IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 26, 2018, accepted August 29, 2018, date of publication September 13, 2018, date of current version October 17, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2869627

Detection of Double Compression for HEVC Videos With Fake Bitrate

XIAOYUN LIANG[1], ZHAOHONG LI[1], YIYUAN YANG[2], ZHENZHEN ZHANG[3], AND YU ZHANG[4], (Member, IEEE)

[1]School of Electronic Science and Technology, Beijing Jiaotong University, Beijing 100044, China
[2]School of Computer Science and Engineering, Beihang University, Beijing 100083, China
[3]Beijing Institute of Graphic Communication, Beijing 102600, China
[4]Qualcomm Research, Beijing 100013, China

Corresponding author: Zhaohong Li (zhhli2@bjtu.edu.cn)

This work was supported by the National Youth Fund under Grant 61702034.

ABSTRACT Bitrate is an important criterion for digital video quality. Forgers prefer to up-convert bitrate of videos without improving the video quality at all, especially for videos over the Internet. Therefore, proposing an effective algorithm to expose fake bitrate videos becomes an important issue in digital video forensics. In this paper, a novel method based on prediction unit (PU) partition types is proposed to detect high efficient video coding (HEVC) videos with fake bitrate. PU partition type is one unique syntactic unit of HEVC and could be reflected by compression history, while it has not been addressed in the topic of fake bitrate detection. The proposed method adopts support vector machine classifiers with feature vectors formed by histograms of PU partition types in the first P-frame in each group of pictures (GOPs). The performance of the algorithm is compared with state-of-the-art algorithms on videos of various resolutions and bitrates. The experimental results show that the proposed method can identify the fake bitrate videos with high accuracy, and is robust to frame-deletion, copy-paste, and shifted GOP structure attacks.

INDEX TERMS Video forensics, high efficiency video coding (HEVC), fake bitrate, prediction unit (PU).

I. INTRODUCTION

Tampered videos may lead to serious moral, ethical and legal consequences if they are considered as evidences. Currently, with the help of video conversion tools, such as AVC [1], AVS [2], and FFmpeg [3], falsifiers can up-convert the bitrate of a video without introducing any additional information about the video content in order to attract more users and gain benefits. In fact, the quality of these videos has not been improved. In this case, the claimed high bitrate video might actually have poor visual quality, and we call this high bitrate "fake bitrate". When the fake bitrate videos are abused, it will not only mislead the users, but also lead to a big waste of storage space. Therefore, it is necessary to develop identification techniques to detect the fake bitrate videos in the field of digital video forensics.

In the past, many effective methods have been proposed in digital video forensics while most of them are aimed at detecting tampered videos with encoding standard preceding HEVC. Such as Sun *et al.* [4] proposed an algorithm to detect MPEG double compression, Bian *et al.* [5] detected fake bitrate MPEG-2 videos and extended their algorithm to the new coding schemes H.264/AVC [6].

In 2013, the latest High Efficiency Video Encoding (HEVC) standard was ratified and published by both ITU-T and ISO/IEC [7], [8]. HEVC standard is really suitable for encoding High-Definition (HD) video that the falsifiers who want to be benefited from the HD video with fake-bitrate have already begun to take action. In the process of making a fake bitrate video, the encoded video is decompressed before up-converting the bitrate and is recompressed after, respectively. The re-encoding may be different from that in the original video, in terms of video coding format and/or video coding parameters. Hence, the fake bitrate video is compressed twice at least. Therefore, detecting whether the video has been recompressed is a key step in detecting fake bitrate video. However, very limited work has been reported on detecting HEVC recompressed videos [9]–[15]. For detecting double compressed HEVC video under different quantization parameter (QP) values, reference [9] analyzed changes in DCT coefficients caused by quantization. The optimal

Markov feature was constructed in [10] by averaging the difference between single and double recompressed videos, and the co-occurrence matrix of DCT coefficients was applied in [11]. Reference [12] contributed to the Multimedia Forensics mission with a method to detect double AVC/HEVC encoding under different QP. By analyzing the effects of QP on the distributions of DCT coefficients and TU size, reference [13] proposed a method to detect double compressed HEVC video with different QP. In order to detect the recompressed HEVC videos under different GOP size, reference [14] provided a new algorithm based on the Sequence of Number of Prediction Unit of its Prediction Mode (SN-PUPM). Furthermore, it can estimation the GOP size of the first compression.

For detecting double compressed HEVC video under different bitrate, our team previously [15] proposed a method based on combining the horizontal co-occurrence matrixes of DCT coefficients (PhoDCT) and the horizontal co-occurrence of PU types in I-frame (PhoPUTs). The DCT coefficient is a characteristic common to all encoders, and the PU type is a unique characteristic of HEVC. Reference [15] combines these two characteristics to obtain great results, but the results of only using the HEVC unique characteristic (PhoPUTs) are not really good (in the experimental section). Since the PU type of P-frame includes symmetric and asymmetric forms, there are a total of 25 types, while the PU type of I-frame has only 5 types, that is, the PU type of P-frame is more abundant, so this paper is concerned with the PU types of P-frame. At the same time, it is found that histogram statistics can effectively represent the distribution of PU types. Therefore, the simple histogram is used to count the number of PU types in P-frame. The experimental results show that the classification accuracy of the proposed algorithm is really high.

Besides, one major reason for people to recompress a video is to tamper its content, such as frame-deletion and copy-paste. If the proposed features for fake bitrate recompression detection do not have robustness against tampering, it lost the practical significance, while almost all the existed fake bitrate detection work did not consider this case. Besides, considering the structure of GOP may affect the algorithm, so the recompression detection methods need to be robust to the attack of shifted GOP structure. This paper presents an algorithm for detecting recompressed HEVC videos with fake bitrate and it is robust to frame-deletion, copy-past and shifted GOP structure attacks.

The main contribution of this paper is as follows. First, we propose an efficient method to identify recompressed videos with different fake bitrates. Second, it is the first time, to the best of our knowledge, that the unique characteristic of HEVC PU type in first P-frames is exploited in fake bitrate detection. Third, we consider the test of the robustness against frame-deletion, copy-paste and shifted GOP structure attacks. Experimental results show that the proposed method could distinguish fake bitrate videos from original ones and has good robustness against attacks.

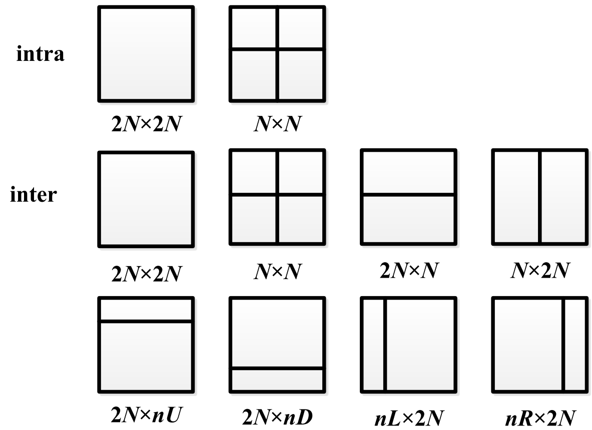

FIGURE 1. Partition of PU in intra and inter prediction.

The rest of the paper is organized as follows. We briefly review the basics of PU partitioning in HEVC compression in Section 2 and continue with the analysis of PU partitioning versus recompression bitrate of single and double compressed videos in Section 3. In Section 4, we show experimental results. Section 5 concludes the paper.

