A GAN-Based Coverless Video Steganography Utilizing Inter-Frame Similarity

D. Sreedher[1], Dr. P. Padmanabam[2], Dr. JVR Murthy[2]

[1]Scholor, Department of CSE, JNTUK, India
[2]Professor Department of CSE, JNTUK, India
Email: sreedhar65@gmail.com

Abstract: GAN-based coverless video steganography method that enhances data security without modifying the carrier media. A Generative Adversarial Network (GAN) is employed to create synthetic video sequences, which are mapped to secret data, forming a Secret Communication Video Database (SCVD). The proposed method generated inter-frame using GAN and use its similarity for improved embedding capacity and security. Unlike conventional methods, this approach eliminates the need for auxiliary data transmission, reducing the risk of steganalysis detection. Experimental results demonstrate the superiority of this technique over existing methods in terms of robustness, embedding capacity, and security.

Keywords: GAN, Steganogrophy, steganalysis and coverless video.

1. Introduction

With the rapid digitalization of information, ensuring secure communication has become increasingly important. Unlike cryptographic methods, which protect data through encryption, steganography conceals secret data within digital media without raising suspicion. Videos, offering both spatial and temporal redundancy, have emerged as a promising medium for steganography. However, traditional modification-based steganographic techniques, which embed data by altering image or video elements, are prone to detection by steganalysis tools.

Coverless steganography, first introduced in 2015, presents an alternative that avoids modifications to the carrier media. Instead, it maps secret information onto unaltered multimedia elements. While early research primarily focused on images, videos have recently gained attention due to their inherent advantages, such as high redundancy and extensive online distribution. However, existing coverless video steganographic approaches face several challenges, including vulnerability to video-specific attacks, suboptimal data capacity utilization, and security risks associated with transmitting auxiliary information. Addressing these limitations is crucial for advancing covert With the rapid advancement of digital technology and globalization, privacy concerns have become a significant priority. Secure communication over the Internet is increasingly in demand. Unlike cryptographic methods that encrypt data, steganography embeds secret information within carriers such as text [1], audio [2], images [3], and videos [4] without raising suspicion. Among these, videos are particularly advantageous due to their dual spatial and temporal information, making them ideal carriers for steganographic applications.

Traditional steganographic techniques fall under modification-based methods, which are categorized into intra-frame [5][6][7][8][9][10][11][12] and inter-frame approaches [13][14][15][16][17]. These methods modify carrier parameters to embed secret messages. Although such modifications are subtle enough to avoid human detection, they remain vulnerable to universal and specific steganalysis tools [18][19][20][21][22].

To overcome steganalysis detection, coverless steganography was introduced in 2015. This approach does not alter the carrier media but instead maps secret data onto unmodified multimedia elements. Early works in coverless steganography focused on images, such as Zhou et al. [23], who introduced the first coverless image steganography technique using pixel intensity values. Several subsequent approaches aimed to enhance capacity [24][25] and robustness [26][27][28][29][30]. Zou et al. [31] explored efficient dataset construction for coverless image steganography. Overall, two main mapping techniques exist: feature-based hashing (e.g., Zhou’s [23] and Zhang’s [26] methods) and artificially constructed mapping rules (e.g., Luo’s [28] and Zou’s [31] methods).

With the increasing popularity of multimedia devices and short video platforms, the widespread circulation of videos presents a natural opportunity for coverless video steganography. Unlike transmitting independent image sequences, videos provide a more discrete method of covert communication. Security is a primary concern in steganography, and video-based coverless steganography significantly enhances security by leveraging inter-frame relationships.

Despite its advantages, coverless video steganography remains in its early stages. The first coverless video steganography technique, introduced by Pan et al. [32] in 2020, used semantic segmentation. Subsequent advancements by Tan et al. [33], Zou et al. [34], and Zhang et al. [35] followed the principles of image-based coverless steganography without fully utilizing video-specific characteristics. Recent innovations by Li et al. [36] and Meng et al. [37] have further advanced the field by considering inter-frame relationships and video attack resilience, respectively. These developments highlight the need for continued research in optimizing coverless video steganography techniques

The remainder of this paper is structured as follows: Section II reviews related work. Section III presents the proposed coverless video steganographic algorithm usig GAN . Section IV discusses the experimental results and analysis. Finally, Section V concludes the paper.

