 https://doi.org/10.3311/PPee.20881

Periodica Polytechnica Electrical Engineering and Computer Science, 67(2), pp. 216–228, 2023

Adaptive and Efficient Hybrid In-loop Filter Based on Enhanced Generative Adversarial Networks with Sample Adaptive Offset Filter for HEVC/H-265

Vanishree Moji[1*], Mathivanan Murugavelu[1]

[1] Department of Electronics and Communication Engineering, ACS College of Engineering, Mysore Road, Bengaluru 560 074, Karnataka, India
[*] Corresponding author, e-mail: vanishreemoji4750@gmail.com

Received: 25 July 2022, Accepted: 13 December 2022, Published online: 16 March 2023

Abstract

In this manuscript, an Adaptive and Efficient Hybrid In-loop Filter based on Enhanced Generative Adversarial Network Deblocking Filter (EGANDF) with Sample Adaptive Offset filter (EGANDF-SAO-HEVC) is proposed for High Efficiency Video Coding (HEVC)/H-265. In this, the proposed hybrid in-loop filter involves EGANDF and Sample Adaptive Offset (SAO) filter that lessens the blocking artifacts caused by block-wise processing for coding unit (CU), which is mainly used for improving the video quality. Initially, EGANDF is proposed for HEVC/H-265 for removing blocking artifacts along low computation. Here, the output of EGANDF is given to the SAO filter for reducing ringing artifacts by diminishing high-frequency components during quantization. Thus, the proposed method efficiently reduces artifacts for improving video quality performance. The proposed EGANDF-SAO-HEVC method is implemented in the working platform of HEVC reference software with MATLAB. Finally, the proposed EGANDF-SAO-HEVC model has attained 27.26%, 29.65%, 12.45% higher accuracy, 33.56%, 31.8%, 28.7% higher sensitivity, 34.7%, 33.5%, 32.6% higher specificity, 46.92%, 35.7%, 41.3% lower MSE, 25.7%, 29.7%, 35.6% higher PSNR, and 25.6%, 28.9%, 13.6% higher SSIM for using basketball video sequence when compared to the existing methods.

Keywords

blocking artifacts, deblocking filter, generative adversarial network, high efficiency video coding, in-loop filter, sample adaptive offset

1 Introduction

Any coding method that works on block-based prediction and transform coding produces discontinuity at block boundaries in the restoring signal [1]. Hence, the observable discontinuity at the block boundary is known as blocking artifacts [2]. Such blocking artifacts are arising based on coarse quantization method and block transform of error prediction method [3]. Further, motion compensated prediction produces visible discontinuity closer to block boundary, while predict the adjoining blocks for current frame [4]. Also, variety of compression methods brings different classes of deblocking artifacts [5]. By deeming the compression in JPEG, discontinuity in adjacent 8 × 8 block experience blocking artifacts. Moreover, higher frequency coarse-quantization modulesundergo the effects of blurring and ringing [6].

Similarly, discontinuity near the block boundary predicts adjoining blocks that is different [7, 8]. Here, various methods utilized for minimizing the blocking artifacts, such as in-loop, post-filtering [9, 10]. Moreover, in-loop filtering method is implemented in H.265 video coding standards, sometimes it is implemented in display buffer [11]. Moreover, filters are recognized along liberty of designing algorithm for the necessity of particular application [12–16]. Also, the in-loop filters works in encoder loops and decoder loops [17, 18].

HEVC is the latest video coding standards attains 50% bit rate reduction that is assessed with prior H.264/AVC standard [19–21]. While executing the block based video coding, these lossy compression methods creates different artifacts, such as blurring, distortion, ringing, and contouring effects on output frames, especially at less bit rates [22]. To less those compression artifacts, HEVC adopts with in-loop filtering method, such as DF and SAO is used as encoder and decoder. Whereas, using deblocking filters in HEVC uses in-loop filtering technique, blocking artifacts are occurred, so computational complexity

Cite this article as: Moji, V., Murugavelu, M. "Adaptive and Efficient Hybrid In-loop Filter Based on Enhanced Generative Adversarial Networks with Sample Adaptive Offset Filter for HEVC/H-265", Periodica Polytechnica Electrical Engineering and Computer Science, 67(2), pp. 216–228, 2023.
https://doi.org/10.3311/PPee.20881

is increased; this reduces the overall performance of the architecture [23–29].

To address the above-mentioned problems, new objective function is formulated in this work that considers adaptive and efficient hybrid in-loop filter for HEVC/H-265. Thus, EGANDF-SAO-HEVC is proposed for removing artifacts effectively.

The main contributions of this manuscript are:

- In this work, Adaptive and Efficient Hybrid In-loop Filter Based on Enhanced Generative Adversarial Network with SAO filter is proposed for HEVC/H-265.
- Here, hybrid in-loop filter is intended for reducing compression artifacts by block-wise quantization and at the same time obtains bit-rate savings by restoring the damaged frame as nearer to the original. It contains DF and SAO filter [23].
- Deblocking filter decreases the blocking artifacts caused using block-wise processing for CU. By this, it improves the video quality [23, 25].
- But normal DF produces blocking artifacts and leads to high computation. To overcome this issue, Enhanced Generative Adversarial Networks [30] based DF is proposed in this research work for HEVC/H-265 that eliminates the blocking artifacts along low computations.
- Moreover, blocking artifacts are identified effectively using pre-processing unit called 1st convolution layer [23, 25].
- Then, the output of EGANDF is given to SAO filter. This SAO filter reduces the ringing artifacts by lessening high-frequency components during quantization [23, 26].
- The proposed method is implemented in the working platform of HEVC reference software with MATLAB.
- Then, the performance of EGANDF-SAO-HEVC method is analyzed with evaluation metrics, like MSE, PSNR, accuracy, sensitivity, specificity, recall, precision, and structural similarity index measurements (SSIM).
- Then, comparison of the proposed EGANDF-SAO-HEVC evaluation metrics is analyzed with various existing methods, like enhanced deep convolutional neural networks in HEVC (EDCNN-HEVC) [23], Content-aware CNN in HEVC (CACNN-HEVC) [24], and CNN based deblocking filter in SHVC (CNN-DF-SHVC) [25] respectively.

Remaining manuscript is organized as: the related works are reviewed in Section 2, Section 3 illustrates about the proposed method, Section 4 proves the results and discussion, Section 5 concludes this manuscript.

2 Related works

Various research mechanisms were previously presented in the literature related to HEVC. A few recent works are exhibited here.

Pan et al. [23] have suggested EDCNN for HEVC. Here, the presented in-loop filtering in HEVC lessens the artifacts. Also, aneffective in-loop filtering framework based EDCNN inspired by deep learning model was presented for enhancing the in-loop filtering performance in HEVC. Here, CNN model involves the processes, like normalization method, network learning ability and loss function were investigated. Finally, EDCNN effectively eliminates the artifactsusing weighted normalization method, feature information fusion block, accurate loss function but mean square error was very high.

Jia et al. [24] have suggested a CACNN for in-loop filtering in higher efficacy video coding. Here, the main focus of CNN technology was integrated by image restoration for enabling video coding performance in HEVC. Subsequently, coding tree unit (CTU) was preserved in independent region for further processing, andpresented content-aware multimodal filtering model wasrecognizedthrough the restoration of various regionsbyvarious CNN methods. Thus, presented method was donebased on the discriminative network of CACNN model, which are reducingartifacts but accuracy was lesser.

