# A CNN-Based HEVC Video Steganalysis Against DCT/DST-Based Steganography

Zhenzhen Zhang[1], Henan Shi[2], Xinghao Jiang[2(✉)], Zhaohong Li[3], and Jindou Liu[3]

[1] School of Information Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China

[2] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
xhjiang@sjtu.edu.cn

[3] School of Electronic and Information Engineering, Beijing JiaoTong University, Beijing 100044, China

**Abstract.** The development of video steganography has sparked ever-increasing concerns over video steganalysis. In this paper, a novel steganalysis approach against Discrete Cosine/Sine Transform (DCT/DST) based steganography for High Efficiency Video Coding (HEVC) video is proposed. The distortion of DCT/DST-based HEVC steganography and the impact on pixel value of HEVC videos is firstly analyzed. Based on the analysis, a convolutional neural network (CNN) is designed. The proposed CNN is mainly composed of three parts, i.e. residual convolution layer, feature extraction and binary classification. In the feature extraction part, a steganalysis residual block module and a squeeze-and-excitation (SE) block are designed to improve the network's representation ability. In comparison to the existing steganalysis methods, experimental results show that the proposed network performs better to detect DCT/DST-based HEVC steganography.

**Keywords:** Video steganalysis · Steganography · DCT/DST · HEVC · CNN

## 1 Introduction

Steganography is an art and science of covert communication. It conveys secret information without arousing any suspicion, and has been widely used by military institutions, government departments, financial institutions, and so on. However, the abuse of steganography may bring potential hazards to the safe of society. Steganalysis, the counter measure to steganography, solves the problem and detects the presence of secret information in the cover, which can prevent the leakage of confidential information, reveal illegal information, and is becoming a hot topic in the area of information security.

With the rapid development of Internet and advanced video compression technique, digital video has becoming one of the most popular media in peoples' lives, which makes video steganalysis drawing more and more attention of researchers. In the existing video steganalysis technology, most steganalysis algorithms [1–4] use handcrafted


P. Gladyshev et al. (Eds.): ICDF2C 2021, LNICST 441, pp. 265–276, 2022.
https://doi.org/10.1007/978-3-031-06365-7_16

classification features, i.e. the extraction of classification features in these methods rely solely on the experience of relevant researchers. Sheng et al. [1] utilized the change rate of Prediction Unit (PU) partition types before and after re-compression to detect intra-prediction mode based HEVC video steganography. Zarmehi et al. [2] estimated the cover frames and computed features both from video frames and residual matrix. Huang et al. [3] designed Combination-based Group Proportion and Difference (CGPD) classification features. Zhai et al. [4] analyzed the common statistical characteristics of the motion vector consistency (MVC) and constructed an effective universal feature set for steganographic methods.

Deep learning technology has been widely used in computer vision and other fields in recent years. Compared with traditional handcrafted features, convolutional neural network (CNN) can extract multi-dimensional abstract features by self-learning and do not depend on the experience of experts, which makes deep learning show stronger advantages. However, the majority of steganalysis methods based on deep learning are proposed for digital image steganography, fewer are focus on digital video steganography. Liu et al. [5] designed the first CNN, named Noise Residual Convolutional Neural Network (NR-CNN) for H.264 steganography. Then Huang et al. [6] constructed a novel feature extraction neural network model, which can estimate the embedding rate of steganography. Furthermore, HEVC, as the new and promising video coding standard, owns much fewer steganalysis methods than its corresponding steganography. Therefore, in this paper, we will focus on HEVC videos, and design new CNN-based steganalysis method.

