectures for designing and optimizing digital image and video processing techniques.

The use of objective metrics has as main difficulty in the case of FR-VQA that it is required to have the original video and the processed video, which usually conspires against its deployment in real-time or online systems, in addition to representing a higher computational cost. Several objective video quality measures that are limited to the differences between the original and degraded sequences can be listed and apply complex calculations based on the human visual system [27]. In this sense, an analysis of the progress achieved in developing objective NR-VQA algorithms using available computational technologies is required. In the current developments, objective quality measures predominate, so it is very convenient to develop new systems or algorithms that contribute to obtaining them automatically. Subjective evaluations are ruled out as they are impracticable since, in many cases, it is necessary to process huge volumes of audiovisual information in real-time.

It is necessary to consider that in most cases, the original video recorded in ideal conditions is not available so that it can be a reference for the quality evaluation. On the contrary, the applications usually only have access to the video sequence of a particular flow without additional concerns. In this case, it is imperative to emphasize the study of NR-VQA methods since evaluating the degradations added to the frame sequence from the capture stage is necessary. This is based on the evaluation of possible affectations due to several external factors that have been previously described. The independence of deep comparisons between frames becomes a significant advantage for the process [28].

This study focuses primarily on the analysis of NR-VQA methods because it is necessary to continue developing increasingly efficient ways that do not require references to achieve an analysis without comparison and taking advantage of the capabilities of new computational infrastructures. The fundamental purpose of the objective NR-VQA methods is to make it possible to determine the quality of each of the frames similarly as experts perceive it. The use of subjective methods based on the video content analysis, including specific encoding and transmission parameters that make it possible to search for degradations such as block effects and blurriness.

3.2. Objective methods of video quality assessment without a reference using Deep Learning

NR-VQA methods based on Deep Learning are a current trend in the field of VQA. It is necessary to consider that it is required to process large video streams that are usually transmitted through data networks, online systems, or real-time systems. For this reason, an additional requirement is to minimize the execution time of the VQA process, including training the deep learning neural networks that support it. Figure 1 shows a deep learning neural network as an overview of a bio-inspired process that is recurrent as a platform in the field of NR-VQA methods.

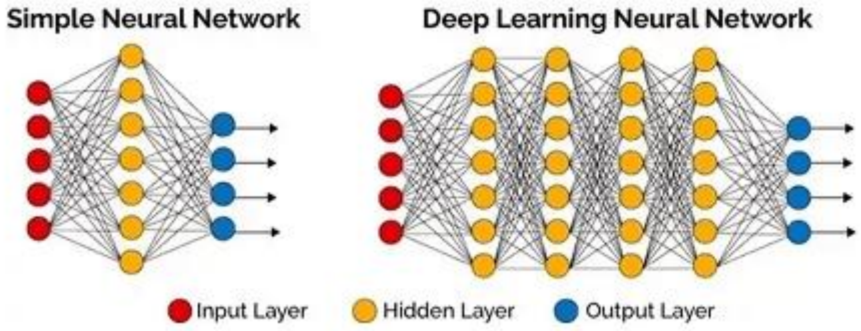

Figure 1: Simple and Deep Learning Neural Network.

To achieve automation of the quality assessment process under time-constrained conditions, it is necessary to employ parallel processing techniques to optimize high-performance computational infrastructures to achieve greater effectiveness [29]. Knowledge initially generated in the field of artificial intelligence is reused. NR-VQA methods are linked to machine learning [30] and include supervised learning that allows the resolution of known problems by using a set of labeled data to train VQA algorithms.

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

TELEMÁTICA
REVISTA DIGITAL de las Tecnologías
de la Información y las Telecomunicaciones

Unsupervised learning methods work by themselves to discover the structure of unlabeled data, attempting to eliminate human intervention in the output validation process. This learning model approach is very practical for realizing accurate video quality estimation. Deep learning is part of machine learning methods and is based on concepts associated with artificial neural networks (ANN-Artificial Neural Network) [31]. Deep learning makes it possible to develop bio-inspired processes that emulate the functioning of the human brain in terms of learning directly from the data. This is achieved without the need to extract features manually, which, combined with unsupervised learning models, enables the development of highly effective and accurate systems.

These networks are beneficial for various recognition and classification processes based on the search for patterns with defined purposes. Machine learning performs automatic data analysis, the construction of analytical models using theoretical knowledge developed in the field of artificial intelligence that enable systems to "learn" from data, and patterns and make decisions minimizing human intervention. The models built can produce results with a high level of accuracy, which are also scalable. These algorithms are designed from the integration of methods capable of processing a large volume of input data in real time that can be accelerated by using parallel designs operating based on high-performance computational means [32].

Algorithms in the context of DNN are structured in interconnected layers of processing that emulate the human brain's connections, precisely enabling "machine learning". NR-VQA objective methods have evolved significantly; they result from the implementation of algorithms that allows the construction of models and the discovery of connections capable of making more accurate decisions. The multilayer perceptron is an artificial neural network (ANN-Artificial Neural Networks) formed by multiple layers that can solve nonlinear problems [33]. It can be fully or locally connected. It is posited to be fully connected when each output of a neuron in a layer is input to all neurons in layer i+1. It is locally connected when each neuron in the layer constitutes the input of a subset of neurons in layer i+1.