Dhanalakshmi and Nagarajan [25] have presented a CNNDF for scalable HEVC in H-265. Here, a CNNDF removes the blocking artifacts and provides low computation. Initially, the CNN framework identifies the input frame samples in video sequences in blocking artifacts were identified using pre-processing unit. Additionally, the normalization processed base layer to enhancement layer for preserving sharpness in video throughеliminatingthe artifacts. Finally, artifacts are removed by CNNDF model that provides better accuracy but PSNR value was low.

Singhadia et al. [26] have suggested a fast integrated DF and sample-adaptive-offset (SAO) parameter estimation architecture for HEVC. Here, the suggested model performs the hardware-effective executionusing integrated DF and SAO in HEVC. Further, it delivers rate–distortion performance compared with HEVC standard and

reports MSE, SSIM, and multi-scale SSIM index for 4K video structures.Finally, the suggested method consumes less power, area, energy. However, the model attained minimum accuracy value for reducing artifacts.

Baldev et al. [27] have presented directional streaming hardware architecture for DF of HEVC. Also, the architecture utilized adaptive parallel as well as pipeline processing strategies. It was processedusing less power with higher performance applications that involvesbroadcasting, virtual reality, and so on. Also, restructured block size was used for removing dependency from neighboring blocks. Thus, DF model of HEVC effectively reduced the artifacts but the accuracy of the model was low.

Dhanalakshmi and Nagarajan [28] have presented the combined residual network (CResNet) in-loop filterbased on scalable extension of HEVC. Here, the presented combined residual network in-loop filter uses layer information that waspresented in localized temporal domain for blocking the visualartifacts, such as blogging, ringing. In particular, block datawas connected with currentblocks and associated blocks of spatial and temporal base layer reference frames were measured for enhancing in-loop filtering. Also, rate distortion optimization (RDO) model utilized in-loop filter for sensing control flagand CTU. Thus, CResNet attained better reduction in bit rate and improved PSNR but the accuracy of the model was low.

Dhanalakshmi and Nagarajan [29] have presented group-normalized deep convolutional neural network (gDCNN) to improve performance of SHVC in-loop filter. Initially, problems encountered during traditional CNN modeling include normalization process, learning ability, and loss functions. Subsequently, gDCNN was presented effectively to remove artifacts based on statistical analysis, which are achieved through group-wise normalization method, feature extraction process, fusion method, and precision loss function. Finally, gDCNN provides improved outcome in PSNR. However, the mean square error of gDCNNmodel was higher than other models.

3 Proposed EGANDF-SAO-HEVC methodology

In this manuscript, the Adaptive and Efficient Hybrid In-loop Filter is designed based on Enhanced Generative Adversarial Networks for HEVC/H-265. Here, the Adaptive and Efficient Hybrid In-loop Filter includes DF and SAO filter. To reduce the blocking artifacts, EGANDF is proposed for HEVC/H-265. Initially, EGANDF detects the blocking artifacts by pre-processing unit and the features are extracted from the video frames. Moreover, the network model is trained by LRA for performing nonlinear mappings.

Additionally, EGANDF attains low computation based on max-pooling layers and soft-max layers. Subsequently, output of EGANDF is given to the sample adaptive offset (SAO) filter, which decrease the ringing artifacts by diminishing high-frequency components during quantization. While, EGANDF is applied to the samples that is placed at the boundaries of Prediction Unit (PU) or Transform Unit (TU) and sample adaptive offset is applied effectively with all samples by satisfying certain conditions to less the mean sample distortion region. Fig. 1 shows the block diagram of the proposed EGANDF-SAO-HEVC methodology.

3.1 Enhanced Generative Adversarial Networks based deblocking filter (EGANDF)

In this manuscript, an Enhanced Generative Adversarial Networks based deblocking filter (EGANDF) is proposed for reducing blocking artifacts. The aim of EGANDF is to reduce the artifacts that areoccurred by block-based coding. The proposed EGANDF method is the combined form of deep convolutional generative adversarial networkand modified CNN that rebuildsloss of quality that is suffered using blocks of reference frames is somewhat same asthe blocks of original frame. Initially, the deep convolutional generative adversarial network generates the vibration signals in multiple fault conditions under sample imbalanced situation. Subsequently, the modified convolutional neural network enhances the feature learning performance of detection by removing artifacts. By this, generated samples attained effective employment, accuracy is higher for artifacts diagnosis.

3.1.1 Sample generation in EGANDF

In this step, the input samples are generated by EGANDF that is located at PU or TU boundaries. Here, EGAN deblocking filter uses the same process in all layers for performing up-sampling process in generator (G) and down-sampling in discriminator (D). Additionally, the generator part involves Rectifier Linear Unit (ReLU) and batch normalization. Also, output layer includes activation function as hyperbolic tangent (Tanh). Moreover, input samples are generated and sampled from the uniform distribution that is controlled to [−1, 1]. Subsequently, pseudo-samples are generated with fixed length after learning internal distribution of inventive vibration signals. Finally, pseudo-samples are generated along original input samples into discriminator for its training. Additionally, the batch normalization of EGANDF is applied in discriminator to diminish gradient vanishing. In this, video dataset