Video steganography usually hide data in the compression process. According to the hiding location of secret data, video steganography can classified into 5 categories, i.e. DCT/DST based [7, 8], motion vector based [9, 10], intra prediction mode based [11, 12], inter-block partition type based [13, 14], and bitstream based [15, 16] steganography. DCT/DST based steganography is a common category, which makes its corresponding steganalysis method a research topic. However, compared with other steganalysis, there are limited number of DCT/DST domain HEVC video steganalysis, especially motion vector and inter-block partition type HEVC steganalysis. Shi et al. [17] developed a combination feature of Special Frames Extraction (SFE) and a temporal to spatial transformation to detect DCT/DST based HEVC steganography. On one hand, the classification features used in [17] are handcrafted. On the other hand, the classification accuracy can be further improved. Thus in this paper, we will analyze DCT/DST HEVC steganography and develop a novel and effective steganalysis method based on CNN.

The rest paper is organized as follows. Section 2 analyzes the distortion of DCT/DST-based HEVC steganography and its impact on pixel value of HEVC videos. Section 3 describes the designed CNN, and the experiment results and analysis are given in Sect. 4. Finally, the paper is concluded in Sect. 5.

## 2 Analysis of Pixel Change in DCT/DST-Based HEVC Steganographic Video

### 2.1 Analysis on Intra-frame Distortion of DCT/DST-Based HEVC Steganography

As a new generation video coding standard, HEVC brings in various new techniques to improve compression efficiency. One of the innovations is the flexible partition of blocks. In HEVC, Coding Tree Unit (CTU) is the basic processing unit, and each coding picture is firstly divided into non-overlapping CTUs. Then each CTU is further sub-divided into coding units (CU) following the quadtree structure. When coding each CU, flexible prediction unit (PU) and transform unit (TU) is adopted to get the optimal rate distortion. PU and TU decide the prediction mode and transform size of the CU, respectively.

Figure 1 shows DCT/DST-based steganography in HEVC coding process, and the red block denotes the position of DCT/DST-based steganography. DCT/DST-based steganography often modifies the quantized transform coefficients, and thus closely tied to TUs. HEVC supports four TU sizes: $4 \times 4$, $8 \times 8$, $16 \times 16$, and $32 \times 32$. For $4 \times 4$ sized TU, DST is adopted to implement transformation, and DCT is adopted for other sized TUs. For simplicity, $4 \times 4$ sized TU is taken as an example to analyze the distortion of steganography in this paper.

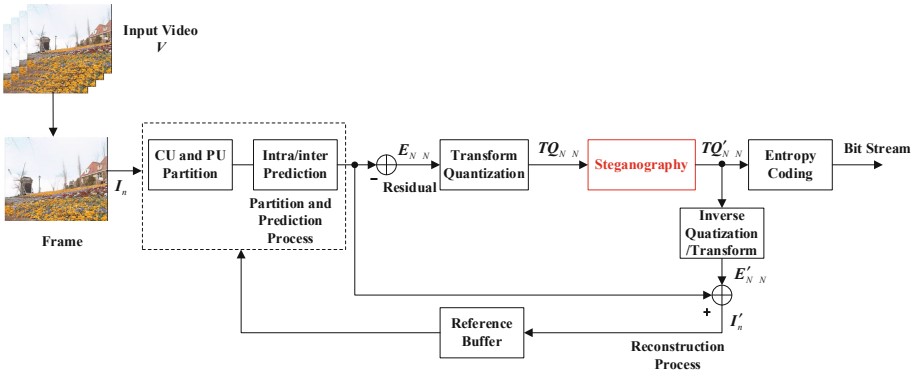



**Fig. 1.** DCT/DST-based steganography in HEVC coding process.

For a given video $\boldsymbol{V}$, it can be considered as a frame sequence $\boldsymbol{I}_n$, $(n = 1, 2, \cdots, T)$ with the length of $T$ frames. For each frame $\boldsymbol{I}_n$, partition and prediction process is performed successively, and then multiple residual block $\boldsymbol{E}_{N\times N}$ can be obtained, where $N \times N$ means the size of the PU block. Take $4 \times 4$ residual block $\boldsymbol{E}_{4\times4}$ as example, after prediction, DST transform and quantization is following. Let $\boldsymbol{TQ}_{4\times4}$ denotes the quantized DST of $\boldsymbol{E}_{4\times4}$, it is calculated as Eq. (1), where $Q$ means the quantization step determined by the Quantization Parameter (QP), $\boldsymbol{M}$ is shown in Eq. (2), and $\boldsymbol{M}^T$ is the transpose matrix of $\boldsymbol{M}$.

