Signal Processing 166 (2020) 107269

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Detection of fake high definition for HEVC videos based on prediction mode feature

Yang Yu[a,b], Haichao Yao[a,b], Rongrong Ni[a,b,*], Yao Zhao[a,b]

[a] Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
[b] Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China

ARTICLE INFO

Article history:
Received 4 June 2019
Revised 21 August 2019
Accepted 24 August 2019
Available online 25 August 2019

Keywords:
Video forensics
HEVC
Fake high definition
Prediction modes

ABSTRACT

A high definition (HD) video usually means a good visual quality, and the video with low quantization parameter (QP) or high bit rate has a good video definition. However, forgers prefer to directly re-encode lower definition videos using lower QP or higher bit rate without video quality improvement to pretend to be HD videos. Therefore, fake HD video detection is necessary in video forensics. In this paper, a novel method is proposed to detect the fake HD for high efficient video coding (HEVC) videos based on prediction mode feature (***PMF***). According to our analysis, the HEVC prediction mode scheme of the subsequent encoding could be influenced by the previous lower quality encoding, hence the frequencies of prediction unit with all types of prediction modes can be used to detect the fake HD videos. Firstly, a 4-D features are extracted from Planar, DC, H0 direction and V0 direction intra prediction modes. Secondly, a 6-D features are extracted from Skip, Merge and AMVP inter prediction modes. Finally, these two feature sets are combined into the ***PMF*** to detect fake HD videos and further estimate their original QPs and bit rates. Experimental results show that the performance of the proposed method outperforms state-of-the-art works.

1. Introduction

With the availability of portable video capture devices such as mobile phones, the digital video is enriching our daily life. The greatest advantage of digital video is that it provides viewers with a high-quality viewing environments compared to digital images, making the digital videos more and more popular in various fields. However, with the popularization of inexpensive and easy-to-use video editing tools, it is easier for forgers to tamper with the original videos. The forged digital videos could have tremendous effects on politics, economy, law enforcement, and other fields. Therefore, video forensics has attracted wide research interests in the field of information security.

Up to now, many effective methods have been reported in digital video forensics. In early studies, active video forensic techniques such as digital video watermarking [1,2] were proposed to validate video evidence. A major drawback of active techniques is that watermark data have to be embedded during the recording process. Contrary to active methods, passive video forensic techniques were proposed only using inherent traces left in any tampering process. For detecting video tampering in temporal domain, Wang et al. [3] detected video frame deletion by analyzing the periodic properties of P-frame prediction errors in MPEG-2 video. Gironi et al. [4] proposed a method based on the variation of prediction footprint to detect insertion and deletion of whole frames in digital videos. Aghamaleki et al. [5] proposed a approach to detect and localize video forgery in the temporal domian. Feng et al. [6] adopted a study of the statistical characteristics of the most common interfering frames such as relocated I-frames, and then developed a new fluctuation feature based on frame motion residuals to identify frame deletion points. For detecting video tampering in spatial domain, Zhang et al. [7] detected video inpainting based on ghost shadow artifacts. Kim et al. [8] proposed a modality fusion method designed for combining spatial and temporal fingerprint information to improve video copy detection performance. For video double compression detection, Yao et al. [9] proposed a method that analyzed the periodic features of the string of data bits and the skip macroblocks for all I-frames and P-frames in a double-compressed H.264/AVC video. Jiang et al. [10] analyzed degradation mechanisms during recompression to detect double compression with the same coding parameters. Elrowayati et al. [11] analyzed the changes in the quantized residual DST coefficients and the values of intra-prediction modes for detection of double-compression in HEVC, even when the same QP is used in

* Corresponding author at: Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China.
E-mail address: rrni@bjtu.edu.cn (R. Ni).

https://doi.org/10.1016/j.sigpro.2019.107269
0165-1684/

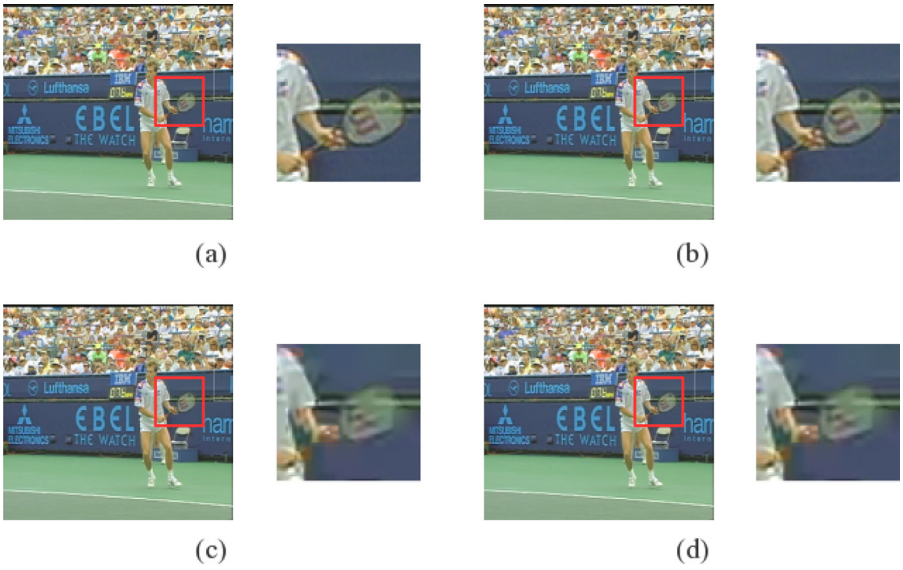

Fig. 1. Visual quality comparisons of true HD videos (top) and fake HD videos (bottom). (a) QP 10. (b) bit rate 1 Mbps. (c) The video with QP of 10 is converted from the video with QP of 45.(d) The video with bit rate of 1 Mbps is converted from the video with bit rate of 200 Kbps.

both compression processes. Jiang et al. [12] proposed a method to detect double HEVC compressed videos with the same coding parameters based on intra prediction mode that analyzed the quality degradation mechanism and considered the source of error in intra coding.

Video quality forgery is a common kind of video tampering operation. In general, video quality is affected by the video definition. A high definition (HD) video usually means that it has a good visual quality. The better the visual quality is, the more popular the video becomes on sharing websites. However, in order to make the lower definition videos more attractive and earn advertising revenue, forgers usually prefer to directly re-encode them using the parameters of HD videos without any improvement as if they were HD videos natively. For a given resolution, the encoding parameter QP or bit rate are important factors in evaluating video definition, and the video with low QP or high bit rate has a good video quality. Therefore, forgers usually modify QP or bit rate in the process of making the fake HD videos with the help of some video tools, such as FFmpeg [13] and x265 [14]. Specifically, the fake HD video is the claimed low QP video that is actually converted from the video with higher QPs, or the claimed high bit rate video is actually converted from the video with lower bit rates. However, when a video is re-encoded, only the information about the most recent encoding could be obtained from the bit stream. In this case, as shown in Fig. 1, although the true HD video (a) (b) and the fake HD video (c) (d) have the same QP and bit rate, they actually have greatly difference in video definition. It will mislead users and cause economic losses. High Efficiency Video Coding (HEVC) [15,16] is the new and increasingly popular generation video compression standard. Therefore, the appearance of fake HD videos in HEVC format is a new security issue. Therefore, the detection of fake HD for HEVC videos is necessary in digital video quality forensics.

However, just a few methods have been proposed in video quality forensics. Vinh et al. [17] proposed an adaptive motion compensated interpolation algorithm for frame rate-up conversion in multimedia applications. Bian et al. [18] analyzed the periodic properties of inter-frame similarity to detect the fake frame-rate videos. Ding et al. [19] proposed a blind forensics approach for the identification of various motion-compensated frame rate up-conversion techniques based on residual signals. Sun et al. [20] detected fake bitrate MPEG videos by analyzing DCT coefficients. Bian et al. [21] detected fake bitrate MPEG-2 videos and extended their algorithm to the H.264/AVC coding schemes in [22]. For the new HEVC standard, Costanzo et al. [23] proposed a forensic technique to detect double AVC/HEVC transcoding under different QPs. Bian et al. [24] exploited the statistics of prediction units (PUs) to detect transcoded HEVC videos from AVC format. Li et al. [25] detected double HEVC video compression with different QP based on transform unit (TU) size and quantized DCT coefficients. Huang et al. [26] proposed a method to detect double compression for HEVC videos with different QPs based on the co-occurrence matrix of DCT coefficients. Liang et al. [27] detected HEVC videos with fake bit rates based on PU partition types.

In this paper, we propose a novel method to detect the HEVC videos with fake HD by analyzing the prediction modes. According to our analysis, we found that the HEVC prediction modes scheme of the subsequent encoding could be influenced by the previous lower quality encoding. Therefore, the frequencies of PU with intra and inter prediction modes are distinguishable between the fake HD videos and the true ones. First, a 4-D feature set and a 6-D feature set are extracted from HEVC intra prediction modes and inter prediction modes, respectively. Then, these two feature sets are combined into the proposed prediction mode feature (***PMF***). Next, the ***PMF*** is fed to the ensemble classifier to detect the fake HD videos. Finally, we further estimate the original QPs and bit rates. Experimental results show that the performance of the proposed method outperforms state-of-the-art works [26,27].