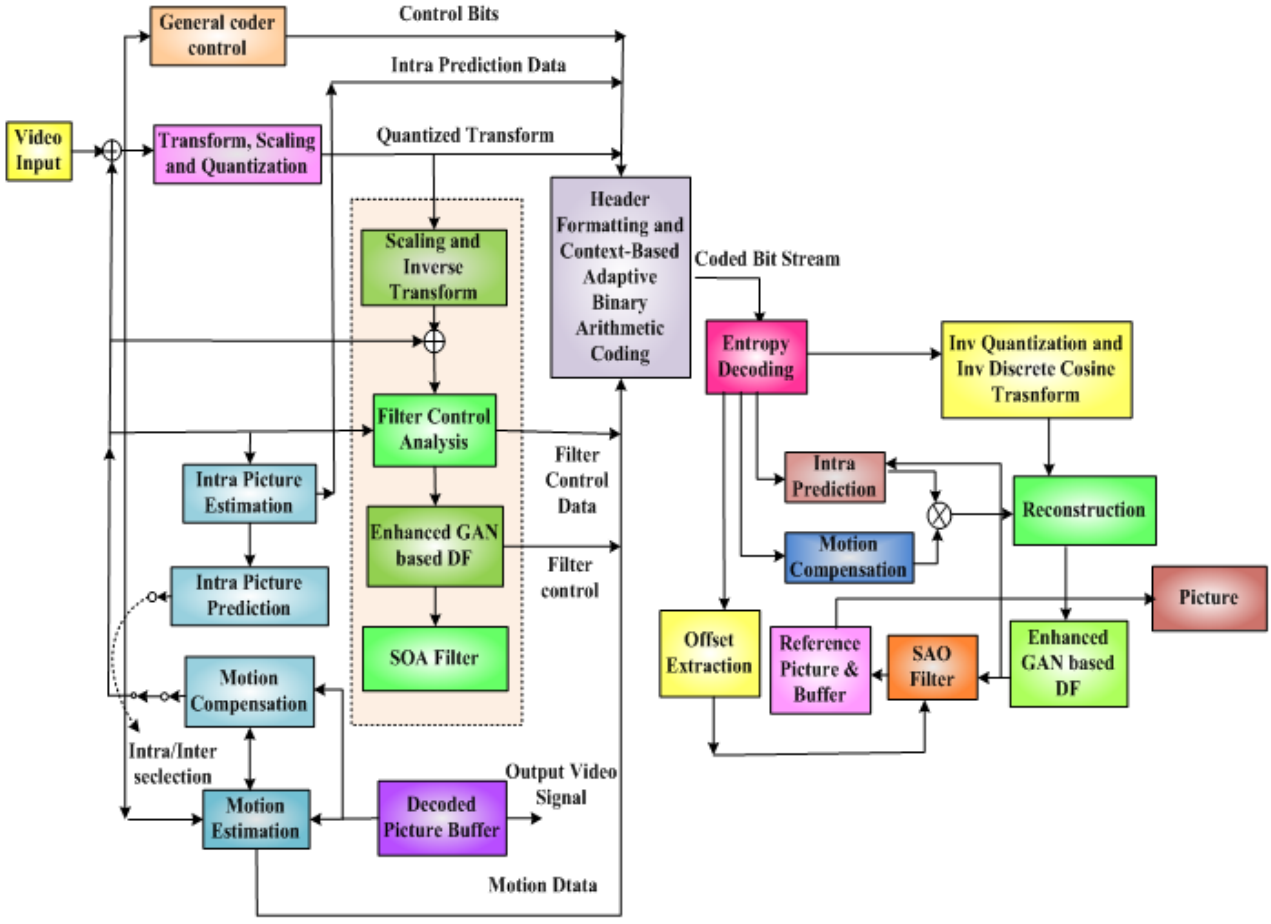


Fig. 1 Block diagram of the proposed EGANDF-SAO-HEVC methodology

is developed that involves 39 video sequences. Here, the initial convolutional layer takes 64 × 64 blocks of less quality frames, such as input, kernel size as 7 × 7, stride value as 2 and 4 for base layer and enhancement layer. Therefore, resultant layersize as 8 × 8, kernel size as 3 × 3, stride value of 1, 2 for base layer including enhancement layer. After the convolution procedure, ReLU is employed for performing nonlinear mapping. Hence, rectified linear unitsis engaged with 1 pixel value for the feature maps in every convolution layers. Likewise, except fully connected layer, every convolutional layers are applied stride 4 and kernel size as 5 × 1. For discriminator, original and created vibration signals along label data are represented as input data. After experiencing 4 convolution layers, input samples are down-sampled as less resolution feature maps along strong semantic data. Then, feature maps along flatten layer, fully connected layer, and the network outputs the final discriminant outcome. Thus, from baselayer to enhancement layer, batch normalization (BN) is utilized in improved generative adversarial networks framework for preserving sharpness in video by eliminating artifacts produced based on inter-layer prediction and quantization

3.1.2 Artifact detection in first convolution layer

The convolution layer shows great potential for extracting the intrinsic features from original vibration signals. For detecting blocking artifacts, one dimensional convolution model is adopted with baseline method for finding the features of discriminative time-series signals. Feature vectors along various channels from pooling layer are flattened with 1-D vector.

In this, the pooling layer is utilized for reducing the visual artifacts based on the reduction of spatial dimension results in low computational power byreducing the dimension. Furthermore, to effectively maintain the training process, important aspect values are needed to be extracted. In pooling operation, the average pooling is stated as average values of each pixel enclosed with image patch and kernel. Similarly, max pooling is defined as the maximum value of each pixel enclosed with image patch and kernel. Thus, the pooling operation performs reduction of noise by diminishing the spatial dimension, which can denoise the frame. Subsequently, the weight parameters of feature optimization are acquired with the flattened vector. Also, extract the features from the frames in video

related to patch by engaging kernel along stride to scan whole frames in video. Then, feature extraction is done in the convolution and max pooling layers that is flattened for feature vector, which is given in Eq. (1):

$$\boldsymbol{k\_fea}_i = F\left[\max conv\left(k_i\right)\right], \tag{1}$$

where the current input samples are represented as (k_i), F denotes the flatten operation, *conv* is the function of convolution operation, and max denotes the max pooling function. Additionally, high-level feature vector of training set ($\boldsymbol{k\_fea}^T$) is mentioned in Eq. (2):

$$\boldsymbol{k\_fea}^T = \left\{k\_fea_1^T, k\_fea_i^T, ..., k\_fea_{b\cdot N_T}^T\right\}, \tag{2}$$

where $\boldsymbol{k\_fea}^T$ denotes the feature vector that denotes the spatial feature depiction of (i^{th}) sample, the block boundaries are denoted as b, and total number of training samples are denoted as N_T. Subsequently, block boundaries b of (i^{th}) balanced input samples are detected by extracting the feature vectors, which is given in Eq. (3):

$$C_b = \frac{1}{N_T}\sum_{i=1}^{N_T} \boldsymbol{k\_fea}_{b,i}^T, \tag{3}$$

where C_b is the detection parameter of (b) block boundaries, and $\boldsymbol{k\_fea}_{b,i}^T$ represents the feature vector.

Therefore, block boundaries are identified by enhancing the feature learning ability to improve the artifacts reduction. Thus, the feature vector of total input samples by identified block boundaries are based on Eq. (4):

$$\boldsymbol{k\_fea}_{b,i}^T = C_b, i \in N_T. \tag{4}$$

The weight matrix ($\boldsymbol{\varphi}$) of the samples are attained by the calculation of two feature matrixes, which is expressed in Eq. (5):

$$\boldsymbol{\varphi} = \left(\boldsymbol{k\_fea}'_{b,i} * \boldsymbol{k\_fea}_{b,i} + \delta \boldsymbol{I}\right)^{-1} \boldsymbol{k\_fea}'_{b,i} * \boldsymbol{k\_fea}^T, \tag{5}$$

where δ represents the biased coefficient, $\boldsymbol{I}$ denotes the identity matrix. Thus, optimized features of the block boundaries are identified by the calculation of weight matrix.

3.1.3 Removing blocking artifacts

In EGANDF, the initial layers are utilized for calculating feature matrix of input training samples for detecting blocking artifacts at block boundaries. Then, feature matrix $\boldsymbol{k\_fea}_{b,i}$ of each training samples are attained by the process of hierarchical feature extraction in the initial layers. Additionally, the combined class labels x_{label}^T of extracted features are utilized for removing blocking artifacts. Particularly, fully connected layers extracts needed content from previous feature maps and combines to form higher-level representation that is same as matrix multiplicationmeans feature transformation. Thus, the nonlinear transform is based on the activation unit RLUby utilizing 2 fully connected layers in EGANDF that tells whether the block has artifacts/not.

Here, EGANDF is processed in the convolution layer after the completion of feature fusion for channel transformation. Hence, the long shortcut connection is from the original input without processing convolution layer to final output, whichisprovenby exception of shortcut connections in blocks to attain previous mapping relationship. Also, the proposed in-loop filtering method is embedded in HEVC reference softwarefor improvingencoding performance of HEVC. Finally, Enhanced generative adversarial networks based DF reduced the blocking artifacts. Additionally, EGANDF model attains low computation using max-pooling layers and soft-max layers in EGAN. Subsequently, output of EGANDF is given to SAO filter.

3.2 Sample Adaptive Offset (SAO) filter

Sample adaptive offset is placed after EGANDF and it belongs to in-loop filters, which specifies samples after EGANDF. Here, the concept of sample adaptive offset filter classifies reconstructed pixels into various types, attainingoffset for every kind, and adds with everykinds of sample that is achieved on largest coding unit basis in HM 9.0 (reference software of HEVC). Moreover, the main objective of the Sample adaptive offset model is less the ringing artifacts by diminishing high-frequency components during quantization.