II. BASICS OF PU PARTITIONING IN HEVC

The video coding layer of HEVC employs the same hybrid approach (inter/intra picture prediction and 2-D transform coding) used in all video compression standards since H.261 [16]. The key design elements in HEVC provide a flexible partitioning framework to adapt coding units (CUs) to the content details. Based on quadtree structure, a coding tree unit (CTU) can be partitioned iteratively into four CUs. Each CU has an associated partitioning into PUs and a tree of transform units (TUs).

Different types of PU partitions can be established based on the prediction mode for the CU, subject to certain size constraints. The types of PU partitions in intra and inter prediction are shown in Fig. 1. A coding block (CB) of size 2N×2N can be symmetrically split into one or four prediction blocks (PBs) for intra picture prediction. In addition, a CB can be split into two symmetric or asymmetric PBs when inter picture prediction mode is established. The PB size of possible PU partitioning of inter picture prediction mode in HEVC is summarized in Table 1. In order to facilitate the histogram statistics for features later, index 1-25 are used to mark these 25 different PU partition types, respectively.

III. PU PARTITIONING IN HEVC VIDEOS

A. FEATURE SELECTION

In a GOP, the PU partitioning in I-frames and P-frames is different. This is because intra picture prediction is mainly established in I-frames, but inter picture prediction is only adopted for P-frames. Thus, there are only symmetric PBs in I-frames, while both symmetric and asymmetric PBs exist in P-frames. We can observe more diverse modes of splitting CBs into PBs in P-frames than in I-frames. When a video is

TABLE 1. PU partitioning types in HEVC.

ID	Size	ID	Size	ID	Size	ID	Size
1	4×4	8	16×12	15	32×24	22	64×48
2	8×8	9	16×4	16	32×8	23	64×16
3	4×8	10	12×16	17	24×32	24	48×64
4	8×4	11	4×16	18	8×32	25	16×64
5	16×16	12	32×32	19	64×64		
6	16×8	13	32×16	20	64×32		
7	8×16	14	16×32	21	32×64		

recompressed with a fake bitrate, we can expect more changes of PU partitioning in a P-frame than in an I-frame due to the diversity, the actual comparison data will be mentioned in the experimental section. Therefore, we take PU partitioning in P-frame as the classification feature to identify HEVC fake bitrate videos.

In the process of HEVC, quantization, reconstruction, rate control and loop filtering are indispensable steps, which inevitably introduce irreversible quantization errors and reconstruction errors, thus the encoded video will lose some of the details and the original video couldn't be fully restored. It means that the input of original video is different from the input of recompressed video. Besides, the rate control algorithm will control the quantization step and the PU division types at any time, hence, the PU partition type in recompressed video would be different from that in original video when the recompressed bitrate is different from the firstly compressed bitrate. However, this difference is only obvious when the recompressed rate is the fake bitrate, the specific reason is shown in the theoretical analysis of section B and the specific data description will be displayed in section C. Therefore, we suppose that the PU partition types of P-frame can be used as a feature to detect fake bitrate recompressed videos from original videos.

B. THEORETICAL ANALYSIS AND MODELING

During the encoding of a video sequence $\boldsymbol{V}$ with HEVC at a bitrate of r, the PU partitioning in each P-frame is determined as follows. For the nth P-frame, the amount of bits allocated by the rate control process $\rho(\cdot)$can be written as $b_n^{(r)} = \rho(n; \boldsymbol{V}; r)$. The PU partitioning process $\pi(\cdot)$ selects splitting of CBs into PBs. For each P-frame, there are 25 different PU partitioning types. Denote $p_{n,k}^{(r)}$ as the number of PBs in the kth PU partitioning type in the nth P-frame, where $k = 1, 2, \ldots, 25$. For brevity, let $\boldsymbol{p}_n^{(r)} = (\boldsymbol{p}_{n,1}^{(r)}, \boldsymbol{p}_{n,2}^{(r)}, \ldots, \boldsymbol{p}_{n,25}^{(r)})^T$. The PU partitioning can thus be represented as $\boldsymbol{p}_n^{(r)} = \pi(\boldsymbol{F}_n; b_n^{(r)})$, where $\boldsymbol{F}_n$ is the nth P-frame.

Consider the following video forgery process. A raw video $\boldsymbol{V}$, consists of a series of image frames, expressed as (1), is encoded to an HEVC video $\boldsymbol{H}_1$ at a bitrate r_1. A falsifier first decodes the compressed video $\boldsymbol{H}_1$ to a YUV sequence $\hat{\boldsymbol{V}}$, expressed as (2), and then encodes $\hat{\boldsymbol{V}}$ to a recompressed HEVC video $\boldsymbol{H}_2$ at a bitrate r_2.

$$\boldsymbol{V} = \{\boldsymbol{F}_n | n = 1, 2, \ldots, \mathrm{N}\} \tag{1}$$

$$\hat{\boldsymbol{V}} = \{\hat{\boldsymbol{F}}_n | n = 1, 2, \ldots, \mathrm{N}\} \tag{2}$$

Notice that the video encoding process can be modeled by $[\mathrm{DCT}(\boldsymbol{F}_n - \boldsymbol{C}_n)/Q_p]$, where $\mathrm{DCT}(\cdot)$ stands for the discrete cosine transform (DCT) in the video frame coding process and $\mathrm{IDCT}(\cdot)$ is its inverse transformation, $[\cdot]$ is the rounding operator, $\boldsymbol{C}_n$ denotes the reference frame of $\boldsymbol{F}_n$, and Q_p is the quantization step size. The decoding process can be written as

$$\begin{aligned}\hat{\boldsymbol{F}}_n &= \mathrm{IDCT}([\mathrm{DCT}(\boldsymbol{F}_n - \boldsymbol{C}_n)/Q_p] \times Q_p) + \boldsymbol{C}_n \\ &\approx \mathrm{IDCT}([\mathrm{DCT}(\boldsymbol{F}_n)/Q_p] \times Q_p) \\ &\quad - \mathrm{IDCT}([\mathrm{DCT}(\boldsymbol{C}_n)/Q_p] \times Q_p) + \boldsymbol{C}_n \\ &= \boldsymbol{F}_n + \boldsymbol{E}(\boldsymbol{F}_n) - \boldsymbol{E}(\boldsymbol{C}_n\end{aligned} \tag{3}$$

where $\hat{\boldsymbol{F}}_n$ is the decompressed frame of $\boldsymbol{F}_n$, $\boldsymbol{E}(\boldsymbol{F}_n)$ and $\boldsymbol{E}(\boldsymbol{C}_n)$ are the quantization error of $\boldsymbol{F}_n$ and $\boldsymbol{C}_n$ under the given quantization step Q_p, respectively. In the second compression, the amount of bits allocated for the nth P-frame is

$$b_n^{(r_1,r_2)} = \rho(n; \hat{\boldsymbol{V}}, r_2) \tag{4}$$

The PU partitioning is

$$\boldsymbol{p}_n^{(r_1,r_2)} = \pi(\hat{\boldsymbol{F}}_n; b_n^{(r_1,r_2)}) \tag{5}$$

