Deep learning-based switchable network for in-loop filtering in high efficiency video coding

Helen K. Joy[1], Manjunath R Kounte[2]
[1]School of Electronics and Communication Engineering, REVA University, Bengaluru, India
[2]Department of Electronics and Computer Engineering, School of Electronics and Communication Engineering, REVA University, Bengaluru, India

Article Info

Article history:

Received Sep 6, 2022
Revised Dec 11, 2022
Accepted Dec 21, 2022

Keywords:

Coding tree unit
Convolutional neural network
Deep learning
High efficiency video coding
In-loop filtering
Video coding

ABSTRACT

The video codecs are focusing on a smart transition in this era. A future area of research that has not yet been fully investigated is the effect of deep learning on video compression. The paper's goal is to reduce the ringing and artifacts that loop filtering causes when high-efficiency video compression is used. Even though there is a lot of research being done to lessen this effect, there are still many improvements that can be made. In This paper we have focused on an intelligent solution for improvising in-loop filtering in high efficiency video coding (HEVC) using a deep convolutional neural network (CNN). The paper proposes the design and implementation of deep CNN-based loop filtering using a series of 15 CNN networks followed by a combine and squeeze network that improves feature extraction. The resultant output is free from double enhancement and the peak signal-to-noise ratio is improved by 0.5 dB compared to existing techniques. The experiments then demonstrate that improving the coding efficiency by pipelining this network to the current network and using it for higher quantization parameters (QP) is more effective than using it separately. Coding efficiency is improved by an average of 8.3% with the switching based deep CNN in-loop filtering.

Corresponding Author:

Helen K. Joy
School of Electronics and Communication Engineering, REVA University
Bengaluru, 560064, India
Email: helenjoy88@gmail.com

1. INTRODUCTION

In recent years, as the usage of video, is in high demand, the advancement needed in video codec is also highly prioritized. The transition of video codec to a hybrid model and smart video codec is the trend of research now. As the codec mostly follows block-based coding the chances of ringing effect and visible block structure are high. The in-loop filter in high efficiency video coding (HEVC) is used to alleviate this effect. The compression artifact reduction is a matter of concern in this. HEVC specifies two in-loop filters, deblocking and sample adaptive offset (SAO), which considerably increase the subjective quality of decoded video sequences as well as compression efficiency by improving the quality of the reconstructed images. The SAO primarily corrects ringing artifacts produced by massive transformations and quantization, as well as sample value offsets in specific sections of an image generated by coding of motion vectors, while the deblocking filter reduces discontinuities on the block borders. Optionally, one or two of these filtering processes can be applied before placing the reconstructed image in the decoded picture buffer (DPB). The de-blocking filter (DBF) is employed in the same way as in H264/ advanced video coding (AVC) [1], although the DBF has been simplified in terms of decision making and filtering. SAO is a nonlinear amplitude mapping filter that works with DBF data. The purpose of SAO is to improve signal amplitude reconstruction through lookup table mapping. For each computed tomography urography (CTU), HEVC provides two types of SAO operations: band offset (BO) and edge offset (EO). Both SAO kinds add an offset value to the sample, which is picked from a lookup table dependent on the local gradient. Comparing HEVC deblocking to H.264/AVC, the complexity has been greatly decreased. Additionally, the HEVC deblocking filter may be parallelized down to 8 8-sample blocks. If SAO is used after complete picture reconstruction, then data output to the application must wait until the reconstruction and deblocking of the entire image is completed before beginning the SAO process. This will be a significant problem in low latency applications. The H.264/AVC deblocking is more of a bottleneck in the decoder implementation than the HEVC deblocking, which offers a better trade-off between computational cost and coding efficiency due to its reduced computational complexity and strong parallelizability.

Researches focus on the reduction of an artifact by the traditional [2] signal processing method. H.265/AVC, HEVC uses a de-blocking filter [3] to reduce the effect of block-based prediction. In HEVC SAO is used after the de-blocking filter to reduce the ringing effect. Above all these the room for development of the filtering was more, if machine learning techniques are used to solve this issue is the next area of research. It hit to start with convolutional neural network (CNN) based [4], [5] super-resolution CNN (SRCNN), very deep super-resolution (WDSR) wide activation super-resolution. These networks were trained to solve the loop filtering issues, when deep learning came to improvise the in-loop filtering issue the focus was mainly on visual enhancement [6], and compression artifact reduction by maintaining the coding efficiency.