Additionally, two types of SAO are assumed in HEVC, they are, Edge offset (EO) and band offset (BO). Therefore, in edge offset, sample classification is comparison among current samples and neighboring pixels. Also, in band offset, classification of samples is related to sample values. Therefore, current largest coding unit reuses the SAO parameters from left largest coding unit/upper largest coding unit by SAO merging modeless the data that is to be coded using entropy coding.

The detailed description regarding the types of SAO are given as follows:

- Edge offset (EO):
  In this, edge offset uses 4 one dimensional direction patterns for pixel classification, such as edge offset-0 class as horizontal manner, edge offset-1 class as vertical manner, edge offset-2 class as 135° diagonal manner, edge offset-3 class as 45° diagonal manner.

Based on these patterns, 4 edge offset classes are mentioned and every class is corresponds to 1 pattern. For a given edge offset class; every sample is categorized into one of 5 kinds by likening its own value along 2 neighboring samples' value. If current sample not belongs with any of thesekindsof classes, then SAO is not applied to it.

- Band offset (BO):
  Band offsetexecutes the sample classification related with the sample's of own value. Also, it categorizesevery pixels of largest coding unit into multiple bands, where every band has pixels in same intensity interval in HM 9.0. Thus, the idea of band in edge offset is similar with edge offset method. Hence, pixel intensity range is split in the same wayas 32 uniform bands is from 0 to higher values (for example, 255 for 8-bit pixels), every band contains offset by themselves. Moreover, only 4 offsets of 4 consecutive bands,starting band position arecoded through entropy coding that is transmitted to decoder. Here, the aim for choosing only 4 bands is the region of sample range that is quite restricted after regions lessens from the picturequad tree partitions with coding tree blocks.
- SAO merging:
  Current largest coding unit reuse SAO parameters (SAO type and 4 offsets) from left largest coding unit (merge-left)/upper LCU (merge-up). If current largest coding unit chooses merge-left or merge-up, eachSample Adaptive Offset data of current largest coding unit are reused from left/upper largest coding unit. Thus, the SAO merging mode reduces side datathat is coded using the entropy coding effective.
- Fast distortion estimation:
  Sample Adaptive Offset type selection and offsets are executed on fast distortion estimation technique related to the region of largest coding unit in HM 9.0. Also, the fast distortion estimation technique is the executionof Sample Adaptive Offset that needs to add the offsets to pre-Sample Adaptive Offset sample (post-deblocking filter samples) for generating post-Sample Adaptive Offset sample and again measure distortion among original sample and post-Sample Adaptive Offset sample. Consider the pixel positions as $p_1(t)$ and original pixel as $p(t)$ input video sequence with set of pixels (t). Therefore, these pixels are represented as SAO types and distortion among original pixels and pre-Sample Adaptive Offset pixels is measured using Eq. (6):

$$d_{\text{pre}} = \sum_t \left(p_1(t) - p(t) \right)^2 . \tag{6}$$

The distortion among original pixels and post-Sample Adaptive Offset pixels is calculated using Eq. (7):

$$d_{\text{post}} = \sum_t \left(p_1(t) - \left(p(t) + \text{Offset} \right)^2 \right). \tag{7}$$

Subsequently, the delta distortion of the model is described in Eq. (8):

$$\Delta d = d_{\text{post}} - d_{\text{pre}} = nO^2 - 2OS , \tag{8}$$

where n denotes the total number of pixels in the given set, S represents the total quantity of differences among pre-SAO pixels and original pixels that is given in Eq. (9):

$$S = \sum_t \left(p_1(t) - p(t) \right). \tag{9}$$

Subsequently, delta rate-distortion cost of the model is described using Eq. (10):

$$\Delta J = \Delta d + \lambda E , \tag{10}$$

where Lagrange multiplier is denoted as λ, and the estimated bits of side data for specified Sample Adaptive Offset type are represented as E. Subsequently, the calculation of λ is accomplished using entropy coding that represents the time-consuming process in HM 9.0. Finally, the proposed EGANDF-SAO-HEVC approach model has effectively reduced the blocking artifacts and ringing artifacts in the compressed videos. Thus, proposed EGANDF-SAO-HEVC approach has efficiently improves the video quality.

4 Results and discussion

Here, the performance of EGANDF-SAO-HEVC approach is evaluated. The proposed EGANDF-SAO-HEVC method is simulated by utilizing the working platform of HEVC reference software with MATLAB. Here, simulations are carried out in PC with Intel Core i5, 2.50 GHz central processing unit, 8 GB random access memory, and Windows 7. The performance of EGANDF-SAO-HEVC method is calculated using performance metrics. The attained outcomes are likened with exiting methods to evaluate the efficiency of the proposed method.

4.1 Dataset description

In this work, the video database is developed along 39 video sequences as datasets. In this, 1st frame for 32 video sequences are taken as training set and randomly

chosen video sequenceas validation set, which are Basket Ball Drill, Basket Ball Pass, Race Horses C, Race Horses D, Basket Ball, Blowing Blubbles and Race Horse. Table 1 tabulates the random frame analysis of some of the test video sequences.

Fig. 2 (a) represents input basketball, blowing bubbles, and race horse images from HEVC test sequences utilizing EGANDF-SAO. The obtained MSE images are shown in Fig. 2 (b). Finally, the artifacts reduction using proposed EGANDF-SAO-HEVC approach is shown in Fig. 2 (c).

Fig. 3 shows the PSNR value of test video sequence by proposed EGANDF-SAO-HEVC method. In this, PSNR value for test video sequences as basketball drill, basketball pass, race horses C, and race horses D are analyzed. Here, PSNR values of the test sequences are increased by improving the bit rate value. In this, basketball drill has attained high PSNR as 47 dB and race horses D has attained low PSNR value as 40.7 dB.