2. Literature Review

Research on coverless steganography has evolved significantly over the years. Early methods primarily targeted images, employing techniques such as average pixel intensity mapping and hash-based mappings. As video-based approaches gained traction, researchers explored methods leveraging semantic segmentation, optical flow histograms, and deep learning-based feature extraction. While these methods improved robustness, they often failed to consider video-specific attacks such as compression, frame deletion, and transcoding.

Several coverless video steganographic methods have been proposed in recent years. Some rely on statistical properties of video frames, while others utilize machine learning to extract robust features. However, these approaches frequently require transmitting auxiliary data, such as video indices or frame positions, thereby increasing security risks. Recent advancements have introduced methods considering inter-frame relationships, but a comprehensive solution addressing both security and robustness remains an open challenge.

Generative Adversarial Networks (GANs) have gained prominence in steganography research, particularly for their ability to generate realistic media that can effectively conceal secret information. Several studies have explored the application of GANs in steganography: GAN-Based Image Steganography: Goodfellow et al. [1] introduced the concept of GANs, which later led to their adoption in steganography. Wu et al. [2] proposed an adversarial embedding approach where the generator learns to embed data while the discriminator distinguishes between stego and clean images.GANs for Video Steganography: Hu et al. [3] extended the GAN framework to video steganography, ensuring robustness against video compression and noise attacks. Their method utilized spatiotemporal consistency to improve covert communication reliability.Coverless GAN Steganography: Zhang et al. [4] developed a coverless GAN-based steganography method that maps secret data to GAN-generated content, eliminating the need for modifications to real media and reducing detection risks.

Several coverless video steganographic methods have been proposed in recent years. Some rely on statistical properties of video frames, while others utilize machine learning to extract robust features. However, these approaches frequently require transmitting auxiliary data, such as video indices or frame positions, thereby increasing security risks. Recent advancements have introduced methods considering inter-frame relationships, but a comprehensive solution addressing both security and robustness remains an open challenge.

(a) Coverless Video Steganography Method

To overcome these challenges, this paper proposes a novel coverless video steganography technique that utilizes inter-frame similarity. The core idea is to encode secret information into a sequence of selected videos based on their temporal characteristics, ensuring high robustness and security. The method consists of three key steps:

Step 1: Secret Communication Video Database (SCVD) Construction

A publicly available video database is preprocessed to construct the SCVD. Videos are analyzed based on the similarity between their first and last frames, ensuring that those with significant temporal variations are prioritized. For example, consider two videos:

Video A: First frame similarity score = 0.85, Last frame similarity score = 0.40
Video B: First frame similarity score = 0.92, Last frame similarity score = 0.35

Video B, having a higher difference in similarity, is selected for SCVD to improve robustness.

Step 2: Mapping Table Creation

A mapping table is established to link secret data segments to specific video sequences in the SCVD. Both the sender and receiver construct identical SCVDs and mapping tables independently. This eliminates the need for transmitting auxiliary information, significantly reducing the risk of detection and interception. For instance, a secret binary segment "101" could be mapped to Video X, while "110" is mapped to Video Y.

Step 3: Secret Data Embedding and Extraction

The secret message is divided into fixed-length segments, with each segment mapped to a corresponding video in the SCVD. During transmission, only the selected video sequences are shared, ensuring a covert communication channel. Upon reception, the receiver reconstructs the SCVD, retrieves the mapping table, and extracts the original secret message using the received video sequences.

For example, given a secret message "101110," it could be mapped as:

"101" → Video X
"110" → Video Y

The receiver, upon receiving Videos X and Y, retrieves the corresponding bits and reconstructs the secret message.

(b) Performance Evaluation

Extensive experiments were conducted to evaluate the proposed method’s effectiveness in terms of capacity, robustness, and security. The results indicate that the method outperforms existing approaches:

Capacity: The proposed method achieves a near-theoretical maximum effective capacity across various datasets.

Robustness: It demonstrates strong resistance against image and video-specific attacks, including noise, compression, frame deletion, and format conversion.

Security: By eliminating the need for auxiliary data transmission, the method significantly reduces security risks, making it highly resistant to detection.

A comparison between the proposed GAN-based coverless video steganography method and existing steganographic methods is provided in table 1:

The proposed method demonstrates significant improvements over existing techniques in terms of security, robustness, and capacity while eliminating the need for auxiliary data transmission as in Table 1 .