Layers are classified into three types:

- Input layers: Introduce the input patterns into the network; no processing occurs in these neurons.
- Hidden layers: Neurons whose inputs come from earlier layers, and their outputs are passed to neurons in later layers.
- Output layer: Neurons whose output values correspond to the outputs of the entire network.

Deep learning usually requires a huge amount of data to learn and is structured as a hierarchical element of machine learning. This learning can be: supervised, unsupervised or semi-supervised, depending on the level of user intervention in decision making. Different types of algorithms in deep learning run data through various levels of neural network algorithms [34]. Usually, data science and programming are usually responsible for feature extraction and understanding a deep learning model.

Deep learning is employed for constructing and training a neural network based on a set of machine learning algorithms. This technic model's high-level data abstractions using computational architectures that support multiple and iterative nonlinear transformations of data expressed in matrix or tensor form [35].

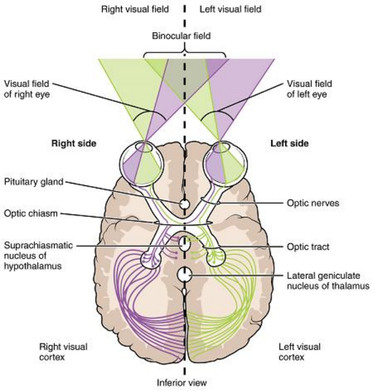

Figure 2: Scheme of visual cortex.

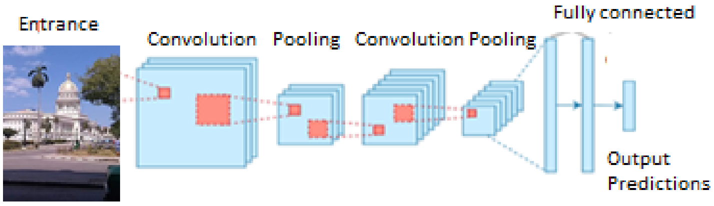

Figure 3: Convolutional Neural Network of VQA.

The study of emerging NR-VQA objective methods is critical since they have been shown to have high accuracy in recognition and classification tasks. These methods have been developed using deep neural networks (DNN-Deep Neural Network) and specifically based on a type of artificial neural network with supervised learning, called convolutional

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

neural network (CNN) [36]. In this case, the neurons corresponding to receptive fields in a very similar way to the neurons of a biological brain's primary visual cortex (V1). This type of network is a variation of multilayer perceptron whose application is realized using two-dimensional arrays, which are very effective in computer vision.

CNNs consist of multiple layers of convolutional filters of one or more dimensions. After each layer a nonlinear causal mapping is usually added, the essential feature of which is the idea of feedback. In nonlinear causality cause and effect can be connected directionally between two or more elements or systems, and a result can have multiple reasons. It is important to specify that training a convolutional neural network is a complex process that requires an analysis of the behavior of the network and how the hyperparameters influence the final result. This process is associated with a high computational cost, so a high computational capacity is required for its effective execution. It should be considered that in the training process, the mathematical operations to be performed by the CNNs involve a large number of iterations to achieve the convergence of the network.

The NR-VQA objective methods based on DNNs, are part of the first phase of the video preprocessing stage and precede the realization of subsequent modules that impact the quality improvement process in various ways. On each input frame, the analysis of the parameters associated with video quality is performed, such as: video format, resolution, bitrate, number of frames per second. The convolution operator aims to filter the input image in a previously trained kernel. This convolution transforms the data so that certain features become dominant in the output image by having a higher numerical value assigned to the pixels representing it. The max-pooling operation searches for the maximum value within a sample window and transfers this value as a summary of the features over that area. As a result, the data size is reduced by a factor equal to the size of the sample window over which it operates.

New video analysis solutions employ these methods for detecting specific distortions and noise more effectively than conventional video analysis methods. These emerging methods can reuse the performance of dissimilar high-performance computing environments in which multiple processing cores. Graphics processing units or tensor processing units (TPUs) can converge as application-specific integrated circuitry to accelerate proprietary artificial intelligence (AI) algorithms [37]. As a whole, these processing units achieve a high acceleration of algorithms when it is required to process large volumes of information that contribute to the learning, recognition, and identification of related patterns in other datasets.

Most VQA methods are based on analyzing the behavior of the combination of five fundamental features: contrast, color, blur, spatial information and temporal information. VQA modules are usually part of the preprocessing stage; specifically, NR-VQA methods also precede the process of enhancing degraded videos and increasing the resolution of lossless videos, among others. In the most recent five years in computer vision, several objective NR-VQA methods based on DNN have been proposed [38], which have been shown to have superior accuracy and lower weights in video recognition and classification tasks.