Denote $b_n^{(r_2)}$ as the amount of bits allocated for the nth P-frame in the encoding of $\boldsymbol{V}$ with a bitrate of r_2, and let $\boldsymbol{p}_n^{(r_2)}$ be the PU partitioning. We have

$$b_n^{(r_2)} = \rho(n; \boldsymbol{V}, r_2) \tag{6}$$

$$\boldsymbol{p}_n^{(r_2)} = \pi(\boldsymbol{F}_n; b_n^{(r_2)}) \tag{7}$$

Using $\mathrm{D}(\cdot)$ to represent the difference between $\boldsymbol{p}_n^{(r_1,r_2)}$ and $\boldsymbol{p}_n^{(r_2)}$. Based on the above formulas, we can see that the difference between the PU partition types of recompressed video and single compressed video is

$$\begin{aligned}&D(\boldsymbol{p}_n^{(r_1,r_2)}, \boldsymbol{p}_n^{(r_2)}) \\ &\quad = D(\pi(\hat{\boldsymbol{F}}_n; b_n^{(r_1,r_2)}), \pi(\boldsymbol{F}_n; b_n^{(r_2)})) \\ &\quad = D(\pi(\hat{\boldsymbol{F}}_n; \rho(n; \hat{\boldsymbol{V}}, r_2)), \pi(\boldsymbol{F}_n; \rho(n; \boldsymbol{V}, r_2))) \\ &\quad = D(\pi(\hat{\boldsymbol{F}}_n; \rho(n; \{\hat{\boldsymbol{F}}_n | n = 1, 2, \ldots, \mathrm{N}\}, r_2)), \\ &\qquad\qquad \pi(\boldsymbol{F}_n; \rho(n; \{\boldsymbol{F}_n | n = 1, 2, \ldots, \mathrm{N}\}, r_2)))\end{aligned} \tag{8}$$

It indicates that the main factor causing the difference between $\boldsymbol{p}_n^{(r_1,r_2)}$ and $\boldsymbol{p}_n^{(r_2)}$ is their input $\hat{\boldsymbol{F}}_n$ and $\boldsymbol{F}_n$. From (3) we can get the quantization error $\hat{\boldsymbol{F}}_n - \boldsymbol{F}_n \approx \boldsymbol{E}(\boldsymbol{F}_n) - \boldsymbol{E}(\boldsymbol{C}_n)$, that is to say, the difference between $\hat{\boldsymbol{F}}_n$ and $\boldsymbol{F}_n$ is mainly caused by the quantization error. Simultaneously, we know that in the video encoding process, the smaller bitrate, the bigger the quantization step Q_p and the bigger the quantization error, which means bigger difference between $\boldsymbol{p}_n^{(r_1,r_2)}$ and $\boldsymbol{p}_n^{(r_2)}$. Therefore, if $r_2$is fixed, when r_1 is much smaller than r_2 (the fake bitrate case), the bigger the difference between $\boldsymbol{p}_n^{(r_1,r_2)}$

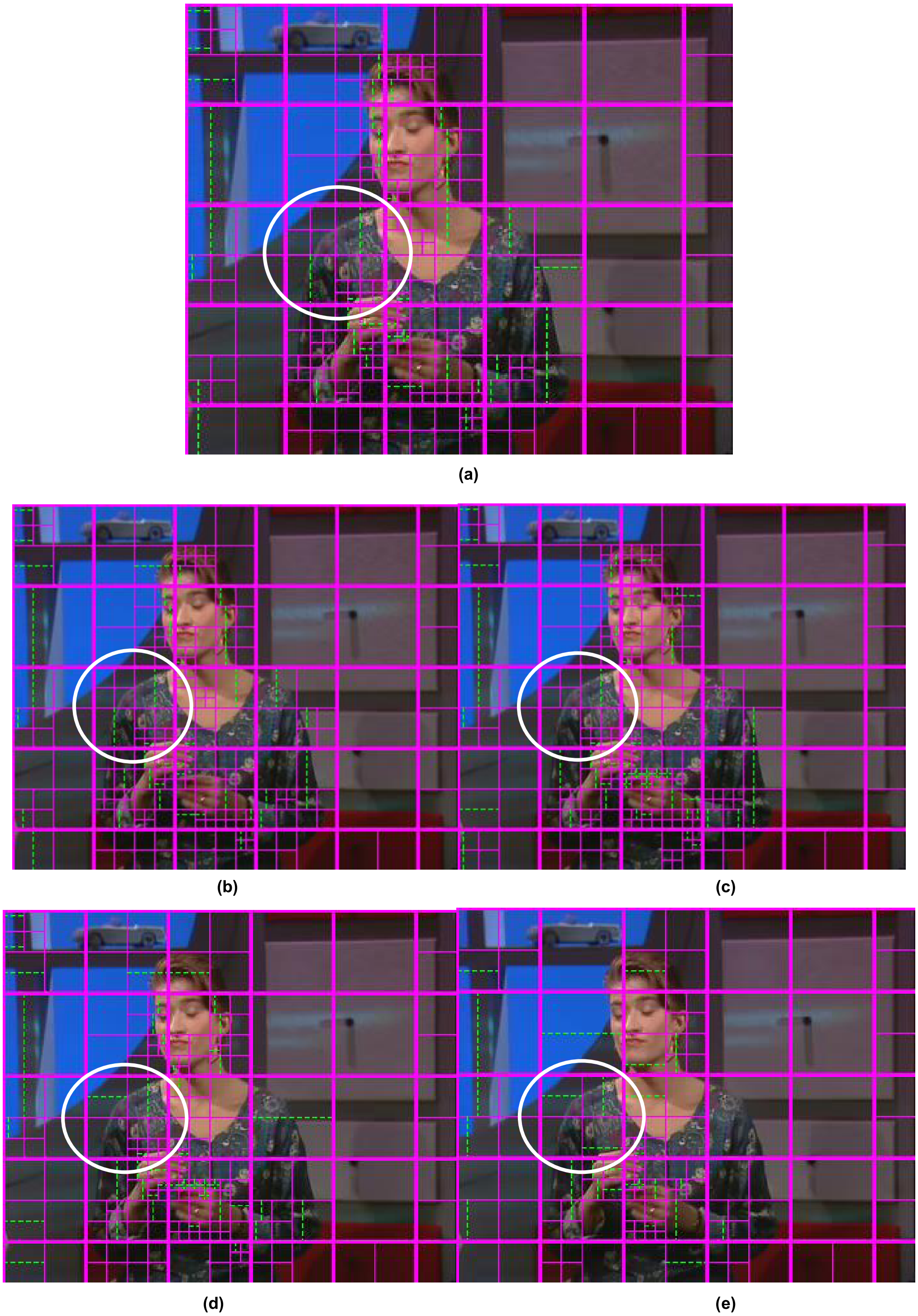

FIGURE 2. PU types of first-P-frame in video sign_irene_cif.yuv before and after double compression. (a) PU types with 300k bitrate. (b) PU types with 300k_300k bitrate. (c) PU types with 500k_300k bitrate. (d) PU types with 200k_300k bitrate. (e) PU types with 100k_300k bitrate.

and $\boldsymbol{p}_n^{(r_2)}$ is. When $r_1 \geq r_2$, the quantization error is really small, $\boldsymbol{p}_n^{(r_1,r_2)}$ and $\boldsymbol{p}_n^{(r_2)}$ are very similar. In such case, it's difficult to identify the recompressed video from original ones.

