Privacy-Preserving Encryption of Surveillance Videos supporting Selective Decryption

Chan Hyeong Cho, Hyun Min Song, and Taek-Young Youn*

Department of Industrial Security, Dankook University, Yongin, Korea
chochan0418@gmail.com, {taekyoung, hyunminsong}@dankook.ac.kr

Abstract

When utilizing video for security or criminal investigation purposes, issues of data leakage and privacy infringement arise. These concerns are associated with the disclosure of stored videos and the potential breach of privacy for individuals involved in the investigative process. To address these challenges, this paper introduces a method that defines Regions of Interest (ROI) as facial areas, and extracts feature values from these ROIs to produce a unique key for encryption. By employing ROI anonymization, the proposed approach prevents personal data leakage from external threats. Additionally, through selective decryption, it ensures the protection of individual privacy during the video access and analysis processes. Experimental results and security assessments validate the efficacy of the proposed method in safeguarding the confidentiality of surveillance footage.

Keywords: Privacy-preserving, Video encryption, Selective decryption, Surveillance video

1 Introduction

With the advancement of image and video processing technologies and the progress in storage device technologies, video devices such as cameras and camcorders have become more integrated into our lives. Notably, video processing devices like CCTV and IP cameras are widely utilized by individuals and public institutions to monitor and capture video footage. These videos can serve as crucial evidence for identifying and apprehending criminals in case of criminal activities. However, there are risks associated with the potential exposure of private information in case of external attacks or when videos are utilized for criminal investigations, which may unintentionally reveal personal privacy. Consequently, the protection of personal information captured in videos is of utmost importance. To address these concerns, the process of de-identification is employed to ensure that personal privacy is not compromised in the captured videos. There are two main methods for protecting privacy in videos: 1) Whole Frame De-identification, and 2) ROI (Region of Interest) De-identification. Whole Frame De-identification involves encrypting the entire video frame to protect privacy. However, when using encrypted videos, all frames need to be decrypted for practical use, thereby failing to protect the privacy of the subjects effectively. Additionally, the computational overhead of encrypting the entire frame can be relatively high. ROI De-identification, on the other hand, focuses on protecting specific regions of interest within the image, such as facial regions. Techniques like privacy masking are employed, where certain areas of the ROI are blurred or obscured. While this method is efficient and incurs low performance overhead, it alters the original image and requires the preservation of the original image for potential restoration. Another approach for ROI De-identification is ROI Encryption, where the ROI is recognized and encrypted. Compared to encrypting the

The 7th International Conference on Mobile Internet Security (MobiSec'23), Dec. 19-21, 2023, Okinawa, Japan, Article No. 4

*Corresponding author: Department of Industrial Security, Dankook University, Yongin, 16890, Rebulic of Korea, Email: taekyoung@dankook.ac.kr entire frame, this method reduces computational operations, and as only the ROI is encrypted, the actions of the subjects can still be identified. Moreover, privacy is preserved, and the encrypted video can be restored through decryption.

To ensure the protection of subjects' privacy when using ROI Encryption for video investigation, it is vital to use different keys for encrypting all recognized ROIs. Encrypting all ROIs with the same key could allow unauthorized access to non-related subjects' ROIs using the decryption key provided for investigation. Proper key management, key generation methods, and key assignment methods must be defined to ensure that each ROI is encrypted with a unique key.

In this paper, we propose an encryption method that generates Object-Only keys based on the features of recognized ROIs. The approach suggested in this paper provides a solution to privacy concerns related to unrelated individuals that previous research results have not addressed when utilizing encrypted videos. By employing the encryption method proposed in this paper, it becomes possible to enable selective decryption, addressing issues that arise when the decrypting entity attempts to re-identify and utilize encrypted videos. Section 2 provides an overview of the proposed model and related works aimed at addressing privacy concerns. Section 3 presents detailed explanations of the model structure and system processes developed to address the threat model presented in Section 2. In Section 4, we conduct security evaluations and experimental evaluations for the proposed method.5 During the research process, several challenges were encountered, and it is crucial to propose potential solutions for these issues to enhance the effectiveness and applicability of the proposed system.Finally, Section 6 summarizes the results of this research.