Considering the literature, the conclusion driven was a single stand-by network the combination of the existing network along with the CNN-based model can be considered as a better design. As the quantization parameters (QP) type of frame is a dependent factor. The design should be a switching one using the goodness of the existing network and the goodness of CNN model. The network complexity is also a factor to be considered. Compatibility is also a main factor to be taken care. Incorporating all these points into consideration the proposed method designs a switchable loop filter model that needs a new CNN based network and existing in loop filter, and the switching is decided by the quantization parameter. The model provides better results in both subjective and objective analysis.

The sections of this paper are the following: section 2 begins with the motivation for the paper, followed by the contributions. The proposed method and the database used for the process are described in section 3. Here, the connection between the existing and proposed networks is also explained. Section 4 compares the proposed method to the current systems, and the results are tabulated. The final section of the paper concentrates on the conclusion, effectiveness, and benefits of the topic.

2. BACKGROUND AND MOTIVATION

The in-loop filtering in HEVC with deblocking and SAO is affected by compression artifacts up to a level. The design of a proper deep neural network (DNN) for the network was a concern and to place it in encoder or decoder side will give more efficiency was also a discussion considering all these scenarios the literature survey can be split into 2 main areas: i) state of art method by replacing the existing in-loop filtering with a novel approach and ii) design a switchable loop filter method using the existing network by adding DNN to enhance the performance.

Variable filter residue learning was introduced to enhance the post-processing of the frame. The visual quality is enhanced and ease of network is preserved. In loop filter CNN also was a network model designed to replace the in-loop filtering by CNN trained model. In loop filtering using residual highways CNN (RHCNN) and content-based loop filtering, residual highway CNN has two unit's residual highway unit and convolution layer. This will not affect the total network of in-loop filters i.e., De blocking filter and SAO. It performs an operation based on QP values, each QP [7] value are divided into various brands content based loop filtering was also a method for multiple dimension where it performed efficiently. Multi-frame in loop filter (MIF) was introduced for visual quality enhancement, still there was room for improvement in this field, some CNN approach [8] was using CNN model [9] to be placed in between de blocking filter and SAO. The multichannel long short-term dependency residual n/w was one among them is used for HEVC. Lai and Wang [10] introduced multi-stage attention convolutional neural networks (MACNNs) for HEVC in-loop filtering as a replacement for loop-filtering. It has a three-stage design fashion, where various artifacts are removed in different stages and all building blocks can concentrate on specific types of artifacts. The 6.5% Bjontegaard delta (BD) rate reduction shows that it performs better in loop filtering than HEVC. By simultaneously maintaining the best trade-off between performance gain and computational cost, deeper networks can be built in the future and can leverage the self-attention mechanism.

A CNN and image/video processing combine to produce extremely good results and achieve greater success in image/video processing. The goal is to implement content-aware CNN-based loop filtering [11] for HEVC. The proposed model will perform quantitative analysis in multiple dimensions to make it more efficient and valuable for content-aware CNN-based loop filtering [12]. The CTU is treated as a separate region in this case. In this case, the CNN model uses a discriminative network to analyze the resolution of different regions with different CNN models. This is how the content aware multimodal filtering mechanism works. Based on the discriminative neural network, the deep learning model aids in the adaptive selection of the content-based region. The need of DNN arises here to improve the visual quality as well as design the system with minimum computational complexity and better coding gain. The researches in literature/previous researches have proved that the double enhancing effect of the frame may affect the visual quality. QP is another feature needs to be taken care as it is directly related to the compression artifact, so a general model to all QP may not be a better option. Considering the group of pictures (GOP), I, P, and B frame the encoding [13] varies in each frame. Each of this will be considered separately to design a better network with high efficiency.

Considering all these criteria a switchable or pipelined network is designed. The network performs based on the frame type and QP. This switchable system pipelined with the HEVC should perform with less computational complexity without affecting the system, performance measure should be considered in peak signal to noise ratio (PSNR), visual quality check, computation complexity by measuring time of encoding and coding performance by bitrate.

$$f = conv(i, k) \tag{1}$$

$$F = (f_{1,} f_{2,} f_{3,} \dots \dots f_n) \tag{2}$$

$$\omega_p = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} f_P(i, j) \tag{3}$$

$$s_p = conv_2(conv_1(\omega_p, k)) \tag{4}$$

Here the idea behind the model is to extract the input image with high quality by alleviating the effect of compression artifact [14]. Initially, the aim is to extract the feature map from the input. The best network to extract the feature of image/video is a convolutional neural network as the kernels have the ability to extract various different features [15].