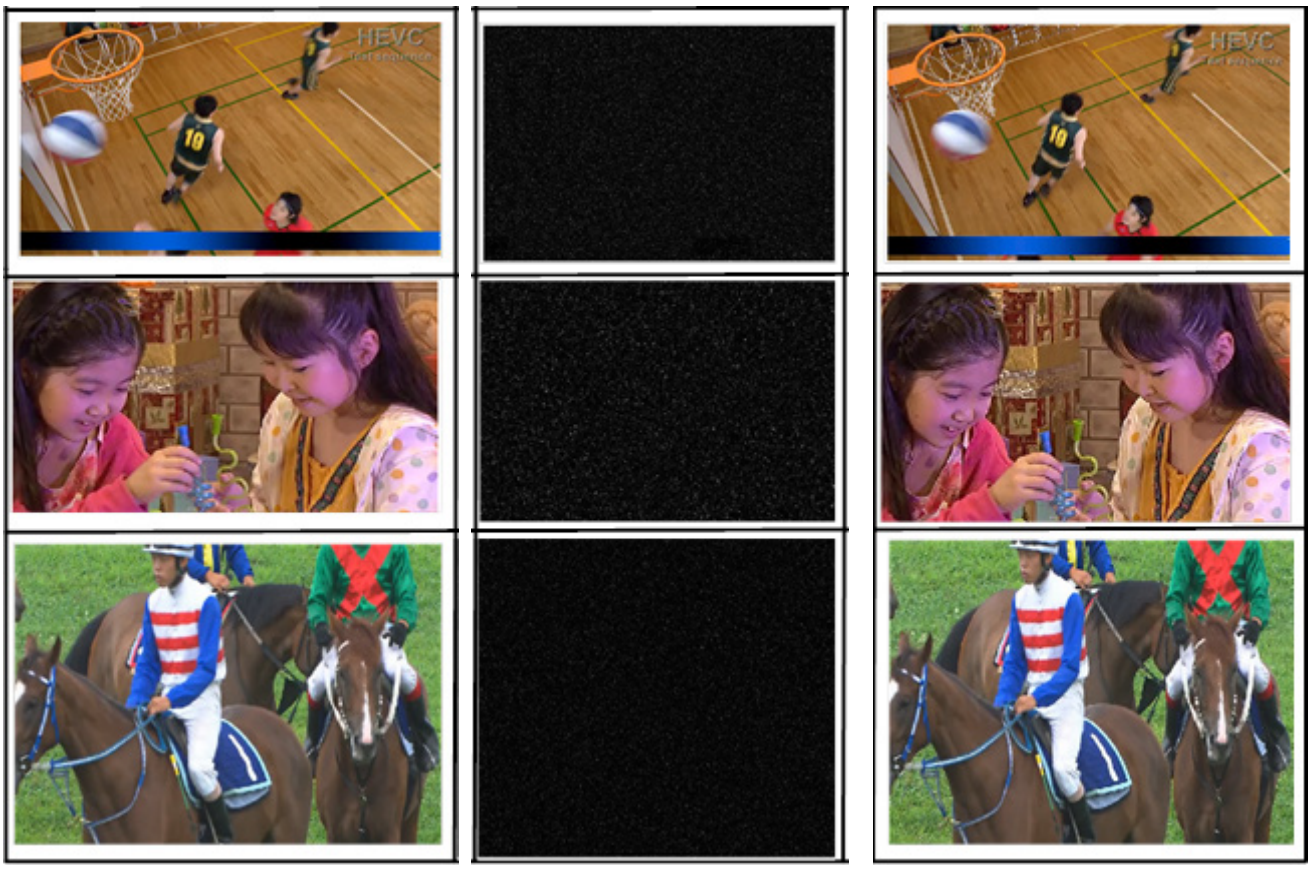
Table 1 Random frame analysis

Test video sequence	PSNR	Bit rate in bps	Encoding time in sec	Decoding time in sec
"Basket Ball Drill"	40.875	10454	101.23	2.5
"Basket Ball Pass"	40.316	967.52	5421.3	6.33
"Race Horses C"	41.654	1123.2	99.243	1.56
"Race Horses D"	41.345	865.25	124.21	4.562

Fig. 4 shows the SSIM value of test video sequence by proposed EGANDF-SAO-HEVC approach. In this, SSIM value for test video sequences as basketball drill, basketball pass, race horses C, and race horses D are analyzed. In this, the SSIM value of basketball drill, and basketball pass are increased by increasing the value of bit rate but the race horses C, and race horses D SSIM values are stable.

4.2 Performance metrics

Here, the efficacy of EGANDF-SAO-HEVC method is analyzed with the performance metrics, such as accuracy, precision, recall, sensitivity, specificity, MSE, PSNR and SSIM. Subsequently, to compute the confusion matrix values, True positive, False positive, True Negative, False Negative is considered.



(a) (b) (c)

Fig. 2 Artifacts removal using proposed EGANDF-SAO-HEVC method, (a) input images from HEVC test sequences utilizing EGANDF-SAO, (b) the obtained MSE images, (c) artifacts reduction using proposed EGANDF-SAO-HAVC approach

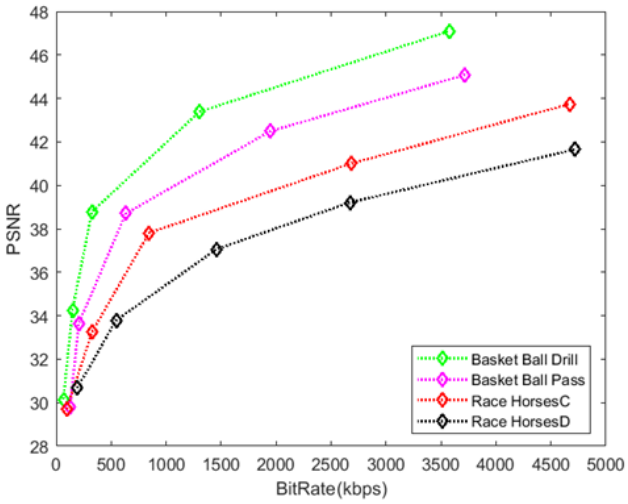


Fig. 3 PSNR value of test video sequence by EGANDF-SAO-HEVC approach

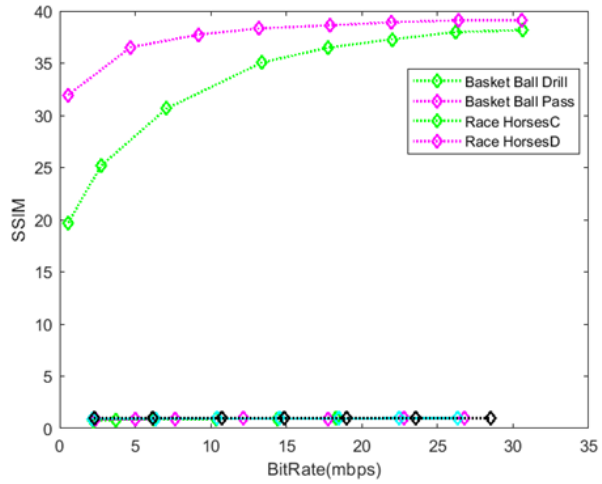


Fig. 4 SSIM value of test video sequence by EGANDF-SAO-HEVC approach

4.2.1 Computation of accuracy

Accuracy is identified the efficacy of the proposed EGANDF-SAO-HEVC method while removing blocking artifacts from samples, which is calculated using Eq. (11):

$$Acc(A) = \frac{T_P + T_N}{T_P + F_P + T_N + F_N}, \tag{11}$$

here (T_P) implies true positive, (F_P) implies false positive, (T_N) implies true negative, (F_N) implies false negative.

4.2.2 Computation of precision

Precision measures the efficacy of the proposed EGANDF-SAO-HEVC approach during artifact detection are evaluated using Eq. (12):

$$P_r = \frac{T_P}{T_P + F_P}. \tag{12}$$

4.2.3 Computation of sensitivity

Sensitivity is utilized to identify the presence of blocking artifacts in the video sequences, which is calculated using Eq. (13):

$$Se = \frac{T_P}{T_P + F_N}. \tag{13}$$

4.2.4 Computation of specificity

Specificity identifies the presence of ringing artifacts in the video sequences, which is calculated using Eq. (14):

$$Sp = \frac{T_N}{T_N + F_P}. \tag{14}$$

4.2.5 Computation of recall

Recall has the ability of proposed deblocking filterreduce total blocking artifacts in the video sequences that is calculated using Eq. (15):

$$R_e = \frac{T_P}{T_P + F_N}. \tag{15}$$

4.2.6 Mean Square Error (MSE)

The calculation MSE denotes error among signals that is expressed in Eq. (16):

$$\text{MSE} = \frac{1}{hw}\sum_{k=0}^{h-1}\sum_{l=0}^{w-1}\left(g_p(k,l) - g'_p(k,l)\right)^2, \tag{16}$$

where input and output signals are represented as h and w respectively. Initial values of input signal is denoted as g'_p and the terminal value of the signal is denoted as g_p.

4.2.7 Peak Signal to Noise Ratio (PSNR)

It is defined as the ratioto measure the quality among original images and compressed images, which is calculated using Eq. (17):

$$\text{PSNR} = 10 \times \lg\left(\frac{255^2}{\text{MSE}}\right). \tag{17}$$

4.2.8 Structural Similarity Index Matrix (SSIM)

It is used to measure the difference among 2 similar imageries. Additionally, Structural Similarity Index Matrix is used on the gradient of images.