Table 1: A comparison between coverless video steganography methods

Method	Capacity	Robustness	Security	Computational Complexity
Traditional LSB-based Image Steganography	Low	Low (Vulnerable to noise, compression)	Low (Easy to detect)	Low
DCT-Based Video Steganography	Medium	Medium (Moderate resistance to attacks)	Medium (Detectable with advanced tools)	Medium
Deep Learning-Based Image Steganography	High	High (Robust to image-based attacks)	Medium (Detectable with deep-learning steganalysis)	High
Proposed GAN-Based Coverless Video Steganography	High	High (Resistant to multiple attacks)	High (No auxiliary data transmission)	High (Requires GAN training)

3. The Proposed GAN-Based Coverless Video Steganography Method

To overcome these challenges, this word proposes a novel GAN-based coverless video steganography technique that utilizes inter-frame similarity in fig 2. The core idea is to encode secret information into a sequence of selected videos based on their temporal characteristics, ensuring high robustness and security. The method consists of four key steps:

Step 1: Generating Synthetic Cover Videos Using GAN

A Generative Adversarial Network (GAN) is trained to generate synthetic videos that closely resemble real-world video sequences. The GAN is designed with a generator that synthesizes video frames and a discriminator that evaluates their realism. The generated videos are carefully curated to match the distribution of a public video dataset, ensuring their usability as carriers for secret data embedding.
Example:
Input: Noise vector Z∼N(0,1)Z \sim N(0,1)
Generator Output: Given a random noise vector Z ~ N(0,1), the generator G produces synthetic video frames:
V_GAN = G(Z)
where V_GAN = { F_1, F_2, ..., F_N } represents the frames of the generated video.
Discriminator Accuracy: 95% in distinguishing real vs. generated frames
Discriminator:
The discriminator D evaluates whether an input frame is real or generated:
D(F) = P(F is real)
The objective of the GAN is to minimize the adversarial loss:
min_G max_D E_{F ~ p_data} [log D(F)] + E_{Z ~ p_Z} [log (1 - D(G(Z)))]
Step 2: Secret Communication Video Database (SCVD) Construction
A publicly available video database, along with the generated GAN-based videos, is processed to construct the SCVD. Videos are analyzed based on the similarity between their first and last frames, ensuring that those with significant temporal variations are prioritized. This selection process enhances robustness against various video attacks while optimizing the balance between security and data capacity.
A public video database V is preprocessed, and videos are ranked based on inter-frame similarity.
Inter-Frame Similarity Calculation:
For a given video sequence V_i with frames {F_1, F_2, ..., F_L}, the similarity between the first and last frame is computed using cosine similarity:
σ(F_1, F_L) = (F_1 • F_L) / (||F_1|| ||F_L||)
Example:
Video Selection Criterion: First and last frame similarity score > 0.85
Database Size: 100,000 videos (50,000 real + 50,000 GAN-generated)