Here 64 sequential filters are used to extract the features of the input. The filters are of 3×3 kernel size. Filtering is repeated till the optimal number when features are extracted. The more filtering may increase the calculation complexity and the model will not be serving the purpose. Here as experimented the filter size is set to an optimum value of 64. Its repeatedly done in 15 stages initially for feature extraction. Once that is retrieved the next aim is to use that extracted feature map to fit into the given frame by using extract weight. Adjusting the weight to get the correct feature to alleviate artifact is main focus. It is done by CNN network with 2 convolution networks. The network will compute the weight for each neuron by considering the input. Taking into consideration, the loss function, the network will train the model to reach global minima by stochastic gradient decent method. After converging to global minima, the weights are reparametrized by the network. This is the way the network performs weighted normalization. It can be represented as follows. If *y* is the output of the network and *i* is the input with *a* bias *b*,

$$y = \phi(w.i + b) \tag{5}$$

were *w* is the weight vector and ϕ is the non-linearity element. The main aim is to update weight. So,

$$w = \left(\frac{s}{||v||}\right).v \tag{6}$$

Here *w* is the stochastic gradient decent weight *s* is the scalar parameter and *v* is the parameter vector. The aim of (6) is to make w independent from *v* i.e., independent from the parametric vector by reacting global minima. This gives the reparametrized updated *w* for the network.

Step 3 is to add up the updates. The bias is also added by (7).

$$y_p = s_p.f_p + b_p \tag{7}$$

Here we consider *i* as input. The convolution of input frame *i* gives us the feature *f*. the *F* in (2) represents the feature map obtained from various CNN networks and w_p obtained. Using that the initial weights obtained from feature map and multiple kernels in between the next stage seeks for updating the weights to minimize artifacts. The weight adjusting block of network will provide updated weights that is added to bias and input to obtain the output is the reconstructed frame. The output frame will be the updated version of the image added up with the extracted features.

3. PROPOSED METHOD

The proposed model as shown in Figure 1 aims at compression artifact removal [16] by maintaining the coding efficiency [17]. By modelling the deep CNN adaptive filter with 64 kernels and with a weight initializing and weight adjusting network for the frames I, P, and B. But considering the characteristics of P and B frame as it has the information of the previous frame and next frame, the chance of enhancement [18], [19] can be doubled and can cause the worst effect in the visual quality [20] of the output, considering this the frame B is not fit for the deep CNN adaptive filter, instead it can be sent directly through the existing in loop filter [21] in HEVC. Here a pipeline model is designed to solve this issue. The GOP with I, P, and B frames will be sorted and the I and P will be taken care by the deep CNN adaptive filter and the B frame that has the information of both previous and next frame will be sent to the loop filtering section of HEVC. This method provides better results compared to the non-pipelined model. Each frame in the deep CNN adaptive filter is labelled in frame level/CTU level. Figure 2 gives an overview of the proposed model. Here the input is given to a sequence of convolution layers with kernel size as experimented (3×3×64) will extract the input features. These are sent to next CNN layers for weight normalization, then the information is combined and fused to extract or reconstruct the output frame. The fusion block helps in this. The features are shaped to required size and added with the input to extract the reconstructed output. The network always tries to minimize the error by weight normalization method as in (8).

