Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 2, pp. 1264–1275, DOI 10.12694/scpe.v25i2.2658

QUALITY ENHANCEMENT WITH FRAME-WISE DLCNN USING HIGH EFFICIENCY VIDEO CODING IN 5G NETWORKS

VIJAYA SARADHI DOMMETI[*], M. DHARANI[†], K. SHASIDHAR[‡], Y DASARATHA RAMI REDDY[§], AND T. VENKATAKRISHNAMOORTHY[¶]

Abstract. In the present situation, applications related to multimedia are discovered to be comfortable with the use of video. The number of end consumers who use video continues to rise every day. People are presently searching for videos with better quality among the ones that are currently available there. This results in the launch and dissemination of HD (high definition) videos. Ultra high-definition (UHD) videos are becoming more and more popular as a result of this advancement and need. However, as video communication keeps expanding, there is an upsurge in network traffic because of the limited bandwidth, especially among smart cities. Different advancement codecs have been suggested to deal with the data stream to overcome this hazardous circumstance. However, the fact that modern UHD videos have huge amounts of data makes the available codecs even more complicated. UHD videos can be processed with the latest improvement codec, H.265/High-efficiency video coding (HEVC). Nevertheless, it is impacted by increased power consumption and intricate calculations. Limitations in the codec's functionality confine its use to specific applications, preventing its application in wireless, mobile, or portable settings. Hence, this research concentrates on implementing frame-level quality enhancement through a deep learning network known as FQE-Net. The deep learning convolutional neural network (DLCNN) is specifically crafted to manage films with resolutions up to 16K. Its primary objectives include reducing complexity, minimizing artifacts, enhancing the efficiency of the HEVC codec, and compacting energy consumption. To achieve superior efficiency, it is imperative to replace the DWT transforms within the HEVC codec with a DLCNN model. Additionally, incorporating the Content Block Search Algorithm for Motion Estimation and Compensation, alongside filtering techniques like Sample Adaptive Filter and Deblocking Filter, becomes essential. The simulation results showed that the suggested FQE-Net performed better than the conventional techniques.

Key words: Convolutional neural networks for deep learning, high-efficiency video codec, and ultra high definition.

1. Introduction. End customers expect films with good resolution these days. Consequently, the demands of the end user are being met with difficulty by the multimedia firm and researchers [1]. The average user doesn't care about the technical details of the movie sequence; all that matters to them is watching a video of the highest quality without any buffering [2]. Because of this, a poll in [3] indicates that it is challenging to use present technology to offer a video of competitively good quality at a lower bit-rate and a higher enhancement ratio. Nevertheless, the surge in demand for superior-quality products is propelled by several factors, including the introduction of expanded communication bandwidths in networks such as LTE and 5G, the proliferation of smart devices, advancements in high-resolution display systems, market growth, elevated user standards, and innovations in technology. These days, it's common to access and upgrade 4K/8K-UHD video contents for distant applications [5]. However, as image quality rises to UHD, the previous H.264/AVC codec's coding efficiency runs out of bandwidth. These days, it's common to access and upgrade 4K/8K-UHD video contents for distant applications [5]. However, as image quality rises to UHD, the previous H.264/AVC codec's coding efficiency runs out of bandwidth. A new architecture is proposed to address the erotic scenario and make HEVC appropriate for both future UHD videos and already available HD videos [6]. The suggested codec is made with minimal resource consumption, minimal computational complexity, optimal latency time, and a lower bit-rate combined with better video quality. Video is becoming increasingly important with the introduction of new electronic devices like smartphones, HDTVs, multimedia systems, video surveillance, and more [7, 8]. These devices carry out several functions, including high-definition video conferencing, web browsing, sharing

[*]Department of ECE, Malineni Lakshmaiah womens Engineering college, Guntur, A.P., India. (saradhi.gitam@gmail.com)
[†]Department of ECE, Mohan Babu University, Tiruapthi, Andhra Pradesh, India (dharani405@gmail.com)
[‡]Department of ECE, Lords Institute of Engineering & Technology, Hyderabad, India (shashignitc2015@gmail.com)
[§]Department of CSE, Chaitanya Bharathi Institute of Technology, Proddatur, A.P., India (dasradh@gmail.com)
[¶]Department of ECE, Sasi Institute of Technology & Engineering, Tadepalligudem, AP, India (tvkmsvu@gmail.com)

social entertainment videos, and web monitoring transmission and distribution of video materials. One of the answers to the Digital Era [10], which has greatly impacted most industries, including communications, arts, entertainment, marketing, and media, is disruptive technology. This risky evolution is a byproduct of the modern period, which is distinguished by a dramatic shift from analog to digital technology. The speed at which digital content—including music, video, and data—is produced and accessed dictates how fast digital gadgets can process data. It is entirely up to us to decide what, if any, comes next. High-quality digital video has special opportunities and problems for analysis and visualization due to its extensive applications in domains such as storing data, internet streaming, monitoring, broadcasting, and conference calling [11].

1. FQE-Net's implementation to improve the video quality of HEVC standards based on H.265.
2. Using the DLCNN model will reduce complexity, energy compaction, and artifacts, and offer enhanced HEVC codec efficiency for handling videos up to 16K resolution.
3. The Motion Estimation and Compensation algorithms use the Content Block Search Algorithm; filtering techniques like the Sample Adoptive Filter and Deblocking Filter are also used.