$$\boldsymbol{TQ}_{4\times4} = \left(\boldsymbol{M}\boldsymbol{E}_{4\times4}\boldsymbol{M}^T\right) \times \frac{1}{Q} \tag{1}$$

$$\boldsymbol{M} = \begin{bmatrix} P & Q & R & S \\ R & R & 0 & -R \\ S & -P & -R & Q \\ Q & -S & R & -P \end{bmatrix} \tag{2}$$

In DCT/DST-based steganography, $\boldsymbol{TQ}_{4\times4}$ will be modified. Directly modifying $\boldsymbol{TQ}_{4\times4}$ would cause error propagation, so measures must be taken to avoid error propagation in DCT/DST-based steganography. Equation (3) shows the hiding rule of reference [18], where $m$ denotes the secret information to be hided, and $\boldsymbol{TQ'}_{4\times4}$ means the quantized DST coefficients after modification.

$$\begin{aligned} \Delta\boldsymbol{TQ}_{4\times4} &= \boldsymbol{TQ'}_{4\times4} - \boldsymbol{TQ}_{4\times4} \\ &= \begin{bmatrix} m & 0 & -m & m \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{3}$$

After steganography module, $\boldsymbol{TQ'}_{4\times4}$ will be inverse quantized and inverse transformed to get reconstructed residual block $\boldsymbol{E'}_{4\times4}$ which will be used to obtain reference pictures for prediction, as shown in Eq. (4).

$$\begin{aligned} \boldsymbol{E'}_{4\times4} &= \boldsymbol{M}^{-1}\left(\boldsymbol{TQ'}_{4\times4} \times Q\right)\left(\boldsymbol{M}^{T}\right)^{-1} \\ &= \boldsymbol{M}^{-1}\left(\left(\boldsymbol{TQ}_{4\times4} + \Delta\boldsymbol{TQ}_{4\times4}\right) \times Q\right)\left(\boldsymbol{M}^{T}\right)^{-1} \\ &= \boldsymbol{E}_{4\times4} + \boldsymbol{M}^{-1}\left(\Delta\boldsymbol{TQ}_{4\times4} \times Q\right)\left(\boldsymbol{M}^{T}\right)^{-1} \end{aligned} \tag{4}$$

From Eq. (4), we can see that there exist an error $\Delta\boldsymbol{E}_{4\times4}$ between the residual block $\boldsymbol{E}_{4\times4}$ and its reconstruction version $\boldsymbol{E'}_{4\times4}$, and it can be obtained by Eq. (5).

$$\begin{aligned} \Delta\boldsymbol{E}_{4\times4} &= \boldsymbol{E'}_{4\times4} - \boldsymbol{E}_{4\times4} \\ &= \boldsymbol{M}^{-1}\left(\Delta\boldsymbol{TQ}_{4\times4} \times Q\right)\left(\boldsymbol{M}^{T}\right)^{-1} \\ &= Q \times m \times \begin{bmatrix} 0 & 0 & 3PR & 0 \\ 0 & 0 & 3QR & 0 \\ 0 & 0 & 3R^2 & 0 \\ 0 & 0 & 3RS & 0 \end{bmatrix} \end{aligned} \tag{5}$$

$\Delta\boldsymbol{E}_{4\times4}$ is caused by the modification introduced by the steganography. We can see that the error is tiny, but it will still cause distortion in pixel domain, because the error will be further integrated to the reference frame and have an impact on the following coding pictures. In the next subsection, we will test the effect of DCT/DST-based steganography on pixel values.