4.3 Comparative analysis of performance metrics

Here, the efficiency of EGANDF-SAO-HEVC method under performance metrics, like accuracy, precision, recall,

sensitivity, specificity, MSE, PSNR and SSIM are compared with existing methods, like EDCNN-HEVC [20], CACNN-HEVC [21], and CNN-DF-SHVC [22]. Table 2 shows that the comparative analysis outcomes for all performance metrics.

Fig. 5 shows the accuracy measurement of the EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, accuracy of the proposed EGANDF-SAO-HEVC method has attained 27.26%, 29.65%, and 12.45% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, the accuracy of the proposed EGANDF-SAO-HEVC method has attained 19.62%, 21.43%, and 17.9% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, accuracy of the proposed EGANDF-SAO-HEVC method has attained 15.7%, 19.75%, and 11.57% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

Fig. 6 shows the sensitivity measurement of the EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, the sensitivity of the proposed EGANDF-SAO-HEVC method has attained 33.56%, 31.8%, and 28.7% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, the sensitivity of the proposed EGANDF-SAO-HEVC method has attained 21.45%, 33.5%, and 23.5% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, sensitivity of the proposed EGANDF-SAO-HEVC method has attained 24.6%, 26.87%, and 22.4% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

Fig. 7 shows the specificity measurement of EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, specificity of the proposed EGANDF-SAO-HEVC approach has attained 34.7%, 33.5%, and 32.6% better than the existing CNN-DF-SHVC,

Table 2 Accurate value of each metrics

Performance metrics	Video sequences	CNN-DF-SHVC	EDCNN-HEVC	CACNN-HEVC	EGANDF-SAO-HEVC (Proposed)
Accuracy (%)	Basketball	0.73	0.68	0.85	0.962761
	Blowing bubbles	0.79	0.71	0.79	0.98121
	Race horse	0.82	0.75	0.81	0.998435
Sensitivity (%)	Basketball	0.67	0.76	0.86	0.977304
	Blowing bubbles	0.69	0.64	0.79	0.965524
	Race horse	0.73	0.71	0.85	0.963549
Specificity (%)	Basketball	0.69	0.73	0.76	0.984468
	Blowing bubbles	0.64	0.69	0.73	0.980169
	Race horse	0.68	0.78	0.75	0.997392
Precision (%)	Basketball	0.69	0.75	0.67	0.923077
	Blowing bubbles	0.67	0.72	0.68	0.923077
	Race horse	0.71	0.78	0.72	0.923077
Recall (%)	Basketball	0.629	0.65	0.68	0.964962
	Blowing bubbles	0.687	0.67	0.75	0.95034
	Race horse	0.651	0.63	0.73	0.957887
MSE	Basketball	6.7	4.8	5.8	2.01
	Blowing bubbles	6.4	4.9	5.6	2.05
	Race horse	6.3	5.1	5.7	2.04
PSNR	Basketball	33.8	28.9	23.5	45.12
	Blowing bubbles	37	25	27	45.1
	Race horse	36	27	29	45.07
SSIM	Basketball	0.73	0.68	0.83	0.97
	Blowing bubbles	0.75	0.67	0.75	0.96
	Race horse	0.76	0.69	0.76	0.96

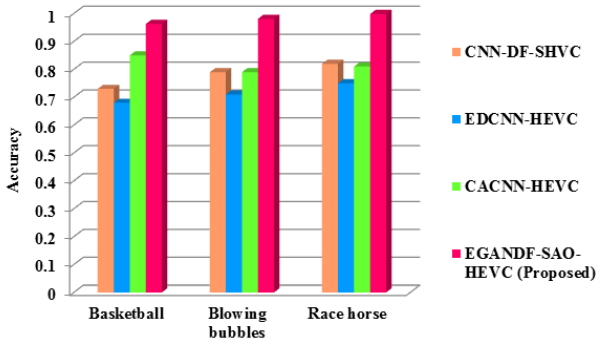


Fig. 5 Comparison of accuracy

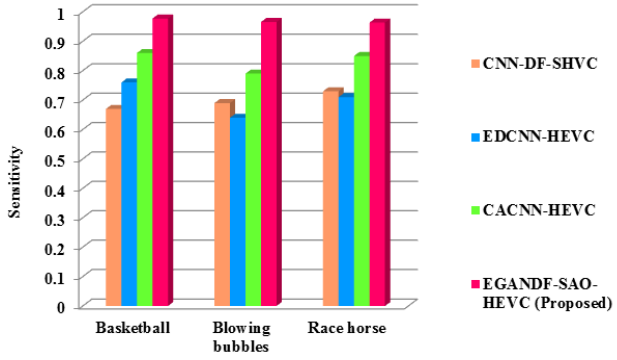


Fig. 6 Comparison of sensitivity

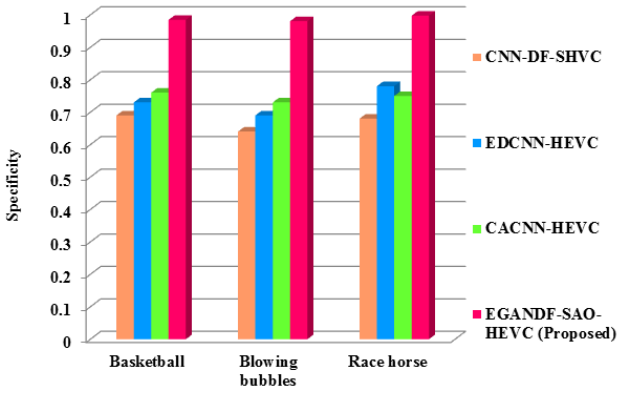


Fig. 7 Comparison of specificity

EDCNN-HEVC and CACNN-HEVC methods respectively. For blowing bubbles video sequence, specificity of the proposed EGANDF-SAO-HEVC method has attained 33.8%, 34.79%, and 35.67% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, specificity of the proposed EGANDF-SAO-HEVC method has attained 28.8%, 19.7%, and 23.7% better than the existing CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC methods respectively.

Fig. 8 shows the precision measurement of EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, the precision of the proposed

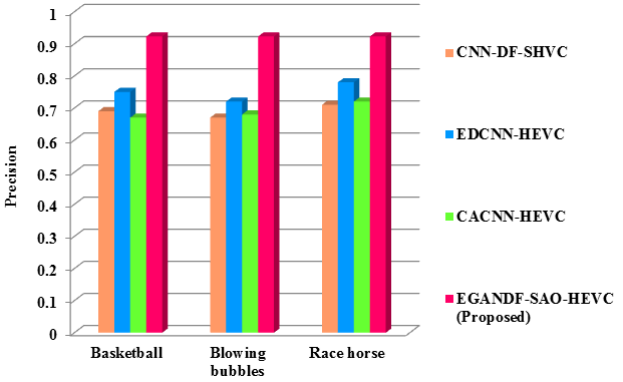


Fig. 8 Comparison of precision

EGANDF-SAO-HEVC method has attained 32.45%, 23.46%, and 28.7% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, the precision of the proposed EGANDF-SAO-HEVC method has attained 31.8%, 24.5%, and 27.8% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, the precision of the proposed EGANDF-SAO-HEVC method has attained 32.5%, 27.6%, and 28.9% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