The rest of the paper is organized as follows. In Section 2, we briefly review the HEVC codec and its prediction modes. In Section 3, the prediction mode feature is analyzed, and then the detection model is presented in detail. We further estimate the original QPs and bit rates based on ***PMF***. Section 4 shows the experimental results and corresponding analysis. Finally, the conclusion of this paper along with future challenges are drawn in Section 5.

2. Prediction modes in HEVC

HEVC applies hybrid video coding scheme (inter/intra picture prediction and 2-D transform coding) used in most prior video coding standards. Similar to the AVC standard, HEVC still processes all pictures of the input video sequence in block units. Specially,

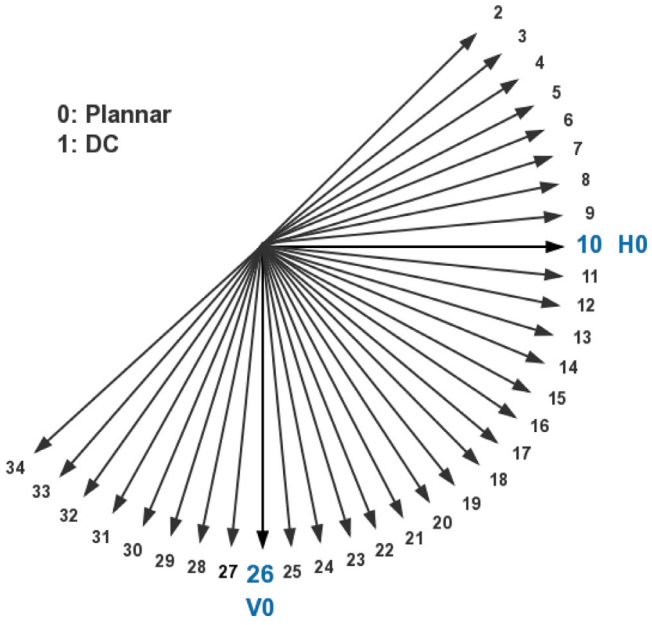

Fig. 2. Intra prediction modes for luminance components in HEVC.

HEVC partitions each picture into flexible coding tree units (CTU) whose size can be 16 × 16, 32 × 32, 64 × 64. Each CTU consists of one luma coding tree block (CTB) and two chroma Coding Tree Blocks at the same position. Based on quad-tree structure, each CTU can be further partitioned into coding units (CUs) and each CU can be partitioned (or not be partitioned) into smaller PUs and TUs. To better understand our method, the prediction modes and the PU are introduced below.

2.1. HEVC intra prediction

In the HEVC intra prediction, a current PU is compared to a reference PU that is already encoded in the same picture to calculate prediction values. Five sizes of PUs, which are 64 × 64, 32 × 32, 16 × 16, 8 × 8 and 4 × 4, can be selected in HEVC intra prediction. In order to better adapt to various textural features, HEVC introduces 35 prediction modes to encode luminance components, including DC, planar and 33 angular modes. Planar and DC modes are suitable for regions with homogeneous texture. In the 33 angular directional modes, V0 and H0 represent the vertical and the horizontal directions respectively, and the remaining angular directional modes offset in these two directions. As shown in Fig. 2, the 0th and the 1th modes represent Plannar and DC modes, respectively, and the 10th and 26th modes represent H0 direction and V0 direction prediction modes, respectively.

2.2. HEVC inter prediction

In the HEVC inter prediction, the motion vectors (MVs) of adjacent PUs in the spatial domain and the temporal domain have strong correlations. A current PU is compared to a reference PU from previously encoded pictures for calculating prediction values to save the coded bits of MV. The difference between the current PU and the reference PU is the prediction residual. HEVC supports three modes to predict the MV, namely advanced motion vector prediction (AMVP), Merge and Skip modes. During the AMVP encoding process, a current PU generates the motion vector prediction value (MVP), and then the MV is obtained by calculating the motion vector difference value (MVD), which is the difference between the MV of the current PU (MVC) and the MVP. The MVD and prediction residual are encoded in the AMVP mode. The MVD is not encoded in the Merge and Skip modes, and the MV are directly obtained from the adjacent encoded PU. To further increase coding efficiency, the prediction residual can be given up in the Skip mode In these three MV prediction modes of HEVC, the most complex is the AMVP mode, and the simplest is the skip mode.

Based on the above observations, we will analyze the HEVC prediction modes artifacts in the following Section 3, and extract some effective features to detect fake HD videos and further estimate their original bit rates and QPs.

3. Proposed method

In this section, we find that fake HD videos present obvious coding artifacts in the intra and inter prediction modes. For more details and the theoretical analysis, refer to the following Section 3.1. In Section 3.2, we present the framework of proposed detection model. In Section 3.3, we propose a method to further estimate the original bit rate and QP for a given fake HD video.

3.1. Feature analysis

3.1.1. Feature of HEVC intra prediction

The 35 prediction modes are designed to adapt to various texture characteristics in the HEVC intra prediction. More specifically, different intra prediction modes can be selected for regions of different texture complexity. The Planar mode uses two linear filters in the horizontal and vertical directions, and then calculates the average of the two prediction values as the prediction value of the current pixel. In the DC mode, the current prediction value can be obtained from the average of the reference pixels on the left and above. Therefore, the Planar and DC modes are suitable for regions with homogeneous texture. In the fake HD videos, the previous lower quality encoding reduces the texture complexity of images, thus the HEVC uses more Planar and DC modes that are suitable for regions with homogeneous texture. Moreover, if the texture is simpler, the 33 angular directional modes will utilize more H0 direction and V0 direction prediction modes and reduce the use of other angular directional modes. Two examples are shown here to verify this phenomenon. We checked the frequency of Planar, DC, H0 direction and V0 direction prediction modes over all I-frames of pristine videos and fake HD videos, respectively. The content of all videos is different, including simple content with simple texture and slow movement (*akiyo, bridge_far, highway, mother_daughter, silent*), as well as complex content with rich texture and quick movement (*bus, mobile, stefan, crew, football*). We found that the frequencies of these four types of intra prediction modes in the fake HD videos are significantly higher than that in the pristine videos, as shown in Fig. 3.

Based on the above analysis, a 4-D feature is extracted from all the I-frames of a video sequence. In the following, we will define the feature of HEVC intra prediction mode as $\boldsymbol{f_{intra}} = [P, D, V0, H0]$, which is an array with the frequencies of four types of intra prediction modes as mentioned above.

3.1.2. Feature of HEVC inter prediction

HEVC inter prediction can select appropriate MV prediction modes for regions of different level correlations. For the fake HD videos, the previous lower quality encoding reduces the spatial and temporal differences between blocks and thus changes the inter prediction modes scheme of HEVC. That is HEVC prefers a simpler MV prediction mode because the difference is reduced. Here are two examples to further explain this phenomenon. We extracted the frequency of all types of MV prediction modes over all P-frames and B-frames of pristine videos and fake HD videos

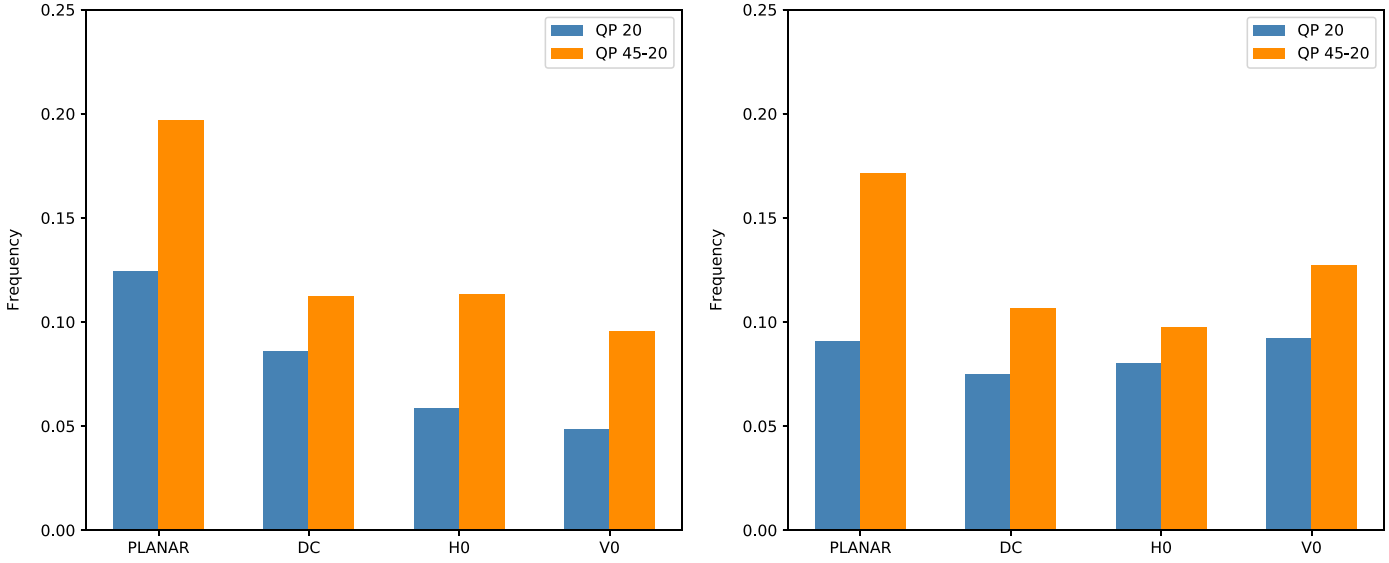

Fig. 3. Average frequency of intra prediction modes in I-frames of simple content videos (left) and complex content videos (right).