The authors of [12] talked about using the FPS (Fast Predictive Search) algorithm for diamond searches. In this case, adding an initial search prediction is necessary. The search point in the first search step of our suggested technique is initially upgraded [13]. Here, the algorithm can locate the fundamental search point precisely and avoid the ineffective global search process. To reduce complexity, PIME (Predictive Integer Motion Estimation) is used in conjunction with a joint architecture algorithm to build an efficient high-motion estimation design. Based on analysis of potential search directions, this approach notably reduces the number of search points. The interpolation filtering procedure uses a processing unit and FME (Fractional Motion Estimation) [15]. For the computations, scheduling using cascade form is used for fractional motion estimation and integer motion estimation. By decreasing BW, gate count, and memory size, Z-scan [16] using indexed addressing simplifies the cache controller. A multi-rate encoding approach was presented by the authors in [17]; it is utilized to lower the encoding difficulty with several spatially distinctive resolutions. Finding similarities between the block structures of the various resolutions recorded using the high-resolution approach is the first step in the process. The resulting block structure is used to accelerate the encoding process [18]. The encoding of a low-resolution representation is accelerated by using a prefabricated block structure. To achieve individualized HEVC encoding with RD performance (Rate-Distortion), greater attention to the video content is needed. Based on simulation results, it is possible to cut the encoding time of flat-resolution video by 50 % on average without compromising rate-distortion performance [19].

The authors of [20] presented the Intra Mode and Inter Prediction decision approach. You should apply both tactics. A prediction with depth LCU technique with any adaptation is used in the third method, which is a generalized version of the first two Fast Intra Mode Decision methods [21]. We have developed four efficient early termination options to help you wrap up the RDO procedure quickly. Encoding times can be shortened by 54.4 %–68.54 %, based on the performance penalty. Based on a quality analysis of the material, the Fast Intra Prediction for Screen material Coding technique was first presented in the study [22]-[23] The size of the current CU is essentially determined by its per-pixel data rate. Every CU is a part of this class. We are expanding the range of data stored by nearby CUs and collocated CUs in order to increase speed even more [24]. According to the trial results, the recommended approach can encode video information with an average cost reduction of 44.92 percent and with the least amount of picture quality loss. The creators of [25]-[26] improved the HEVC's HM 16 tool was the reason for the process's good encoding efficiency and 1 % increase in the BD rate. Reference 27 outlines a method that streamlines the calculation and expedites reinforcement learning. The authors of [28] suggest using ABCO, or adaptive block coding, as a method of built-in prediction. Following enhancement, a volumetric bit rate The spatial relationship between close and distant blocks ensures a volumetric bit rate when it has been improved. For cost savings, using a method that adaptively nominates block and sub-block coding orders is advised [29]. In opposition to HEVC interfering with set blocking orders. The potential for improving predictions in the edge region using adaptive means is the Advocate method in this work. Utilizing the suggested algorithm, the HM9.2 software produced 1.3 % outcomes and 4.4 % bit savings. The authors of [30]-[31] used a motion estimation approach to compensate for frame rate, which lessens computational complexity. The recommended method produces clearer images than baseline techniques, increasing PSNR by 1.3 dB. One potential method for finding exact motion vectors with minimal processing complexity was the

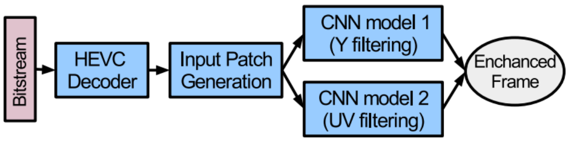

Fig. 2.1: FQE-Net proposal

Multi-Directional Motion assessment [32]. The Adaptive Motion Vector Smoothing Algorithm was proposed in the MVS phase as a different name for the AGMVS algorithm [33] in a way that modifies the motion vectors' failure by particular values in the in the unified framework. In light of the subjective judgment, the image is quite clear, but the objective examination resulted in a high PSNR.

The authors of [34] proposed using the Fast Coding Unit (FCU) choice for the HEVC encoding process, involving an offline approach for the Trained Supporting Vector Machines. By using the suggested approach, the proposed algorithm's complexity can be lowered by up to 48 % at a Bjontegard Delta Bit Rate of 0.48 [35]. 44 percent reduction in loss of RA (Random Access) setup using 0.62 % BD-BR. With a 41 percent decrease in LDB (Low Delay B) configuration and a 0.6 % BD-BR, the HEVC encoding complexity was efficiently reduced by a fast coding design approach, and the rate distortion optimization yielded competitive results. In reference [36], a new depth level and an inter-mode prediction method were introduced in an alternative to traditional SHVC. We first use interlayer correlation to predict potential depth ranges. The next step is to distinguish between squared and non-squared predictions within the depth candidate prediction mode [37, 38]. Gaussian distribution function is used for the early termination residual distribution. enhanced speed and efficiency of coding. The most computationally demanding activity and one that is extremely resilient to accuracy losses is the video encoding process with Motion estimation (ME) stage, which includes IME and FME techniques, as described by the authors in [39]. This paper provides an approximation of energy efficiency using motion estimation architectures that are supported by FME and IME [40].

2. Proposed Method. The proposed quality improvement system, depicted in Figure 2.1, employs the DLCNN-based filtering technique after fully decoding the current frame. Figure 1 illustrates this HEVC architecture modification. The two types of patches that are generated as the first step in the suggested technique filter the Y channel and ones that filter the U and V channels. The process then uses the two models to produce the increased chroma and luminance components, Ue ,Ve and Ye respectively, for the enhanced frame, Ye,Ue,Ve.

Block-based video coding is a widely used technique in the improvement process. Videos in this process can be categorized based on how many frames they contain. The most popular video encoding utilized by Block-based video coding is seen in Figure 2.2. Each frame in a particular frame sequence is divided into smaller parts using video coding blocks, from which the elements of frame 1 can be automatically predicted. Fundamentally, the motion estimation block generates a prediction before the application of the motion compensation block, or alternatively, the motion compensation block initiates an inter-prediction prior to the motion estimation block. Conversely, inter-prediction empowers the initial block of the slice algorithm to determine whether each frame should be intra-coded or not.