Fig. 9 depicts the recall measurement of the EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, the recall of the proposed EGANDF-SAO-HEVC approach has attained 34.7%, 32.6%, and 35.8% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, the recall of the proposed EGANDF-SAO-HEVC approach has attained 36.7%, 32.9%, and 35.7% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, the

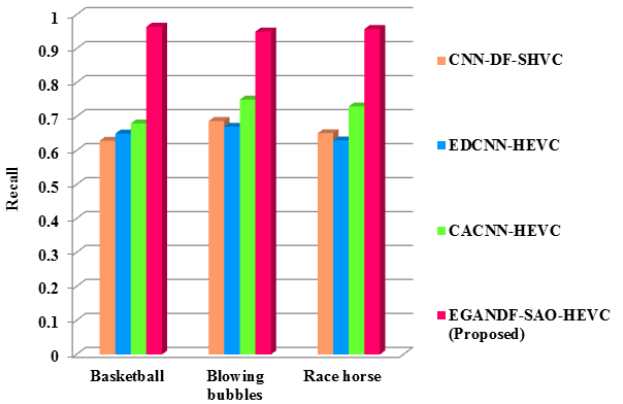


Fig. 9 Comparison of recall

recall of the proposed EGANDF-SAO-HEVC method has attained 38.9%, 34.6%, and 39.9% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

Fig. 10 shows the MSE measurement of EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, MSE of the proposed EGANDF-SAO-HEVC method has attained 46.92%, 35.7%, and 41.3%lower than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, the MSE of the proposed EGANDF-SAO-HEVC method has attained 45.7%, 38.9%, and 41.6% lower than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC methods. For race horse video sequence, MSE of the proposed EGANDF-SAO-HEVC method has attained 48.5%, 41.4%, and 45.7% lower than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

Fig. 11 depicts the PSNR measurement for EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, PSNR of the proposed EGANDF-SAO-HEVC approach has attained 25.7%, 29.7%, and 35.6% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, PSNR of the proposed EGANDF-SAO-HEVC approach has attained 21.7%, 31.6%, and 37.8% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, PSNR of the proposed EGANDF-SAO-HEVC approach has attained 20.9%, 28.9%, 34.6% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

Fig. 12 shows the SSIM measurement of EGANDF-SAO-HEVC method with existing methods. For basketball video sequence, SSIM of the proposed EGANDF-SAO-HEVC approach has attained 25.6%, 28.9%, and 13.6% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For blowing bubbles video sequence, SSIM of the proposed EGANDF-SAO-HEVC approach has attained 27.8%, 32.6%, 21.7% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively. For race horse video sequence, the SSIM of the proposed EGANDF-SAO-HEVC approach has attained 26.7%, 31.5%, 27.8% higher than the existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC respectively.

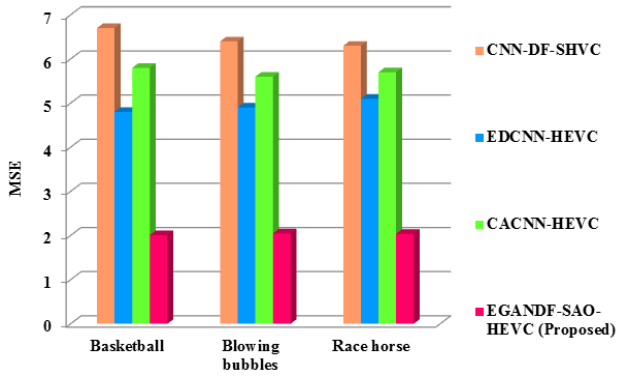


Fig. 10 Comparison of MSE

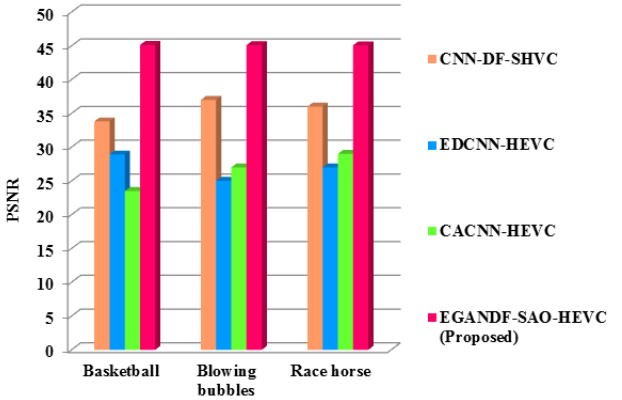


Fig. 11 Comparison of PSNR

5 Conclusion

In this work, the in-loop filtering model based on Enhanced Generative Adversarial Network with Sample Adaptive Offset filter (EGANDF-SAO-HEVC) effectively reduced the artifacts for HEVC encoded videos. Here, EGANDF is used for removing blocking artifacts, and SAO filter is utilized to reduce the ringing artifacts. Finally, the proposed method efficiently reducesartifacts for improving video quality performance. Here, the implementation of the proposed method is done in working platform

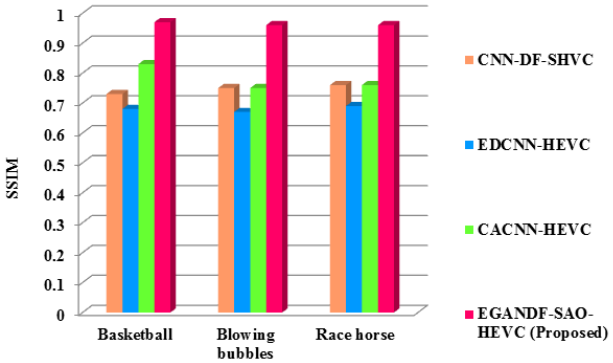


Fig 12 Comparison of SSIM

of HEVC reference software with MATLAB. The proposed EGANDF-SAO-HEVC filter has achieved 19.62%, 21.43%, 17.9% higher accuracy, 21.45%, 33.5%, 23.5% higher sensitivity, 45.7%, 38.9%, 41.6% lower MSE, 21.7%, 31.6%, 37.8% higher PSNR, and 27.8%, 32.6%, 21.7% higher SSIM for blowing bubbles video sequence than existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC methods respectively. Also, the proposed EGANDF-SAO-HEVC filter has achieved 15.7%, 19.75%, 11.57% higher accuracy, 24.6%, 26.87%, 22.4% higher sensitivity, 48.5%, 41.4%, 45.7% lower MSE, 20.9%, 28.9%, 34.6% higher PSNR and 26.7%, 31.5%, 27.8% higher SSIM for race horse video sequence existing methods, like CNN-DF-SHVC, EDCNN-HEVC and CACNN-HEVC methods respectively.