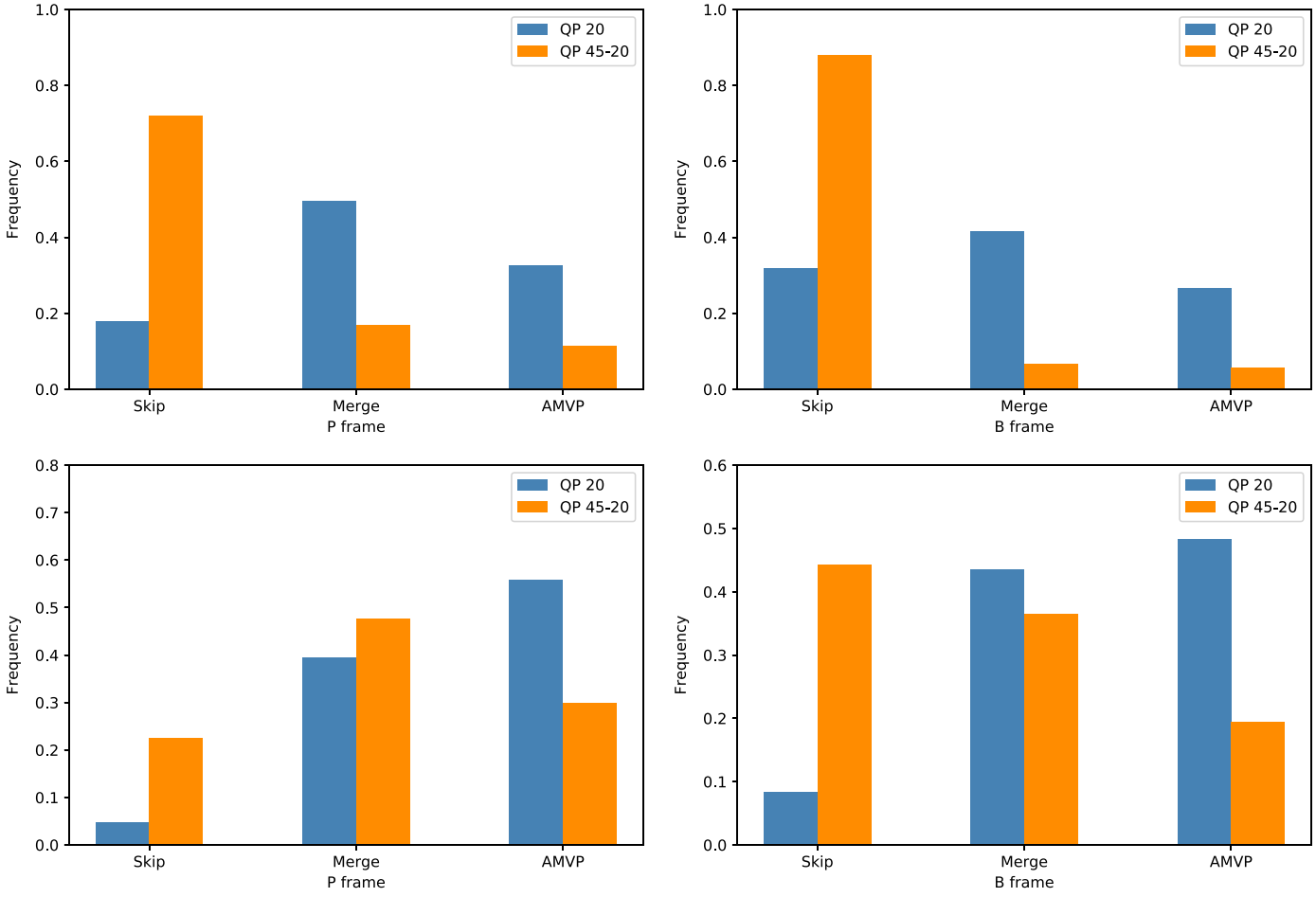

Fig. 4. Average frequency of inter prediction modes of simple content videos (top) and complex content videos (bottom).

with different contents. For the fake HD videos with simple content, HEVC uses the simplest Skip mode for MV prediction more frequently rather than the most complex AMVP and more complex Merge modes, as shown in Fig. 4. For the complex content videos, more Skip mode and the relatively simple Merge mode are adopted, and the use of the most complex AMVP mode is reduced, as shown in Fig. 4.

Simialr to the feature extraction from intra prediction mode, a 6-D feature is extracted from all P-frames and B-frames of a video sequence. And, the feature of HEVC inter prediction mode is denoted as $\boldsymbol{f_{inter}} = [\boldsymbol{f_{Pinter}}, \boldsymbol{f_{Binter}}] = [S_P, M_P, A_P, S_B, M_B, A_B]$, which is the array with the frequencies of three types of inter prediction modes in P-frames and B-frames. The features extracted from

the intra prediction mode $\boldsymbol{f}_{intra}$ and inter prediction mode $\boldsymbol{f}_{inter}$ are combined into the final proposed 10-D feature ***PMF***.

3.1.3. Theoretical analysis

The prediction modes can distinguish between the fake HD videos and true videos in HEVC format as mentioned above, and the theoretical analysis is shown in this section.

An original YUV sequence of length N could be expressed as (1).

$$\boldsymbol{Y} = \{F_1, F_2, \ldots, F_N\} \tag{1}$$

where $\boldsymbol{F_n}$ represents the nth uncoded frame.

$\boldsymbol{Y}$ is encoded by HEVC with QP Q_1 or bit rate B_1 to obtain a low definition video $\boldsymbol{V}_1$. In the process of making a fake HD video, $\boldsymbol{V}_1$ is first decompressed to a YUV sequence could be expressed as (2). Then $\boldsymbol{Y}^{'}$ is re-encoded by HEVC with QP Q_2 or bit rate B_2 to obtain a fake HD video $\boldsymbol{V}_2$. In the second compression, the prediction mode can be denoted as (3).

$$\boldsymbol{Y}^{'} = \{F_1^{'}, F_2^{'}, \ldots, F_N^{'}\} \tag{2}$$

$$\begin{aligned} \boldsymbol{PM_n}^{(Q_1,Q_2)} &= \rho\left(\boldsymbol{PU}_{\boldsymbol{F_n'}}(x,y), Q_2\right) \\ \boldsymbol{PM_n}^{(B_1,B_2)} &= \rho\left(\boldsymbol{PU}_{\boldsymbol{F_n'}}(x,y), B_2\right) \end{aligned} \tag{3}$$

where $\boldsymbol{F}_n^{'}$ represents the nth frame of the decompressed YUV sequence, $\rho(\cdot\ ,\ \cdot)$ represents the prediction mode process in all frames, $\boldsymbol{PU}_{\boldsymbol{F}_n'}(x,y)$ denotes the current PU to encode $\boldsymbol{F}_n^{'}$, and $Q_1 > Q_2$, $B_1 < B_2$.

As for a true HD video, $\boldsymbol{Y}$ is directly encoded by HEVC with QP Q_2 or bit rate B_2 to obtain $\boldsymbol{V}$, and the prediction modes can be denoted as:

$$\begin{aligned} \boldsymbol{PM_n}^{(Q_2)} &= \rho(\boldsymbol{PU}_{\boldsymbol{F_n}}(x,y), Q_2) \\ \boldsymbol{PM_n}^{(B_2)} &= \rho(\boldsymbol{PU}_{\boldsymbol{F_n}}(x,y), B_2) \end{aligned} \tag{4}$$

where $\boldsymbol{PU}_{\boldsymbol{F_n}}(x,y)$ denotes the current PU to encode $\boldsymbol{F}_n$.

Then, the difference $\boldsymbol{D}(\cdot,\ \cdot)$ between prediction modes of fake HD videos and true videos can be described as follows:

$$\begin{aligned} &\boldsymbol{D}(\boldsymbol{PM_n}^{(Q_1,Q_2)}, \boldsymbol{PM_n}^{(Q_2)}) \\ &= \boldsymbol{D}(\rho(\boldsymbol{PU}_{\boldsymbol{F_n'}}(x,y), Q_2), \rho(\boldsymbol{PU}_{\boldsymbol{F_n}}(x,y), Q_2)) \end{aligned} \tag{5}$$

$$\begin{aligned} &\boldsymbol{D}(\boldsymbol{PM_n}^{(B_1,B_2)}, \boldsymbol{PM_n}^{(B_2)}) \\ &= \boldsymbol{D}(\rho(\boldsymbol{PU}_{\boldsymbol{F_n'}}(x,y), B_2), \rho(\boldsymbol{PU}_{\boldsymbol{F_n}}(x,y), B_2)). \end{aligned} \tag{6}$$

In Eqs. (5) and (6), the difference is caused by $\boldsymbol{F}_n^{'}$ and $\boldsymbol{F}_n$.

The encoding process to obtain $\boldsymbol{V}_1$ is

$$\boldsymbol{E_n} = [DCT(\boldsymbol{PU}_{\boldsymbol{F_n}}(x,y) - \boldsymbol{RPU}_{\boldsymbol{F_n}}(x,y))/Q_1] \tag{7}$$

where $\boldsymbol{E}_n$ represents the nth encoded frame in $\boldsymbol{V}_1$, $\boldsymbol{R}PU_{F_n}$ denotes the reference PU of PU_{F_n}, $DCT(\cdot)$stands for the discrete cosine transform in the video frame coding process, and $[\cdot]$ is the rounding operator.