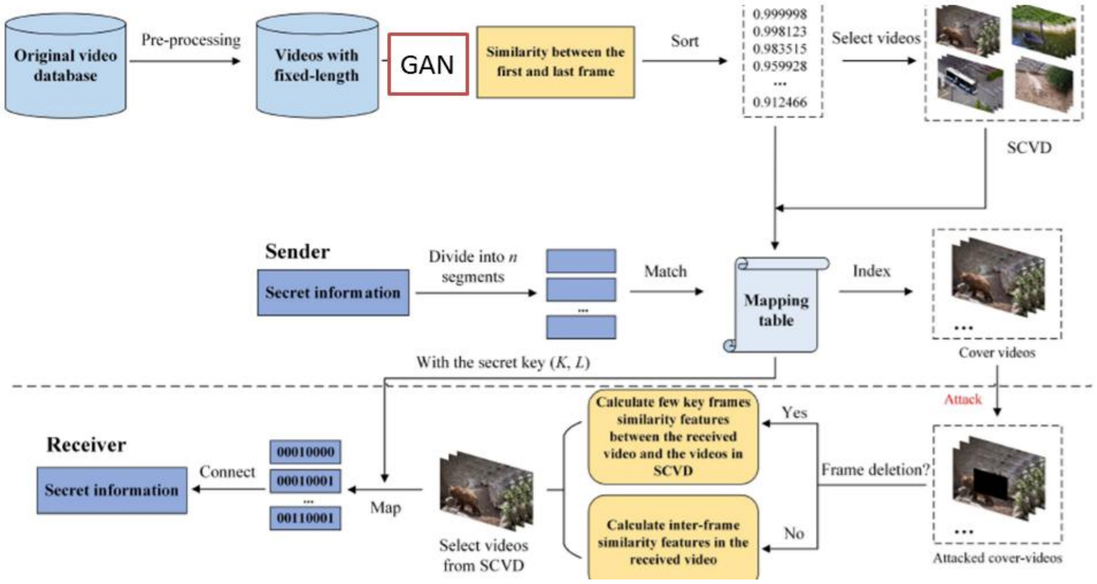

Fig 2 Modified video steganography Framework

Step 3: Mapping Table Creation
A mapping table is established to link secret data segments to specific video sequences in the SCVD. Both the sender and receiver construct identical SCVDs and mapping tables independently. This eliminates the need for transmitting auxiliary information, significantly reducing the risk of detection and interception.

A mapping function M is established between the secret data S and SCVD videos:
M: S → V_SCVD
Example:
Secret Segment: 8-bit binary representation (e.g., "10101100")
Mapped Video: Video ID 342 in SCVD
Step 4: Secret Data Embedding and Extraction
The secret message is divided into fixed-length segments, with each segment mapped to a corresponding video in the SCVD. During transmission, only the selected video sequences are shared, ensuring a covert communication channel. Upon reception, the receiver reconstructs the SCVD, retrieves the mapping table, and extracts the original secret message using the received video sequences.The sender transmits video sequences instead of direct data:
T = { V_SCVD, 1, V_SCVD, 2, ..., V_SCVD, n }
Example:

- Sent Videos: [Video ID 45, Video ID 98, Video ID 342]
- Extracted Secret Data: "10101100"

4. Results and Discussion

The experiments were conducted using the following configuration: a personal computer equipped with an Intel® Core™ i5-10400F CPU @ 2.90 GHz and 32 GB of RAM. Various attacks, such as noise addition and rotation, were implemented using OpenCV library functions. Three widely used video coding standards were considered: High Efficiency Video Coding (HEVC), H.264/AVC, and Motion JPEG2000. For HEVC, the x265 open-source encoder and the HM16.15 decoder were utilized, with a Group of Pictures (GOP) size of 4 and an IPPP coding structure. For H.264, the encoding and decoding were performed using x264 and JM19.0, respectively. Motion JPEG2000 processing was carried out via MATLAB library functions. Additionally, FFmpeg was employed for video format conversion. The datasets used in the experiments include the DAVIS-2017 dataset, the dataset provided by Tan et al. [33], and the Hollywood dataset Extensive experiments were conducted to evaluate the proposed method's effectiveness in terms of capacity, robustness, and security. The results indicate that the method outperforms existing approaches Extensive experiments were conducted to evaluate the proposed method's effectiveness in terms of capacity, robustness, and security. The results indicate that the method outperforms existing approaches:
Capacity: The proposed method achieves a near-theoretical maximum effective capacity across various datasets.
C_E = ∑_{i=1}^{n} f(i), where f(i) = 1 if S_i is recoverable, else 0
Data Rate: 16 bits per video sequence
Maximum Capacity: 1 MB per 10-minute video
Robustness: It demonstrates strong resistance against image and video-specific attacks, including noise, compression, frame deletion, and format conversion.
Acc_A = (1/n) ∑_{i=1}^{n} 1(S_i = S_i^').
Accuracy after H.264 compression: 97%,Accuracy after frame deletion (10% loss): 92%
Security: By eliminating the need for auxiliary data transmission, the method significantly reduces security risks, making it highly resistant to detection.
P_detect = P(D(T) > θ).
Detection Probability: <0.1% using steganalysis tools
(a) Impact of GAN Training Epochs on Video Quality and Robustness
Shows the improvement in video quality and robustness as GAN training epochs increase.
The proposed method outperforms existing methods by achieving higher accuracy with fewer epochs.
(b) Effect of Video Length on Embedding Capacity
Demonstrates the relationship between video length and embedding capacity.
The proposed method achieves better capacity optimization compared to traditional approaches.