Prior to reaching the motion estimation and compensation block, the Motion Vectors undergo another round of entropy encoding. The present frames undergo subtraction from the Motion-Compensated frames (yielding the residual information) to generate the residual frame blocks. The rest of the frame blocks are quantized and changed before going into the entropy-based encoding stage. The identical set of projected data is available to both the encoder & the decoder. An encoder frequently incorporates a decoding loop as well, which uses the information at hand to piece together the original frames. In Images 2 and 3, dashed lines indicate the decoding blocks. Second, before the quantized data can be decoded, the decoder must pass it through inverse transformation blocks and inverse quantization.

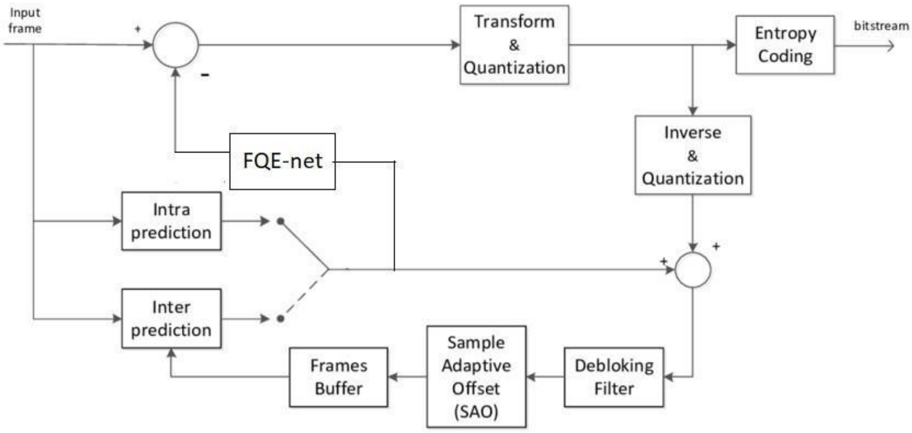

Fig. 2.2: The HEVC block diagram

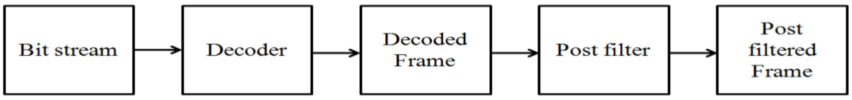

Fig. 2.3: Blocking artifacts reduction

2.1. Looping filters in HEVC. To decrease the effect of artifacts that obstruct highways, there are two general approaches that can be applied. They go by the labels post-filtering and loop-filtering, respectively. Figure 2.3 illustrates how the Deblocking filter is employed in post-filtering, which follows the decoder and makes use of the decoded parameters. It uses the display buffer to carry out its operations along with the main code loop. The frame needs to be filtered and encoded in the reference frame buffer before being sent to the monitor. To properly apply the post filter, the decoder might require an extra buffer. The deployment will not include an increase in bit rate or a change in encoding technique. It is completely optional to apply a post filter because none of the applicable standards require it. Loop filtering, which happens inside the encoding loop, will result in deblocking. This is shown graphically in Figure 2.4.

To correct the motion in frames, initially frames are sent to filtering primarily. For the decoder to carry out filtering in the same way as the encoder, a standard conformant decoder is needed. Each CTU's output is handled independently during decoding, and A reference frame buffer holds the processed data. Both the encoding and decoding processes employ the same filtering. There's no need for additional decoder buffering. However, this complicates incorporation into commercial code packages. When Post and Loop filtering are examined side by side, their advantages and disadvantages become evident.

One critical phase in block-based video encoders is motion estimation (ME), which plays a vital role in intelligently predicting Motion Vectors (MVs) for each block in every frame. This process potentially reduces the overall bit rate of a video stream. We utilize motion estimation for this purpose. Achieving precise motion estimation within a frame leads to a higher enhancement ratio, as it generates less entropy information for the remaining frame blocks. The first thing ME needs to do is figure out which block from the newest and most recent decoded frame (the area of interest, also called the search window) most closely fits every block in the reference frame. to determine the motion content of each block in each frame. Figure 2.5 depicts the notion of ME. The ME approach finds the best feasible match by first creating a distinct Search Window (SW) based on a given cost function for every block in the present frame. The final results generated by the ME are the coordinates of the optimum MV and the corresponding cost. The Lagrangian multiplier DMV, the distortion D, the bit-rate required to encode the motion vector R, the Lagrangian cost function J, the search window,

