IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 11, 2021, accepted February 2, 2021, date of publication February 8, 2021, date of current version February 17, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3057605

Intelligent Video Data Security: A Survey and Open Challenges

JIN-YONG YU, YUJUN KIM, AND YOUNG-GAB KIM, (Member, IEEE)
Department of Computer and Information Security, and Convergence Engineering for Intelligent Drone, Sejong University, Seoul 05006, South Korea

Corresponding author: Young-Gab Kim (alwaysgabi@sejong.ac.kr)

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant funded by the Korean Government (MSIT) under Grant 2019-0-00231, and in part by the Development of artificial Intelligence-Based Video Security Technology and Systems for Public Infrastructure Safety.

ABSTRACT A range of video contents and technology have provided convenience to humans, with real-time video applications—such as surveillance applications—able to contribute to increasing public safety by reducing physical crimes. The development of video technology has made it possible to achieve an improved quality of life. However, this technology can also be exploited and lead to security issues such as physical and digital crimes. Unfortunately, security breaches are increasing in complexity and frequency, making current countermeasures insufficient to prevent them. Given recent trends, we recognize the need for security technology to respond to advanced video crimes. Intelligent security is one of the methods that can be used to respond to these issues. Although research on video data security has been actively conducted, not enough studies have been published on video data security that also addresses intelligent security. Specifically, a classification system for research on video data security has not been provided, and no systematic analysis has been conducted for advanced research. Thus, the purpose of this is to fill in these gaps in existing research. This study offers a classification of research on video data security based on the collection and analysis of related works. Moreover, this study presents an analysis of research on video data security technologies combined with intelligent technologies based on SLR methodology.

INDEX TERMS Video data security, intelligent security, user access control, visual security, video data validation.

I. INTRODUCTION

Video applications (e.g., video broadcasts, movies, surveillance, and video recordings) have made life more convenient for humans. The use of video technology has increased exponentially, making it an essential medium in our lives. Video has the ability to depict real scenes visually, featuring information related to geolocation, specific circumstances, and the visual identification of humans. This kind of information, which can be collected from video sources, is called visual data and it may have a different value depending on the environment in which it was collected. For example, in a private space such as a house or a corridor, there is a high probability that important private information will be included in visual data because a video collector could define what you want to observe. On the other hand, in crowded environments such as terminals, airports, and parks, it does not contain relatively important information because it collects a large number of unspecified visual data. However, if multiple collectors are deployed, various information may be collected due to collectors' phase, angle, and performance [1]. This is easily demonstrated in environments where multiple cameras are deployed targeting the same space. When visual data are collected from these environments, there is often a greater amount of sensitive information than expected.

More recently, it has become easier to access videos that include the identity and private information of strangers and even Deepfake videos. Sometimes, due to a creator's mistake, videos result in unintentional exposure. On the other hand, sometimes videos are created with malicious intentions. Once a video that features private or important information is shared, it is very difficult to stop the spread due to the nature of current social networks where everything is shared, and content can go ''viral.'' Moreover, we should not overlook the possibility of video content being misused. Unfortunately, as video technology develops, criminal methods to

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Sanchez.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

use this technology maliciously are becoming more diverse and intelligent. Although some organizations—such as the ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission)—have established standards for video data security to improve social safety and resilience, this is not enough. Even at the national level, policies such as the GDPR (General Data Protection Regulation) have been established to protect people's privacy; however, they are not considered to be thorough countermeasures to prevent advanced crime. Therefore, we have to urgently devise more effective countermeasures to prevent these types of crimes. Among potential methods to protect video data, intelligent security is the most promising, and it can be defined as a combination of intelligent technologies, such as adaptive and awareness, with security that can actively respond to various threats. It also can enhance the efficiency of overall system functions such as management, transport, and searching [2]. Intelligent security is essential when considering system performance efficiency as well as adequate responses to cunning crime [3].

Despite these urgent issues, there are few studies that deal with intelligent security in video data, especially when compared with other kinds of research on video data security. Additionally, it is difficult to identify and counter security issues because research on video data security has not been clearly classified. When considering the video data security requirements [4], it is possible to infer the research area that should be studied intensively. Still, to prevent security issues, it is necessary to clarify the research area of video data security research. For example, security issues related to visual privacy—such as video data sniffing and video tampering including unauthorized observation—have increased on an annual basis [5]–[7]. There are many potential reasons to explain why privacy issues are increasing; for instance, the increasing number of cameras, the ability to enhance the resolution of visual sensors, and the performance of video algorithms [8]. However, these are some of many possibilities and they do not lead to an accurate diagnosis. Thus, a systematic classification scheme of research dealing with video data security is required to promptly counter emerging threats. In order to propose a detailed taxonomy for video data security research, we analyzed the existing literature on video data security in this study. The study's main contributions are as follows:

- We classified the research area dealing with video data security into the proposed taxonomy. Although a taxonomy already existed for some research areas, this may be the first attempt to classify the research area focused on video data security.
- Unlike the existing video data security studies, the latest research trends were identified by analyzing intelligent video data security studies intensively.
- We identified the challenges to be addressed through various literature on video data security and offered suggestions for future work on this topic by representing it as a scenario.