$$e = |y^{\wedge} - y| \tag{8}$$

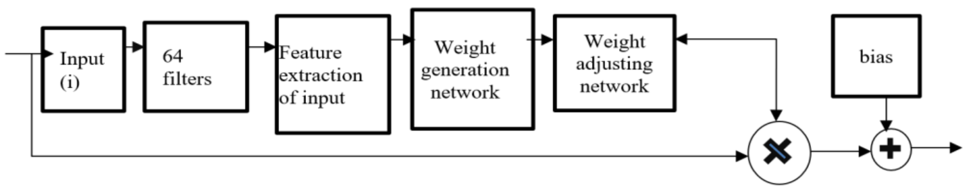

Figure 1. The framework of deep learning based in loop filtering

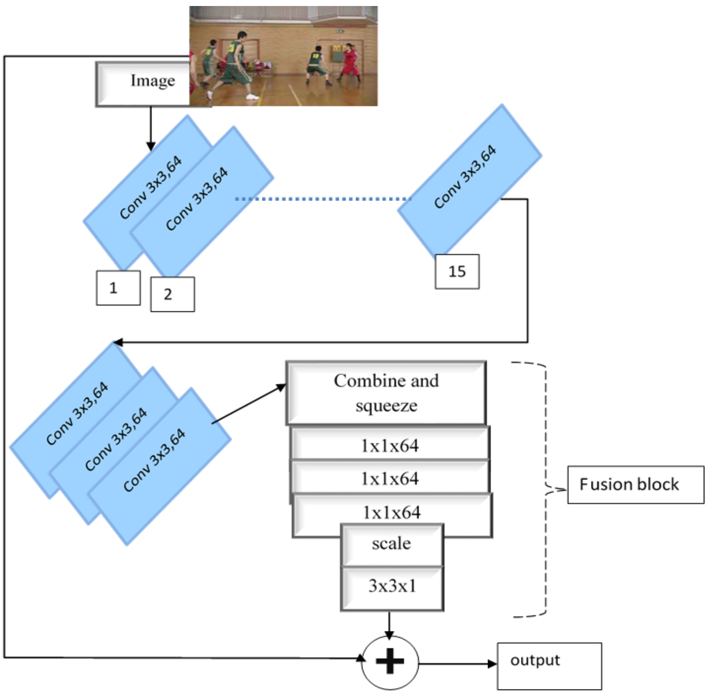

Figure 2. The design of deep learning based in loop filtering with series of CNNs and combine and squeeze network

The low-density P configuration (LDP), random access (RA), all intra (AI) configuration are considered separately. For LDP the deep CNN filter [22] is applied only to selective frames out of all frames in it. The B frames will have the future information and previous frame info where P frame has only future info. From all GOP the frames with B are switched to normal loop filter while the I and P are sent to deep CNN filter. The experiment result shows that it improvises not only quality but bit rate too. In RA, a similar strategy is performed. The frame is detected as I/P/B the B frames will be switched and I and P follow the proposed method. Along with this, the other effects of blurring, ringing [23], and blocking [24] are also reduced by the proposed method by ensuring the visual quality.

3.1. Database establishment

The database consists of 330 raw videos and out of that 200 frames are for training. The HM.16.9 will be set as default, the deblocking [25] and SAO are turned off and it is encoded for all base QP values 22, 27, 32, 37. This is applied for I, P, and B frame training. Totally 2,034 frames constitute the database of intra training [26], [27] and 2,628 for P and B frames. For I frame the frames are selected randomly from the reconstructed sequence. The QP offset in HM software is kept as 0. For the training process of I frame 2 frames are randomly chosen from the reconstructed output video. For P and B, the training procedure remains same but the training is done separately. To do it the sequence is initially encoded using the LDP and random-access techniques is HM software. Frames are allotted for all QPs frames are split into patches and compared in inter prediction, for testing purpose 18-20 test sequence is chosen from the database to evaluate HEVC performance. A generalized and patch-based scheme is used for P and B frame. This helps to reduce the training complexity. The HEVC software HM.16.9 version is used for the testing and training as default configuration.

The low QP has less amount of artifact compared to large QP. The main aim is to reduce compression artifact by reducing computational complexity and improving coding efficiency. Taking into consideration these facts a switchable model of loop filtering is designed and pipelined with original HEVC as shown in Figure 3. The frame types also vary as we know I frame holds more of texture and direction and prediction information whereas P and B frame hold motion estimation information. So, the switching process will be considered in this manner. If the frame is I with any range of QP the original HEVC loop filtering can be opted and for P, B frames with higher QP too goes with original HEVC in loop filtering. The frames with QP 32, 37 which is P or B switches to the proposed model. This provides a better coding efficiency and less complexity.