NR-VQA models have been in full development for many companies in the telecommunications sector dedicated to content distribution, such as YouTube and Facebook; however, they have not been deployed in many other required contexts. The vast majority of companies in the telecommunications sector do not currently employ NR-VQA model-based methods, which are highly adaptable to the solution of problems characterized by a high degree of difficulty, diversity, and real-time processing requirements.

The authors of this article have focused on NR-VQA methods based on CNN, whose standard and distinctive features are very useful for the development of new methods using the theoretical foundations discussed above. In the following sections, a review of eight objective NR-VQA methods designed by various international research groups to form the basis for developing proprietary NR-VQA strategies for processing multiple real-time video streams is presented. These methods have been considered relevant based on the analysis of their experimental validation results [39].

DeVIQ: A deep model of NR-VQA

The NR-VQA (DeVIQ-Deep Video Quality) deep model employs a pre-trained classification DNN by creating hierarchical sub-images. It can process Ultra High Definition (UHD-Ultra High Definition) content with a 3840 x 2160 pixels resolution. It can be trained using a large dataset generated from an FR-VQA. The proposed model achieved a high prediction quality measurement, comparable to an FR-VQA method. The authors have considered it as a proof of concept for efficient video quality estimation in video streams.

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

DeViQ has managed to handle the most recent quality models, is applicable to images and videos, uses the combination of patches (subsection of the input frame) using CNN to reduce computational complexity. The DNN is pre-trained for reference-free feature extraction using a full reference metric that enables the automatic construction of the training database, therefore, a large dataset for training can also be dispensed with, although it should be noted that these patches do not allow managing global image connections [40].

The general architecture of the model consists of 3 stages: The first stage performs feature extraction using a DNN with a hierarchical patching method. The second is a feature selection algorithm, and the third is quality score calculation using a random forest model, a widespread machine learning technique with utility generation capability for diverse problems. The pre-trained DNN computes a feature vector of probabilities similar to what can be perceived after searching a given image from a cognitive point of view. The search and classification of a known pattern in an unknown image is performed, and each probability of the calculated vector represents an indicator for the relationship of the known pattern concerning the object.

In conjunction with the classification DNN, VMAF (Video Multimethod Assessment Fusion) score is employed, an objective full-reference video quality metric developed by Netflix in cooperation with the University of Southern California. VMAF predicts quality based on a reference and distorted video stream. This metric evaluates the quality of different video codecs, encoders, encoding configurations, or streaming variants. In essence, it has been shown that the DeViQ system can employ a NR-VQA model over a FR-VQA using a DNN, which provides high effectiveness and can be extended from employing a full reference model that can be subsequently used to train DeViQ itself.

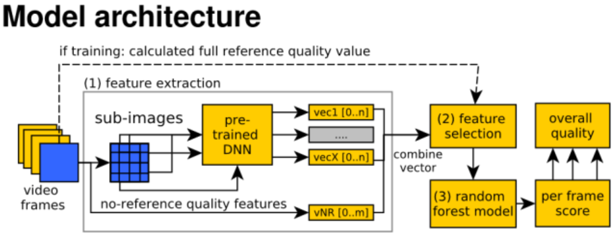

Figure 4: Arquitectura del modelo DeVIQ.

NR-VQA based on the temporal concentration of features obtained with CNNs

This method is based on a feature concentration algorithm using a CNN for evaluating the video sequence. First, the features of each frame are extracted using pre-trained CNN networks, then the feature vectors of each frame are concentrated. Finally, the video-level features are temporally concentrated and mapped onto a subjective quality score with a pre-trained support vector regressor (SVR).

It can be summarized that the framework developed for NR-VQA is based on the features obtained from the pre-trained CNNs (Inception-V3 [41] and Inception-ResNet-V2 [42] transfer learning, temporal pooling, and regression. The main contribution of the method is to achieve temporal pooling of feature vectors at the frame level, avoiding manual feature derivation. The authors demonstrated that with the features extracted from the pre-trained network, it is possible to perform more effective representations for video quality tasks using NR-VQA methods [43].

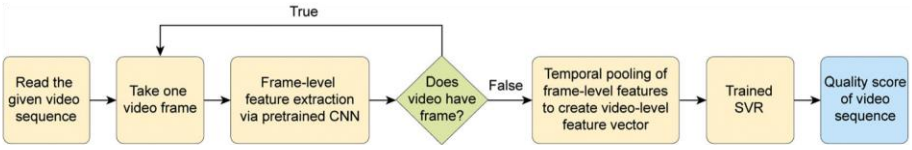
True
Read the given video sequence
Take one video frame
Frame-level feature extraction via pretrained CNN
Does video have frame?
False
Temporal pooling of frame-level features to create video-level feature vector
Trained SVR
Quality score of video sequence

Figure 5: DeVIQ model architecture.

The algorithm was implemented in MATLAB R2018b, employing the Deep Learning Toolbox functions consisting of Neural Network Toolbox and Statistics and Machine Learning Toolbox. It was trained and tested using KoNViD-1k [44], a natural video quality database containing 1200 sequences with a quality score. Its execution was performed very well, approximately 0.1 in terms of Pearson's linear correlation coefficient (PLCC) and Spearman's rank-ordered correlation (SROCC). The method was also validated using the LIVE-VQA database. The authors used a compute node with an octa-core i7-7700K CPU and a NVidia Titan X GPU.