C. FEATURE ANALYSIS AND EXAMPLE DESCRIPTION

Take the single compressed video with bitrate $r_2 = 300$kbps and the double compressed video with first bitrate $r_1 = 100kbps, 300kbps, 500kbps$ and second bitrate

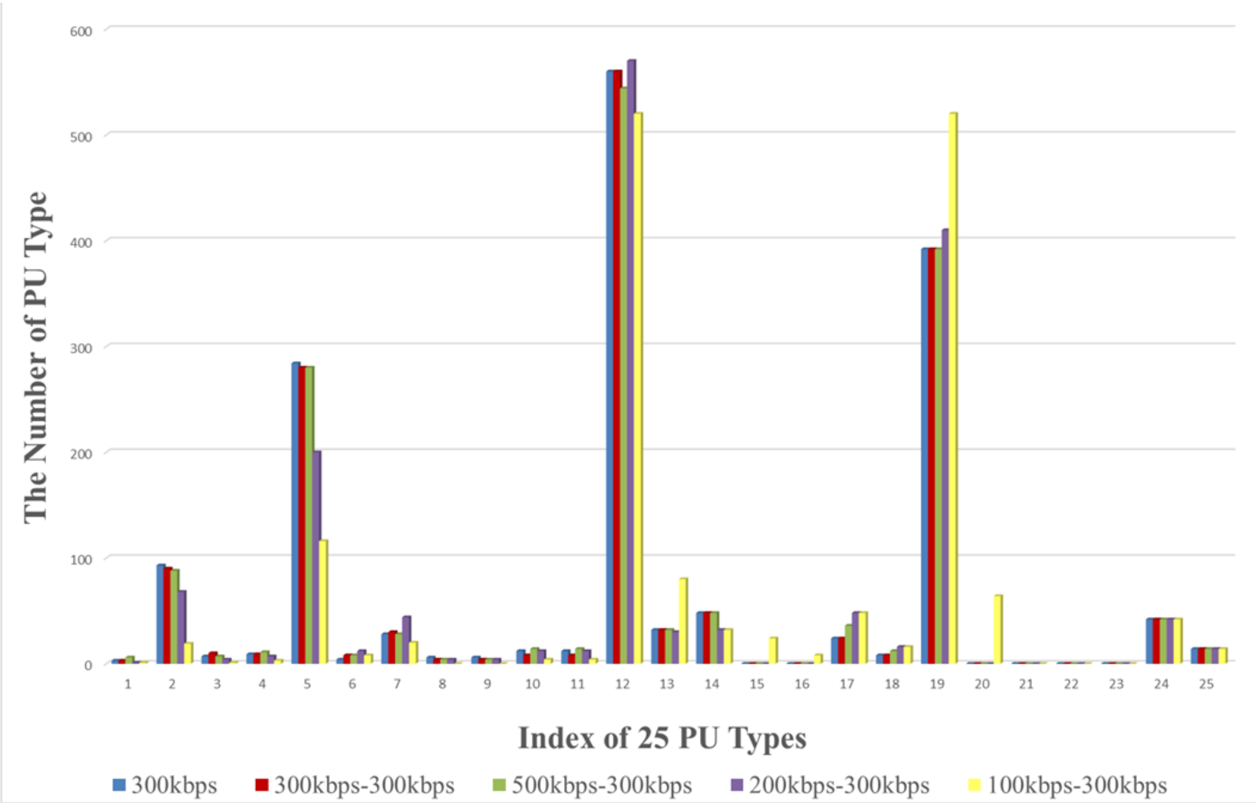

FIGURE 3. The number of PU types for Fig. 2 (a-e).

$r_2 = 300$kbps respectively as examples. Fig. 2 shows the PU partition types of the same first P-frame in sign_irene_cif.yuv single compressed video and its corresponding recompressed video. Take a 64×64 block surrounded by the white circle in Fig. 2 as an example. In single compressed video with bitrate $r_2 = 300$kbps (Fig. 2 (a)), the number of PU partition types is {16×32: 4, 16×16: 6, 8×8: 6, 8×4: 4}. When $r_1 \geq r_2$, the number of PU partition types $p_n^{(r_1,r_2)}$ in recompressed video with bitrate $r_1 = 300$kbps, $r_2 = 300$kbps (Fig. 2 (b)) and the recompressed video with bitrate $r_1 = 500$kbps, $r_2 = 300$kbps (Fig. 2 (c)) are {16×32: 4, 16×16: 6, 8×8: 6, 8×4: 4} and {16×32: 2, 32×32: 1, 16×16: 6, 8×8: 6, 8×4: 4} respectively, obviously they are really similar to that of single compressed video. For the fake bitrate videos with bitrate $r_1 = 200$kbps, $r_2 = 300$kbps (Fig. 2 (d)), the number of PU partition types are {32×32: 1, 16×32: 2, 32×16: 2, 16×16: 2, 8×8: 7, 8×4: 2} which is quite different from the single compressed video. Furthermore, it is much bigger difference between the number of PU partition types of the single compressed video and the fake bitrate videos with bitrate $r_1 = 100$kbps, $r_2 = 300$kbps (Fig. 2 (e)) {32×32: 1, 16×32: 2, 32×16: 2, 32×24: 1, 32×8: 1}. This phenomenon fully conforms to the above theoretical analysis

In order to visualize the change of PU type in the same frame between single and recompressed video at different bitrate, we make a histogram of each PU type in Fig. 2 (a-e) and show them in Figure 3. The histogram illustrates when r_2 is fixed ($r_2 = 300$kbps), the greater r_1 is, the smaller the difference of the number of each PU type between double compressed frame with bitrate $r_1\_r_2$ and single compressed frame with bitrate r_2 is. When r_1 is going up near to r_2, the difference is so small and can be ignored. This is why when $r_1 \geq r_2$, most algorithms, including [9]–[15], cannot detect the recompressed videos, and so do our algorithm. In this case, it requires specialized algorithms for detecting recompressed videos with the same coding parameters. For example, Chen *et al.* [17] introduced a novel method based on the statistical feature of macroblock mode (MBM) which consists of macroblock type and motion vector in P-frames to identify recompressed MPEG videos with the same QS, [18] concatenated the statistical features of rounding and truncation errors from the intra-coding process and the macroblock-mode based features from the inter-coding process to detect double MPEG compression with the same coding parameters, and experiments are performed with the same bitrate, the same QS, and the same GOP size. Under the same QPs, Jia *et al.* [19] directed to the detection of recompressed HEVC video based on the partitioning of prediction units. Aghamaleki *et al.* [20] proposed a method to detect double compressed MPEG videos with the same quantization matrix and synchronized GOP structure. In practice, the falsifier generally up-convert the bitrate of video to attract