2 Preliminaries

2.1 Models

In this section, we present the models for our proposed technique.

2.1.1 System Model

The proposed approach in this paper involves CCTV and IP cameras, which are installed by individuals or institutions for video recording. After capturing videos, the entities in control of the installed devices generate object-only keys for each ROI (Region of Interest) using the keys possessed by entities. These keys are then used to encrypt the corresponding ROIs, and the encrypted videos are stored in storage. If the stored videos are later utilized for criminal investigation purposes, the encrypted ROIs need to be decrypted. The party requesting the decryption provides the encrypted video to the decryption entity, and the decryption entity requests the object-specific key for the ROI they wish to decrypt. The entity in control of the device receives the request and sends the requested object-only key to the decryption entity, enabling the utilization of the video for criminal investigation purposes.

2.1.2 Threat Model

Privacy infringement in recorded videos can be categorized into two situations:

- **Active Adversary:** In case the videos stored in the storage are leaked by external attackers, it can lead to privacy breaches for all subjects in the recorded videos.

- **Passive Adversary:** When the recorded videos are utilized as evidence for criminal investigations, there is a risk of privacy infringement for non-related subjects as the investigation may involve non-relevant personal information.

To prevent privacy leakage by external attackers, specific ROIs are defined and recognized using machine learning, followed by encryption. In this paper, the ROIs targeted for encryption are facial regions, which are identifiable by others. To protect the privacy of subjects during the investigation process, the encryption involves object-only keys based on the facial characteristics of the subjects, enabling individual encryption of each ROI.

2.2 Related Works

As people's awareness of the importance of privacy increases, researchers have explored various methods to protect the privacy in recorded videos. In [1, 2], the approach used is full encryption of entire frames to safeguard privacy. By encrypting all pixel information in the captured frames, the video ensures the impossibility of identifying the subjects' privacy. The method proposed in these studies involves utilizing a chaos maps-based encryption approach for complete de-identification, which efficiently encrypts all parts of the video data and provides a high level of security. The advantage of full video encryption lies in its robust data protection and complete security. However, this strong security comes with the drawback of increased processing and bandwidth requirements, making encryption and decryption computationally intensive. On the other hand, [3, 4] propose techniques such as blurring and mosaic to de-identify ROIs. Through these techniques, certain parts of the video are modified or removed, making it difficult to recognize the original information. This process protects personal information and makes it challenging to recover the original image. The blurring and mosaic masking techniques blur the original image, making it unrecognizable. Additionally, they employ methods of removal and modification to delete or overwrite pixel values in the original image, making it difficult to reconstruct the original image. Due to these characteristics, a separate storage for non-de-identified videos is required when utilizing the videos. In [5, 6, 7], the authors propose ROI encryption methods where only the Regions of Interest (ROIs) are recognized and encrypted. Unlike entire video encryption, ROI encryption focuses on identifying and encrypting specific regions within the video. Due to the relatively smaller encryption scope compared to entire video encryption, ROI encryption requires less computational effort. Additionally, when viewing the video, the regions outside the ROI remain unencrypted, allowing the actions to be identified. In [6, 8], they propose a method using a Gaussian mixture model and Histogram of Oriented Gradients (HOG) feature extraction to select irregularly shaped ROIs. On the other hand, [9] presents an effective ROI recognition technique and employs chaos map-based scrambling for encryption.

To efficiently utilize recorded videos while ensuring privacy protection, adopting the ROI (Region of Interest) encryption approach is crucial. However, encrypting only the ROIs may lead to privacy breaches for unrelated third parties when they use the video for purposes not originally intended during recording. This issue arises due to the following reasons: 1)When using a single key to encrypt the recognized ROIs, unauthorized access to the decryption key could potentially enable the decryption of all ROIs. In such situations, if the encrypted video is used in investigations, it could result in privacy breaches for unrelated individuals as all ROIs could be decrypted. 2)Encrypting each recognized ROI with a different key poses challenges in achieving effective encryption and managing multiple keys.