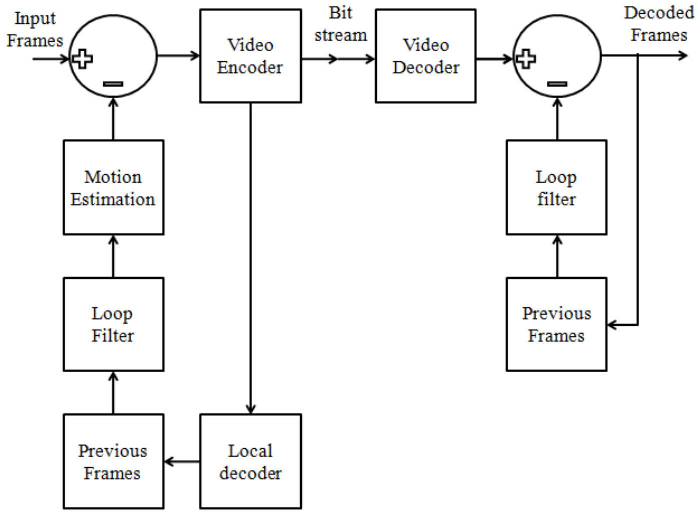

Fig. 2.4: Artifacts block reduction using loop filter

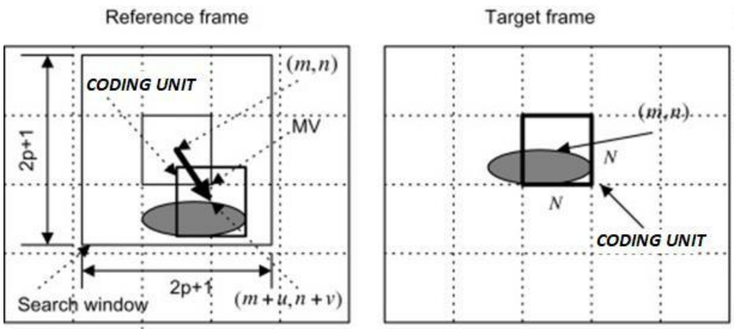

Fig. 2.5: A Process Illustration for Motion Estimation

and the optimal motion vector MV(x,y) can all be used to characterize the ME issue. Implementing an MV encoder with a bit rate of R makes it easier to solve the ME problem. The SAD or SSD distortion properties (Sum of Squared Difference) are utilized by numerous applications.

The SAD distortion feature is commonly utilized and that can

$$MV(x_i, y_j) = min(J_c oct(X + i, Y_j)(X + i, Y))cSW \tag{2.1}$$

$$J_c oct = D + h_M VR \tag{2.2}$$

$$D = SAD(x, y) = \sum_{i=1}^{M} \sum_{j=1}^{M} |C(X_i, Y_j) - R(X_i, Y_j)| \tag{2.3}$$

We can use MXN for the block's pixel size and R for a reference block, and C for the current block to get an estimate of the size of the active block (3). Video encoders for ME that are block-based frequently use Block Matching Algorithms (BMAs). When a problem arises in the software system, the ME technique, also

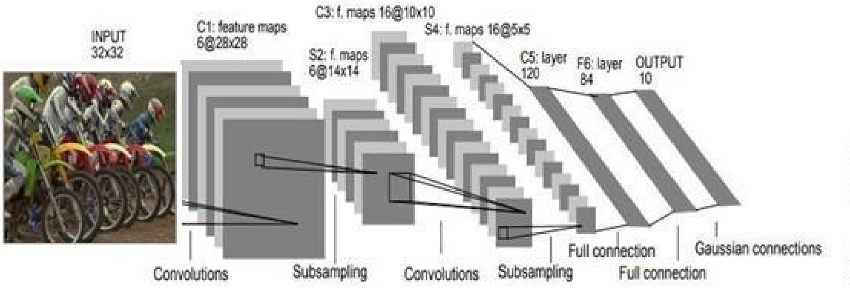

Fig. 2.6: Model of DLCNN

called the Full Search algorithm, looks into each block individually. Searching through every block in the entire software system is more challenging for the encoder. Effective ME approaches are used to bypass blocks that are less likely to have the optimal MV due to the limitations of video encoding time. Nevertheless, the final video quality (MVs) may degrade if you opt for a quick search approach instead of one that takes longer to find the best frames. Therefore, a quick and effective ME algorithm is required to reduce ME complexity while preserving an adequate ratio of enhancement and high-quality output video.

The suggested architecture for the DLCNN, illustrated in Figure 2.6, comprises convolutional blocks (CB), deconvolutional blocks (DB), attention-based shared weights blocks (ASB), attention-based shared weights residual blocks (ASRB), multi-head attention blocks (MHA), low-resolution feature fusions (LFF), and high-resolution feature fusions (HRFF). During its first stages of development, the DLCNN was inspired by three concepts: First, there's the attention mechanism that fine-tunes the feature maps using channel and spatial attention; The second idea is weight sharing, which employs the same convolutional layer twice in a single layer block; the third idea is a novel multiresolution fusion of features block design, which uses a design block to fuse current-resolution feature maps with either low- or higher-resolution feature maps.

There are three feasible patch resolutions to execute the DLCNN model: full, half, and quarter. This also includes the establishment of new hubs. The network is now better able to obtain feature maps from the input patch and forecast the specifics of the refinement at every level thanks to the addition of these additional nodes. The DLCNN helps all three patch kinds (hw, h2w2, and h4w4). To create the three feature maps, the input patch must be processed in the first stage of DLCNN. This method uses the entire patch for processing as opposed to the normal method, which only uses the input patch. In contrast to the conventional method that solely relies on full-patch resolution, the DLCNN algorithm gathers refined information from half, quarter, and full-patch resolutions. The typical approach disregards resolutions smaller than the entire patch in its standard setup. This deviation from the norm is essential.

3. Results and Discussions. The suggested method employed the DLCNN configuration, which produced superior visual and RD performance compared to alternative configurations. An improved visual result in random access setup is shown by the growing RD performance metrics, such as PSNR and bit rate. Additionally, resolution is used to gauge the suggested HEVC's visual analysis. The suggested technique reduced the quantity of transitional frames in the video while maintaining resolution throughout. The video's total memory usage decreases when the number of frames is decreased, but the resolution remains unchanged.

A range of video clips, such as People Street, BQTerrace, Basketball Pass and Cactus are used to assess the subjective performance of the proposed technique. These are high-definition videos, and the recommended method preserved the videos' sizes throughout. Table 3.1 displays the properties of videos in various sequences. The suggested parameters, including multiple test sequence types with varying QP and resolutions, are displayed in Table 3.1. Every aspect of human perception is dependent on rate-distortion optimization; prediction techniques, along with each of intra, random access, and low delay techniques, are crucial to this process.