The remainder of this study is organized as follows. In Section II, we introduce our literature review methodology and describe our process for extracting specific studies In Section III, we classify the research area dealing with video data security based on analyzing the existing literature. Also, we analyzed studies combining video data security and intelligent technology for identifying the trends. In Section IV, we discuss significant challenges to securing video data and suggest future directions and unresolved issues. In Section V, we present our conclusions.

II. METHODOLOGY

The remainder of this study is organized as follows. In Section II, we introduce our literature review methodology and describe our process for extracting specific studies In Section III, we classify the research area dealing with video data security based on analyzing the existing literature. Also, we analyzed studies combining video data security and intelligent technology for identifying the trends. In Section IV, we discuss significant challenges to securing video data and suggest future directions and unresolved issues. In Section V, we present our conclusions.

A. GOALS AND RESEARCH QUESTION

As stated earlier, this study's goal is to classify the research area dealing with video data security specifically and to systematically review the literature that combines intelligent technology with video data security. In this end, we encountered a lot of literature and extracted reliable studies to cite in our paper.

1) RESEARCH QUESTION

In order to enhance this study's feasibility and accuracy, we focused on extracting reliable studies from the existing literature. Thus, locating reliable studies was a crucial aspect of this literature review. The study's research question served as the backbone for conducting a systematic review.

RQ1. How will reliable literature be identified and collected?

RQ2. What advantages can be derived from locating and reviewing reliable literature?

Regarding RQ1, there is still significant interest in how to locate and collect reliable literature in survey-based studies [13]. Usually, survey or review studies should define specific criteria for identifying reliable literature through objective indicators. For instance, in this study, we selected prominent journals as a source of reliable literature. However, journals dealing only with specialized fields are a special case because they can measure poorly based on the selected indicators. Although objective indicators do not always determine whether a study is reliable or not, they can still be used to judge the quality of a journal objectively, as Table 1 shows. We selected four engines to search for journal rankings (b) and four digital libraries (a) related to the field of video data security, as Table 2 shows. In addition, we considered the aims and scope of each journal as well

TABLE 1. The selected journals, taking into account Aims & Scope and Objective Indicators.

Title	Quartile	Cite Score	JIF	SJR	SNIP	H Index	Eigenfactor Score
IEEE Transactions on Pattern Analysis and Machine Intelligence		19.67	17.73	3.764	9.231	326	0.06883
IEEE Transactions on Image Processing		9.63	6.790	1.809	3.838	242	0.06145
IEEE Transactions on Information Forensics and Security		9.03	6.211	1.364	3.634	95	0.01953
IEEE Transactions on Circuits and Systems for Video Technology		6.04	4.046	0.983	2.743	154	0.01628
Image and Vision Computing		3.58	2.733	0.633	1.541	118	0.00695
Pattern Recognition		7.35	5.898	1.363	3.012	180	0.03033
Expert Systems with Applications		6.36	4.292	1.190	2.696	162	0.03629
IEEE Transactions on Intelligent Transportation Systems		7.98	5.744	1.412	3.497	112	0.02108
Information Sciences		6.90	5.524	1.620	2.636	154	0.05080
IEEE Transactions on Multimedia		7.41	5.452	1.220	2.951	108	0.01984
IEEE Transactions on Dependable and Secure Computing	Q1	6.09	6.404	0.898	3.030	59	0.00344
ACM Transactions on Multimedia Computing Communications and Applications		3.75	2.870	0.569	1.769	38	0.00251
Multimedia Tools and Applications		2.33	2.101	0.335	1.038	52	0.01176
EURASIP Journal on Image and Video Processing	Q2	2.36	1.737	0342	1.238	19	0.00228
Security and Communication Networks		1.81	1.376	0.311	n/a	30	0.00421

TABLE 2. Sources for locating journals and existing literature.

Source	URL
IEEE Xplore	http://ieeexplore.ieee.org
ScienceDirect	https ://www.sciencedirect.com
Springer Link	https ://link.springer.com
ACM Digital Library	https ://dl.acm.org
Google Scholar	https://scholar.google.co.kr/

(a)

Source	URL
Scimago Journal & Country Rank	https://www.scimagojr.com
Guide2Research	http://www.guide2research.com
Journal Citation Reports	https://jcr.clarivate.com
Letpub	https://www.letpub.com

(b)

as the specific categories (e.g., computer vision, artificial intelligence, and multimedia security, etc.) required in this study.