To overcome these limitations, this paper proposes a method of partially encrypting ROIs by utilizing a key for ROI encryption based on both the main key and ROI characteristics.

Method	ROI De-indentification	ROI Re-identification	Selecticve Decryption
ROI masking	Possible	Impossible	Impossible
ROI Encryption	Possible	Possible	Impossible
Proposal Method	Possible	Possible	Possible

Table 1: Comparison with related works

The proposed approach suggests efficient key management and the generation of Object Only keys with appropriate efficiency. Existing research results have been confined to methods of encrypting ROIs, lacking the ability to address privacy infringement issues arising during the utilization of encrypted videos. However, by employing the encryption method proposed in this paper, it becomes possible to address privacy infringement problems occurring during the utilization of encrypted videos.

3 Design

In this section, we address the active adversary and passive adversary scenarios by applying encryption to ROIs and supporting selective decryption. To achieve selective decryption, we explain the process of generating object-only keys for each ROI based on its characteristics during the encryption process. It is important to note that this paper focuses on the ROI encryption method with selective decryption to counter active and passive adversaries, and the enhancement of face detection and feature extraction through machine learning is not the primary research target.

To protect the privacy of the subjects, several steps are taken. Firstly, the range of ROIs is defined and recognized using machine learning. ROIs are defined as facial regions that can be recognized by others, and Dlib's face detection feature is utilized for recognition. To distinguish the recognized faces, Dlib's 68-Point Shape Predictor is employed to assign a unique ID for each ROI and track consecutive targets across frames. Based on the recognized characteristics of ROIs, object-only keys are generated. These keys are derived using a hash algorithm with the main key and the ID value. By adopting this approach, various object-only keys can be generated from the main key and ID alone. Each ROI is then encrypted using its corresponding object-only key.

By adopting this encryption method, only the necessary object-only keys generated during the encryption process are provided to the entity decrypting the video. This ensures that the decrypting entity can only decrypt the ROIs of relevant individuals, preventing the decryption entity from accessing ROIs of unrelated individuals. This approach addresses privacy concerns related to unrelated individuals that were previously not addressed in the utilization of encrypted videos.

3.1 ROI Recognition and Object-only Key Generation

In this paper, ROIs are defined as regions where a person can recognize themselves, particularly facial regions. There are several methods for recognizing ROIs, and we utilize a machine learning-based approach. Using machine learning for facial recognition offers advantages such as automation, high accuracy, and real-time processing. To achieve ROI recognition, we utilize Dlib's face detection feature [10]. Dlib's face detection is a library that accurately and rapidly detects faces using the Histogram of Oriented Gradients (HOG) algorithm. To distinguish the recognized ROIs, we use Dlib's 68-Point Shape Predictor [11], which is a technology that detects facial landmarks. This technology identifies 68 specific points on a face, enabling a precise understanding of the shape and structure of the face. Using this predictor, we generate unique IDs for object differentiation based on the recognized facial features, in the order of recognition. The generated IDs are then used to create object-only keys. These keys should not allow inference of the main key and enable the encryption entity to generate multiple object-only keys from a single main key efficiently. Moreover, since the encryption algorithm uses fixed key values, the size of the object-only keys should align with the key size supported by the encryption algorithm. Considering these characteristics, an object-only key generation algorithm using a hash algorithm is proposed.

The object-only key generation algorithm is implemented using a hash algorithm, which ensures unpredictability and reduces collisions. Unpredictability guarantees that the decryption entity cannot infer the main key. The fixed-length property of the hash algorithm allows the generation of keys with the correct key size supported by the encryption algorithm. Additionally, the fast processing feature of the hash algorithm saves time during key generation, expanding the system's scope of use.

Furthermore, to increase the entropy and enhance the security of the object-only key, additional information such as the device information, recording date, and entity information can be utilized in the object-only key generation process.