References

[1] Kopperundevi, P., Prakash, M. S., Ahamed, S. R. "A high throughput hardware architecture for deblocking filter in HEVC", Signal Processing: Image Communication, 100, 116517, 2022. https://doi.org/10.1016/j.image.2021.116517

[2] Chopp, H. H., Banerjee, S., Cossairt, O., Katsaggelos, A. K. "Removing Blocking Artifacts in Video Streams Using Event Cameras", [preprint] arXiv, arXiv:2105.05973v1, 12 May 2021. https://doi.org/10.48550/arXiv.2105.05973

[3] Yeh, C.-H., Lin, C.-H., Lin, M.-H., Kang, L.-W., Huang, C.-H., Chen, M.-J. "Deep learning-based compressed image artifacts reduction based on multi-scale image fusion", Information Fusion, 67, pp. 195–207, 2021. https://doi.org/10.1016/j.inffus.2020.10.016

[4] Rashmi, B. S., Nagendraswamy, H. S. "Video shot boundary detection using block based cumulative approach", Multimedia Tools and Applications, 80(1), pp. 641–664, 2021. https://doi.org/10.1007/s11042-020-09697-6

[5] Singh, A., Singh, J. "A content adaptive method of de-blocking and super-resolution of compressed images", Multimedia Tools and Applications, 80(7), pp. 11095–11131, 2021. https://doi.org/10.1007/s11042-020-10112-3

[6] Jagtap, N. S., Thepade, S. D. "High-quality image multi-focus fusion to address ringing and blurring artifacts without loss of information", The Visual Computer, 2021. https://doi.org/10.1007/s00371-021-02300-5

[7] Wu, Y., Li, X., Zhang, Z., Jin, X., Chen, Z. "Learned block-based hybrid image compression", IEEE Transactions on Circuits and Systems for Video Technology, 32(6), pp. 3978–3990, 2022. https://doi.org/10.1109/TCSVT.2021.3119660

[8] Yuan, Z., Liu, H., Mukherjee, D., Adsumilli, B., Wang, Y. "Block-based Learned Image Coding with Convolutional Autoencoder and Intra-Prediction Aided Entropy Coding", In: 2021 Picture Coding Symposium (PCS), Bristol, UK, 2021, pp. 1–5. ISBN 978-1-6654-3078-4 https://doi.org/10.1109/PCS50896.2021.9477503

[9] Zhang, S., Fan, Z., Ling, N., Jiang, M. "Recursive residual convolutional neural network-based in-loop filtering for intra frames", IEEE Transactions on Circuits and Systems for Video Technology, 30(7), pp. 1888–1900, 2020. https://doi.org/10.1109/TCSVT.2019.2938192

[10] Ma, D., Zhang, F., Bull, D. R. "MFRNet: a new CNN architecture for post-processing and in-loop filtering", IEEE Journal of Selected Topics in Signal Processing, 15(2), pp. 378–387, 2021. https://doi.org/10.1109/JSTSP.2020.3043064

[11] Ding, D., Kong, L., Wang, W. Zhu, F. "A progressive CNN in-loop filtering approach for inter frame coding", Signal Processing: Image Communication, 94, 116201, 2021. https://doi.org/10.1016/j.image.2021.116201

[12] Kuanar, S., Rao, K. R., Conly, C., Gorey, N. "Deep learning based HEVC in-loop filter and noise reduction", Signal Processing: Image Communication, 99, 116409. 2021. https://doi.org/10.1016/j.image.2021.116409

[13] Shajin, F. H., Rajesh, P., Thilaha, S. A. "Bald eagle search optimization algorithm for cluster head selection with prolong lifetime in wireless sensor network", Journal of Soft Computing and Engineering Applications, 1(1), pp. 1–7, 2020.

[14] Rajesh, P., Shajin, F. H., Rajani, B., Sharma, D. "An optimal hybrid control scheme to achieve power quality enhancement in micro grid connected system", International Journal of Numerical Modelling: Electronic Networks, Devices and Fields, 35(6), e3019, 2022. https://doi.org/10.1002/jnm.3019

[15] Shajin, F. H., Rajesh, P., Raja, M. R. "An efficient VLSI architecture for fast motion estimation exploiting zero motion prejudgment technique and a new quadrant-based search algorithm in HEVC", Circuits, Systems, and Signal Processing, 41(3), pp. 1751–1774, 2022. https://doi.org/10.1007/s00034-021-01850-2

[16] Rajesh, P., Shajin, F. H., Cherukupalli, K. "An efficient hybrid tunicate swarm algorithm and radial basis function searching technique for maximum power point tracking in wind energy conversion system", Journal of Engineering, Design and Technology, 2021. https://doi.org/10.1108/JEDT-12-2020-0494

[17] Xu, Q., Jiang, X., Sun, T., Kot, A. C. "Detection of transcoded HEVC videos based on in-loop filtering and PU partitioning analyses", Signal Processing: Image Communication, 92, 116109. 2021. https://doi.org/10.1016/j.image.2020.116109

[18] Sun, W., He, X., Chen, H., Xiong, S., Xu, Y. "A nonlocal HEVC in-loop filter using CNN-based compression noise estimation", Applied Intelligence, 2022. https://doi.org/10.1007/s10489-022-03259-z

[19] Li, T., Xu, M., Zhu, C., Yang, R., Wang, Z., Guan, Z. "A deep learning approach for multi-frame in-loop filter of HEVC", IEEE Transactions on Image Processing, 28(11), pp. 5663–5678, 2019. https://doi.org/10.1109/TIP.2019.2921877

[20] Kopperundevi, P., Prakash, M. S. "Methods to develop high throughput hardware architectures for HEVC Deblocking Filter using mixed pipelined-block processing techniques", Microelectronics Journal, 105413, 2022. https://doi.org/10.1016/j.mejo.2022.105413

[21] García-Lucas, D., Cebrián-Márquez, G., Díaz-Honrubia, A. J., Cuenca, P. "Accelerating the CU partitioning decision in an HEVC-JEM transcoder", Multimedia Tools and Applications, 79(3–4), pp. 2047–2067, 2020.
https://doi.org/10.1007/s11042-019-08326-1

[22] Xu, Q., Jiang, X., Sun, T., Kot, A. C. "Detection of HEVC double compression with non-aligned GOP structures via inter-frame quality degradation analysis", Neurocomputing, 452, pp. 99–113, 2021.
https://doi.org/10.1016/j.neucom.2021.04.092

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

[24] Jia, C., Wang, S., Zhang, X., Wang, S., Liu, J., Pu, S., Ma, S. "Content-aware convolutional neural network for in-loop filtering in High Efficiency Video Coding", IEEE Transactions on Image Processing, 28(7), pp. 3343–3356, 2019.
https://doi.org/10.1109/TIP.2019.2896489

[25] Dhanalakshmi, A., Nagarajan, G. "Convolutional Neural Network-based deblocking filter for SHVC in H.265", Signal, Image and Video Processing, 14(8), pp. 1635–1645, 2020.
https://doi.org/10.1007/s11760-020-01713-4

[26] Singhadia, A., Minhazuddin, M., Mamillapalli, M., Chakrabarti, I. "A fast integrated deblocking filter and sample-adaptive-offset parameter estimation architecture for HEVC", Microprocessors and Microsystems, 85, 104317, 2021.
https://doi.org/10.1016/j.micpro.2021.104317

[27] Baldev, S., Rathore, P. K., Peesapati, R., Anumandla, K. K. "A directional and scalable streaming deblocking filter hardware architecture for HEVC decoder", Microprocessors and Microsystems, 84, 104029, 2021.
https://doi.org/10.1016/j.micpro.2021.104029

[28] Dhanalakshmi, A., Nagarajan, G. "Combined spatial temporal based In-loop filter for scalable extension of HEVC", ICT Express, 6(4), pp. 306–311, 2020.
https://doi.org/10.1016/j.icte.2020.04.006

[29] Dhanalakshmi, A., Nagarajan, G. "Group-normalized deep CNN-based in-loop filter for HEVC scalable extension", Signal, Image and Video Processing, 16(2), pp. 437–445, 2022.
https://doi.org/10.1007/s11760-021-01966-7

[30] Aggarwal, A., Mittal, M., Battineni, G. "Generative adversarial network: An overview of theory and applications", International Journal of Information Management Data Insights, 1(1), 100004, 2021.
https://doi.org/10.1016/j.jjimei.2020.100004