The decoding process to obtain $\boldsymbol{Y}^{'}$ is

$$\boldsymbol{F_n'} = IDCT(\boldsymbol{E_n} \times Q_1) + \boldsymbol{RPU}_{\boldsymbol{F_n}}(x,y) \tag{8}$$

Based on the above analysis, there is quantization error in the HEVC encoding process. Therefore, $\boldsymbol{F_n'} = \boldsymbol{F_n} + \boldsymbol{\Delta}$, where $\boldsymbol{\Delta}$ denotes the quantization error under Q_1 or B_1 in the first compression of the fake HD video.

Therefore, $\boldsymbol{F_n'} - \boldsymbol{F_n} = \boldsymbol{\Delta}$, the quantization error leads to the difference between $\boldsymbol{F}_n^{'}$ and $\boldsymbol{F}_n$. That is to say, the main factor causing the difference between the prediction modes in the fake HD video and the true video is the quantization error, and the larger $\boldsymbol{\Delta}$ leads to the larger difference. In general, the quantization error is larger when the bit rate is lower and the QP is higher in the video encoding process. According to our previous description, the HD video usually have high bit rate B_2 or low QP Q_2. If B_2 or Q_2 is fixed, when the $B_1 < B_2$ or $Q_1 > Q_2$ (the fake HD videos), $\boldsymbol{\Delta}$ is large, and the difference between the prediction modes in the fake HD videos and the true videos is larger. Conversely, if $B_1 \geq B_2$ or $Q_1 \leq Q_2$, $\boldsymbol{\Delta}$ is really small. In this case, it is difficult to detect the fake videos based on the prediction modes.

In summary, the prediction modes are distinguishable between the fake HD videos and true videos in HEVC format.

3.2. Detection model

Based on the above analysis, we found that the frequencies of HEVC prediction modes are distinguishable between the pristine HEVC videos and fake HD videos. To exploit the artifacts, we extract a 10-D feature vector from a query HEVC video. The framework of the proposed detection model in this paper is given in Fig. 5. The frequencies of HEVC intra prediction modes from I frames and inter prediction modes from P frames and B frames are calculated to generate 4-D feature $\boldsymbol{f}_{intra}$ and 6-D feature $\boldsymbol{f}_{inter}$. After extracting $\boldsymbol{f}_{intra}$ and $\boldsymbol{f}_{inter}$, these two sets of features are connected to obtain the final 10-D detection feature of the input videos, $\boldsymbol{PMF} = [\boldsymbol{f}_{\boldsymbol{intra}}, \boldsymbol{f}_{\boldsymbol{inter}}]$. Moreover, ensemble classifier is employed for classification to obtain detection results.

3.3. Estimation of original QP and original bit rate

In this section, to further estimate the original QPs and bit rates for given fake HD videos, we propose a method that is based on the concept of video compression idempotency introduced in [28]. Idempotency of lossy coding indicates that a video coding scheme is applied to a video twice with the same parameters, it produces roughly the same output as applied once. Hence, we can re-encode the fake video with different parameters and verify whether the results are similar to the input video. If this is the case, it is likely that the parameter used for re-encoding correspond to the original parameter used to the fake video. Specifically, the method we proposed is as follows:

Step 1. Input the fake HD video $\boldsymbol{V}$.

Step 2. Extract the ***PMF*** of $\boldsymbol{V}$.

Step 3. Check the type of conversion. In general, the video tools encode the videos by fixed QPs or fixed bit rates. If the QP of each PU in the video is the same, it is a fixed QP compressed video. If not, it is a fixed bit rate compressed video.

Step 4. If it is converted from an original higher QP version, extract QP of $\boldsymbol{V}$ as Q, re-encode $\boldsymbol{V}$ with higher $Q_1^i = (Q, Q+1, Q+2, \ldots, 51)$ to obtain $\boldsymbol{V}_{Q_1^i}$, and re-encode $\boldsymbol{V}_{Q_1^i}$ with Q to obtain $\boldsymbol{V}_{Q_2^i}$, $i = 1, \ldots, length(Q_1^i)$.

Step 5. If it is converted from an original lower bit rate version, extract bit rate of $\boldsymbol{V}$ as B, re-encode $\boldsymbol{V}$ with lower bit rates $B_1^i = (B_1^1, \ldots, B_1^k)$ to obtain $\boldsymbol{V}_{B_1^i}$, where $B_1^1 - B_1^{i+1} = T, B_1^1 = B$, T is the descending step size, and $i = 1, \ldots, length(B_1^i)$. Then re-encode $\boldsymbol{V}_{B_1^i}$ with B to obtain $\boldsymbol{V}_{B_2^i}$.

Step 6. Extract the $\boldsymbol{P}MF^i$ of $\boldsymbol{V}_{Q_2^i}$ or $\boldsymbol{V}_{B_2^i}$.

Step 7. Measure the euclidean distance $\boldsymbol{D}$ between ***PMF*** and $\boldsymbol{P}MF^i$.

Step 8. Estimate the original QPs and bit rates for given fake HD videos based on the distances $\boldsymbol{D}$.

3.3.1. Estimation of original QPs

Based on the concept of video compression idempotency, for a fake HD video $\boldsymbol{V}$, the $\boldsymbol{D}$ will be minimum in correspondence of the Q_1^i that coincides with the original Q_1.

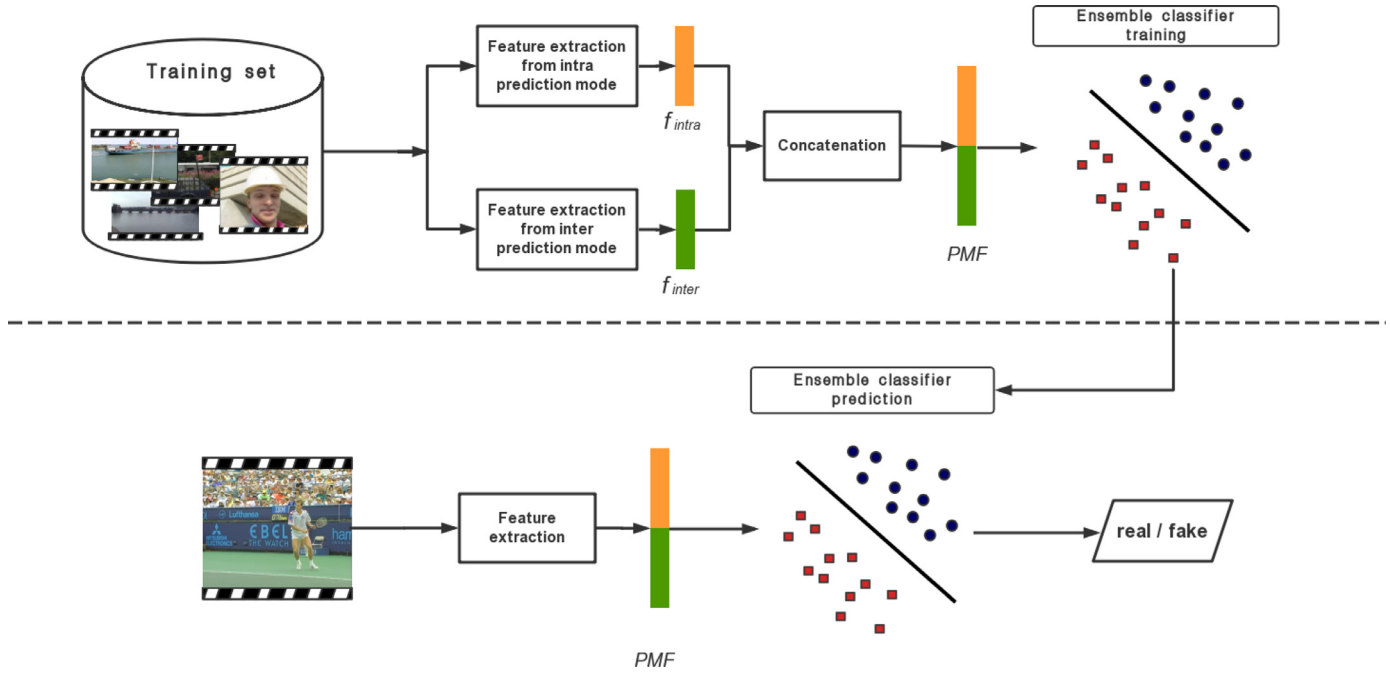

Fig. 5. Framework of proposed detection model.

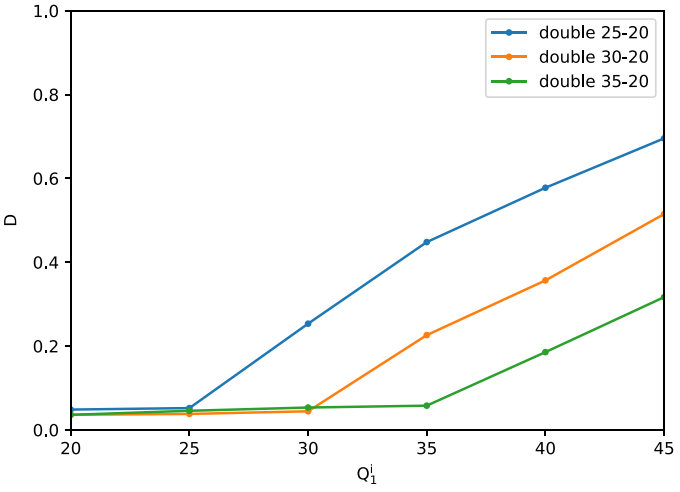

Fig. 6. Distance between the fake video $\boldsymbol{V}$ and the re-encoded videos $\boldsymbol{V}_{Q_1^i}$.