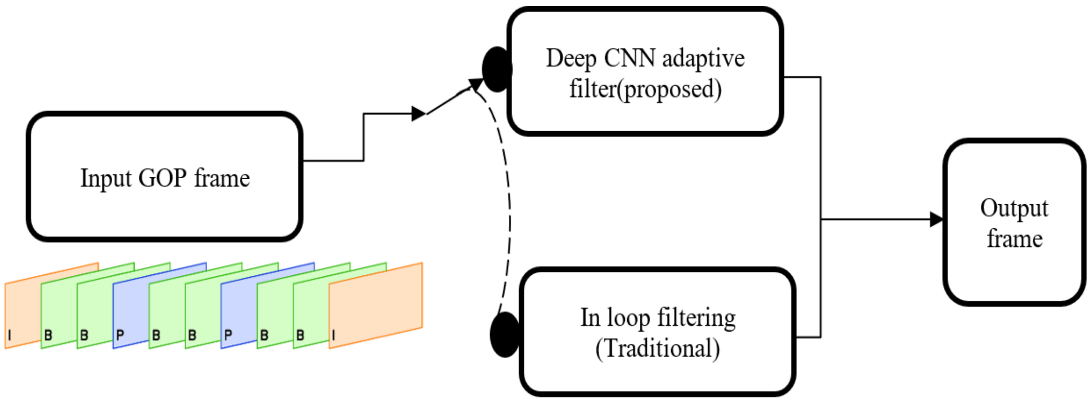

Figure 3. The switchable design pipelined in the original HEVC

4. COMPARISON OF PROPOSED METHOD WITH EXISTING TECHNIQUE

Each frame in the data set of 330 is split into two training set, validation and testing set. Out of 330, 200 will be used for training and rest for testing and validation. The learning rate is 10^{-4} with a step rate of 0.5. the optimizer used is adaptive moment estimation (Adam). For training process, the frame is split into patches of 35×35 and a group of 64 random patches with different QP are chosen for training. To test the model each frame from test is given to the network and to the original HEVC. The error rate in both can be evaluated. The designed model id pipelined to the original HEVC; the system is used then for comparison with existing system.

The performance of the proposed system can be proved by comparing it with the similar techniques. Here our proposed system is compared with some existing technique like variable-filter-size residue-learning convolutional neural network (VRCNN), long short-term memory network (LSTM) model placed between de-blocking and SAO. The Tables 1 and 2 compare the bit rate saving i.e., the BD-bit rate saving of the VRCNN and the proposed scheme. The BD-rates noted here are in percentage and each value shown in Tables 1 and 2 shows the dip or decrement or the bitrate saving. This table shows the coding efficiency id improved in a better way. Analyzing the Tables 1 and 2, BD-rate the difference is clearly seen that's is because of the network design and the switchable property. Here not only the all intra, low-delay P, RA also performs equally for the switchable mechanism. The proposed system is compared for efficiency and coding efficiency with the existing techniques. The results are tabulated in Table 1 for VRCNN and Table 2 for LSTM technique. The average bit rate saving shows that the proposed method is efficient. For VRCNN the proposed model has an improvement of 9.9% reduction in artificial intelligence (AI) and 8% for latent Dirichlet allocation (LDA) and 7.6% for random access. For its 8.7%-bit rate saving for AI 8.3% for LDP and 7.7% for random access. This clearly shows the improvement in bit rate and thus proves that the model improves coding efficiency.

Table 1. BD-Bitrate comparison for coding efficiency check (VRCNN vs switchable scheme)

Sl.no	Class	Name	VRCNN [2]			Our Scheme		
			AI (%)	LDP (%)	RA (%)	AI (%)	LDP (%)	RA (%)
1	A	PeopleOnstreet Traffic	5.4	1.8	3.1	10.2	5.7	7.2
			5.6	1.8	3.5	11.2	6.8	7.48
2	B	BasketballDrive	2.2	1.3	1.0	8.4	7.5	6.4
3		BQterrace	2.6	6.0	4.7	5.4	11.4	9.2
4		Cactus	4.54	3.8	5.1	8.2	7.1	7.98
5		Kimono	2.5	1.8	0.6	8.4	5.6	5.4
6		ParkScene	4.4	0.9	1.8	8.5	3.5	4.7
7	C	BasketballDrill	7.0	3.1	1.1	14.9	9.3	6.1
8		BQmall	5.2	3.6	3.6	10.3	8.2	7.6
9		PartyScene	3.6	2.4	1.2	6.4	4.8	4.1
10		RaceHorses	3.6	2.1	2.7	6.2	3.9	4.9
11	D	BasketballPass	5.0	3.7	3.1	11.3	8.9	7.9
12		BlowingBubbles	4.9	2.6	2.2	8.5	4.9	4.2
13		BQSquare	4.0	4.5	0.8	8.1	9.3	6.6
14		RaceHorses	7.1	3.9	5.3	10.7	6.1	7.6
15	E	FourPeople	7.0	5.7	7.0	14.7	13.7	13.8
16		Johnny	5.9	4.8	3.4	13.6	14.6	12.4
17		KristenAndSara	6.7	6.7	6.2	13.5	12.6	12.7
	Average		4.8	3.4	3.0	9.9	8.0	7.6

Table 2. Coding performance comparison between LSTM and proposed switchable loop filtering mechanism