### 2.2 Analysis of Pixel Value Changes in DCT/DST-Based Steganography

In order to test the effect of DCT/DST-based steganography on video pixel values, reference [19] which modifies quantized DST to hide information was used as the test steganography. In the experiment, four YUV videos with different resolution and scenes were used as test samples. The videos were compressed with and without steganography, respectively, and the pixel value change ratio between these two video versions was calculated. Please note that only the luminance component was tested in the experiment. As shown in Eq. (6), $P(x, y)$ and $P'(x, y)$ represent the luminance value at row x and column y in one picture with and without steganography, respectively. $F(x, y)$ means the pixel value change ratio caused by the steganography. The larger the $F(x, y)$, the greater the influence of steganographic algorithm on the video pixel value is.

$$F(x, y) = \frac{|P(x, y) - P'(x, y)|}{255} \times 100\% \tag{6}$$

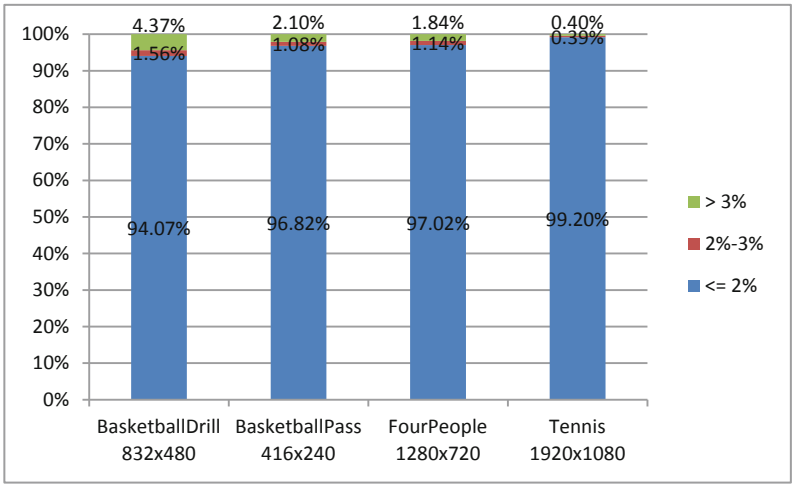



**Fig. 2.** The pixel change ratios of four tested videos.

To intuitively display the results, the percentages of pixels with change rate $F(x, y)$ below 2%, between 2% and 3%, and greater than 3% in each tested video were calculated and shown in Fig. 2. We can get that the percentages of pixels with change rate below 2%, between 2% and 3%, and greater than 3% are 96.78%, 1.04%, 2.18% at average, respectively. The result indicates that the change rate of nearly 97% pixels is below 2%, and the pixel change is tiny after steganography. However, the tiny change inspired us to explore CNN to learn more abstract features to distinguish DCT/DST-based steganography.

## 3 Proposed Steganalysis Network

The steganalysis network proposed in this paper is shown in Fig. 3. Single channel of gray images with size of 128 × 128 are taken as input of the network. According to the

shape of input, the shape of each layer is also marked in Fig. 3. The proposed network is mainly composed of three parts, i.e. high pass filter convolution layer, feature extraction and binary classification, and the three parts are marked by black dotted bordered box in Fig. 3. The high pass filter convolution layer is marked in yellow in Fig. 3, and is used to remove the influence of image content to the training network. The feature extraction part consists of five convolution layers, three pooling layers, an integration of steganalysis residual block module and a squeeze-and-excitation (SE) block. The binary classification part includes a full connection layer and softmax layer.