Three examples are shown here to verify this phenomenon. First, a video sequence (*akiyo*) is encoded twice with $Q_1 = \{25, 30, 35\}$ followed with fixed $Q = 20$ to obtain the fake HD videos. Next, each video is re-encoded with increasing $Q_1^i = (20, 21, 22, 23, \ldots, 51)$ followed with Q, and $i = (1, 2, \ldots, 32)$. Finally, we compute the $\boldsymbol{D}$ between the $\boldsymbol{PMF}$ of the fake videos and the $\boldsymbol{PMF}^i$ of the re-encoded videos. As illustrated in Fig. 6, there is an obvious inflection point for each fake definition video. The distance is very small and close to 0 before the inflection point, after this point, the $\boldsymbol{D}$ starts increasing rapidly. Therefore, the rule we adopt to locate the inflection point is that if $k(Q_1^i) \leqslant T_1$ and $k'(Q_1^i) \geqslant T_2$, the Q_1^i is the inflection point, where $k(Q_1^i)$ is the slope of Q_1^i to Q_1^1, $k'(Q_1^i)$ is the slope of Q_1^i to Q_1^{i+1}, and T_1 and T_2 will be given in Section 4.1. We denote Q_1^f as the inflection point, where $f \in i$. The value of inflection point Q_1^f corresponds to the original Q_1. When $Q_1^i < Q_1$, the first compression is not strong enough to alter the second compression, thus the re-encoded videos are similar to the original videos. When $Q_1^i \geqslant Q_1$, the distance starts growing.

The original QPs coincide with the critical values for the fake HD videos. Therefore, these critical values can be used to estimate the original QPs.

3.3.2. Estimation of original bit rates

Similar to the estimation of original QP, we apply a similar approach based on the concept of video compression idempotency to estimate the original bit rate.

Three examples are shown here too. First, we encode a video sequence (*akiyo*) twice with $B_1 = \{300\,\text{K}, 500\,\text{K}, 800\,\text{K}\}$ bps followed with fixed $B = 1\text{Mbps}$ to obtain the fake HD videos. Next, we re-encoded each video with decreasing $B_1^i = (1\,\text{M}, 900\,\text{K}, 800\,\text{K}, \ldots, 300\,\text{K}, 200\,\text{K}, 100\,\text{K})$ bps followed with B, and $i = (1, 2, \ldots, 10)$. Finally, we computed the distance between the $\boldsymbol{PMF}$ of the fake videos and the $\boldsymbol{PMF}^i$ of the re-encoded videos. Fig. 7 illustrates the results, there is an obvious inflection point in every fake one, which means that prior to the point, the distance is very small until the B_1^i is higher than the original B_1, while it increases significantly after passing the point. Similarly, the rule to locate the inflection points is that if $k(B_1^i) \leqslant T_3$ and $k'(B_1^i) \geqslant T_4$, the B_1^i is the inflection point, where $k(B_1^i)$ is the slope of B_1^i to B_1^1, $k'(B_1^i)$ is the slope of B_1^i to B_1^{i+1}, and T_3 and T_4 will be given in Section 4.1. We denote B_1^f as the inflection point, where $f \in i$.

The original bit rates coincide with the critical values for the fake HD videos. Therefore, these critical values can be used to estimate the original bit rates.

4. Experimental results

4.1. Experiment setup

The dataset used for experiments is constructed by 61 original YUV sequences [29] with four kinds of resolutions: QCIF (*akiyo, bowing, carphone, bridge_close, bridge_far, bus, city, claire, coastguard, container, mother-daughter, grandma, hall _objects, silent, highway, husky, miss_am, news, salesman, crew, deadline,*

Table 1
Details of parameters in video sets and experiments.

Parameters	QCIF Video Set	CIF Video Set	720P Video Set	1080P Video Set
Encoder	x265	x265	x265	x265
Decoder	HM16.15 TAppDecoder	HM16.15 TAppDecoder	HM16.15 TAppDecoder	HM16.15 TAppDecoder
Q1	{15, 20, 25, 30, 35, 40, 45}	{15, 20, 25, 30, 35, 40, 45}	{15, 20, 25, 30, 35, 40, 45}	{15, 20, 25, 30, 35, 40, 45}
Q	{10, 20}	{10, 20}	{10, 20}	{10, 20}
B1	{100 K, 200 K, 300 K} (bps)	{300 K, 500 K, 700 K} (bps)	{5 M, 10 M, 15 M} (bps)	{10 M, 15 M, 20 M} (bps)
B	500 Kbps	1 Mbps	30 Mbps	40 Mbps
GOP Size	15	15	15	15
Training/Test Set Ratio	4:1	4:1	4:1	4:1

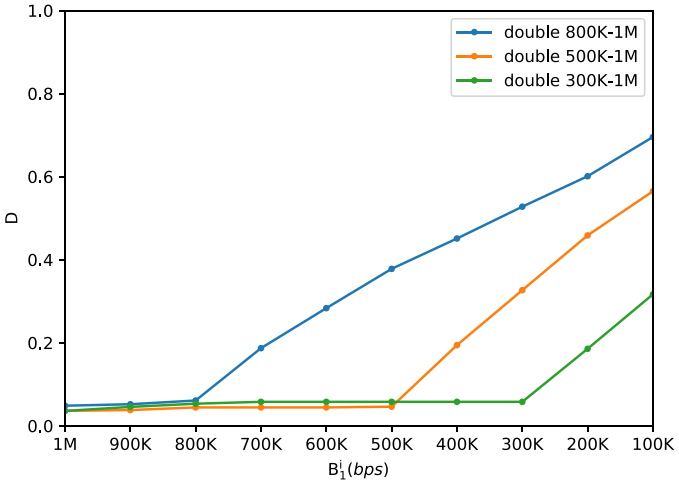

Fig. 7. Distance between the fake video $\boldsymbol{V}$ and the re-encoded videos $\boldsymbol{V}_{B_2^i}$.

flower_garden, football(a), football(b), foreman, gall-eon, ice), CIF (*akiyo, bus, container, flower, foreman, hall, highway, mobile, mother_daughter, news, paris, silent, stefan, tempete, waterfall, bridge_far, ice*), 720P (*ducks_take_off, Johnny, FourPeople, in_to_tree, mobcal, old_town_cross, parkrun, park_ joy, stockholm*) and 1080P (*Kimono, rush_field_cuts, sunflower, BasketballDrive, pedestrian_area, station2, Cactus, ParkScene, duck-s_take_off*).

For each sequence, we cut it into several nonoverlapping sequences with 100 frames. Consequently, 133 QCIF YUV sequences, 70 CIF YUV sequences, 48 720P YUV sequences and 37 1080P YUV sequences are produced. The content of all videos covers a variety of scenes, including quick/slow movement, indoor/outdoor scenes, distance/close shot and rich/simple texture. We compressed the videos by codec x265, and GOP size is set as 15 and the GOP configuration is IBBPBBPBBPBBPBBI. Frequency of intra or inter prediction modes are extracted during the first and second decompression with HM16.15. To guarantee acceptable video quality in many applications, the values of compression quantization parameter Q of all video sets are 10 or 20. The bit rates B of QCIF videos are set as 500 Kbps, and for CIF, 720P, 1080P video sets, the bit rates B are set as 1 Mbps, 30 Mbps and 40 Mbps, respectively. The ratio of videos in training and test sets is set to be 4:1 for each video set, and the training set and test set are randomly selected. In our experiments, 10 experiments are performed and the average result is taken each time. The ensemble classifier used for experiments adopts the default settings [30]. The Thresholds T_1, T_2, T_3 and T_4 need to be set up. We randomly select 500 fixed QP compressed videos in our video datasets with four kinds of resolutions and calculate $k(Q_1^f)$ and $k'(Q_1^f)$ on these videos, as shown in Fig. 8. From the results, we set T_1 and T_2 as 0.006 and 0.015 respectively. Similarly, we randomly select 500 fixed bit rate compressed videos with four kinds of resolutions to calculate $k(B_1^f)$ and $k'(B_1^f)$, as illustrated in Fig. 9. From the results, T_3 and T_4 are set as 0.00015 and 0.001 respectively. Details of parameters in video sets, together with parameters in experiments are summarized in Table 1.

4.2. Detection of fake HD videos

In this section, we evaluate the performance of the proposed method on the detection of fake HD videos in all video sets. In order to better demonstrate the superiority of the proposed method, we compare it with the existing methods on detection of fake HD videos [26,27]. For comparison, their methods are applied to our video sets. The comparative results are shown in Table 2(a)–(d). From the results, all the accuracies of our method are higher than 90% and outperform [26] and [27] in all cases, which means that the proposed method can effectively detect the fake HD videos.