Sl.no	Class	Name	LSTM [9]			Our Approach		
			AI (%)	LDP (%)	RA (%)	AI (%)	LDP (%)	RA (%)
1	B	BasketballDrive	4.40	7.8	6.9	5.75	7.5	6.3
2		BQterrace	3.70	6.9	6.3	4.37	9.5	6.4
3		Kimono	6.20	8.7	7.1	7.30	8.8	7.4
4	C	BasketballDrill	6.00	9.2	8.8	13.1	9.3	8.9
5		BQmall	4.30	7.7	6.8	8.7	7.9	6.9
6		RaceHorses	4.70	8.3	7.7	5.1	8.6	8
7	D	BasketballPass	6.60	8.4	7.5	9.5	8.5	7.8
8		BlowingBubbles	5.70	6.6	5.3	6.63	6.6	5.5
9		BQSquare	4.30	5.7	5.1	7.77	5.8	5.4
10	E	FourPeople	4.07	9.8	9.2	12.81	10.0	9.5
11		Johnny	8.70	10.9	10.5	11.9	10.9	10.9
12		KristenAndSara	8.06	10.3	10.1	12	10.3	10.7
	Average		5.3	8.3	7.6	8.7	8.6	7.8

For examining the subjective analysis of the video frames PSNR is taken into consideration and plotted as graph for HM, VRCNN, RHCNN, and proposed method in Figure 4. The graph compares the PSNR with configurations Figure 4(a) AI, Figure 4(b) LDP, Figure 4(c) RA of proposed method with the existing systems, [2], [8], [9]. The results clearly show the improvement in PSNR. Table 2 shows the comparison of the proposed switchable system with long short-term memory-based loop filtering. Here the higher resolutions are not considered. Still the performance metrics look great by saving the bitrate with better BD-rate. The calculated rates are noted in Table 2.