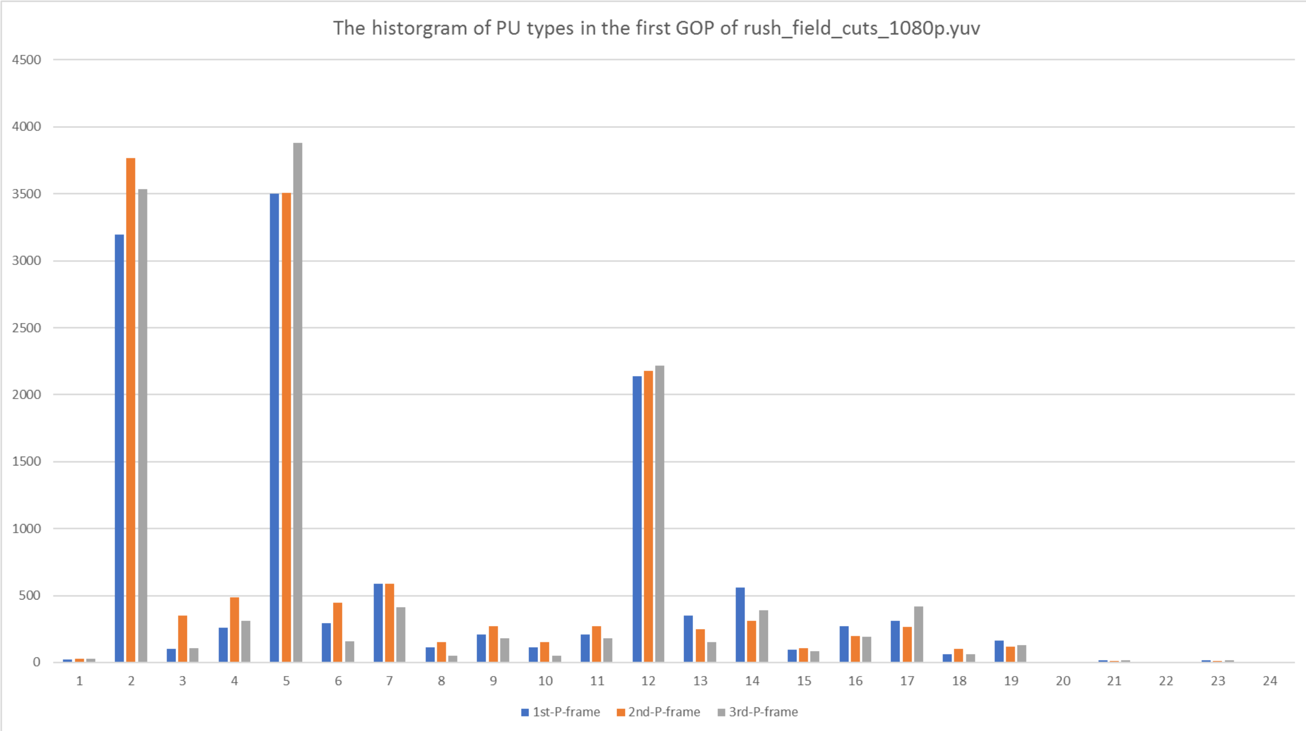

FIGURE 4. **The historgram of PU types in the first GOP of rush_field_cuts_1080p.yuv.**

more audience, and do not down-convert the bitrate of video, so in this paper the proposed algorithm is only focus on detecting recompressed videos with fake bitrate.

In general, Fig. 2-3 are completely consistent with the above theoretical analysis. Therefore, it can be determined that the proposed feature is valid for detecting recompressed HEVC video with fake bitrate.

D. THE PROPOSED TARGET FEATURE

Fig. 4 shows the histrograms of PU types of all P-frames in the first GOP of single compressed video rush_field_cuts_1080p.yuv with 8Mbps bitrate and GOP size of four. It indicates that the histrograms of PU types of consecutive P-frames are similar to each other. Considering the feature dimension, we only take the histogram of PU types in the first P-frame in each GOP as our final feature. Based on the analysis in subsection B, we propose to construct an SVM classifier using the histogram of PU partitioning (HPP) features to identify HEVC fake bitrate videos. The PU partitioning of the first P-frame in each GOP is obtained by an HEVC bitstream analysis tool [21]. Notice that the PU partitioning is established hierarchically based on the quadtree structure. To better highlight the difference between single and double compressed video in PU partition types, we flatten the multi-level structure of PU splitting. We divide each PB into one or more 8×8 basic units. Each basic unit is labeled by the ID associated with the parent PB type as defined in Table 1. In Fig. 5, we illustrate the labeling by an example. The 32×32 CU partitioned as shown in Fig. 5(a) is represented by the 8×8 matrix in Fig. 5(b), where each entry in the matrix is the label of the corresponding 8×8 basic unit. The 25-dimensional HPP feature, $\boldsymbol{p}_n$, is extracted from the first P-frame in the nth GOP of a video after flattening and labeling. For a video with N GOPs, the averaged HPP feature $\bar{\boldsymbol{p}} = \frac{1}{N}(\sum_{n=1}^{N}\boldsymbol{p}_n)$ is employed as the target classification feature.

We use three forms (histogram, line chart, graph) to represent the classification feature HPP of three groups of videos respectively, each group contains a single video and a corresponding recompressed video with fake bitrate. For convenience, we just take HPP with 5 typical PU types (64×64; 32×32; 16×16; 8×8; 4×4) for example in Fig. 6. Clearly to see that the HPP between original and recompressed fake bitrate video is so different that can be taken for a classification feature to distinguish them.

IV. EXPERIMENTAL RESULTS

Four kind of resolutions are adopted to construct the video sets: QCIF, CIF, 720p, and 1080p, the corresponding yuv sequences are: QCIF (akiyo, bridge_close, carphone, claire, coastguard, container, foreman, grandma, hall, highway, miss_america, mobile, mother_daughter, news, salesman, silent, suzie), CIF (bowing, bridge_far, bus, city, crew, deadline, flower, football, harbour, husky, ice, pamphlet, paris, sign_irene, soccer, tempete, waterfall), 720p (mobcal_ter, parkrun_ter, shields_ter, stockholm_ter, ducks_take_off, FourPeople, Johnny, KristenAndSara,

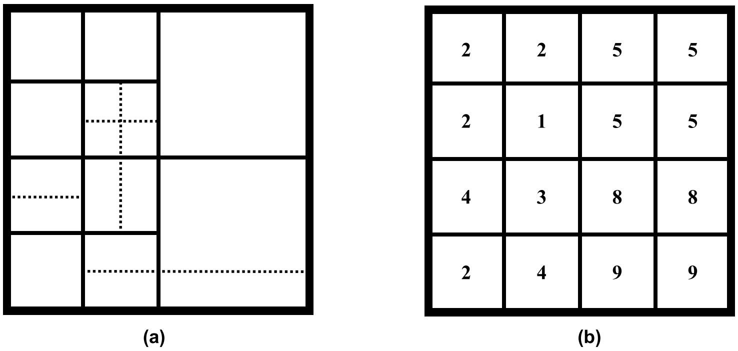

FIGURE 5. **Illustration of labeling PU Partitioning. (a) 32×32 CU. (b) The labels of 8×8 units.**

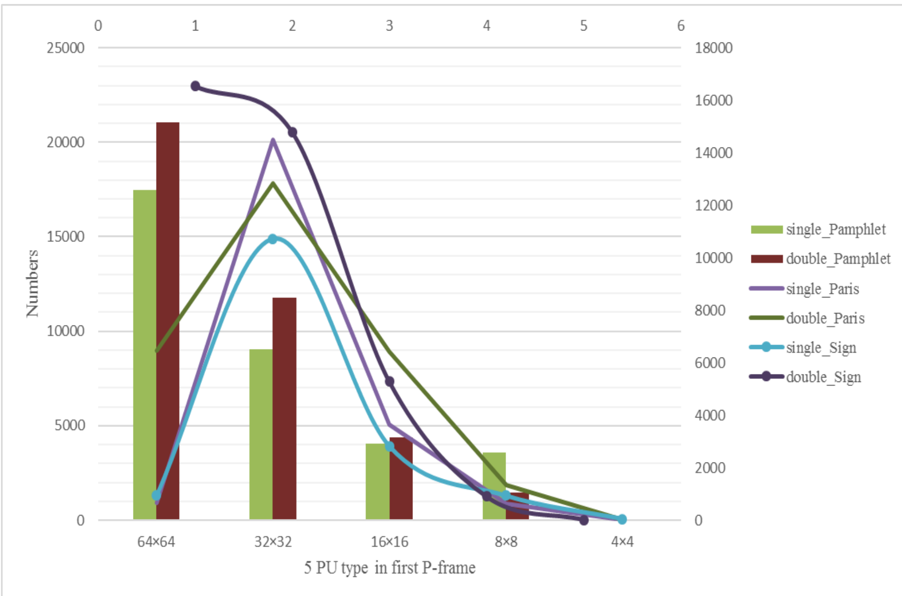

FIGURE 6. **The number of 5 PU types in first P-frame for their single and three corresponding recompressed videos.**