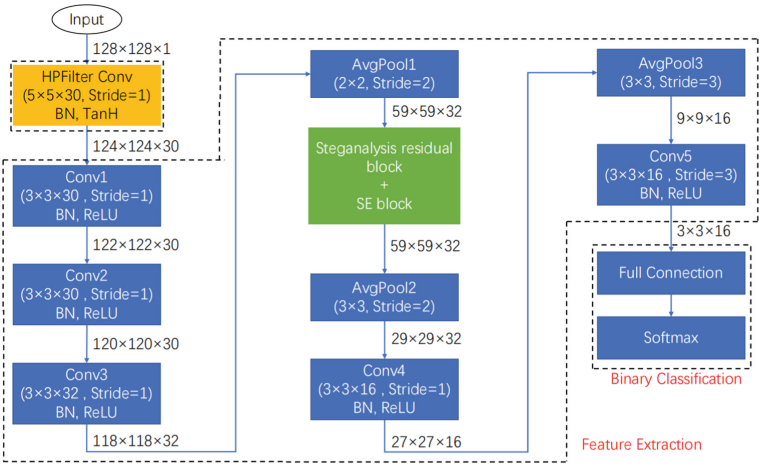



**Fig. 3.** The proposed steganalysis network structure.

### 3.1 High Pass Filter Convolution Layer

The yellow block in Fig. 3 denotes the high pass filter convolution layer which is consisted of 30 convolution kernel with size of $5 \times 5$. The initial value of each weight of the convolution kernel is set according to high pass filter core in SRM [20] which is suitable for pixel-based steganalysis. The introduction of the 30 high pass filters makes the proposed network focus on the noise caused by the steganography, rather than the content of the image itself, and thus set a good starting point for the training process of the proposed network.

In addition, it is a crucial work to select an appropriate activation function since it introduces nonlinear factors into the neural network, and thus makes the network can approximate various nonlinear models. In this paper, TanH is adopted as activation function of the high pass filter convolution layer. Compared with common used Relu activation function, it won't set the negative value of samples as zero and lead to information loss, and much more suitable for capturing features for tiny changes caused by the steganography.

### 3.2 Steganalysis Residual Block

The green block in Fig. 3 is the integration of steganalysis residual block and SE block. Steganalysis residual block adopts the ResNet structure. For convolutional neural network, generally speaking, the deeper the network is, the better the classification accuracy is. But deepening the network always results in problems such as gradient disappearance, and causes performance degradation of the network. Therefore, ResNet is designed to solve the problem, and we introduce it to the proposed network, and the structure is shown in Fig. 4. Furthermore, the steganalysis residual block has another function. It is used to further remove the video content and makes its output closer to the secret information.

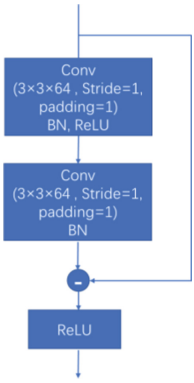
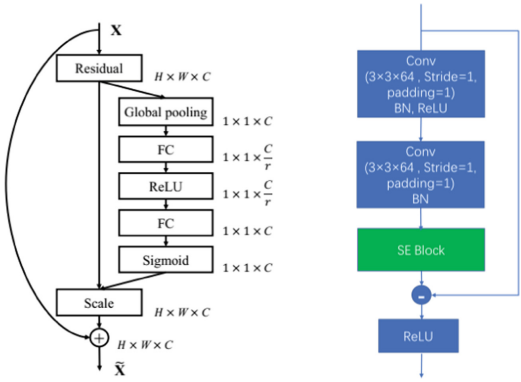



**Fig. 4.** The structure of the steganalysis residual block.

### 3.3 Squeeze-and-Excitation (SE) Block

Traditional CNNs are trained to capture spatial correlations between features, while Hu et al. [21] devised a new CNN unit from a different aspect and termed as SE block. SE block can explicitly model the interdependencies between each channel of convolutional features, and efficiently improve the representation quality of CNN with slight computational burden. Thus we adopt the SE block to our CNN to improve the network's representation quality. The structure of SE block and the position of the SE block in our CNN are displayed in Fig. 5. Please note that according to the illustration of SE block's application in ResNet in reference [21], the SE block should be integrated to the steganalysis residual block, rather than independent from it.

In conclusion, for a $128 \times 128$-sized gray image, it is firstly sent into the high pass filter layer which including 30 convolution kernel, and $128 \times 128 \times 30$ feature maps can be obtained as the output of the high pass filter layer. Then the $128 \times 128 \times 30$ feature maps are taken into the feature extraction part. According to the flow shown in Fig. 3, after experiencing a series of convolution layer, pooling layer, especially the steganalysis residual block module and SE block, $3 \times 3 \times 16$ feature maps can be obtained as the output of feature extraction part. Then $3 \times 3 \times 16$ feature maps are put into binary classification part, and the decision whether it is DCT/DST-based steganography is given as the output of the network.