In the following algorithm, we describe the pseudo-code for the object-only key generation algorithm, which uses ”mainkey||id||additional information” to create the object-only key by applying a hash function to the value.

Algorithm 1 Object Only Key Generation Algorithm

```
Require: key, object's ID, additional information
input = key || Object's ID || additional information
Object only key = hash algorithm({input})
return Object only key
```

3.2 ROI encryption

To encrypt the recognized ROIs, the system uses the generated object-specific keys. In the video, the ROIs are encrypted on a bounding box basis. When encrypting bounding boxes of undetermined sizes in the video frames, it is essential to use an encryption algorithm where the plaintext size and the generated ciphertext size are the same. If the sizes of the plaintext and generated ciphertext differ, using an encryption algorithm that does not account for this discrepancy, such as one involving padding, may lead to issues where the encrypted data size exceeds the size of the original bounding box, causing the encrypted data to extend beyond the bounds of the bounding box. To prevent such issues, this paper utilizes the CTR mode, which is a type of block cipher mode, for encryption. By using CTR mode, the lengths of the encrypted and decrypted data are the same, regardless of the size of the data being encrypted. This ensures that even bounding boxes of various sizes can be effectively encrypted without causing any issues related to the bounding box size.

3.3 Partial Decryption for Investigation

The decryption phase involves obtaining the decrypted video of the desired subjects from the video where ROIs are encrypted. To achieve this, we perform the reverse process of encryption and identify the positions of the encrypted ROIs. The decryption process proceeds as follows: 1)The decryption entity receives the object-only key and the ROI location of the 2)desired subject from the encryption entity. 3)The decryption entity uses the object-only key to decrypt the corresponding ROI. The decrypted video containing only the ROIs of the desired subject is obtained. By performing this process, selective decryption can be achieved. Selective decryption allows the protection of the privacy of non-related subjects while obtaining the decrypted video of only the ROIs belonging to the desired subject.

4 Analysis

In this section, various statistical tests and measurements are used to analyze the proposed approach. The tested videos from [12] are used as surveillance videos. The proposed approach was simulated in Google Colab's default mode for experimentation and executed in Jupyter Notebook for real-time environment. The hardware used for the experiments includes an Intel(R) Core(TM) i7-12700F CPU, 16 GB memory, and Windows 11 operating system. The AES-CTR encryption algorithm was used. The table below provides information about the sample videos used in the analysis: [1]

Name	Frame	Resolution
akiyo	300	QCIF
sign_irene	540	QCIF
Johnny	600	720p
FourPeople	600	720p
vidyo1	600	720p

Table 2: Test video information

4.1 Security

To counteract both active and passive adversaries, it is crucial to evaluate the security and effectiveness of selective decryption for encrypted videos. In this section, we conduct assessments, including RGB analysis, edge detection analysis, and Analysis of ID Assignment, to ensure the appropriateness of ROI ID allocation for selective decryption. Through these evaluations, we aim to guarantee the safety and integrity of the encrypted video, effectively protecting it from potential threats posed by attackers.

4.1.1 RGB analysis

The proposed method in this paper aims to provide security against active and passive adversaries. This section addresses the security issues related to encryption of RGB values for active

[1] In [9], experiments were conducted on a total of six videos. The experiments aimed to evaluate ROI encryption with varying numbers of individuals appearing in the videos. In our study, we decided to conduct experiments on five videos featuring clear distinctions with 1 to 4 individuals, in line with the characteristics of our proposed approach.

adversaries and the security of ID assignment for passive adversaries. Figure 1-(a) illustrates the RGB value analysis of original and encrypted ROIs for three test videos (akiyo, Sign Irene, and Johnny) at specific frames. These images demonstrate that the intensity distribution of encrypted ROIs is uniform compared to the original ROIs. Therefore, the proposed method can prevent active adversaries from discerning any patterns in the original ROIs.