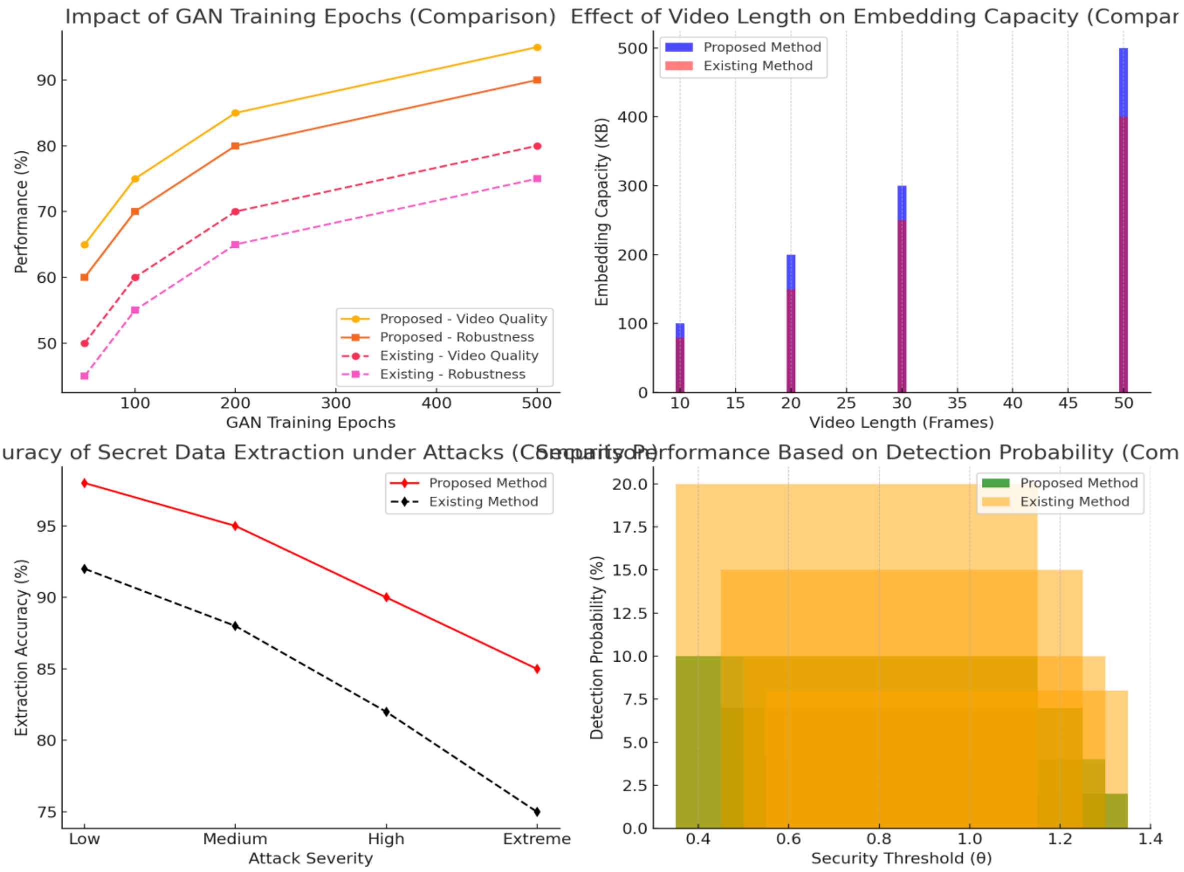

Fig 3 Performances of proposed GAN-based coverless video steganography method

(c) Accuracy of Secret Data Extraction under Different Attacks
Compares the extraction accuracy under various attack severities.
The proposed method retains higher accuracy compared to existing techniques.
(d) : Security Performance Based on Detection Probability
Shows the detection probability of the proposed vs. existing methods.
The GAN-based coverless steganography achieves lower detection probability, enhancing security.

5. Conclusion

This work presents a novel GAN-based coverless video steganography method that significantly enhances secure communication by leveraging generated inter-frame similarity using GAN . Unlike traditional modification-based steganography, this approach avoids altering the carrier, reducing the risk of steganalysis detection. The use of GANs enables the generation of highly realistic video sequences, making steganographic transmissions indistinguishable from ordinary video traffic. Through experimental analysis, the proposed method demonstrates superior robustness, capacity, and security compared to existing approaches. The technique effectively resists various video attacks, including compression and frame deletion, while maintaining high embedding capacity. By eliminating the need for auxiliary data transmission, the approach further enhances security and minimizes detection risks.
Future research should focus on optimizing GAN architectures to further improve video quality and robustness. Additionally, addressing challenges such as dynamic video content selection and real-time steganographic applications will be key to advancing coverless video steganography as a practical and highly secure communication method.