Strong baseline for image and video quality assessment

The method is an efficient, high-performance unified model that enables the evaluation of videos and images and achieves comparable or superior performance to other existing methods. A model that achieves an excellent complexity to performance ratio is proposed. Therefore, it may be helpful in deployment in industrial environments. Based on the method, three derived training models have been proposed that are directly employed in video quality with refinement of the dataset [45].
An architecture that works for FR-VQA and NR-VQA models is proposed that realizes the application of a low-weight network as a backbone of the proposed model to achieve efficiency of the inference process. The output characteristics of the last convolution layer of the backbone feed the GAP-Global Average Pooling (GAP) module [46]. It has two layers fully connected with a DNN that has 1024 hidden nodes that make it possible to flatten the extracted features, representing the GAP inputs. For the evaluation of the video or image quality, each frame is fed to the model to obtain the final quality score as a result of averaging the quality of all the frames.

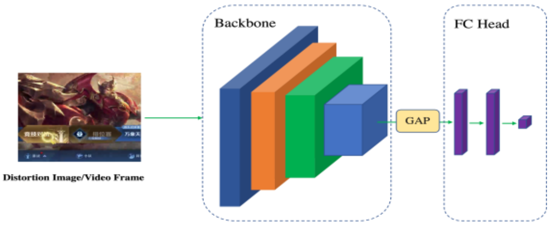

Figure 6: Network architecture of the NR model.

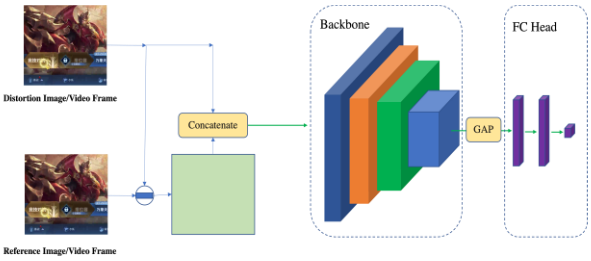

Figure 7: Network architecture of the FR model.

DeepBVQA

The Deep Blind Video Quality Assessment (DeepBVQA) method is a highly accurate and reliable deep learning scheme. The quality assessment is performed considering various spatial and temporal parameters. In the first step, the video is also processed as a sequence of independent frames, from extracting spatial and temporal characteristics or quality indicators for each frame. The vector or function corresponding to the extracted parameters is obtained [47]. In the feature aggregation stage, the set of features of the frames extracted from the distorted video using a CNN is combined into a function. Finally, the video feature is trained to predict the quality score in the regression stage. Experimental results show that the performance of this model is very effective and highly efficient for video quality assessment.

The deep learning DeepBVQA scheme method achieves more reliable measurement. The CNNs are employed for the extraction of the spatial samples of each video in the VQA and the new features are proposed for the temporal samples. The performance achieved by this method is identified as one of the best NR-VQA that have been developed so far. The performance of the proposed model concerning the VQA models shows the standard PLCC and SROCC as quality criteria of these VQA models on the LIVE database and the CSIQ video quality database. The high-performance VQA algorithm is handy for achieving high-quality video displays comparable with FR and RR metrics [48].

The authors of the method evaluated the database to demonstrate the generality of the proposed model. After a Deep-VQA learning model using 80% of the training data from the LIVE VQA database, the quality scores are inferred on the CSIQ video database using the parameters of the training model. The DeepBVQA method contributes to NR-VQA of distorted videos without the need to employ pre-trained CNN models that enable the extraction of spatial and temporal features to obtain a better evaluation. According to the authors of the method, it is necessary to continue studying the methods of temporal feature extraction using the Deep Learning model and the adaptive temporal evaluation method considering the characteristics of the video content.

Method for evaluating the quality of video captured in natural environments

In evaluating the quality of videos captured in natural environments, the complexity of the distortions and the diversity of the video contents come together. In this case, the number of existing distortion datasets is entirely different. The performance improvement of this quality assessment model is based on human perception, specifically in dependence on the content and temporal memory effects of HVS. Exploration is performed with a training strategy with mixed

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

datasets of a simple VQA model using multiple datasets. The proposed unified framework includes three stages: the first is the relative quality assessor, and the second is the nonlinear mapping and perceptual scale alignment of the specific dataset for the prediction of comparable quality, perceptual quality, and subjective quality.

The sampling algorithm extracts temporal redundancy by selecting a set of representative frames. These frames are transferred to two low-weight CNNs that encode each frame's attribute quality and semantic content. The frame-level features are aggregated within the video features and mapped into a quality score using SVR. The authors conducted experiments using four large-scale users generated content (UGC) video databases that made it possible to demonstrate the accuracy of the proposed method. On the other hand, the method is considered efficient; as reported by the authors, it was executed with a speedup of 185 FPS on an NVIDIA GPU card on an Intel i7-770 CPU for 1080p videos [50].