Regarding RQ2, we exerted significant effort to collect studies because the purpose of survey-based studies is to make diverse and reliable information accessible to readers. Moreover, citations of valid studies help to make a study of this kind more reliable. This is the most advantage of locating and reviewing reliable literature. Thus, we adopted a systematic strategy to search for and analyze research on video data security.

B. STUDY SELECTION

Study selection is important in survey-based studies because it can enable us to identify primary studies that provide direct evidence regarding the primary research questions. Table 2 presents this study's list of journal sources used to extract reliable literature. This list served as the cornerstone of the search process, which is described below.

1) SEARCH STRATEGY

The search process was organized as shown in Fig. 1. We used search engines (i.e., IEEE Xplore, ScienceDirect, Springer

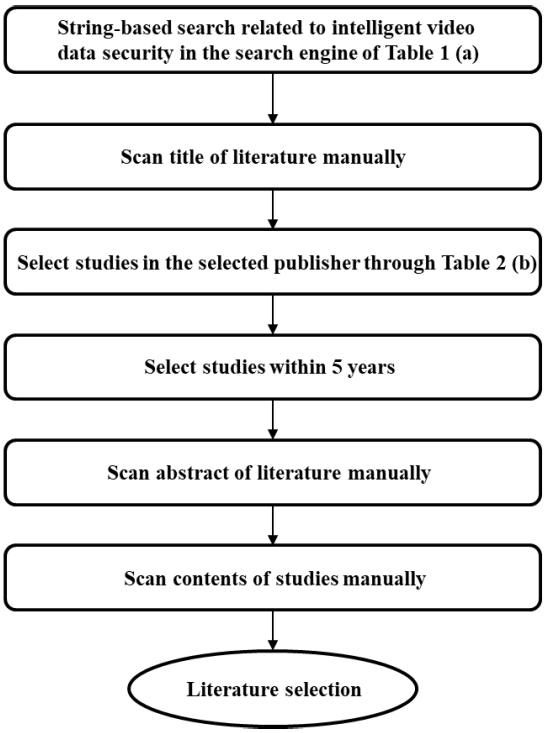

FIGURE 1. Search strategy for reliable literature selection.

Link, ACM Digital Library, and Google Scholar) to locate relevant literature. In phase 1, journal search engines were searched without restrictions based on video data security keywords such as video data security, video privacy, and video forgery. These keywords were also combined with surveillance, authentication, and intelligent security. It was also possible to acquire video-related expert security terminology from the literature and use it as part of the search process. This first phase provided a comprehensive sample of relevant works. In phase 2, specific studies were extracted from the selected journals. Although this condition can help to identify reliable literature, some constraints restricted the search because studies not published in journals were excluded. Moreover, it is sometimes difficult to determine that the extracted literature is of high quality. We discuss this problem in more depth in Section II-C. In phase 3, we sorted the articles based on date of publication, focusing on those published in the last five years. This was done to ensure that the review focused on the latest security technologies taking into account the rapid development of the video field. In phase 4, studies' titles were used to confirm the validity of the literature for this study. The criteria for inclusion and exclusion are described below:

- Does the study contain terms related to security such as encryption and access control?
- Are the terms are used within the research area dealing with video security or image security?

In phase 5, we analyzed the studies' abstracts. This phase allowed us to gather more detailed information about the literature that was not available in the previous phase. The criteria for inclusion and exclusion for this phase are described below:

- Does the abstract include topics or technical terminology that match the paper's title?
- Does the study propose appropriate countermeasures based on real challenges and/or is the security goal clear?
- Does the study address specific areas of video data security?

The above criteria are intended to exclude survey-based literature and include more specialized literature. Finally, in phase 6, we determined whether the literature fit in this study given the study's aim of reviewing research on video data security combined with intelligent security. The criteria for inclusion and exclusion for this phase are presented below:

- Does the study include future-oriented video data security combined with intelligent technologies such as adaptive technology, context-aware technology, and machine learning?

This search process was intended to not only locate literature dealing with a suitable topic in the context of the present study but also to identify the specific research area of video data security. Although filtration processes enabled us to extract some literature, we were able to review a wide range of literature related to video security through this search process. Section III discusses this in more detail.

C. QUALITY ASSESSMENT

We developed a set of quality assurance questions to evaluate the selected research. Each question was given a score to quantitatively evaluate the literature. The questions for quality evaluation are listed below:

QA1. Is the study cited in other literature?

QA2. Does the study target an obvious and specific area of video data security?

QA3. Does the study provide a comparison with or evaluation of other studies?

QA4. Does the study recognize the importance of integrating the field of video data security with intelligent technology?