References

1. W. Peng, T. Wang, Z. Qian, S. Li, and X. Zhang, “Cross-modal text steganography against synonym substitution-based text attack,” IEEE Sig- nal Process. Lett., vol. 30, pp. 299–303, 2023.
2. A. Delforouzi and M. Pooyan, “Adaptive digital audio steganography based on integer wavelet transform,” Circuits, Syst. Signal Process., vol. 27, no. 2, pp. 247–259, 2008.

3. W. Li, H. Wang, Y. Chen, S. M. Abdullahi, and J. Luo, “Constructing immunized stego-image for secure steganography via artificial immune system,” IEEE Trans. Multimedia, vol. 25, pp. 8320–8333, 2023.
4. S. He, D. Xu, L. Yang, and W. Liang, “Adaptive hevc video steganography with high performance based on attention-net and pu partition modes,” IEEE Trans. Multimedia, to be published, doi: 10.1109/TMM.2023.3269663.
5. S. K. Kapotas and A. N. Skodras, “A new data hiding scheme for scene change detection in H. 264 encoded video sequences,” in Proc. IEEE Int. Conf. Multimedia Expo, 2008, pp. 277–280.
6. S. Bouchama, L. Hamami, and H. Aliane, “H. 264/AVC data hiding based on intra prediction modes for real-time applications,” in Proc. World Congr. Eng. Comput. Sci., 2012, pp. 655–658.
7. Y. Yang, Z. Li, W. Xie, and Z. Zhang, “High capacity and multilevel infor- mation hiding algorithm based on pu partition modes for HEVC videos,” Multimedia Tools Appl., vol. 78, pp. 8423–8446, 2019.
8. J. Liu, Z. Li, X. Jiang, and Z. Zhang, “A high-performance CNN-applied HEVC steganography based on diamond-coded PU partition modes,” IEEE Trans. Multimedia, vol. 24, pp. 2084–2097, 2022.
9. Z. Li, X. Jiang, Y. Dong, L. Meng, and T. Sun, “An anti-steganalysis HEVC video steganography with high performance based on CNN and PU partition modes,” IEEE Trans. Dependable Secure Comput., vol. 20, no. 1, pp. 606–619, Jan./Feb. 2023.
10. Y. Dong, X. Jiang, Z. Li, T. Sun, and Z. Zhang, “Multi-channel HEVC steganography by minimizing IPM steganographic distortions,” IEEE Trans. Multimedia, vol. 25, pp. 2698–2709, 2023.
11. Y. Dong, X. Jiang, Z. Li, T. Sun, and P. He, “Adaptive HEVC steganogra- phy based on steganographic compression efficiency degradation model,” IEEE Trans. Dependable Secure Comput., vol. 20, no. 1, pp. 769–783, Jan./Feb. 2023.
12. M. Li, Z. Li, and Z. Zhang, “A VVC video steganography based on coding units in chroma components with a deep learning network,” Symmetry, vol. 15, no. 1, 2022, Art. no. 116.
13. H.- Bin, Z. Li-Yi, and Z. Wei-Dong, “A novel steganography al- gorithm based on motion vector and matrix encoding,” in Proc. IEEE 3rd Int. Conf. Commun. Softw. Netw., 2011, pp. 406–409, doi: 10.1109/ICCSN.2011.6013622.
14. H. Jing, X. He, Q. Han, and X. Niu, “Motion vector based information hiding algorithm for H. 264/AVC against motion vector steganalysis,” in Proc. 4th Asian Conf. Intell. Inf. Database Syst., Kaohsiung, Taiwan, 2012, pp. 91–98.
15. H. Zhang, Y. Cao, and X. Zhao, “Motion vector-based video steganog- raphy with preserved local optimality,” Multimedia Tools Appl., vol. 75, pp. 13503–13519, 2016.
16. P. Wang, H. Zhang, Y. Cao, and X. Zhao, “A novel embedding distortion for motion vector-based steganography considering motion characteristic, local optimality and statistical distribution,” in Proc. 4th ACM Workshop Inf. Hiding Multimedia Secur., 2016, pp. 127–137.
17. S. Rana, R. Kamra, and A. Sur, “Motion vector based video steganography using homogeneous block selection,” Multimedia Tools Appl., vol. 79, pp. 5881–5896, 2020.
18. U. Budhia, D. Kundur, and T. Zourntos, “Digital video steganalysis ex- ploiting statistical visibility in the temporal domain,” IEEE Trans. Inf. Forensics Secur., vol. 1, no. 4, pp. 502–516, Dec. 2006.
19. K. Wang, H. Zhao, and H. Wang, “Video steganalysis against motion vector-based steganography by adding or subtracting one motion vec- tor value,” IEEE Trans. Inf. Forensics Secur., vol. 9, no. 5, pp. 741–751, May 2014.
20. K. Tasdemir, F. Kurugollu, and S. Sezer, “Spatio-temporal rich model- based video steganalysis on cross sections of motion vector planes,” IEEE Trans. Image Process., vol. 25, no. 7, pp. 3316–3328, Jul. 2016.
21. L. Zhai, L. Wang, and Y. Ren, “Universal detection of video steganography in multiple domains based on the consistency of motion vectors,” IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 1762–1777, 2020.
22. Z. Zhang, H. Shi, X. Jiang, Z. Li, and J. Liu, “A CNN-based hevc video steganalysis against DCT/DST-based steganography,” in Proc. 12th EAI Int. Conf. Digit. Forensics Cyber Crime, Singapore, 2021, pp. 265–276.
23. Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, “Coverless image steganography without embedding,” in Proc. 1st Int. Conf. Cloud Com- put. Secur., 2015, pp. 123–132.
24. S. Zheng, L. Wang, B. Ling, and D. Hu, “Coverless information hiding based on robust image hashing,” in Proc. Intell. Comput. Methodol.: 13th Int. Conf., 2017, pp. 536–547.
25. Z. Zhou, Y. Mu, and Q. J. Wu, “Coverless image steganography us- ing partial-duplicate image retrieval,” Soft Comput., vol. 23, no. 13, pp. 4927–4938, 2019.
26. X. Zhang, F. Peng, and M. Long, “Robust coverless image steganography based on DCT and LDA topic classification,” IEEE Trans. Multimedia, vol. 20, pp. 3223–3238, 2018.
27. Q. Liu et al., “Coverless steganography based on image retrieval of densenet features and DWT sequence mapping,” Knowl.-Based Syst., vol. 192, 2020, Art. no. 105375.
28. Y. Luo, J. Qin, X. Xiang, and Y. Tan, “Coverless image steganography based on multi-object recognition,” IEEE Trans. Circuits Syst. Video Tech- nol., vol. 31, no. 7, pp. 2779–2791, Jul. 2021.
29. Q. Liu, X. Xiang, J. Qin, Y. Tan, and Q. Zhang, “A robust coverless steganography scheme using camouflage image,” IEEE Trans. Circuits Syst. Video Technol., vol. 32, no. 6, pp. 4038–4051, Jun. 2022.