park_joy, sintel_trailer, vidyo) and 1080p (controlled_burn, ducks_take_off, life, rush_field_cuts, BasketballDrive, BQTerrace, Cactus, beach, Kimono, ParkScene, taishan). In order to increase the size of video bases, one yuv sequence is cut into several sequences with 100 frames without overlap. Finally, 36 QCIF videos, 43 CIF videos, 36 720p videos and 32 1080p videos are obtained. Furthermore, the pixel format for these videos is set as YUV420P. The GOP structure of single and double compressed videos is IPPP, and the rate control is enabled. The robustness of the method to frame-deletion and copy-paste attacks are tested on QCIF, CIF and 720P video bases. In order to verify the efficiency of our proposed method on recompression with shifted GOP structure, we recompressed 1080p videos with GOP size 8 (IPPPPPPP). To guarantee video quality, the values of compression bitrates r_1 and r_2 of QCIF and CIF video base are selected from {100k, 200k, 300k} (bps) and {200K, 300k, 400k} (bps), respectively, and those for 720p video base are selected from {10M, 20M, 30M} (bps) and {20M, 30M, 40M} (bps), respectively, and those for 1080p video base are selected from {8M, 10M, 20M} (bps) and {10M, 20M, 50M} (bps), respectively.

The ratio of videos in training and test sets is set to be 5:1 for each video base. Videos are randomly assigned to training or test set. For the LIBSVM [22] classifier, PolySVC is chosen as the kernel function, Gamma=0.5 and Cost=1. For all recompressed videos, including fake bitrate videos and simple recompression videos, when $r_1 \geq r_2$, the detection accuracy is calculated according to the (9). The entire training and testing procedures are repeated for 20 times and the average AR is treated as the final detection accuracy.

$$\mathrm{AR} = (\mathrm{TNR} + \mathrm{TPR})/2 \tag{9}$$

The classification accuracy AR is defined as the arithmetic mean of the true negative rate (TNR), TNR = TN/(TN + FP),

TABLE 2. Classification accuracy of proposed method in QCIF, CIF, 720P, and 1080P video sets (in percentage). (a) QCIF video set. (b) CIF video set. (c) 720P video set. (d) 1080P video set.

(a)

$r_1 \setminus r_2$ (bps)	200k	300k	400k
100k	**100**	**100**	**100**
200k	58.3	**91.7**	**100**
300k	50.0	50.0	**91.7**

(b)

$r_1 \setminus r_2$ (bps)	200k	300k	400k
100k	**92.9**	**100**	**100**
200k	50.0	**92.9**	**92.9**
300k	50.0	50.0	**92.9**

(c)

$r_1 \setminus r_2$ (bps)	20M	30M	40M
10M	**91.7**	**100**	**100**
20M	58.3	**91.7**	**100**
30M	75.0	**75.0**	**91.7**

(d)

$r_1 \setminus r_2$ (bps)	10M	20M	50M
8M	**91.5**	**100**	**100**
10M	72.0	**93.0**	**100**
20M	81.0	**73.0**	**99.0**

TABLE 3. Classification accuracy comparison: [11], [14], and [15] versus the proposed method (in percentage, QCIF video set).

$r_1 \setminus r_2$ (bps)	[11]	[14]	[15]	HPP	[11]	[14]	[15]	HPP	[11]	[14]	[15]	HPP
	200k				300k				400k			
100k	**88.6**	**66.7**	**100**	**100**	**92.2**	**66.7**	**100**	**100**	**92.5**	**64.4**	**100**	**100**
200k	50.0	50.0	50.6	58.3	**85.3**	**66.3**	**98.8**	**91.7**	**90.8**	**68.8**	**100**	**100**
300k	50.6	50.3	50.0	50.0	48.9	50.0	50.0	50.3	**76.7**	**75.62**	**96.7**	**91.7**

TABLE 4. Classification accuracy comparison: I-frame versus first P-frames (in percentage, QCIF video set).

$r_1 \setminus r_2$ (bps)	I	P
100k \ 200k	**91.7**	**100**
100k \ 300k	**91.7**	**100**
100k \ 400k	**83.3**	**100**
200k \ 300k	**75.0**	**91.7**
200k \ 400k	**75.0**	**100**
300k \ 400k	**75.0**	**91.7**

TABLE 5. Classification accuracy comparison: PhoPUTs of [15] versus the proposed method (in percentage, QCIF video set).

$r_1 \setminus r_2$ (bps)	PhoPUTs	HPP
100k \ 200k	**85.4**	**100**
100k \ 300k	**91.7**	**100**
100k \ 400k	**90.8**	**100**
200k \ 300k	**75.8**	**91.7**
200k \ 400k	**83.3**	**100**
300k \ 400k	**59.6**	**91.7**

and the true positive rate (TPR), TPR = TP/(TP+FN), where TN is the number of videos which are original videos and identified as original videos, FP is the number of videos which are original videos and identified as recompressed videos, TP is the number of videos which are recompressed videos and identified as recompressed videos, FN is the number of videos which are recompressed videos and identified as original videos.

TABLE 6. Classification accuracy comparison: LBP, Markov versus the proposed method (in percentage, QCIF video set).

$r_1 \setminus r_2$ (bps)	LBP	Markov	HPP
100k \ 200k	**90.5**	**94.6**	**100**
100k \ 300k	**91.7**	**95.3**	**100**
100k \ 400k	**91.7**	**97.8**	**100**
200k \ 300k	**75.8**	**79.4**	**91.7**
200k \ 400k	**83.3**	**89.2**	**100**
300k \ 400k	**82.0**	**79.0**	**91.7**

A. FAKE BITRATE DETECTION

Tables 2(a)–(d) report the classification accuracy of the proposed method (HPP based on first P-frames) for identifying recompressed videos from original ones in QCIF, CIF, 720p and 1080p video sets, respectively. The experimental results of the **bold** font in tables are the accuracy of the fake bitrate detection. In the fake bitrate case, the proposed method provides accuracy higher than 90% for detecting recompressed fake-bitrate videos from original videos with various resolutions. When $r_1 \geq r_2$, the classification accuracy drops to around 50% for QCIF and CIF videos. It's worth noting that such phenomenon exists in the most published detection algorithms including the early literatures [9]–[15] and it is in accordance with the theoretical analysis of section III.

Regarding the digital video forensics of HEVC, some algorithms are specific to the same quantization parameter, e.g. references [17]–[20], some algorithms like references [9]–[13] are proposed for different quantization

TABLE 7. Classification accuracy of proposed method in QCIF, CIF, and 720P frame-deletion video sets (in percentage). (a) QCIF frame-deletion video set. (b) CIF frame-deletion video set. (c) 720p frame-deletion video set.