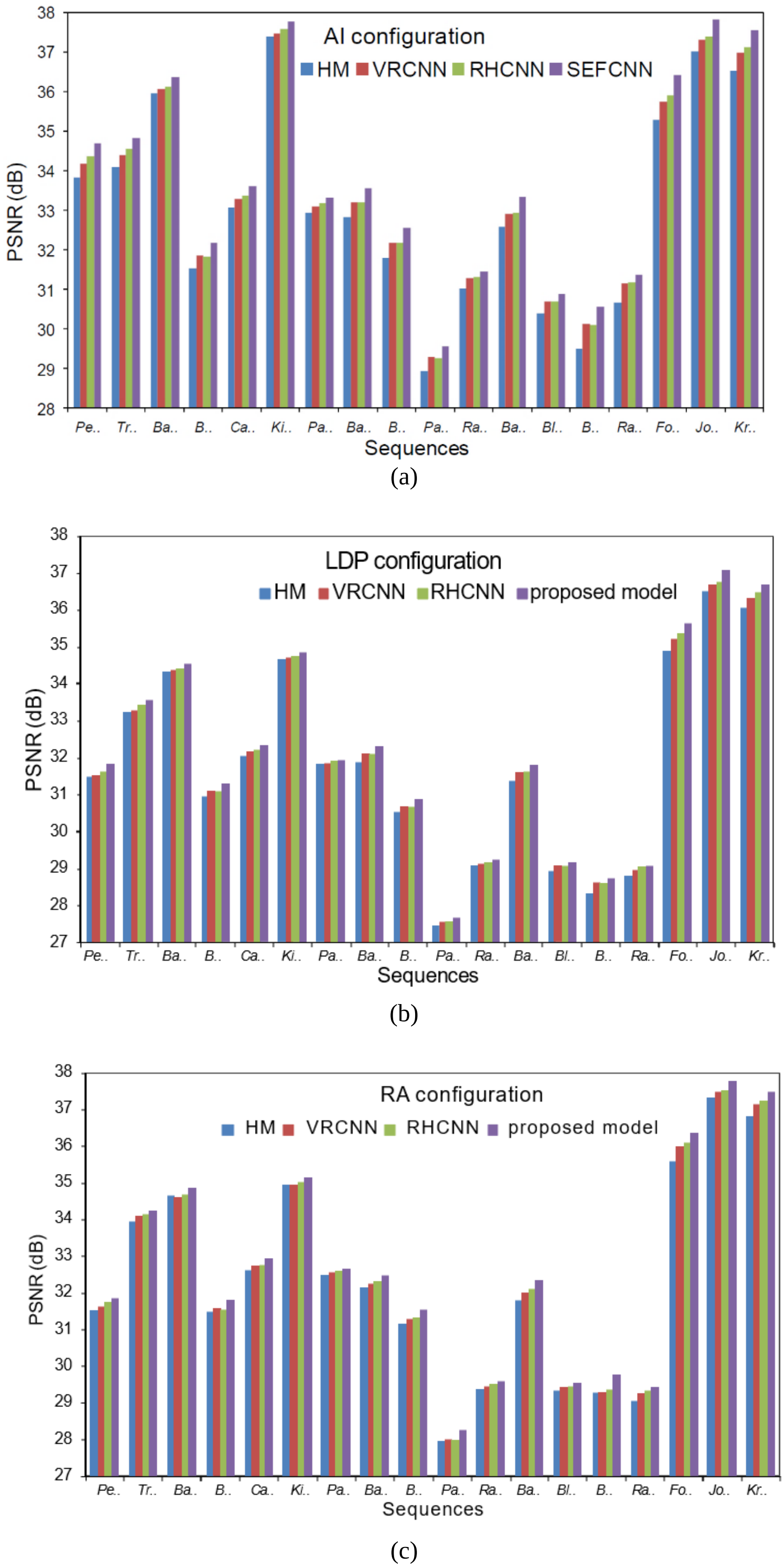

Figure 4. Graph comparing the PSNR with configurations (a) AI, (b) LDP, and (c) RA of proposed method to the existing systems [2], [8], [9]

5. CONCLUSION

Loop filtering in video coding with deblocking and SAO is replaced by CNN architecture in this paper. The switchable CNN method is used here, based on the quantization parameter the proposed method will switch from the original traditional method and the CNN loop filter to give better efficiency. Then using only, the proposed system, the switchable system performs better in real scenario than the existing methods. The proposed system is compared with various existing methods and results are tabulated. The results show that the coding efficiency and the PSNR is improved compared to the existing techniques. The BD-rate calculation shows prove that the switchable system is compatible and efficient with wits dip in the BD-bitrate (%). This shows that the combined system with switchable mechanism performs efficient for existing video codec systems. In future the system design for independent model with compatibility and efficiency can be evolved with this reference model. Refinement should be done in reducing computational complexity and the network structure, betterment in existing method with other network is an area of research.

ACKNOWLEDGEMENTS

The authors like to express their gratitude to REVA University for extending research facilities to carry out this research.

REFERENCES

[1] D. Mukherjee, S. Li, Y. Chen, A. Anis, S. Parker, and J. Bankoski, "A switchable loop-restoration with side-information framework for the emerging AV1 video codec," in *2017 IEEE International Conference on Image Processing (ICIP)*, Sep. 2017, pp. 265–269, doi: 10.1109/ICIP.2017.8296284.

[2] Y. Dai, D. Liu, and F. Wu, "A convolutional neural network approach for post-processing in HEVC intra coding," in *International Conference on Multimedia Modeling*, 2017, pp. 28–39.

[3] H. K. Joy, M. R. Kounte, and B. K. Sujatha, "Design and implementation of deep depth decision algorithm for complexity reduction in high efficiency video coding (HEVC)," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, 2022, doi: 10.14569/IJACSA.2022.0130168.

[4] W.-S. Park and M. Kim, "CNN-based in-loop filtering for coding efficiency improvement," in *2016 IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, Jul. 2016, pp. 1–5, doi: 10.1109/IVMSPW.2016.7528223.

[5] Z. Pan, X. Yi, Y. Zhang, B. Jeon, and S. Kwong, "Efficient In-loop filtering based on enhanced deep convolutional neural networks for HEVC," *IEEE Transactions on Image Processing*, vol. 29, pp. 5352–5366, 2020, doi: 10.1109/TIP.2020.2982534.

[6] X. Song *et al.*, "A practical convolutional neural network as loop filter for intra frame," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 1133–1137.

[7] T. Li, M. Xu, C. Zhu, R. Yang, Z. Wang, and Z. Guan, "A deep learning approach for multi-frame in-loop filter of HEVC," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5663–5678, Nov. 2019, doi: 10.1109/TIP.2019.2921877.

[8] Y. Zhang, T. Shen, X. Ji, Y. Zhang, R. Xiong, and Q. Dai, "Residual highway convolutional neural networks for in-loop filtering in HEVC," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3827–3841, Aug. 2018, doi: 10.1109/TIP.2018.2815841.

[9] X. Meng, C. Chen, S. Zhu, and B. Zeng, "A new HEVC In-loop filter based on multi-channel long-short-term dependency residual networks," in *2018 Data Compression Conference*, Mar. 2018, pp. 187–196, doi: 10.1109/DCC.2018.00027.

[10] P.-R. Lai and J.-S. Wang, "Multi-stage attention convolutional neural networks for HEVC in-loop filtering," in *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, Aug. 2020, pp. 173–177, doi: 10.1109/AICAS48895.2020.9073980.

[11] T. Wang, M. Chen, and H. Chao, "A novel deep learning-based method of improving coding efficiency from the decoder-end for HEVC," in *2017 Data Compression Conference (DCC)*, Apr. 2017, pp. 410–419, doi: 10.1109/DCC.2017.42.