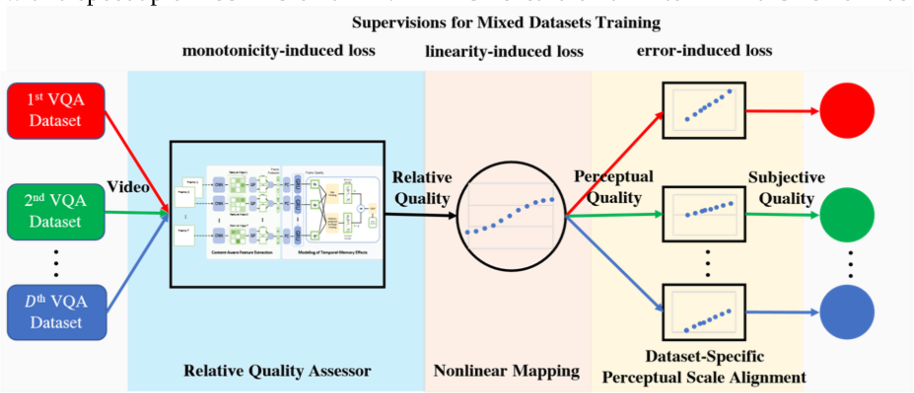

Figure 8: Architecture of the NR-VQA method for videos captured in natural environments.

The method has as a relevant component the use of a mixed dataset training strategy for videos taken in a natural environment. The backbone model is a neural network for characterizing two relevant effects of the human vision system depending on the content and temporal memory effects. The training of datasets is enabled by the design of two losses (induced monotonicity loss and induced linearity loss) that will allow the prediction of relative quality, perceptual quality, and the assignment of perceptual scale levels of the specific dataset for the prediction of subjective quality. This approach enables superior performance and allows verification of the effectiveness of the unified VQA model.

Video quality evaluation method based on Benford's law and perceptual characteristics

The goal of this method is to propose a feature vector for NR-VQA inspired by Bedford's law, which is based on probability theory. It has been found that the probability of the first non-zero digit "n" in a sample of numbers drawn from the real world appears with a probability with the logarithmic distribution. It has allowed demonstrating specifically in this domain that the first digit distributions drawn from different domain transformations of video data result in quality features and can be effectively mapped into perceptual quality scores. The feature vectors extracted from the training video sequences are used to train a learning machine model in state tests, which make it possible to estimate the perceptual quality of videos without prior references.

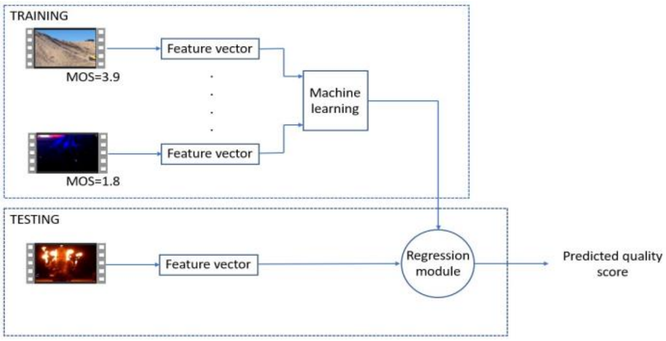

Figure 9: Architecture based on Benford's law.

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

A set of first digit distribution (FDD-First Digit Distribution) and perceptual quality features presented for NR-VQA are analyzed for application within the model. Several features based on the HVS quality criteria are also incorporated. A subset of perceptual features was also applied for reference-free image quality assessment, such as: blur, color, contrast, dark channel features, entropy, phase congruency means, spatial information, temporal information, and natural image quality evaluator (NIQE). A pre-trained convolutional neural network was applied for feature extraction from video frames, while temporal features are modeled from their "manual" shaping [51].

Efficient method for the evaluation of video quality without reference

The efficient method for reference-free video quality assessment proposes a module capable of selecting a predetermined number of frames from the entire video sequence on which the quality assessment will be performed. The quality attributes and the content of the structures are encoded using two low-weight CNNs. Subsequently, the quality scores of the whole video is estimated using an SVR. The proposed method was validated by employing four databases containing user-generated videos (UGC): (CVD2014, KoNVID-1K, LIVE-Qualcomm and LIVE-VQC). The authors of the technique demonstrated its accuracy.

The presented results demonstrate that the proposed method achieves online video quality prediction with low computational cost by employing independent databases and has the potential for generalization with a cross-database setup [52]. It is highly effective and efficient since it removes temporal redundancy by selecting and analyzing a set of representative frames transferred to low-weight CNNs, avoiding processing all the sequence frames. This strategy reduces the process execution time, achieving high accuracy several times more efficiently than other methods.