This could be explained as below: the feature set ***PMF*** in our method was extracted from all types of frames rather than only P frames in [27]. Our feature set $\boldsymbol{f}_{intra}$ was extracted from all intra prediction modes, which allows our method better capture changes of texture. The feature set $\boldsymbol{f}_{inter}$ makes our method more adaptable in capturing changes of spatial domain and temporal domain correlations and enables our feature ***PMF*** to be more adaptable in different video contents. Even the two compression parameters of the fake HD videos are close, our feature set is still effective.

Therefore, our method can effectively detect the fake HD for HEVC videos.

4.3. Estimation of original QP and bit rate

In this section, we further evaluate the performance of proposed method on the estimation of original QPs and bit rates for the fake HD videos in all video sets. The estimation accuracy as shown in Table 3(a)–(d) for QCIF, CIF, 720P and 1080P videos, respectively. It is observed that the accuracies of estimating the original QPs and bit rates are all over 90%. When the two compression parameters of the fake videos differ greatly, namely the video quality of the two compression differs obviously in the fake videos, the accuracies of estimating original parameters are almost over 97%. Besides, the estimation accuracies are more than 90% when the two compression parameters of the fake videos are close, such as the QPs are 20-10, 15-10, and the bit rates 300 Kbps-500 Kbps, 700 Kbps-1 Mbps, 15 Mbps-30 Mbps and 20 Mbps-40 Mbps.

In a word, the performance of our method is satisfactory on the estimation of original QP and bit rate.

4.4. Results on different feature sets

As described in Section 3, the feature sets $\boldsymbol{f}_{Pinter}$ and $\boldsymbol{f}_{Binter}$ are combined into $\boldsymbol{f}_{inter}$. In order to analyze these two feature sets, we exploit $\boldsymbol{f}_{Pinter}$, $\boldsymbol{f}_{Binter}$ and $\boldsymbol{f}_{inter}$ to detect fake HD videos in the CIF and 1080P video sets, respectively. The comparative results of the

Table 2
Performance on fake HD videos detection in different video sets (%). (a) QCIF video set. (b) CIF video set. (c) 720P video set. (d) 1080P video set.

QP From-To	Proposed	Liang's method [27]	Huang's method [26]	QP From-To	Proposed	Liang's method [27]	Huang's method [26]
45-10	**100**	99.25	98.12	45-10	**100**	99.29	97.86
40-10	**100**	98.87	96.99	40-10	**100**	97.86	96.43
35-10	**100**	97.74	95.49	35-10	**100**	97.14	95.71
30-10	**100**	94.74	93.98	30-10	**100**	96.43	92.86
25-10	**100**	93.98	92.48	25-10	**100**	95.71	91.43
20-10	**97.74**	93.98	91.73	20-10	**97.86**	92.86	91.43
15-10	**96.24**	90.98	90.23	15-10	**95.71**	90.71	89.29
QP From-To	Proposed	Liang's method [27]	Huang's method [26]	QP From-To	Proposed	Liang's method [27]	Huang's method [26]
45-20	**100**	98.85	97.74	45-20	**100**	98.57	97.14
40-20	**100**	97.74	95.49	40-20	**100**	97.86	94.28
35-20	**100**	96.99	94.74	35-20	**100**	96.43	92.86
30-20	**99.62**	92.48	90.98	30-20	**98.57**	92.86	91.43
25-20	**97.74**	91.73	90.23	25-20	**96.43**	91.43	90
From-To (bps)	Proposed	Liang's method [27]	Huang's method [26]	From-To (bps)	Proposed	Liang's method [27]	Huang's method [26]
100 K–500 K	**100**	99.25	98.50	300 K-1 M	**100**	99.29	97.86
200 K–500 K	**100**	97.74	93.98	500 K-1 M	**100**	97.14	92.86
300 K–500 K	**98.50**	94.74	90.98	700 K-1 M	**95.71**	91.43	88.57
		(a)				(b)	
QP From-To	Proposed	Liang's method [27]	Huang's method [26]	QP From-To	Proposed	Liang's method [27]	Huang's method [26]
45-10	**100**	100	97.92	45-10	**100**	100	98.64
40-10	**100**	98.96	96.88	40-10	**100**	98.64	97.30
35-10	**100**	96.88	94.79	35-10	**100**	97.30	95.95
30-10	**100**	95.83	93.75	30-10	**100**	95.95	94.59
25-10	**100**	93.75	91.67	25-10	**100**	93.24	91.89
20-10	**95.83**	91.67	90.63	20-10	**95.95**	91.89	90.54
15-10	**90.63**	89.58	88.54	15-10	**91.89**	89.19	87.84
QP From-To	Proposed	Liang's method [27]	Huang's method [26]	QP From-To	Proposed	Liang's method [27]	Huang's method [26]
45-20	**100**	98.96	96.88	45-20	**100**	100	98.64
40-20	**100**	97.92	94.79	40-20	**100**	98.64	95.95
35-20	**100**	95.83	91.67	35-20	**100**	95.95	94.59
30-20	**97.14**	93.75	89.58	30-20	**97.30**	94.59	91.89
25-20	**90.48**	90.63	88.54	25-20	**93.24**	90.54	89.19
From-To (bps)	Proposed	Liang's method [27]	Huang's method [26]	From-To (bps)	Proposed	Liang's method [27]	Huang's method [26]
5 M–30 M	**100**	100	98.96	10 M–40 M	**100**	98.64	97.30
10 M–30 M	**100**	97.92	96.88	15 M–40 M	**100**	97.30	95.95
15 M–30 M	**98.96**	93.75	91.67	20 M–40 M	**97.30**	95.95	91.89
		(c)				(d)	

Table 3
Performance on the estimation of original QP And bit rate in different video sets. (a) QCIF video set. (b) CIF video set. (c) 720P video set. (d) 1080P video set.

QP From-To	Accuracy (%)	QP From-To	Accuracy (%)	QP From-To	Accuracy (%)	QP From-To	Accuracy (%)
45-10	**100**	45-10	**100**	45-10	**100**	45-10	**100**
40-10	**100**	40-10	**100**	40-10	**100**	40-10	**100**
35-10	**100**	35-10	**100**	35-10	**100**	35-10	**100**
30-10	**100**	30-10	**100**	30-10	**100**	30-10	**100**
25-10	**100**	25-10	**98.57**	25-10	**98.96**	25-10	**97.30**
20-10	**97.74**	20-10	**96.43**	20-10	**95.83**	20-10	**95.95**
15-10	**95.49**	15-10	**94.28**	15-10	**94.79**	15-10	**93.24**
From-To (bps)	Accuracy (%)	From-To (bps)	Accuracy (%)	From-To (bps)	Accuracy (%)	From-To (bps)	Accuracy (%)
100 K–500 K	**100**	300 K-1 M	**100**	5 M–30 M	**100**	10 M–40 M	**100**
200 K–500 K	**97.74**	500 K-1 M	**97.86**	10 M–30 M	**96.88**	15 M–40 M	**95.95**
300 K-500 K	**91.73**	700 K-1 M	**91.43**	15 M–30 M	**91.67**	20 M–40 M	**91.89**
(a)		(b)		(c)		(d)	

experiments are presented in Table 4. From Table 4, the results on the feature set $\boldsymbol{f}_{inter}$ outperforms the results on these two feature set in all cases. Therefore, the combination of features from $\boldsymbol{f}_{Pinter}$ and $\boldsymbol{f}_{Binter}$ is effective.

Moreover, the feature sets $\boldsymbol{f}_{intra}$ and $\boldsymbol{f}_{inter}$ are distinguishable between fake and true HD videos. Therefore, we combine these two feature sets as feature vector ***PMF*** in the proposed method. To further analyze these two feature sets, we conducted a comparative experiments on the CIF and 1080P video sets. Table 5 shows the experimental results. From Table 5, the detection results based on $\boldsymbol{f}_{inter}$ is slightly higher than that on $\boldsymbol{f}_{intra}$, indicating that $\boldsymbol{f}_{inter}$ contributes more in our experiments. The reason is that inter prediction modes can keep more information than intra modes. It is also observed that the results on the proposed feature set ***PMF*** are all

Table 4
Detection accuracies on different inter feature sets in different video sets (%). (a) CIF video set. (b) 1080P video set.