The subjective assessments of the two samples' various approaches are displayed in Figures 3.2 and 3.3. In

Table 3.1: Assumption

	Suggested Values
Software Reference	HM 16.8
QP	22,27,32,37
Total Sequence	People on street, Cactus, BQ Terrance, Basketball pass
The Resolution	(2400x1600),(1920x1080),(832x480),(416x240)
Central Processing Unit	Intel Core i7 x 990, 3.47GHz
Operating System	Windows 10 (64-bits)

Fig. 3.1: A few frames from a video

Fig. 3.2: An impartial assessment of sample 1 (a) original frame, (b) HM 11[23], (c) Luo's [25], (d)PVC [34], (e) Proposed FQE-Net

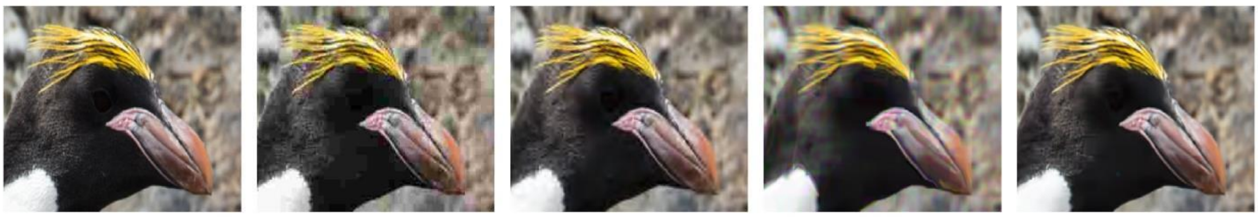

Fig. 3.3: An impartial assessment of sample 2, (a) original frame, (b) HM 11[23], (c) Luo's [25], (d) PVC [34], (e) Proposed FQE-Net

this case, Proposed FQE-Net outperformed traditional methods like HM 11[23], Luo's [25], and PVC [34] in terms of subjective performance.

Table 3.2 displays the results of the block mode tests we ran. When the Motion Vector's magnitude conditions are changed, the simulated analysis produces different outcomes. Table 3.3 illustrates how content-

Table 3.2: Block mode FQE-Net results

Sequence	QP	U-PSNR	V-PSNR	Y-PSNR	Bytes written to file	YUV-PSNR	Time (sec)
PeopleStreet	23	40.78	40.90	42.90	4457095	42.52	68.489
Cactus	23	40.70	40.70	42.34	6673845	42.09	122.70
BQTerrace	23	40.40	40.40	41.72	10947093	41.61	193.779
BasketballPass	23	40.39	40.46	41.80	1164445	41.66	21.63
PeopleStreet	28	36.40	36.40	36.82	3577587	36.67	76.30
Cactus	28	36.50	36.53	37.06	5473656	36.86	118.614
BQTerrace	28	35.39	36.41	36.84	90152893	36.69	190.207
BasketballPass	28	36.39	36.44	36.86	958439	36.71	22.14
PeopleStreet	33	32.49	32.49	31.87	2921654	32.08	64.06
Cactus	33	33.59	32.59	31.91	44438764	32.11	114.741
BQTerrace	33	32.40	32.49	31.87	7354124	32.08	189.50
BasketballPass	33	31.60	32.54	31.90	782213	32.10	22.12
PeopleStreet	37	29.60	29.55	27.97	2346543	27.68	69.38
Cactus	37	29.60	29.58	27.01	35727654	27.71	114.569
BQTerrace	38	29.60	29.55	27.97	5899859	27.67	187.112
BasketballPass	38	29.60	29.56	27.98	628833	27.69	19.91

Table 3.3: Content mode results of FQE-Net

Sequence	QP	U-PSNR	V-PSNR	Y-PSNR	Bytes written to file	YUV-PSNR	Time (sec)
PeopleStreet	23	41.32	40.32	40.61	433795	40.51	15.61
Cactus	23	41.36	40.36	40.65	6585630	40.55	24.93
BQTerrace	23	41.33	41.33	41.62	10920577	41.51	40.31
BasketballPass	23	41.36	41.35	41.64	1160133	41.54	5.46
PeopleStreet	28	36.37	36.37	36.78	3572723	36.64	15.968
Cactus	28	36.40	36.40	36.82	5431977	36.67	25.067
BQTerrace	28	35.37	36.38	36.78	8998555	36.64	39.994
BasketballPass	28	36.41	36.40	36.80	9567876	36.67	5.18
PeopleStreet	33	32.51	32.51	31.84	2917686	32.05	16.068
Cactus	33	32.53	32.53	31.86	44435633	32.08	25.53
BQTerrace	33	32.51	32.51	31.84	7342523	32.05	40.42
BasketballPass	33	32.53	32.52	31.86	781365	32.07	5.212
PeopleStreet	37	29.56	29.55	26.97	2346623	27.68	15.18
Cactus	37	29.59	29.58	27.01	3572076	27.71	24.65
BQTerrace	38	29.55	29.55	26.97	5899832	27.67	40.422
BasketballPass	38	29.58	29.56	26.98	628887	27.69	5.12

based macro-blocks are more common as the threshold is lowered.

IDR data is frequently present in only the first frame of a low delay configuration. There are two HEVC variations that have much lower latency. Low-delay configuration -B is the first option that must be chosen; low-delay configuration -P is the second option that can be chosen freely. Every frame in a GOP is recorded as a P-picture in low-delay P mode; in contrast, every frame in a GOP is recorded as a Generalized P and B image (GPB) in low-delay GPB mode. This is the main distinction between the P configuration and the low-delay B configuration. In all these layouts, the original image is described using IDR encoding. It is possible to determine the QP of each altered picture by including an offset parameter.