Fast and accurate video quality prediction method of user-generated content

The Rapid and Accurate Video Quality Evaluator (RAPIQUE) [53] introduces an efficient and effective video quality model. It performs an evaluation comparable to the most recent models, but has considerably faster run times. This method combines and balances the advantages offered by the statistical characteristics of scenes known as quality-aware and deep convolutional, known as semantic awareness. They allow designing the first general Spatio-temporal bandpass statistical model that is efficient for video quality modeling. The experimental results of RAPIQUE are obtained on large-scale UGC databases to achieve high performance at a relatively low computational cost.

This model derived new spatially efficient statistical models of the scenes and an ensemble model that makes it possible to include normalized temporal bandpass responses. Among the advantages of this method is the practical modeling of video quality problems for real-time and low latency applications. Given the existence of natural scene statistics and deep learning features that capture spatial distortions, predicting the quality of complex UGC videos is useful for a variety of practical applications.

3.3. Advantages of CNN-based NR-VQA methods

As a result of the study of more than 40 emerging NR-VQA methods developed in the last five years, it has been identified that machine learning enables the construction of evaluation models that do not require comparison with pre-existing videos. In all cases, the first step is the extraction of frames from the video sequence to be analyzed, to achieve the final score of the quality of the entire video, using various techniques that in many cases, average the values resulting from the analysis of the performance metrics that are obtained for each frame. It is always possible to retrain the model by adding the false positives or false negatives detected from the corresponding set and other pre-trained CNNs.

The CNNs contribute to the extraction of features from the information in the frames resulting from the decomposition of the input video, this sub-process is very useful in the evaluation process since they make it possible to obtain more accurate information for making decisions regarding video quality [54]. In all cases, the parameters analyzed in the CNN are adjustable for the function to be modeled. The CNN training algorithms constitute an optimization problem that defines the function to be optimized. The values of the network parameters make it possible to find the optimum of the function from a given strategy.

It is important to consider that all the studied methods are usually validated by employing various video databases containing distortions proper to scenes captured in natural or artificial conditions. They have associated approaches that are very useful for designing and implementing new NR-VQA methods more effectively and efficiently. Each of

the methods described contains references by their authors to the articles that gave rise to them and offer assessments that are important for developing and implementing new and increasingly efficient procedures. The increased computational capacity brings additional potentialities to face the new challenges in this field.
A review of more than 70 publications describing methods developed by various groups of experts from international universities and research centers has been carried out. The emphasis has been on eight ways that provide beneficial knowledge for designing and optimizing new, more effective NR-VQA methods based on CNN for processing multiple video streams.

4. CONCLUSIONS

Objective benchmark-free video quality assessment methods are more effective for video quality estimation than conventional methods. Their use is more appropriate in the context of a wide range of applications requiring real-time video frame sequences processing. Video quality assessment is now inextricably linked with machine learning and high performance computing infrastructures. This technological alliance offers potential capabilities for developing methods that can be scaled using parallel programming techniques. Integrating the potentials described by the NR-VQA methods based on CNN is highly efficient and constitutes the basis for the design of increasingly effective VQA processes. These methods permit to extract useful information for the video enhancement process by determining the type of degradation present in each of the frames, including the regions of the image in which these distortions and noises are located. An analysis of the main regression mechanisms currently used as an essential part of the evaluation methods scheme has been carried out, which allows us to conclude that the methods using CNN are the most adequate for identifying and classifying degradations.

5. FUTURE WORK

Design and implement an automatic NR-VQA objective method that significantly reduces the process execution costs, as a result of maximizing the use of the available computational capacity, to achieve real-time processing of multiple video streams.

ACKNOWLEDGMENTS

The authors would like to thank the High-Performance Computing Laboratory for Telecommunications, the Telematics Research Group, and the Department of Telecommunications, belonging to the Universidad Tecnológica de la Habana, the environment in which this research has taken place.

BIBLIOGRAPHY

[1] Zhengzhong Tu, Yilin Wang, Neil Birkbeck, Balu Adsumilli, Alan C. Bovik, "UGC-VQA Benchmarking Blind Video Quality Assessment for User Generated Content", arXiv:2005.14354v2, 2021.
[2] Ngai-Wing Kwonga, Sik-Ho Tsangb, Yui-Lam Chana, Daniel Pak-Kong Luna b, and Tsz-Kwan Leec, "No-Reference Video Quality Assessment Metric Using Spatiotemporal Features Through LSTM", International Workshop on Advanced Imaging Technology (IWAIT), 2021.
[3] Qian Cheng, "Application of Computer Vision Technology in Industrial Automation", ICAIIT 2021, 2021.
[4] "Cisco Annual Internet Report (2018–2023) White Paper", https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html, 2020.
[5] Shahriar Akramullah, "Digital Video Concepts, Methods, and Metrics, A press open Springer", A press open Springer, 2014.
[6] Leandro Marcomini, "The impact of different video resolutions in a feature-based vehicle detection", https://www.researchgate.net/publication/335888020 , Conference Paper XXXIII ANPET Congresso de Pesquisa a Ensino em Transportes, 2019.
[7] Jordan W. Moffett, Judith Anne Garretson Folse, Robert W. Palmatier, "A theory of multiformat communication: mechanisms, dynamics and strategies", Journal of the Academy of Marketing Science (2021) 49:441–461, 2021.
[8] Helard Becerra Martinez, Andrew Hines, Mylène C. Q. Farias, “Perceptual Quality of Audio-Visual Content with Common Video and Audio Degradations”, MDPI, 2021.
[9] Krzysztof Okarma, Piotr Lech, Vladimir V. Lukin, "Combined Full-Reference Image Quality Metrics for Objective Assessment of Multiply Distorted Images", MDPI, 2021.
[10] Derrick Timmerman, Guru Swaroop Bennabhaktula, Enrique Alegre, George Azzopardi, "Video Camera Identification from Sensor Pattern Noise with a Constrained ConvNet", CPRAM 2021, 2021.
[11] Chen, Mettes, Snoek, "Diagnosing Errors in Video Relation Detectors", Computer Vision and Pattern