(a) SE block (from reference [21]) (b) position of SE block in the proposed CNN

**Fig. 5.** Structure of SE block and the position of the SE block in the proposed CNN.

## 4 Experiment Results and Analysis

### 4.1 Data Set and Experimental Setup

In our experimental part, 110 4:2:0 YUV video sequences with resolution of 1920 × 1080 are used as video set. The video set includes 22 different kinds of scenes, and the length of each video sequence is 100 frames. In order to construct a video set that is suitable for CNN, each YUV sequence is separated into non-overlapped subsequences with size of 128 × 128. Then the subsequences are used to construct no-steganography set and stenography set for CNN. For both no-steganography set and stenography set, X265(version 2.8) and HM16.7 are adopted to compress and de-compress each subsequence. The GOP structure is "IPPP", and the adopted QP is 22.

The steganography method we adopted is reference [19] since it is an efficient and classic method in DCT/DST based HEVC steganography. The steganography set and no-steganography set are used as positive samples and negative samples of CNN, respectively. Please note that because reference [19] focuses on the DST coefficients of I picture, only I pictures of each subsequence are selected as positive samples and negative samples in the experiments.

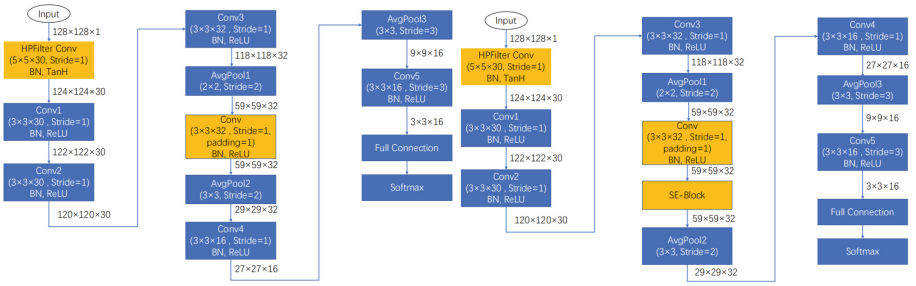
Before inputting samples into the proposed CNN, each sample is required to be decoded into portable graymap file format (PGM) gray image. In addition, in order to maximize the difference between the sample content, not all the 128 × 128 subsequences are used as samples. Since each 1920 × 1080 YUV picture can be divided into 120 128 × 128 blocks, the first 20 128 × 128 blocks of the first I picture are selected, and the next 20 128 × 128 blocks of the second I picture are selected, and the selecting process is repeated until the 120 128 × 128 blocks are selected. Then the selecting process is implemented on each I picture. Finally 55,000 positive samples and 55,000 negative samples can be obtained, among which 49000 are used for training and 6000 for testing.

In this paper, deep learning platform PyTorch is selected to construct the proposed CNN. Except for ResConv in Fig. 3, the weights and offsets of other convolution layers are initialized with Kaiming uniform distribution. Optimizer AdaDelta with rho 0.95,

weight decay $5 \times 10^{-4}$, eps $1 \times 10^{-8}$, and initial learning rate 0.1 is employed. Cross entropy loss function is adopted as the cost function, and batch size and epoch is set as 32, and 200, respectively.

### 4.2 Experimental Results

In this section, the performance of the proposed CNN is illustrated, and the efficiency of the steganalysis residual block and SE block is described. For convenience of description, the CNN without steganalysis residual block and SE block is abbreviated as no-SRB-SE, and the CNN with SE block but without steganalysis residual block is abbreviated as SE-no-SRB. The no-SRB-SE and SE-no-SRB CNN structures are described in Fig. 6.