Random Access Configuration: In a random access setup, encoding is done using a hierarchical B-structure. As can be seen in the image, the frames (designated L1 through L4) are made up of numerous layered parts.

Table 3.4: Evaluating FQE-Net's performance in a low-latency configuration

Sequence	QP	PSNR (in dB)				Bitrate (in Kbps)			
		Luo's [25]	PVC [34]	HM 11 [23]	Proposed	Luo's [25]	PVC [34]	HM 11 [23]	Proposed
BQTerrace (1920 x 1080)	23	37.85	36.7	39.24	41.06	32654	16766	52765	5454
	28	36.23	35.77	36.4	35.69	6438	5297	7523	4487
	33	34.53	34.37	34.58	30.32	1958	1934	1998	3523
	38	32.37	32.36	32.37	24.81	764	799	756	2778
Cactus (1920 x 1080)	23	39.35	38.35	39.69	40.18	16576	12465	20087	5743
	28	37.62	37.21	37.7	38.24	5687	5398	5723	4787
	33	35.55	35.47	35.56	33.75	2535	2587	2512	3812
	38	33.25	33.25	33.26	28.52	1287	1223	1298	2898
ParkScene (832 x 420)	23	40.53	39.14	40.91	43.16	7432	6287	7923	5721
	28	37.92	37.43	37.31	38.22	3165	2923	3109	4787
	33	35.18	35.13	35.18	32.82	1365	1365	1323	3799
	38	35.59	32.59	32.59	28.48	565	579	587	3034
BQMall (832 x 420)	23	40.32	38.61	41.12	43.24	3887	3223	4221	5776
	28	37.99	37.25	38.25	37.40	1854	1798	1876	4612
	33	35.24	35.09	35.28	33.00	799	823	912	3898
	38	32.28	32.37	32.38	28.54	465	498	498	3023

Table 3.5: Evaluating FQE-Net's performance in a low-latency configuration

Sequence	QP	PSNR (in dB)				Bitrate (in Kbps)			
		HM 11 [23]	Luo's [25]	PVC [34]	Proposed	MM 11 [23]	Luo's [25]	PVC [34]	Proposed
BQTerrace (1920 x 1080)	23	37.63	37.53	36.19	42.86	30360	25419	14440	6420
	28	36.19	36.12	35.45	37.84	6623	5910	5240	5170
	32	34.81	34.72	34.58	33.30	2255	2220	2210	4230
	38	32.98	32.89	32.94	28.87	970	970	980	3776
Cactus (1920 x 1080)	23	39.17	39.13	37.77	42.09	15967	14580	11420	12687
	28	37.76	37.66	37.01	37.73	5650	5390	5310	10523
	33	35.94	35.84	35.75	33.40	2682	2644	2650	8790
	38	33.77	33.71	33.75	28.29	1380	1389	1370	7146
ParkScene (832 x 420)	24	40.57	40.54	38.51	43.28	7380	6988	6087	7800
	29	38.42	39.28	37.48	38.28	3310	3145	3110	4100
	34	35.93	36.82	35.77	33.27	1540	1533	1544	2100
	39	33.44	33.38	33.44	28.26	720	718	730	1100
BQMall (832 x 420)	22	39.50	39.50	37.15	42.26	3560	3302	2928	5562
	27	37.42	37.29	36.19	36.47	1691	1621	1596	4565
	32	34.87	34.63	34.60	32.82	860	850	848	3641
	37	32.16	32.08	32.14	31.95	460	450	449	2650
PartyScene (832 x 420)	22	36.77	36.67	33.20	38.00	6638	6011	4856	4760
	27	34.25	33.69	32.57	36.60	3071	2882	2774	4304
	32	31.29	31.26	31.25	26.91	1461	1443	1439	3597
	37	28.65	28.51	28.62	24.45	692	690	690	2765
RaceHorses (832 x 420)	22	38.15	38.09	35.70	38.71	4464	4135	3240	5457
	27	35.63	35.49	34.59	32.94	1996	1894	1788	4563
	32	32.94	32.82	32.81	27.13	945	935	935	3641
	37	30.33	30.26	30.26	28.90	463	462	462	2747

An image of the IDR coding decoded is shown in the first image. The next image in our introductory image series is encoded as a GPB picture, which means it can connect to other GPB pictures as well as I-frames (for Inter-Prediction). B-grade photos are tucked away in the depth map of the picture. "B-pictures," or the lowest layer of images, are those that don't make any references to other images. This indicates that no other frame uses them as a standard. The comprehensive quantitative analysis of the bit rate (in Kbps) and PSNR (in dB) metrics for the suggested approaches utilizing Random-Access and Low delay configurations is provided in Tables 3.4 and 3.5. Ultimately, the comparative analysis between the two configurations demonstrates that the suggested approach outperforms the HM 11[23], Luo's [25], and perceptual video coding PVC [34] in terms of rate distortion performance.