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022
Sitio web:http://revistatelematica.cujae.edu.cu/index.php/tele

Recognition 2021, CPRAM 2021, 2021.

[12] "Perceptual Quality of Audiovisual Content with Common Video and Audio Degradations", MDPI, 2021.

[13] Xuayi Wu, Irene Cheng, Zhenkun Zhou, Anup Basu, "RAVA: Region-Based Average Video Quality Assessment", Sensors, 2021.

[14] Sensors, 2021. "UGC-VQA: Benchmarking Blind Video Quality Assessment for User Generated Content, Computer Vision and Pattern Recognition".

[15] D. Ghadiyaram, J. Pan A.C. Bovik, A. K. Moorthy, P. Panda, and K. C. Yang, "Subjective and objective quality assessment of mobile videos with in-capture distortions" IEEE Int. Conf. on Acoustics, Speech and Signal Proc., 2017.

[16] Nagabhushan Somraj, Manoj Surya Kashi, S. P. Arun, Rajiv Soundararajan, "Understanding the Perceived Quality of Video Predictions", Image and Video Processing, arXiv:2005.00356v5, 2021.

[17] "Recent Advances and Challenges in Video Quality AssessmentVideo Quality Assessment". DOI: 10.12142/ZTECOM.201901002, 2019.

[18] "NTIRE 2021 Challenge on Quality Enhancement of Compressed Video: Methods and Results", NTRIE 2021,2021.

[19] "Models and tools for quality assessment of streamed media", ITU-T P.1204, 2020.

[20] Duin Baek, Mallesham Dasari, Samir R. Das, Jihoon Ryoo, "dcSR: Practical Video Quality Enhancement Using Data-Centric Super Resolution", 17th International Conference on emerging Networking Experiments and Technologies (CoNEXT' 21), 2021.

[21] I P Skirnevskiy1, A V Pustovit1, M O Abdrashitova, "Digital image processing using parallel computing based on CUDA technology", Journal of Physics, 2017.

[22] "Methods for objective and subjective assessment of speech and video quality", P800.1 ITU-T, 2016.

[23] Kjell Brunnström, Anders Djupsjöbacka, Börje Andrén, "Objective video quality assessment methods for Video assistant refereeing (VAR) System", RISE report 2021:30, 2021.

[24] "The state of broadband 2021: People-Centred Approaches for Universal Broadband", ITU-UNESCO Broadband Commission for Sustainable Development ITU-UNESCO, 2021.

[25] Sathya Veera Reddy Dendi, Sumohana S. Channappayya, "No-Reference Video Quality Assessment Using Natural Spatiotemporal Scene Statistics", IEEE Transactions on Image Processing, 2020.

[26] I P Skirnevskiy, A V Pustovit, M O Abdrashitova, "Digital image processing using parallel computing based on CUDA technology", Journal of Physics: Conference Series, 2017.

[27] Anastasia Antsiferova, Alexander Yakovenko, Nickolay Safonov, Dmitriy Kulikov, Alexander Gushin, Dmitriy Vatolin, "Objective video quality metrics application to video codecs comparisons: choosing the best for subjective quality estimation" GraphiCon2021: 31th International Conference on Computer Graphics and Vision, 2021.

[28] LI Dingquan, JIANG Tingting, and JIANG Ming, "Recent Advances and Challenges in Recent Advances and Challenges in Video Quality Assessment", DOI: 10.12142/ZTECOM.201901002, 2019.

[29] Nathaniel J. Tye, Stephan Hofmann, Phillip Stanley-Marbell, "Bridging the Band Gap What Device Physicists Need to Know About Machine Learning", arXiv:2110.05910v2, 2021.

[30] Wei Sun, Tao Wang, Xiongkuo Min, Fuwang Yi, Guangtao Zhai, "Deep learning based full-reference and no-reference quality assessment models for compressed UGC videos" arXiv:2106.01111v1, 2021.

[31] Yasamin Fazliani, Ernesto Andrade, Shahtam Shirani, "Neural network solution for a real-time no-reference video quality assessment of H.264/AVC bitstreams" Multimedia Tools and Applications, Springer, 2021.