(a) no-SRB-SE CNN structure (b) SE-no-SRB CNN structure

**Fig. 6.** no-SRB-SE and SE-no-SRB CNN structure.

The numbers of parameters and detection accuracy of each CNN structure are listed in Table 1. We can see that the no-SRB-SE CNN owns the lowest detection accuracy 91.22%. After adding the SE block to no-SRB-SE CNN, the detection accuracy is increased by 0.38%, and reaches 91.6%. The reason is that SE block can integrate the correlations between each channel of convolutional features to spatial correlations between features, and increase the representation ability of the CNN. Furthermore, by using the steganalysis residual block, the detection accuracy can be further improved by 0.21%, and achieves 91.81%. The result shows the efficiency of the steganalysis residual block.

**Table 1.** The numbers of parameters and detection accuracy of each CNN structure.

| CNN structure | no-SRB-SE | SE-no-SRB CNN | Proposed CNN |
|---|---|---|---|
| Parameters | 43406 | 43534 | 52846 |
| Accuracy | 91.22% | 91.60% | 91.81% |

### 4.3 Comparative Analysis

To evaluate the performance of the proposed CNN, two steganalysis methods to detect DCT/DST based steganography methods are selected for comparison. The first one, called SFE-AU [17], is the latest work to detect DCT/DST based steganography and uses handcrafted classification features. The second one, called NRCNN [5], is the first video steganalysis CNN and is an universal steganalysis method. Table 2 describes the comparative results of the three methods.

**Table 2.** The detection accuracy of the three methods.

| Method | SFE-AU | NRCNN | Proposed CNN |
|---|---|---|---|
| Accuracy | 70.76% | 91.37% | 91.81% |

It can be seen from Table 2 that compared with [17] which uses handcrafted classification features, the detection accuracy of NRCNN and the proposed CNN based on deep learning technology is 20.61% and 21.05% higher, respectively. The results indicate that CNN can capture more abstract steganalysis features than traditional handcrafted features. In addition, the detection accuracy of the proposed method is 0.44% higher than that of NRCNN, indicating that the proposed CNN has stronger steganalysis ability as it can capture features from both pixel domain and channel domain.

## 5 Conclusion

In this paper, targeting on detecting DCT/DST based HEVC steganography, a novel steganalysis algorithm is proposed by introducing the CNN. In the proposed CNN, high pass filter convolution layer consisting of thirty convolution kernels is firstly adopted to reduce the impact of image content on the network. Then steganalysis residual block and SE block are introduced to improve the network's representation ability and get better classification accuracy. In the experimental parts, the latest traditional steganalysis method which uses handcrafted features, and one universal steganalysis method based on CNN are used as comparison methods to evaluate the performance of the proposed method. The comparative results show that the proposed steganalysis performs better than other two existing methods. In our future work, we will further modify the structure of the proposed CNN and improve its detection accuracy.

**Acknowledgement.** This work is funded by the National Key R&D Program of China (2018YFC0831405), Joint Funding Project of Beijing Municipal Commission of Education and Beijing Natural Science Fund Committee (KZ201710015010), The Scientific Research Common Program of Beijing Municipal Commission of Education (No. KM202110015004, No. KM202010015001, No. KM202010015009), and Initial funding for the Doctoral Program of BIGC (27170120003/037, 27170120003/020).