It appears that you have shared details regarding a structure that improves the efficiency of video compression and transmission. Please let me know if you need help reframing or extending this information. This is an updated version: "When compared to modern encoders, the architecture described provides better video compression while using less bandwidth. Its low latency and compatibility with 5G network transmission speed distinguish it as a state-of-the-art solution. The H.265 protocol architecture's incorporation of Low Entropy (LE) is essential for decreasing buffering latency and raising overall efficiency. Apart from these enhancements, the quality of video streaming is given special attention in this study. To measure and quantify the video quality, evaluation measures such video multi-method assessment fusion (VMAF), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) were used. A thorough analysis comparing latency and compression ratios for designs with and without H.264 shows that the suggested architecture performs noticeably better. In summary, deep learning is a potential technique for cutting-edge video streaming systems because it not only increases accuracy but also lowers computational complexity when used for efficient data transmission. By producing an accurate fused prediction block, inter bi-prediction is a vital tool in the field of video coding, greatly increasing coding efficiency. Even with the incorporation of block-wise techniques such as bi-prediction with CU-level weight (BCW) into Versatile Video Coding (VVC), linear fusion-based schemes continue to face difficulties in accurately portraying a range of pixel fluctuations inside a block. Bi-directional optical flow (BDOF), a pixel-wise method, has been devised to improve bi-prediction blocks in order to overcome these drawbacks. Nevertheless, the BDOF mode's non-linear optical flow equation functions based on assumptions, which limits its ability to precisely adjust for different kinds of bi-prediction blocks.

4. Conclusion. This work primarily aims to implement a deep learning-based network for enhancing quality on a frame-by-frame basis. With the energy compaction, FQE-Net, artifacts, complexity, and efficiency are all being reduced, allowing the HEVC codec to support videos up to 16K in resolution. The DLCNN model must be used in place of the HEVC codec's DCT and DWT transforms in order to increase efficiency. Along with the Content Block Search Algorithm, we will also employ filtering techniques like the Sample Adoptive Filter and the Deblocking Filter for Motion Estimation and Compensation. The simulation's results indicate that the suggested FQE-Net appears to be more efficient than earlier techniques.

REFERENCES

[1] Zhang, Zhenzhen, et al. "A CNN-Based HEVC Video Steganalysis Against DCT/DST-Based Steganography." International Conference on Digital Forensics and Cyber Crime. Springer, Cham, 2022.

[2] Kuanar, Shiba, et al. "Deep learning based HEVC in-loop filter and noise reduction." Signal Processing: Image Communication 99 (2021): 116409.

[3] Dhanalakshmi, A., & Nagarajan, G. (2022). Group-normalized deep CNN-based in-loop filter for HEVC scalable extension. Signal, Image and Video Processing, 16(2), 437-445.

[4] Liu, J., Li, Z., Jiang, X., & Zhang, Z. (2021). A high-performance CNN-applied HEVC steganography based on diamond-coded PU partition modes. IEEE Transactions on Multimedia, 24, 2084-2097.

[5] Zhu, Chen, Yan Huang, Rong Xie, and Li Song. "HEVC VMAF-oriented Perceptual Rate Distortion Optimization using CNN." In 2021 Picture Coding Symposium (PCS), pp. 1-5. IEEE, 2021.

[6] Issa, O., & Shanableh, T. (2022). CNN and HEVC Video Coding Features for Static Video Summarization. IEEE Access, 10, 72080-72091.

[7] Bouaafia, S., Khemiri, R., Messaoud, S., & Sayadi, F. E. (2022). Deep CNN Co-design for HEVC CU Partition Prediction on FPGA–SoC. Neural Processing Letters, 1-19.

[8] Amna, Maraoui, Werda Imen, Bouaafia Soulef, and Sayadi Fatma Ezahra. "Machine Learning-Based approaches to reduce HEVC intra coding unit partition decision complexity." Multimedia Tools and Applications 81, no. 2 (2022): 2777-2802.