[32] Nabajeet Barman, Emmanuel Jammeh, Seyed ali Ghora shi, Maria g. Martini, "No-Reference

Video Quality Estimation Based on Machine Learning for Passive Gaming Video Streaming Applications", IEEE Access, 2019.

[33] Hassan Ramchoun, Mohammed Amine Janati Idrissi, Youssef Ghanou, Mohamed Ettaouil, "Multilayer Perceptron: Architecture Optimization and Training", IJIMAI, 2021.

[34] Michael W. Berry, Azlinah Mohamed, Bee Wah Yap, "A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science", Springer, 2020.

[35] Mahmoud A. M Albreem, "Efficient Iterative Massive MIMO Detector Based on Iterative Matrix Inversion Methods", IEEE Access, 2021.

[36] Bayar, B. and Stamm, M. C., "Design principles of convolutional neural networks for multimedia forensics", Electronic Imaging, 2017.

[37] Amir Yazdanbakhsh, Kiram Seshdri, Berkin Akin, James Laudon, Ravi Narayanaswami, "An Evaluation of Edge TPU Accelerators for Convolutional Neural Networks", arXiv:2102.10423v1, 2021.

[38] Markus Utke, Saman Zadtootaghaj, Steven Schmidt, Sebastian Möller, "Towards Deep Learning Methods for Quality Assessment of Computer-Generated Imagery", arXiv:2005.00836v1, 2020.

[39] Junyong You, Jari Korhonen, "Deep Neural Networks for No-Reference Video Quality Assessment", IEEE International Conference on Image Processing (ICIP), 2019.

[40] Steve Göring, Janto Skowronek, Alexander Raake, "DeViQ- A deep no reference video quality model", Technische Universität Ilmenau Germany, 2018.

[41] "Pretrained Inception-V3 Convolutional Neural Network", 2019.

[42] Szegedy C, Ioffe S, Vanhoucke V, Alemi AA, "Inception-v4, inception-resnet and the impact of residual connections on learning". In: AAAI, vol 4, p 12, 2017.

[43] Domonkos Varga, "No-Reference Video Quality Assessment Based on the Temporal Pooling of Deep Features" Neural Processing Letters-Springer, Neural Processing Letters (2019), 2019.

[44] Franz Götz-Hahn, Vlad Hosu, Hanhe Lin, Dietmar Saupe, "KonVid-150k: A Dataset for No-Reference Video Quality Assessment of Videos in-the-Wild", Transactions on circuits and systems for video technology vol. 21, no. 2, feb. 2020, 2020.

[45] Shaoguo Wen, Junle Wang, "A strong baseline for image and video quality assessment", arXiv:2111.07104v1, 2021.

[46] Chris, "What are Max Pooling, Average Pooling, Global Max Pooling and Global Average Pooling, machinecurve", 2020.

[47] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, "Rethinking the inception architecture for computer vision", In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826, 2016.

[48] Sewoong Ahn, Sanghoon Lee, "Deep Blind Video Quality Assessment Based on Temporal Human Perception", IEEE International Conference on Image Processing (ICIP), 2018.

[49] Deepti Ghadiyaram, Chao Chen, Sasi Inguva, Anil Kokaram, "A no-reference video quality predictor for compression and scaling artifacts", Google Inc., 2016.

[50] DingquanLi, Tingting Jiang, dingquanli, TingtingJiang, MingJiang, "Quality Assessment of in the Wild Videos", ACM ISBN978-1-4503-6889-6/19/10, 2019.

[51] Domonkos Varga, "No-Reference Video Quality Assessment Based on Benford's Law and Perceptual Features", Electronics 2021, 2021.

[52] Mirko Agarla, Luigi Celona, Raimondo Schettini, "An Efficient Method for No-Reference Video Quality Assessment", Journal of Imaging, 2021.

[53] Zhengzhong Tu, Xiangxu Yu, Yilin Wang, Neil Birkbeck, Balu Adsumilli, and Alan C. Bovik, "RAPIQUE Rapid and Accurate Video Quality Prediction of User Generated Content", arXiv:2101.10955v2, 2021.

[54] "Deep Blind Video Quality Assessment Based on Temporal Human Perception", 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018

AUTHORS SECTION

Omar Antonio Hernandez Duany. MSc. Signal Processing, Storage and Transmission. Assistant Researcher, Assistant Professor, Head of the Programming discipline and the TeleHPC Laboratory of the Faculty of Telecommunications and Electronics Engineering at CUJAE. ORCID: 0000-0002-0073-1036

Caridad Anías Calderón. PhD in Technical Sciences, Full Professor. Director of the Center for the Study of Technologies and Informatics of the CUJAE, President of the National Commission for Engineering in Telecommunications and Electronics, ORCID: 0000-0002-5781-6938

Jose Alberto Padilla Morani. Telecommunications and Electronics Engineer from CUJAE. Member of the TeleHPC.laboratory. ORCID: update field.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

CONTRIBUTIONS OF THE AUTHORS

The three authors contributed in the research, writing and edition of the current contribution.

Manuscrito recibido: 02-02-2022, aceptado: 07-04-2022