## References


1. Sheng, Q., Wang, R., Huang, M., et al.: A prediction mode steganalysis detection algorithm for HEVC. J. Optoelectron. Laser **28**(4), 433–440 (2017)
2. Zarmehi, N., Akhaee, M.: Digital video steganalysis toward spread spectrum data hiding. IET Image Proc. **10**(1), 1–8 (2016)
3. Huang, K., Sun, T., Jiang, X., Dong, Y., Fang, Q.: Combined features for steganalysis against PU partition mode-based steganography in HEVC. Multimedia Tools Appl. **79**(41–42), 31147–31164 (2020). https://doi.org/10.1007/s11042-020-09435-y
4. Zhai, L., Wang, L., Ren, Y.: Universal detection of video steganography in multiple domains based on the consistency of motion vectors. IEEE Trans. Inf. Forensics Secur. **15**, 1762–1777 (2019)
5. Liu, P., Li, S.: Steganalysis of intra prediction mode and motion vector-based steganography by noise residual convolutional neural network. In: IOP Conference series: Materials Science and Engineering, vol. 719, issue 1, p. 012068. IOP Publishing (2020)
6. Huang, X., Hu, Y., Wang, Y.: Deep neural network detection method for motion-vector-based video steganography. J. South China Univ. Technol. (Nat. Sci. Ed.) **48**(8), 1–9 (2020)
7. Li, H., Wang, H., Wu, H.: Multi-classification information hiding algorithm for H.264/AVC video with high capacity in QDCT domain. J. Optoelectron.·Laser **28**(4), 404–410 (2017)
8. Nguyen, D., Nguyen, T., Hsu, F., et al.: A novel steganography scheme for video H. 264/AVC without distortion drift. Multimedia Tools Appl. **78**(12), 16033–16052 (2019)
9. Rana, S., Kamra, R., Sur, A.: Motion vector based video steganography using homogeneous block selection. Multimedia Tools Appl. **79**(9–10), 5881–5896 (2019). https://doi.org/10.1007/s11042-019-08525-w
10. Yang, J., Li, S.: An efficient information hiding method based on motion vector space encoding for hevc. Multimedia Tools Appl. **77**(10), 11979–12001 (2018)
11. Dong, Y., Sun, T., Jiang, X.: A high capacity hevc steganographic algorithm using intra prediction modes in multi-sized prediction blocks. In: Yoo, C.D., Shi, Y.-Q., Kim, H.J., Piva, A., Kim, G. (eds.) IWDW 2018. LNCS, vol. 11378, pp. 233–247. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11389-6_18
12. Wang, J., Wang, R., Xu, D., Li, W.: An information hiding algorithm for HEVC based on angle differences of intra prediction mode. JSW **10**(2), 213–221 (2015)
13. Yang, Y., Li, Z., Xie, W., Zhang, Z.: High capacity and multilevel information hiding algorithm based on pu partition modes for HEVC videos. Multimedia Tools Appl. **78**(7), 8423–8446 (2019)
14. Tew, Y., Wong, K.: Information hiding in HEVC standard using adaptive coding block size decision. In: IEEE International Conference on Image Processing, pp. 5502–5506. IEEE (2014)
15. Xu, D., Wang, R., Shi, Y.: Data hiding in encrypted H.264/AVC video streams by codeword substitution. IEEE Trans. Inf. Forensics Secur. **9**(4), 596–606 (2014)
16. Zhang, H., Cao, Y., Zhao, X., Haibo, Y., Liu, C.: Data hiding in H.264/AVC video files using the coded block pattern. In: Shi, Y.Q., Kim, H.J., Perez-Gonzalez, F., Liu, F. (eds.) IWDW 2016. LNCS, vol. 10082, pp. 588–600. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-53465-7_44
17. Shi, H., Sun, T., Jiang, X., Dong, Y., Xu, K.: A HEVC video steganalysis against DCT/DST-based steganography. Int. J. Dig. Crime Forensics **13**(3), 19–33 (2021)
18. Chang, P., Chung, K., Chen, J., Lin, C., Lin, T.: A DCT/DST-based error propagation-free data hiding algorithm for HEVC intra-coded frames. J. Vis. Commun. Image Represent. **25**(2), 239–253 (2014)

19. Liu, Y., Liu, S., Zhao, H., Liu, S.: A new data hiding method for H.265/HEVC video streams without intra frame distortion drift. Multimedia Tools Appl. **78**(6), 6459–6486 (2018)
20. Ye, J., Ni, J., Yi, Y.: Deep learning hierarchical representations for image steganalysis. IEEE Trans. Inf. Forensics Secur. **12**(11), 2545–2557 (2017)
21. Hu, J., Li, S., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7132–7141. IEEE (2018)