4.1.2 Edge detection analysis

Figure 1-(b) presents the results of edge detection using the Canny algorithm on the original and corresponding encrypted Regions of Interest (ROIs) for the akiyo, sign irene, and johnny frames of the test video. The displayed outcomes clearly demonstrate substantial disparities between the original and encrypted ROIs, signifying the encryption's efficacy in concealing critical visual details within the video.

4.1.3 Analysis of ID Assignment

The issue of assigning IDs to objects significantly impacts the security against passive adversaries in the proposed approach. Assigning different IDs to the same object may cause inconvenience during the decryption process, but it does not raise security concerns. However, assigning the same ID to different objects can compromise the privacy of objects with the same ID. Therefore, situations where different objects share the same ID should be avoided. ID assignment relies on the threshold value used to determine if the ROI (Region of Interest) between frames is the same. Setting an appropriate threshold value is crucial in taking responsibility for the security aspects of ID assignment.

Figure 1-(c) illustrates the process of verifying if the same object is assigned the same ID by extracting three frames from the video. The results from Figure 9 confirm that objects appearing in the same video are assigned the same ID. Occasionally, different IDs are assigned based on facial expressions or angles, but there were no instances where different objects were assigned the same ID.

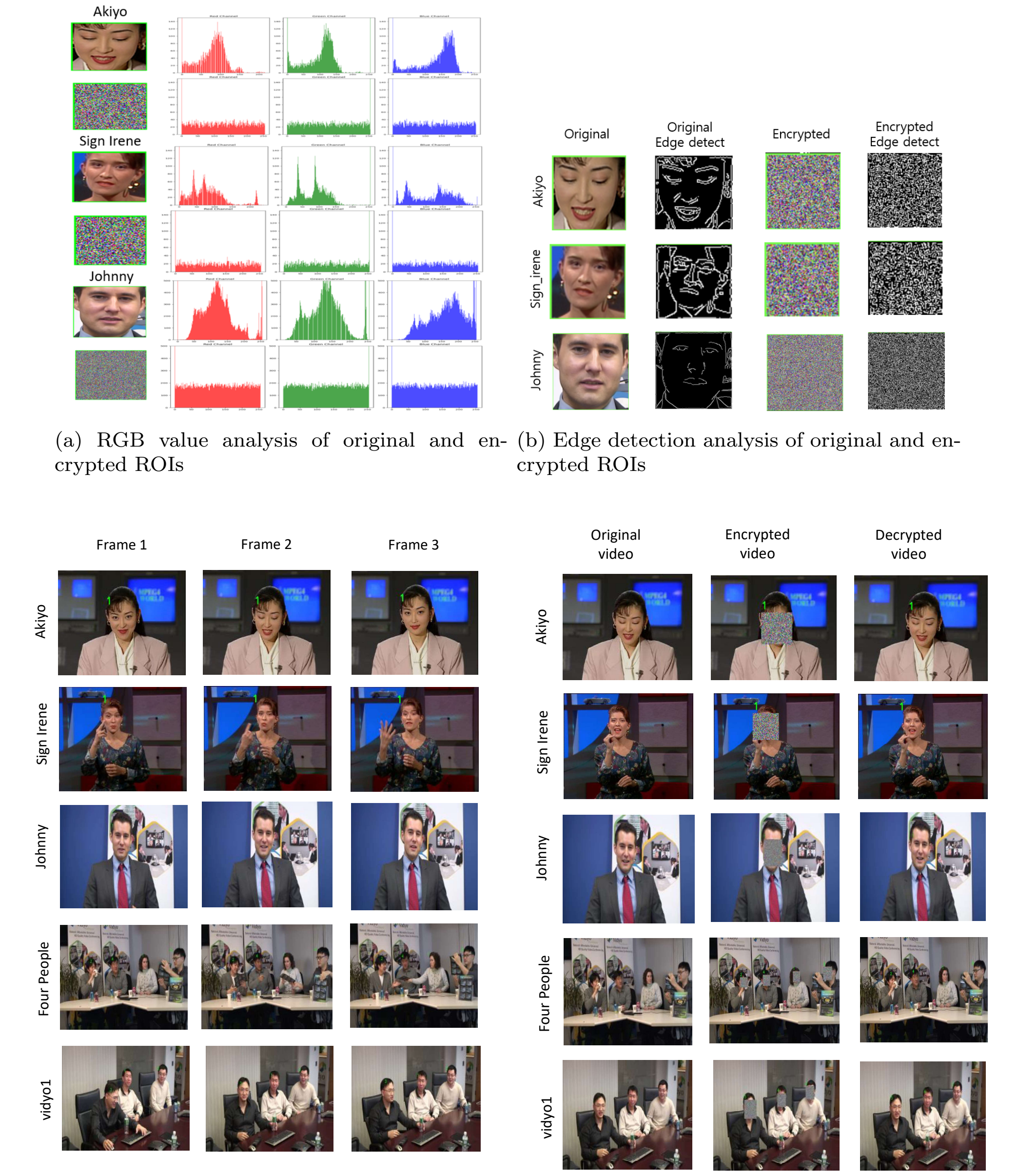
To verify if different faces are assigned different ID values, the [13] dataset was used for ID assignment. This dataset encompasses all common creeds, races, age groups, and profiles, and it also includes some images generated by Generative Adversarial Networks (GANs) to aid in distinguishing between real and generated faces. In total, 7,219 image files were classified into ID-specific folders, resulting in 2894 folders being created, with each folder containing image files assigned the same ID. When an appropriate threshold value was set during the object differentiation process between frames, no issues of assigning the same ID to different faces were observed Overall, the experiment successfully demonstrated the significance of threshold value selection and the successful classification of images based on their respective faces.