QP From-To	f_{Pinter}	f_{Binter}	f_{inter}	QP From-To	f_{Pinter}	f_{Binter}	f_{inter}
45-20	92.86	95.71	**96.43**	45-20	94.59	97.30	**98.64**
40-20	90.71	92.86	**95.71**	40-20	91.89	95.95	**97.30**
35-20	89.29	91.43	**92.86**	35-20	90.54	94.59	**95.95**
30-20	88.57	90	**91.43**	30-20	89.19	90.54	**91.89**
25-20	85.71	87.14	**88.57**	25-20	87.84	87.84	**89.19**
From-To(bps)	f_{Pinter}	f_{Binter}	f_{inter}	From-To(bps)	f_{Pinter}	f_{Binter}	f_{inter}
300 K-1 M	95.71	96.43	**97.14**	10 M–40 M	94.59	97.30	**98.64**
500 K-1 M	94.28	95.71	**96.43**	15 M–40 M	91.89	95.95	**98.64**
700 K-1 M	88.57	89.29	**90.71**	20 M–40 M	90.54	91.89	**95.95**
	(a)				(b)		

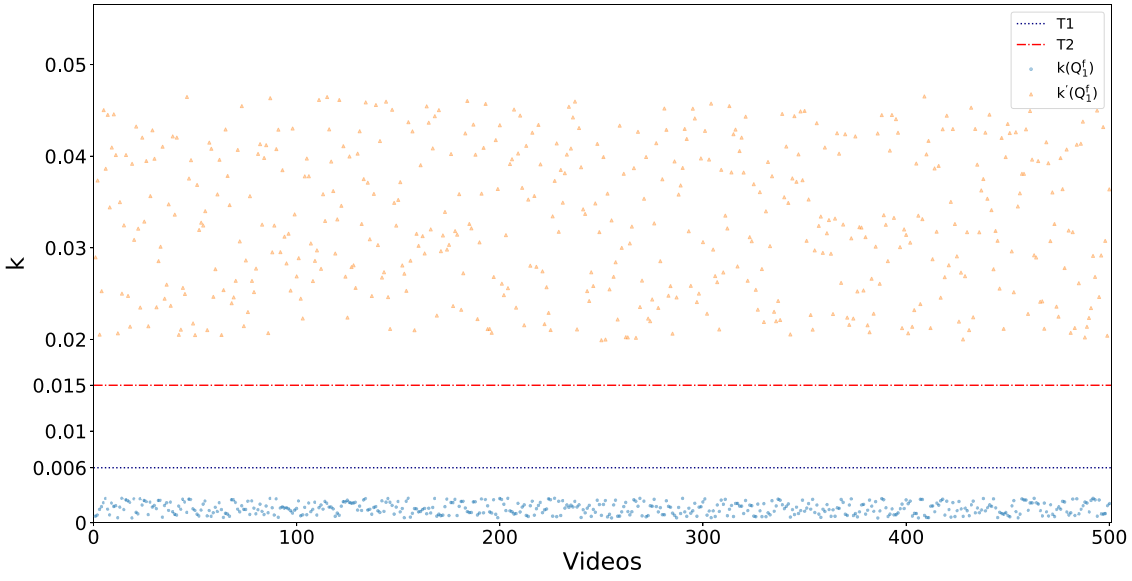

Fig. 8. The values of $k(Q_1^f)$ and $k'(Q_1^f)$ on 500 videos.

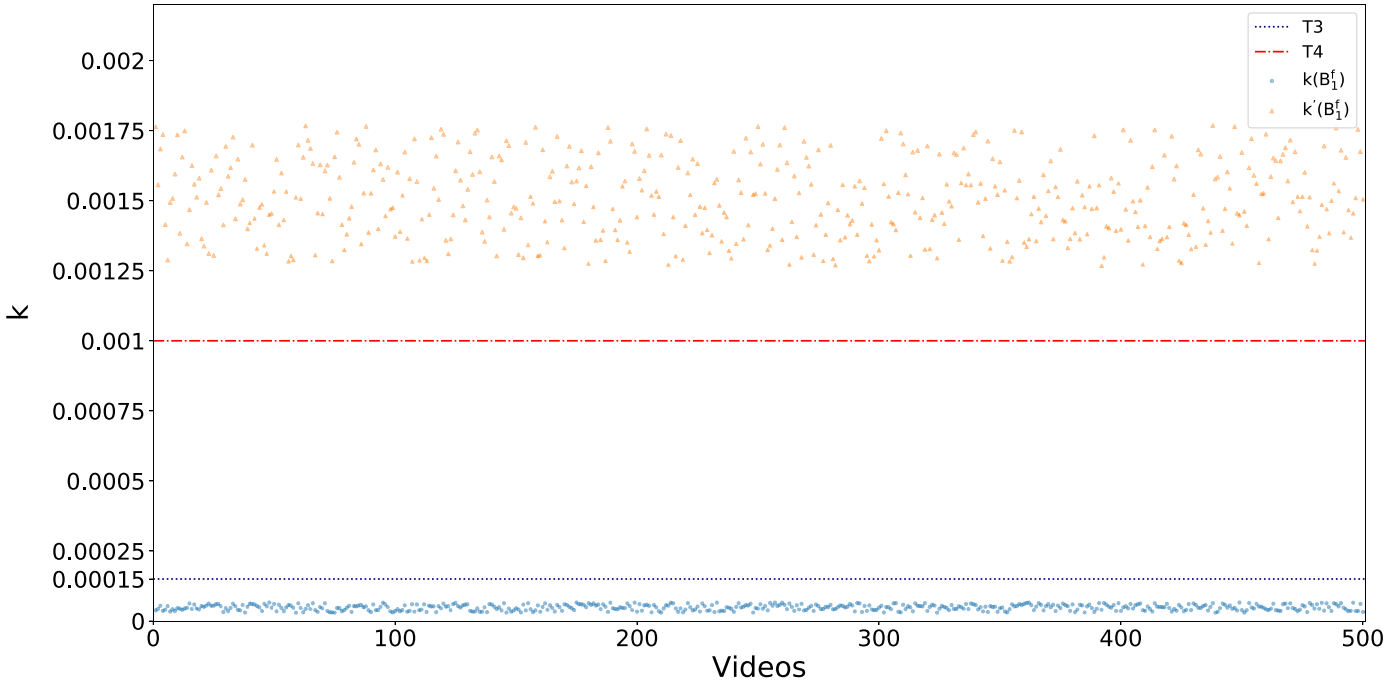

Fig. 9. The values of $k(B_1^f)$ and $k'(B_1^f)$ on 500 videos.

Table 5
Detection accuracies on intra, inter and combined feature sets in different video sets (%). (a) CIF video set. (b) 1080P video set.

QP From-To	f_{intra}	f_{inter}	*PMF*	QP From-To	f_{intra}	f_{inter}	*PMF*
45-20	92.86	96.43	**100**	45-20	94.59	98.64	**100**
40-20	91.43	95.71	**100**	40-20	94.59	97.30	**100**
35-20	90.71	92.86	**100**	35-20	91.89	95.95	**100**
30-20	89.29	91.43	**98.57**	30-20	90.54	91.89	**97.30**
25-20	85.71	88.57	**96.43**	25-20	87.84	89.19	**93.24**
From-To(bps)	f_{intra}	f_{inter}	*PMF*	From-To(bps)	f_{intra}	f_{inter}	*PMF*
300 K-1 M	95.71	96.43	**100**	10 M–40 M	95.95	98.64	**100**
500 K-1 M	94.28	95.71	**100**	15 M–40 M	93.24	97.30	**100**
700 K-1 M	88.57	89.29	**95.71**	20 M–40 M	91.89	95.95	**97.30**
	(a)				(b)		

the best compared to those on the feature set $\boldsymbol{f}_{intra}$ or $\boldsymbol{f}_{inter}$. Therefore, the combination of features from $\boldsymbol{f}_{intra}$ and $\boldsymbol{f}_{inter}$ is necessary to detect fake HD videos.

4.5. Results on different GOP size in double compression

In this section, we further verify the effectiveness of our proposed method on different GOP size in double compression. First, the first GOP size of 1080P video set in double compression is set as 15 (IBBPBBPBBPBBPBBI), and the second GOP size is set as 12 (IBBPBBPBBPBBI). Then, we exploit our method to detect fake HD videos in this 1080P video set and further estimate their original QPs and bit rates. The results are shown in Table 6. From the results, our method is efficient on different GOP size in double compression.

4.6. Results on different video contents

In this section, we further evaluate the performance of the proposed method on different video contents in the CIF video set. Research on different video content is very important because the complexity of the texture, and the speed of motion can significantly affect performance of proposed method. The results of detection accuracy with 18 different video contents are shown in Fig. 10. From the results, the average accuracy of the method we proposed is 96.39%, which means the performance of our method is effective on different content videos. The method shows relatively low performance in videos of *tempete* (93.33%), *bus* (88.89%), *mobile* (91.11%) and *stefan* (86.67%), since they have rich texture and fast movement in the videos. The change in texture and the spatial and temporal differences between blocks is not obvious when the two compression parameters are similar, which reduces the effectiveness of the proposed feature. The method is steady when the video contents involve simple texture and slow movement, especially in the first 5 videos. In a word, the proposed method performs well on different video contents.