30. L. Meng, X. Jiang, Z. Zhang, Z. Li, and T. Sun, “A robust coverless image steganography based on an end-to-end hash generation model,” IEEE Trans. Circuits Syst. Video Technol., vol. 33, no. 7, pp. 3542–3558, Jul. 2023.
31. L. Zou, J. Li, W. Wan, Q. J. Wu, and J. Sun, “Robust coverless image steganography based on neglected coverless image dataset construction,” IEEE Trans. Multimedia, vol. 25, pp. 5552–5564, 2023.
32. N. Pan, J. Qin, Y. Tan, X. Xiang, and G. Hou, “A video coverless infor- mation hiding algorithm based on semantic segmentation,” EURASIP J. Image Video Process., vol. 2020, no. 1, pp. 1–18, 2020.
33. Y. Tan, J. Qin, X. Xiang, C. Zhang, and Z. Wang, “Coverless steganography based on motion analysis of video,” Secur. Commun. Netw., vol. 2021, pp. 1–16, 2021.
34. L. Zou, W. Wan, B. Wei, and J. Sun, “Coverless video steganography based on inter frame combination,” in Proc. Geometry Vis.: 1st Int. Symp., 2021, pp. 134–141.
35. C. Zhang, Y. Tan, J. Qin, and X. Xiang, “Coverless video steganography based on audio and frame features,” Secur. Commun. Netw., vol. 2022, 2022, Art. no. 1154098.
36. R. Li, J. Qin, Y. Tan, and N. N. Xiong, “Coverless video steganography based on frame sequence perceptual distance mapping,” Comput., Mater. Continua, vol. 73, no. 1, pp. 1571–1583, 2022.
37. L. Meng, X. Jiang, Z. Zhang, Z. Li, and T. Sun, “A robust coverless video steganography based on maximum DC