(a)

$r_1 \setminus r_2$ (bps)	200k	300k	400k
100k	**100**	**100**	**100**
200k	83.3	**100**	**100**
300k	83.3	91.7	**100**

(b)

$r_1 \setminus r_2$ (bps)	200k	300k	400k
100k	**92.9**	**92.9**	**100**
200k	85.7	**92.9**	**100**
300k	78.6	85.7	**100**

(c)

$r_1 \setminus r_2$ (bps)	20M	30M	40M
10M	**100**	**100**	**100**
20M	83.3	**100**	**100**
30M	83.3	83.3	**91.7**

parameter and reference [14] is proposed for different GOP, and reference [15] is proposed for different bitrate. In order to better demonstrate the superiority of the proposed algorithm, we compare it with references [11], [14], and [15]. As reference [15] is our previous work, we use the same video set, so the experimental results can be compared directly. For the references [11] and [14], we apply their algorithms to our video sets for comparison. Firstly, using their code to complete feature extraction. The training and testing module are performed using the above demonstrated SVM classifier. In the QCIF video set, we compare the proposed method with the references [11], [14], and [15] respectively, as shown in Table 3. It can be seen that the fake-bitrate detection accuracy of the proposed algorithm is much higher than that of reference [11] and [14], while the accuracy is not as good as that of algorithm [15] in the cases when the bitrates are 200kbps_300kbps and 300kbps_400kbps, but in other cases, the fake-bitrate detection accuracy are all 100%. Therefore, we can conclude that the proposed algorithm has high accuracy for HEVC fake bitrate detection.

In subsection III.A, we know that PU type in P-frame is more abundant than I-frame thus we suppose that applying PU type in P-frame as distinguishing features would be more effective than I-frame when detecting fake bitrate videos. In order to verify this hypothesis, we use the proposed algorithm to test the classification accuracy with the PU type feature in I-frame, and we can get a set of comparative test results shown in Table 4. We can see that the classification accuracy of the PU types feature from first P-frame (bold red font in Table 4) is much higher than that of the PU types feature from I-frame in the fake bitrate case. It demonstrates that the HPP features based on first P-frames are preferred over those extracted from I-frames.

The proposed HPP feature and PhoPUTs feature of reference [15] are both extracted from the HEVC unique characteristic. For demonstrating the superiority of the proposed feature, we compare the classification accuracy of the HPP feature and PhoPUTs feature, as shown in the Table 5. It shows that the HPP features are more effective than the PhoPUTs features of reference [15].

Moreover, we do three sets of experiments to show that the histogram statistics can represent the traces of recompression left on the PU types more effectively. Table 6 shows the classification accuracy of the three sets of feature representation methods on the QCIF video set. The features of the three methods are extracted from the PU types of the first P-frame. HPP means the histogram of the PU types, LBP refers to its LBP operator, and Markov refers to its Markov transition probability matrix. It is obviously that HPP has the highest classification accuracy. That is to say, histogram can represent the characteristic of PU types in first P-frame more effectively.

TABLE 8. Classification accuracy comparison: [15] versus proposed method (in percentage, QCIF frame-deletion video set).

$r_1 \setminus r_2$ (bps)	[15]	HPP
100k \ 200k	**97.1**	**100**
100k \ 300k	**97.5**	**100**
100k \ 400k	**98.3**	**100**
200k \ 300k	**89.2**	**100**
200k \ 400k	**93.3**	**100**
300k \ 400k	**77.1**	**100**

In order to fully evaluate the robustness of the proposed algorithm which is ignored in most papers, we performed recompression detection experiments on frame-deletion, intra-frame copy-paste and shifted GOP structure video sets, respectively. The following three sections will specifically analyze the results of these three robustness experiments.

B. ROBUSTNESS TO FRAME-DELETION

To test the robustness of the proposed method to frame-deletion, we construct a frame-deletion video set for each of the QCIF, CIF and 720p video sets as follows. We decompress the videos compressed at a bitrate of r_1, delete frames 30–59 from each decompressed video, and then recompress them at a bitrate of r_2. Tables 7(a)–(c) summarize the classification accuracy of the recompressed videos with frame-deletion in various resolutions. For fake bitrate recompressed videos, the identification accuracy has been improved. Furthermore when $r_1 \geq r_2$, under the condition of which most of the algorithms are failed, the classification accuracy is above 78%. The results show that the proposed method can provide higher classification accuracy for identifying recompressed frame-deleted videos than detecting recompressed videos without tampering. In subsection II.B, we can get the difference between PU partitioning of single and double compressed videos is

$$D(\boldsymbol{p}_n^{(r_1,r_2)}, \boldsymbol{p}_n^{(r_2)}) = D(\pi(\hat{\boldsymbol{F}}_n; \rho(n; \{\hat{\boldsymbol{F}}_n | n = 1, 2, \ldots, \mathrm{N}\}, r_2)), \pi(\boldsymbol{F}_n; \rho(n; \{\boldsymbol{F}_n | n = 1, 2, \ldots, \mathrm{N}\}, r_2)))$$

TABLE 9. Classification accuracy of proposed method in QCIF, CIF, and 720P copy-paste video sets (in percentage). (a) QCIF copy-paste video set. (b) CIF copy-paste video set. (c) 720P copy-paste video set.

(a)

$r_1 \setminus r_2$ (bps)	200k	300k	400k
100k	**91.7**	**100**	**100**
200k	83.3	**100**	**100**
300k	75.0	83.3	**91.7**

(b)

$r_1 \setminus r_2$ (bps)	200k	300k	400k
100k	**92.9**	**100**	**100**
200k	78.6	**92.9**	**92.9**
300k	78.6	78.6	**92.9**

(c)

$r_1 \setminus r_2$ (bps)	20M	30M	40M
10M	**100**	**100**	**100**
20M	83.3	**100**	**100**
30M	75.0	75.0	**100**

Due to the frame-deletion operation, not only the difference between $\boldsymbol{F}_n$ and $\hat{\boldsymbol{F}}_n$ is increased, but also the rate control process $\rho(\cdot)$ is affected, thereby increasing the difference of PU partition types between frame deleted video and single compressed video Thus, the recompressed frame-deletion video exhibits more significant changes in terms of HPP feature compared to the recompressed video without frame-deletion.

Reference [15] shows its experimental results on the QCIF frame-deletion video set. The comparison results of the proposed method and reference [15] are listed in Table 8. It can be seen that the robustness on frame-deletion of the proposed method is much better than that of reference [15].

C. ROBUSTNESS TO COPY-PASTE TAMPERING

We also test the robustness of the proposed method to copy-paste tampering. A copy-paste video set is constructed for each of QCIF, CIF and 720p video sets as follows. We copy a region of the first frame in a decompressed video and paste it to frames 30–59, and then recompress them at a bitrate of r_2. The copied region size is 30% of area in the first frame. Tables 9(a)–(c) list the classification accuracy for various resolutions. We observe that the proposed method can identify recompressed videos under copy-paste attacking at a higher accuracy than detect recompressed videos without tampering in all cases. This high accuracy is due to the fact that both the change of the frame content and the recompression impact the PU partitioning as illustrated in the discussion of robustness to frame-deletion.

D. ROBUSTNESS TO SHIFTED GOP STRUCTURE

Reference [14] is a recompression detection algorithm specifically for recompression with different GOP structures. In order to facilitate comparison with it, the yuv sequences used here are the same as described above, and the resolution of shifted GOP structure video set is 1920*1080 (1080p) which is same as the reference [14]. The encoding parameters of the 1080p single-compressed video are: GOP structure IPPP, and bitrate r_1. The video set of the shifted GOP structure is obtained with decompressing single-compressed 1080p videos and recompressing them by the parameters: GOP structure IPPPPPPP, and recompressed bitrate r_2. The accuracy of the fake bitrate detection of the proposed method and the reference [14] in unshifted and shifted GOP structure video sets are shown in Table 10. It is obviously that when the recompressed GOP structure is unshifted, the accuracy of the proposed method is much higher than that of reference [14]. Though in the shifted GOP structure video set, the accuracy of the proposed method is not as good as reference [14], but they are all above 93%. Hence the proposed method is also effective to resist the shifted GOP structure attack.

TABLE 10. Classification accuracy comparison: [14] versus proposed method in 1080P unshifted and shifted GOP structure video sets (in percentage).