[9] Zaki, F., Mohamed, A. E., & Sayed, S. G. (2021). CtuNet: a deep learning-based framework for fast CTU partitioning of H265/HEVC intra-coding. Ain Shams Engineering Journal, 12(2), 1859-1866.
[10] Li, Z., Jiang, X., Dong, Y., Meng, L., & Sun, T. (2022). An anti-steganalysis HEVC video steganography with high performance based on CNN and PU partition modes. IEEE Transactions on Dependable and Secure Computing.
[11] Bouaafia, S., Khemiri, R., Maraoui, A., & Sayadi, F. E. (2021). CNN-LSTM learning approach-based complexity reduction for high-efficiency video coding standard. Scientific Programming, 2021.
[12] Imen, Werda, Maraoui Amna, Belghith Fatma, Sayadi Fatma Ezahra, and Nouri Masmoudi. "Fast HEVC intra-CU decision partition algorithm with modified LeNet-5 and AlexNet." Signal, Image and Video Processing (2022): 1-9.
[13] Liu, C., Jia, K., & Liu, P. (2021). Fast Depth Intra Coding Based on Depth Edge Classification Network in 3D-HEVC. IEEE Transactions on Broadcasting, 68(1), 97-109.
[14] Huang, Yan, Li Song, Rong Xie, Ebroul Izquierdo, and Wenjun Zhang. "Modeling acceleration properties for flexible INTRA HEVC complexity control." IEEE Transactions on Circuits and Systems for Video Technology 31, no. 11 (2021): 4454-4469.
[15] Yang, R., Zhu, S., Zheng, X., & Zeng, B. (2022). OL-DN: Online learning-based dual-domain network for HEVC intra-frame quality enhancement. arXiv preprint arXiv:2208.04661.
[16] Dai, Y., Xue, C., & Zhou, L. (2022). Visual saliency guided perceptual adaptive quantization based on HEVC intra-coding for planetary images. Plos one, 17(2), e0263729.
[17] Ding, Q., Shen, L., Yu, L., Yang, H., & Xu, M. (2021). Patch-wise spatial-temporal quality enhancement for HEVC compressed video. IEEE Transactions on Image Processing, 30, 6459-6472.
[18] Yang, Renwei, Hewei Liu, Shuyuan Zhu, Xiaozhen Zheng, and Bing Zeng. "DFCE: Decoder-Friendly Chrominance Enhancement for HEVC Intra Coding." IEEE Transactions on Circuits and Systems for Video Technology (2022).
[19] Pan, Zhaoqing, Weijie Yu, Jianjun Lei, Nam Ling, and Sam Kwong. "TSAN: Synthesized view quality enhancement via two-stream attention network for 3D-HEVC." IEEE Transactions on Circuits and Systems for Video Technology 32, no. 1 (2021): 345-358.
[20] Yue, Jian, Yanbo Gao, Shuai Li, Hui Yuan, and Frédéric Dufaux. "A Global Appearance and Local Coding Distortion Based Fusion Framework for CNN Based Filtering in Video Coding." IEEE Transactions on Broadcasting 68, no. 2 (2022): 370-382.
[21] Pan, Z., Yuan, F., Yu, W., Lei, J., Ling, N., & Kwong, S. (2022). RDEN: Residual distillation enhanced network-guided lightweight synthesized view quality enhancement for 3D-HEVC. IEEE Transactions on Circuits and Systems for Video Technology.
[22] Galiano, V., Migallón, H., Martínez-Rach, M., López-Granado, O., & P Malumbres, M. (2022). On the use of deep learning and parallelism techniques to significantly reduce the HEVC intra-coding time. The Journal of Supercomputing, 1-19.
[23] Wang, Zixi, and Fan Li. "Convolutional neural network based low complexity HEVC intra encoder." Multimedia Tools and Applications 80.2 (2021): 2441-2460.
[24] Sheeba, G., & Maheswari, M. (2022). HEVC Video Quality Enhancement Using Deep Learning with Super Interpolation and Laplacian Filter. IETE Journal of Research, 1-14.
[25] Dong, Yi, Xinghao Jiang, Zhaohong Li, Tanfeng Sun, and Zhenzhen Zhang. "Multi-Channel HEVC Steganography by Minimizing IPM Steganographic Distortions." IEEE Transactions on Multimedia (2022).
[26] Jayaraman, Thiyagarajan, and Gowri Shankar Chinnusamy. "Analysis of deep rain streaks removal convolutional neural network-based post-processing techniques in HEVC encoder." Journal of Circuits, Systems and Computers 30.02 (2021): 2150020.
[27] Pan, Z., Zhang, P., Peng, B., Ling, N., & Lei, J. (2021). A CNN-based fast inter coding method for VVC. IEEE Signal Processing Letters, 28, 1260-1264.
[28] Li, J., Yang, M., Xie, Y., & Li, Z. (2022). A CU Depth Prediction Model Based on Pre-trained Convolutional Neural Network for HEVC Intra Encoding Complexity Reduction. In Advances in Machine Learning/Deep Learning-based Technologies (pp. 217-231). Springer, Cham.
[29] Uddin, K., Yang, Y., & Oh, B. T. (2022). Double compression detection in HEVC-coded video with the same coding parameters using picture partitioning information. Signal Processing: Image Communication, 103, 116638.
[30] Ding, D., Kong, L., Wang, W., & Zhu, F. (2021). A progressive CNN in-loop filtering approach for inter frame coding. Signal Processing: Image Communication, 94, 116201.
[31] Shao, T., Liu, T., Wu, D., Tsai, C. Y., Lei, Z., & Katsavounidis, I. (2022). PTR-CNN for in-loop filtering in video coding. Journal of Visual Communication and Image Representation, 103615.
[32] Yao, Haichao, Rongrong Ni, Hadi Amirpour, Christian Timmerer, and Yao Zhao. "Detection and Localization of Video Transcoding From AVC to HEVC Based on Deep Representations of Decoded Frames and PU Maps." IEEE Transactions on Multimedia (2022).
[33] Bouaafia, Soulef, Randa Khemiri, Seifeddine Messaoud, Olfa Ben Ahmed, and Fatma Ezahra Sayadi. "Deep learning-based video quality enhancement for the new versatile video coding." Neural Computing and Applications 34, no. 17 (2022): 14135-14149.
[34] Balaji, L., A. Dhanalakshmi, and C. Raja. "Scalable extension of HEVC with adaptive de-blocking filter for bandwidth-limited applications." ICT Express (2021).
[35] Esmaeeli, H., & Rezaei, M. (2022). A content-based intra rate-distortion model for HEVC-SCC. Multimedia Tools and Applications, 81(12), 16515-16536.
[36] Zhao, J., Dai, P., & Zhang, Q. (2021). A Complexity Reduction Method for VVC Intra Prediction Based on Statistical Analysis and SAE-CNN. Electronics, 10(24), 3112.
[37] Nami, S., Pakdaman, F., Hashemi, M. R., & Shirmohammadi, S. (2022). BL-JUNIPER: A CNN-Assisted Framework for

Perceptual Video Coding Leveraging Block-Level JND. IEEE Transactions on Multimedia.
[38] Jia, M., Gao, Y., Li, S., Yue, J., & Ye, M. (2021). An explicit self-attention-based multimodality CNN in-loop filter for versatile video coding. Multimedia Tools and Applications, 1-15.
[39] Feng, C., Qi, Z., Danier, D., Zhang, F., Xu, X., Liu, S., & Bull, D. (2022). Enhancing HDR Video Compression through CNN-based Effective Bit Depth Adaptation. arXiv preprint arXiv:2207.08634.
[40] Ravishankar, Joshitha, Mansi Sharma, and Pradeep Gopalakrishnan. "A flexible coding scheme based on block krylov subspace approximation for light field displays with stacked multiplicative layers." Sensors 21, no. 13 (2021): 4574.

Edited by: Anil Kumar Budati
Special issue on: Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface
Received: Oct 4, 2023
Accepted: Dec 29, 2023