4.2 Experimental Evaluation

4.2.1 Functionalities

The proposed system in this paper recognizes the ROI in the video, extracts its features, and uses individual object-specific keys to perform encryption and decryption successfully. Figure 1-(d) presents the encryption and decryption results for five sample videos.

Since the proposed method supports selective decryption to protect the privacy of individuals unrelated to the investigation, it can effectively decrypt the desired ROI with the corresponding ID's object-only key. Moreover, multiple ROIs can be decrypted successfully if each object has its own object-only key.



(a) RGB value analysis of original and encrypted ROIs

(b) Edge detection analysis of original and encrypted ROIs

(c) Objects with the same ID are marked in the top left corner of the recognized face

(d) Demonstration of encryption and decryption for five sample videos with object differentiation

Figure 1: Analysis data

4.2.2 Performance

To measure the performance of the proposed system, the time taken for each step in the encryption process, including ROI detection and feature extraction, object-only key generation, and encryption, was recorded. The data used for measurement was based on the number of recognized ROIs in the five sample videos, with 0 indicating no ROIs detected in the video frame. The measurements were repeated 25 times for each of the five test videos, and the average execution time is presented in the table below. As shown in the table, the process time increases as the number of recognized ROIs increases, with the fastest execution time observed when no objects are recognized.

Number of person	Average execution time
0	0.01679 sec
1	0.09656 sec
2	0.13783 sec
3	0.14449 sec
4	0.14810 sec

Table 3: execution time table

Additionally, the proposed system supports real-time encryption and decryption. The real-time experiment was performed using a Logitech C920 PRO HD webcam as input for real-time video capturing. The system successfully performed ROI detection, feature extraction, object-only key generation, and encryption without any noticeable delay.

5 Further Work

The proposed system uses a probability-based machine learning ROI detection method, which may result in the failure to detect ROIs in some frames out of the entire video. To address this issue, video-specific features can be utilized. In a video, objects typically do not abruptly disappear in consecutive frames but rather maintain continuity. This feature can be leveraged to address the problem of undetected ROIs in certain frames. By examining the preceding and succeeding frames of frames with undetected ROIs and identifying ROIs with the same ID, the system can designate the midpoint of these ROIs as the ROI for the undetected frame.