[12] W. H. Bangyal, K. Nisar, A. A. Bin Ag. Ibrahim, M. R. Haque, J. J. P. C. Rodrigues, and D. B. Rawat, "Comparative analysis of low discrepancy sequence-based initialization approaches using population-based algorithms for solving the global optimization problems," *Applied Sciences*, vol. 11, no. 16, Aug. 2021, doi: 10.3390/app11167591.

[13] R. Yang, M. Xu, Z. Wang, and T. Li, "Multi-frame quality enhancement for compressed video," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6664–6673.

[14] S. Kuanar, C. Conly, and K. R. Rao, "Deep learning based HEVC in-loop filtering for decoder quality enhancement," in *2018 Picture Coding Symposium (PCS)*, Jun. 2018, pp. 164–168, doi: 10.1109/PCS.2018.8456278.

[15] X. He, Q. Hu, X. Zhang, C. Zhang, W. Lin, and X. Han, "Enhancing HEVC compressed videos with a partition-masked convolutional neural network," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 216–220.

[16] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 1646–1654, doi: 10.1109/CVPR.2016.182.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*, 2016, pp. 630–645.

[18] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, and L. Zhang, "Ntire 2017 challenge on single image super-resolution: Methods and results," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 114–125.

[19] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *VCEG-M33*, 2001.

[20] F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, and D. Zhao, "An end-to-end compression framework based on convolutional neural networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 3007–3018, Oct. 2018, doi: 10.1109/TCSVT.2017.2734838.

[21] B. U. V Prashanth, M. R. Ahmed, and M. R. Kounte, "Design and implementation of DA FIR filter for bio-inspired computing architecture," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 2, pp. 1709–1718, Apr. 2021, doi: 10.11591/ijece.v11i2.pp1709-1718.

[22] M. Jeon and B.-D. Lee, "Toward content-aware video partitioning methods for distributed HEVC video encoding," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 3, pp. 569–578, Jun. 2015, doi: 10.11591/ijece.v5i3.pp569-578.

[23] H. K. Joy and M. R. Kounte, "Deep CNN based video compression with lung ultrasound sample," *Journal of Applied Science and Engineering*, vol. 26, no. 3, pp. 313–321, 2022.

[24] H. K. Joy and M. R. Kounte, "Deep CNN depth decision in intra prediction," in *Lecture Notes in Electrical Engineering*, Springer Singapore, 2022, pp. 1–10.

[25] Y. Li *et al.*, "Convolutional neural network-based block up-sampling for intra frame coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2316–2330, Sep. 2018, doi: 10.1109/TCSVT.2017.2727682.

[26] W. H. Bangyal, J. Ahmad, and H. T. Rauf, "Optimization of neural network using improved bat algorithm for data classification," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 4, pp. 670–681, May 2019, doi: 10.1166/jmihi.2019.2654.

[27] W. H. Bangyal *et al.*, "Constructing domain ontology for Alzheimer disease using deep learning-based approach," *Electronics*, vol. 11, no. 12, Jun. 2022, doi: 10.3390/electronics11121890.

BIOGRAPHIES OF AUTHORS

Helen K. Joy received the Bachelor of Engineering in Electronics and Communication Engineering from National college of engineering, Anna University, Chennai, India, and Master of engineering from Satyabhama University, Chennai, India, in 2013, and currently working toward the Ph.D. degree at REVA University. The research interests include video compression, video processing, deep learning. She can be contacted at email: helenjoy88@gmail.com.

Manjunath R Kounte is currently working as Associate Professor and Head, Electronics and Computer Engineering, School of ECE, REVA University, Bangalore, India. He has over 15+ years of Teaching, Research and Administrative experience. He has Completed his Bachelor of Engineering in Electronics and Communication Engineering from VTU, Belagavi, India and Master of Technology in Computer Network Engineering from VTU, Belagavi, India and Ph.D. from Jain University, Bengaluru, India. Dr. Manjunath Kounte is Senior Member, IEEE, and member of various international professional bodies such as IETE, ISTE. He has over published over 15+ Indexed Journal papers, 25+ International Conference papers, 1 book and 3 book chapters. He is reviewer of various reputed International Journals of MDPI, Elsevier, Springer Publishers. His research interests include Machine Learning, Deep Learning, IoT Edge Computing, Internet of Vehicles, Signal Processing, Wireless Communication, and Blockchain Technologies. He is also heading Machine Learning and Blockchain (MLB) research lab in REVA University. He can be contacted at email: manjunath.kounte@gmail.com.