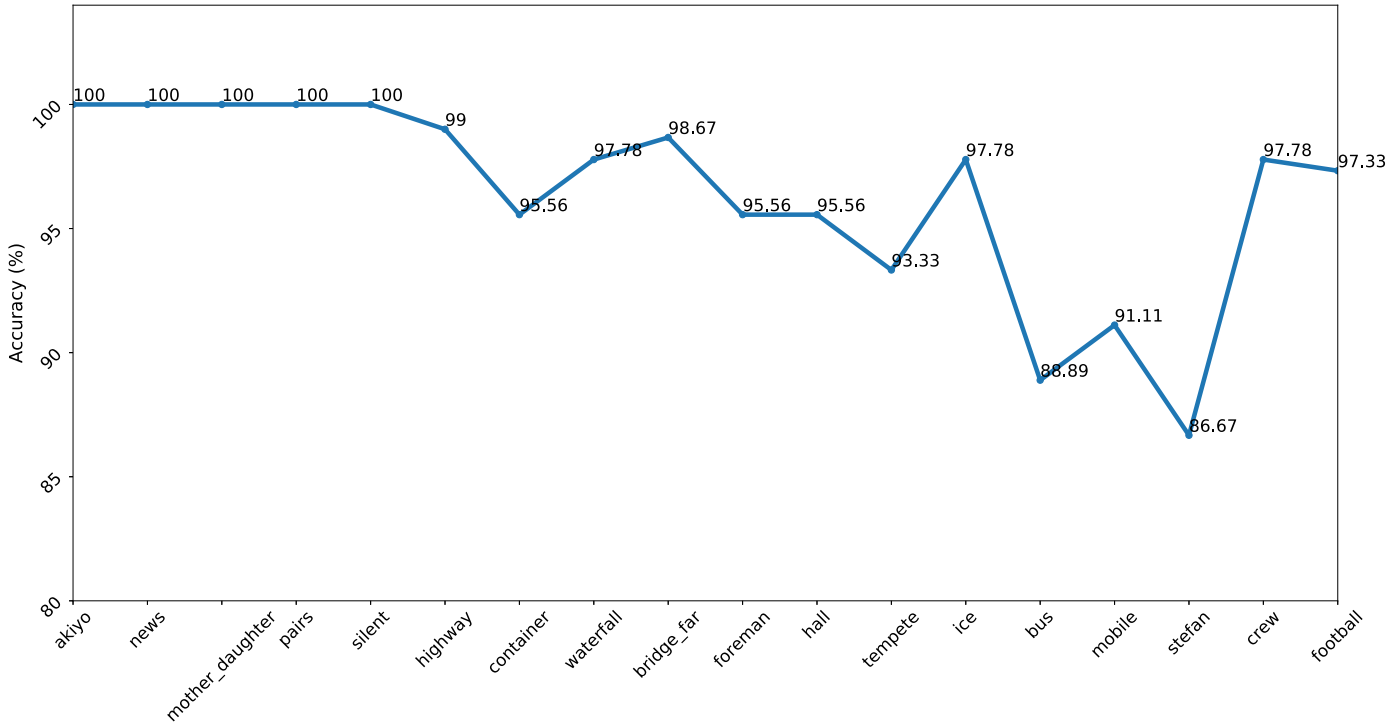

Fig. 10. Fake HD videos detection accuracy with different video contens.

Table 6
Performance on different GOP size in double compression in 1080P video set. (a) Detection accuracies on different GOP size. (b) Estimation accuracies on different GOP size.

QP From-To	Accuracy (%)	QP From-To	Accuracy (%)
45-10	**100**	45-10	**100**
40-10	**100**	40-10	**100**
35-10	**100**	35-10	**100**
30-10	**100**	30-10	**100**
25-10	**100**	25-10	**95.95**
20-10	**94.59**	20-10	**94.59**
15-10	**93.24**	15-10	**91.89**
From-To(bps)	Accuracy (%)	From-To(bps)	Accuracy (%)
10 M–40 M	**100**	10 M–40 M	**100**
15 M–40 M	**100**	15 M–40 M	**94.59**
20 M–40 M	**98.64**	20 M–40 M	**91.89**
(a)		(b)	

5. Conclusion

In this paper, we propose an effective method that exploits the frequencies of HEVC prediction modes to detect fake HD videos, and estimate their original QPs and bit rates. A 4-D feature set and a 6-D feature set are extracted from HEVC intra prediction modes and inter prediction modes, respectively. These two feature sets are combined into a 10-D feature vector which is fed to the classifier. The experimental results show that the performance of the proposed method outperforms state-of-the-art works. In future work, we will consider more complex situations and focus on the effectiveness in detecting the fake HD videos on social networks and video websites.

Declaration of Competing Interest

The authors declared that we have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgments

This work was supported in part by the National Key Research and Development of China (2018YFC0807306), National NSF of China (61672090, 61532005), and Fundamental Research Funds for the Central Universities (2018JBZ001).

References

[1] H. Ling, L. Wang, F. Zou, Z. Lu, P. Li, Robust video watermarking based on affine invariant regions in the compressed domain, Signal Process. 91 (2011) 1863–1875.
[2] A. Cedillo-Hernandez, M. Cedillo-Hernandez, M. Garcia-Vazquez, M. Nakano-Miyatake, H. Perez-Meana, A. Ramirez-Acosta, Transcoding resilient video watermarking scheme based on spatio-temporal HVS and DCT, Signal Process. 97 (2014) 40–54.
[3] W. Wang, H. Farid, Exposing digital forgeries in video by detecting double MPEG compression, in: Proceedings of the 8th Workshop on Multimedia and Security, ACM, 2006, pp. 37–47.
[4] A. Gironi, M. Fontani, T. Bianchi, A. Piva, M. Barni, A video forensic technique for detecting frame deletion and insertion, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014, pp. 6226–6230.
[5] J.A. Aghamaleki, A. Behrad, Inter-frame video forgery detection and localization using intrinsic effects of double compression on quantization errors of video coding, Signal Process Image Commun. 47 (2016) 289–302.
[6] C. Feng, Z. Xu, S. Jia, W. Zhang, Y. Xu, Motion-adaptive frame deletion detection for digital video forensics, IEEE Trans. Circuits Syst. Video Technol. 27 (2017) 2543–2554.
[7] J. Zhang, Y. Su, M. Zhang, Exposing digital video forgery by ghost shadow artifact, in: Proceedings of the First ACM Workshop on Multimedia in Forensics, ACM, 2009, pp. 49–54.
[8] S. Kim, J.Y. Choi, S. Han, M.R. Yong, Adaptive weighted fusion with new spatial and temporal fingerprints for improved video copy detection, Signal Process. Image Commun. 29 (2014) 788–806.
[9] H. Yao, S. Song, C. Qin, Z. Tang, X. Liu, Detection of double-compressed h. 264/AVC video incorporating the features of the string of data bits and skip macroblocks, Symmetry 9 (2017) 313.
[10] X. Jiang, P. He, T. Sun, F. Xie, S. Wang, Detection of double compression with the same coding parameters based on quality degradation mechanism analysis, IEEE Trans. Inf. ForensicsSecur. 13 (2018) 170–185.
[11] A.A. Elrowayati, M. Abdullah, A.A. Manaf, A.S. Alfagi, Tampering detection of double-compression with the same quantization parameter in HEVC video streams, in: 2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), IEEE, 2017, pp. 174–179.
[12] X. Jiang, Q. Xu, T. Sun, B. Li, P. He, Detection of HEVC double compression with the same coding parameters based on analysis of intra coding quality degradation process, IEEE Trans. Inf. ForensicsSecur. (2019) Early Access Article, doi:10.1109/TIFS.2019.2918085.
[13] FFmpeg. Accessed: Jan. 2014. [Online]. Available: http://www.ffmpeg.org/download.html.
[14] X265. [Online]. Available: http://x265.org (libx265).
[15] G.J. Sullivan, J.R. Ohm, W.J. Han, T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, IEEE Trans. Circuits Syst. Video Technol. 22 (2013) 1649–1668.
[16] M. Wien, High efficiency video coding, 2015.
[17] T.Q. Vinh, Y.C. Kim, S.H. Hong, Frame rate up-conversion using forward-backward jointing motion estimation and spatio-temporal motion vector smoothing, in: International Conference on Computer Engineering & Systems, 2010.
[18] S. Bian, W. Luo, J. Huang, Detecting video frame-rate up-conversion based on periodic properties of inter-frame similarity, Multimed Tools Appl. 72 (2014) 437–451.
[19] X. Ding, G. Yang, R. Li, L. Zhang, Y. Li, X. Sun, Identification of motion-compensated frame rate up-conversion based on residual signals, IEEE Trans. Circuits Syst. Video Technol. 28 (2018) 1497–1512.
[20] T. Sun, W. Wang, X. Jiang, Exposing video forgeries by detecting MPEG double compression, in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2012, pp. 1389–1392.
[21] S. Bian, W. Luo, J. Huang, Exposing fake bitrate video and its original bitrate, in: 2013 IEEE International Conference on Image Processing, IEEE, 2013, pp. 4492–4496.
[22] S. Bian, W. Luo, J. Huang, Exposing fake bit rate videos and estimating original bit rates, IEEE Trans. Circuits Syst. Video Technol. 24 (2014) 2144–2154.
[23] A. Costanzo, M. Barni, Detection of double AVC/HEVC encoding, in: 2016 24th European Signal Processing Conference (EUSIPCO), IEEE, 2016, pp. 2245–2249.
[24] S. Bian, H. Li, T. Gu, A.C. Kot, Exposing video compression history by detecting transcoded HEVC videos from AVC coding, Symmetry 11 (2019) 67.
[25] Q. Li, R. Wang, D. Xu, Detection of double compression in HEVC videos based on TU size and quantised DCT coefficients, IET Inf. Secur. 13 (2018) 1–6.
[26] M. Huang, R. Wang, J. Xu, D. Xu, Q. Li, Detection of double compression for HEVC videos based on the co-occurrence matrix of DCT coefficients, in: International Workshop on Digital Watermarking, Springer, 2015, pp. 61–71.
[27] X. Liang, Z. Li, Y. Yang, Z. Zhang, Y. Zhang, Detection of double compression for HEVC videos with fake bitrate, IEEE Access 6 (2018) 53243–53253.
[28] P. Bestagini, S. Milani, M. Tagliasacchi, S. Tubaro, Video codec identification extending the idempotency property, in: European Workshop on Visual Information Processing (EUVIP), IEEE, 2013, pp. 220–225.
[29] YUV video sequences. Accessed: August. 2015. [Online]. Available: http://www.trace.eas.asu.edu/yuv/index.html.
[30] Ensemble classifier. Accessed: September. 2013. [Online]. Available: http://dde.binghamton.edu/download/ensemble/.