$r_1 \setminus r_2$ (bps)	[14]	HPP	[14]	HPP
	Unshifted GOP structure		Shifted GOP structure	
8M \ 10M	**74.5**	**91.5**	**100**	**93.0**
8M \ 20M	**82.5**	**100**	**99.5**	**94.5**
8M \ 50M	**81.0**	**100**	**99.0**	**94.0**
10M \ 20M	**83.0**	**93.0**	**100**	**99.5**
10M \ 50M	**81.5**	**100**	**100**	**96.0**
20M \ 50M	**78.0**	**99.0**	**99.0**	**96.5**

V. CONCLUSION

In this paper, we propose a novel method to detect HEVC re-compressed videos with fake bitrate. Exploiting the inherent flexible partitioning framework in HEVC, we develop a feature for SVM classifier, based on the histogram of the PU partitioning types in the first P-frames. Experimental results show that the proposed method has high classification accuracy in detecting fake bitrate videos with various resolutions. In addition, the proposed method has good robustness against frame-deletion, copy-paste and shifted GOP structure attacks. In future work, we will establish a large-scale dataset with high-resolution videos and fully study other HEVC characteristic, such as CU, TU and intra prediction mode, etc. Besides, we will also consider the impact of different versions of HEVC encoders on the proposed algorithm.

REFERENCES

[1] *AVC Video Converter*. Accessed: Jan. 2014. [Online]. Available: http://www.avcware.com/

[2] *AVS Video Converter*. Accessed: Jan. 2014. [Online]. Available: http://www.avs4you.com/AVS-Video-Converter.aspx

[3] FFmpeg. Accessed: Jan. 2014. [Online]. Available: http://www.ffmpeg.org/download.html

[4] T. Sun, W. Wang, and X. Jiang, "Exposing video forgeries by detecting MPEG double compression," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2012, pp. 1389–1392.

[5] S. Bian, W. Luo, and J. Huang, "Exposing fake bitrate video and its original bitrate," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 4492–4496.

[6] S. Bian, W. Luo, and J. Huang, "Exposing fake bit rate videos and estimating original bit rates," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 12, pp. 2144–2154, Dec. 2014.

[7] *High Efficiency Video Coding*, document Rec. ITU-T H.265, Dec. 2016.

[8] *Information Technology—High Efficiency Video Coding and Media Delivery in Heterogeneous Environments—Part 2: High Efficiency Video Coding*, document ISO/IEC 23008-2:2017, Oct. 2017.

[9] M. Huang, R. Wang, J. Xu, Q. Li, and D. Xu, "Detection of double compression for HEVC videos based on the co-occurrence matrix of DCT coefficients," *J. Optoelectron. Laser*, vol. 26, no. 4, pp. 733–739, 2015.

[10] M. Huang, R. Wang, J. Xu, D. Xu, and Q. Li, "Detection of double compression for HEVC videos based on the Markov feature optimization," in *Proc. 12th China Inf. Hiding Multimedia Secur. Workshop (CIHW)*, Wuhan, China, Mar. 2015, pp. 475–481.

[11] M. Huang, R. Wang, J. Xu, D. Xu, and Q. Li, "Detection of double compression for HEVC videos based on the co-occurrence matrix of DCT coefficients," in *Proc. 14th Int. Workshop Digit. Water-Marking (IWDW)*, Tokyo, Japan, Oct. 2015, pp. 61–71.

[12] A. Costanzo and M. Barni, "Detection of double AVC/HEVC encoding," in *Proc. Signal Process. Conf. (EUSIPCO)*, Aug./Sep. 2016, pp. 2245–2249.

[13] Q. Li, R. Wang, and D. Xu, "Detection of double compression in HEVC videos based on TU size and quantised DCT coefficients," *IET Inf. Secur.*, May 2018, doi: 10.1049/iet-ifs.2017.0555.

[14] Q. Xu, T. Sun, X. Jiang, and Y. Dong, "HEVC double compression detection based on SN-PUPM feature," in *Digital Forensics and Watermarking—IWDW* (Lecture Notes in Computer Science), vol. 10431, C. Kraetzer, Y. Q. Shi, J. Dittmann, and H. Kim, Eds. Cham, Switzerland: Springer, 2017, pp. 3–17.

[15] Z.-H. Li, R.-S. Jia, Z.-Z. Zhang, X.-Y. Liang, and J.-W. Wang, "Double HEVC compression detection with different bitrates based on co-occurrence matrix of PU types and DCT coefficients," in *Proc. ITM Web Conf.*, vol. 12, 2017, Art. no. 01020.

[16] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[17] J. Chen, X. Jiang, T. Sun, P. He, and S. Wang, "Detecting double MPEG compression with the same quantiser scale based on MBM feature," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Shanghai, China, Mar. 2016, pp. 2064–2068.

[18] X. Jiang, P. He, T. Sun, F. Xie, and S. Wang, "Detection of double compression with the same coding parameters based on quality degradation mechanism analysis," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 1, pp. 170–185, Jan. 2018.

[19] R.-S. Jia, Z.-H. Li, Z.-Z. Zhang, and D.-D. Li, "Double HEVC compression detection with the same QPs based on the PU numbers," in *Proc. 3rd Annu. Int. Conf. Inf. Technol. Appl. (ITA)*, 2016, pp. 1–4, Art. no. 02010.

[20] J. A. Aghamaleki and A. Behrad, "Detecting double compressed MPEG videos with the same quantization matrix and synchronized group of pictures structure," *J. Electron. Imag.*, vol. 27, no. 1, p. 013031, 2018.

[21] *Gitl HEVC/H.265 Analyzer*. Accessed: Dec. 9, 2016. [Online]. https://github.com/lheric/GitlHEVCAnalyzer

[22] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. no. 27. Accessed: Aug. 2, 2015. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm

XIAOYUN LIANG is currently pursuing the master's degree with the School of Electronic Science and Technology, Beijing Jiaotong University, China. Her research interest is video forensics.

ZHAOHONG LI received the B.S. degree in communication engineering from Southwest Jiaotong University in 2003 and the Ph.D. degree in traffic information engineering and control from Beijing Jiaotong University in 2008. Since 2008, she has been with the School of Electronic and Information Engineering, Beijing Jiaotong University where she is currently an Associate Professor. Her research interests include video forensics, information hiding, watermarking, and anti-forgery.

YIYUAN YANG is currently pursuing the bachelor's degree with the School of Computer Science and Engineering, Beihang University, China.

ZHENZHEN ZHANG received the B.S. degree in communication engineering from Zhengzhou University in 2009 and the Ph.D. degree in circuits and systems from Beijing Jiaotong University in 2017. After the completion of her Ph.D. program, she has been an Assistant Professor with the Beijing Institute of Graphic Communication. Her research interests include video forensics, video watermarking, and information hiding.

YU ZHANG received the B.S. degree from the Beijing University of Aeronautics and Astronautics in 2003, the M.S. degree from Beijing Jiaotong University in 2006, and the Ph.D. degree from the Beijing University of Posts and Telecommunications in 2009, all in electrical engineering. From 2009 to 2012, he was a Research Staff Member at NEC Laboratories China. From 2011 to 2012, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, the University of California at Davis. Since 2013, he has been with Qualcomm Research, where he is currently a Staff Engineer and a Manager. Since 2013, he has been actively contributing to the research, development, and standardization of 3GPP LTE-Advanced Pro and 5G NR technology. His research interests include several aspects of wireless communications, signal processing, and artificial intelligence.