6 Conclusion

This paper proposes a method for protecting privacy in recorded videos. The approach utilizes Dlib's face detection and 68-Point Shape Predictor to recognize Regions of Interest (ROIs) and extract their characteristic features to generate unique IDs. By combining these IDs with the main key, only object-specific keys are generated to encrypt the ROIs. The proposed method safeguards against external and internal threats, ensuring that only authorized parties with the appropriate permissions can decrypt and access specific ROIs while preserving the privacy of unrelated individuals in the video footage. The suggested method provides a concrete solution to privacy concerns related to unrelated individuals that arise during the utilization of encrypted videos. The effectiveness of the proposed approach is demonstrated through various experiments, including encryption and decryption parallel processing, collision analysis of ID generation, and measurement of execution times. The results indicate that the proposed method successfully decrypts the desired objects from encrypted ROIs, demonstrating robustness against both external and internal attackers. Overall, the proposed method offers an effective solution to protect privacy in recorded videos, enabling secure utilization of video data while safeguarding the rights of individuals involved.

Acknowledgments

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea Government (MSIT) (No. 2021-0-00779, Development of High- Speed Encryption Data Processing Technology That Guarantees Privacy Based Hardware, 50%) and the Institute of Information and Communications Technology Planning and Evaluation(IITP) grant funded by the Korea government(MIST) (No.2022-0-01022, Development of Collection and Integrated Analysis Methods of Automotive Inter/Intra System Artifacts through Construction of Event-based experimental system, 50%).

References

[1] Dong Jiang, Zhen Yuan, Wen-xin Li, and Liang-liang Lu. Real-time chaotic video encryption based on multithreaded parallel confusion and diffusion. *arXiv preprint arXiv:2302.07411*, 2023.

[2] Andrew Senior. *Protecting privacy in video surveillance*. Springer, 2009.

[3] Ralph Gross, Latanya Sweeney, Jeffrey Cohn, Fernando De la Torre, and Simon Baker. Face de-identification. *Protecting privacy in video surveillance*, pages 129–146, 2009.

[4] Geoffrey Letournel, Aurélie Bugeau, V-T Ta, and J-P Domenger. Face de-identification with expressions preservation. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 4366–4370. IEEE, 2015.

[5] Mousa Farajallah, Wassim Hamidouche, Olivier Déforges, and Safwan El Assad. Roi encryption for the hevc coded video contents. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 3096–3100. IEEE, 2015.

[6] Di Xiao, Qingqing Fu, Tao Xiang, and Yushu Zhang. Chaotic image encryption of regions of interest. *International Journal of Bifurcation and Chaos*, 26(11):1650193, 2016.

[7] Wei Song, Chong Fu, Yu Zheng, Lin Cao, Ming Tie, and Chiu-Wing Sham. Protection of image roi using chaos-based encryption and dcnn-based object detection. *Neural Computing and Applications*, pages 1–14, 2022.

[8] Yang Liu and Jindong Zhang. A multidimensional chaotic image encryption algorithm based on dna coding. *Multimedia Tools and Applications*, 79:21579–21601, 2020.

[9] Khalid M Hosny, Mohamed A Zaki, Hanaa M Hamza, Mostafa M Fouda, and Nabil A Lashin. Privacy protection in surveillance videos using block scrambling-based encryption and dcnn-based face detection. *IEEE Access*, 10:106750–106769, 2022.

[10] S Sharma, Karthikeyan Shanmugasundaram, and Sathees Kumar Ramasamy. Farec—cnn based efficient face recognition technique using dlib. In *2016 international conference on advanced communication control and computing technologies (ICACCCT)*, pages 192–195. IEEE, 2016.

[11] Kostiantyn Khabarlak and Larysa Koriashkina. Fast facial landmark detection and applications: A survey. *arXiv preprint arXiv:2101.10808*, 2021.

[12] Xiph.Org. Derf's test media collection. Accessed: Jun. 26, 2022.

[13] Ashwin Gupta. Human faces dataset. https://www.kaggle.com/datasets/ashwingupta3012/human-faces.