Regarding QA1, although not applicable in all cases, being cited in other studies usually indicates the reliability of and interest in a study [14], [15]. We gave each journal a citation score. Regarding QA2, many scenarios can arise given that videos are exposed to different environments. This means that there is a wide range of threats and that different kinds of protection are needed. Therefore, video data security research should specify particular areas that need to be secured. Studies were scored based on the clarity level (i.e., opaque, normal, or clear) regarding the specific area that should be secured. Regarding QA3, given that research on video data security has been conducted for a long time, many similar studies exist. Accordingly, the appeal of research achievements or contributions should be clarified. Thus, we gave a score based on whether existing studies feature evaluation and comparison or not. Regarding QA4, intelligent technology

is the next-generation technology that will lead the field of IT, and it needs to be combined with various other technologies [16], [17]. Since it is important to recognize the need to integrate intelligent technology with other technologies, we scored studies based on whether they clearly recognized this need.

D. DATA EXTRACTION AND DATA ANALYSIS

The purpose of data extraction is to define and pilot a data extraction process to record the information obtained from the primary study and reduce bias. Then, it is possible to synthesize the extracted data in a way that answers the research questions in the data analysis step. The processes of data extraction and data analysis are necessary to provide answers to the research questions. The results of this process herein are as described below:

1) DATA EXTRACTION

The data extracted from each study were as follows:

- **Information regarding studies' external aspects:** title, author, publication date, journal, objective evaluation indicators regarding the journal, journal categories
- **Information regarding studies' internal aspects:** scope of the study (e.g., user access control of the video data, visual security, video data validation), overall study summary, problem and solution presented in study, study type (e.g., survey, research, etc.), evaluation and comparison with other studies

2) DATA ANALYSIS

We extracted the data from the searched literature. On the information regarding studies' external aspects, we took the information such as the latest in the literature, indicators of the influence of the literature, categories of the literature. Furthermore, we took detailed information on the information regarding studies' internal aspects such as the scope of the study, proposed scheme, and new approaches. However, some necessary data could poorly be extracted because security technologies that could be applied to the research area of video data security are not covered in the research area of video data security. As the next section shows, even if the study has not been conducted, we search for literature covering similar topics and extract the required data. In conclusion, the literature's internal information (addressing RQ1) and external information (addressing RQ2) were used as data for answering the research questions.

III. CLASSIFYING AND ANALYZING INTELLIGENT VIDEO DATA SECURITY

This section describes the present study's results. We first located relevant works on intelligent video data security and then analyzed them. In reviewing studies on intelligent video data security, we also included some research on image security because images are closely related to video; in other words, an image is one frame that makes up a video. We were able to design a taxonomy mainframe for video data security based on studies extracted through our search strategy. We established the taxonomy's components by reviewing a range of literature related to video data security (user access control, visual security, video data validation), as Fig. 2 shows.

We designed a taxonomy for video data security based on research and literature collected through search strategies. The collected literature studies authentication, authorization, encryption, visual security, and forgery detection of video data security. Accordingly, research regarding video data security is typically classified into three categories: user access control, visual security, and video data validation. User access control manages access authority by identifying users and granting authority [18], [19]. Visual security protects confidentiality from unauthorized users [20]. The related visual security is associated with the protection of videos' identifying information such as faces and passwords of locked doors [21]. Video data validation guarantees integrity through the verification of video data [22]. A more detailed description of these research areas that are combined with intelligent technology is given in the subsection, and selected studies are sorted by author, title, and year of publication in the study's Appendix. Video data security research trends in each area are presented in Fig. 3, 4, 5 and 6. We also featured an additional research area—the integration of video data security with intelligent technology—as Tables 3, 4, and 5 show.

A. USER ACCESS CONTROL

Access control is not a static security technology, and its usage can be changed depending on particular scenarios. In video data security, access control can be defined as the identification of users and the process of granting authority to access video content. In the video field, requiring authorization means that only users with proper authority are able to view the original data including identification information [23]. Knowledge-based, possession-based, biometric-based, location-based, and hybrid authentication models exist for user identification [24]. Each authentication model can be combined with other models to conduct two- and multi-factor authentication

In this study, we examined various access control models for authorization. The selected literature reviewed herein features access control models for specific situations and scenarios using basic access control models such as discretional access control (DAC), mandatory access control (MAC), role-based access control (RBAC), organization-based access control (OrBAC), and attribute-based access control (ABAC) [25]. Policy-based, token-based, and hybrid models also exist as reference architectures.

We devoted great effort to locating video research focused on intelligent access control in various journals and search engines based on this study's search strategy. However, that area of research was not as advanced as other research areas. For example, we were only able to review a small number of studies on intelligent authentication schemes related to video

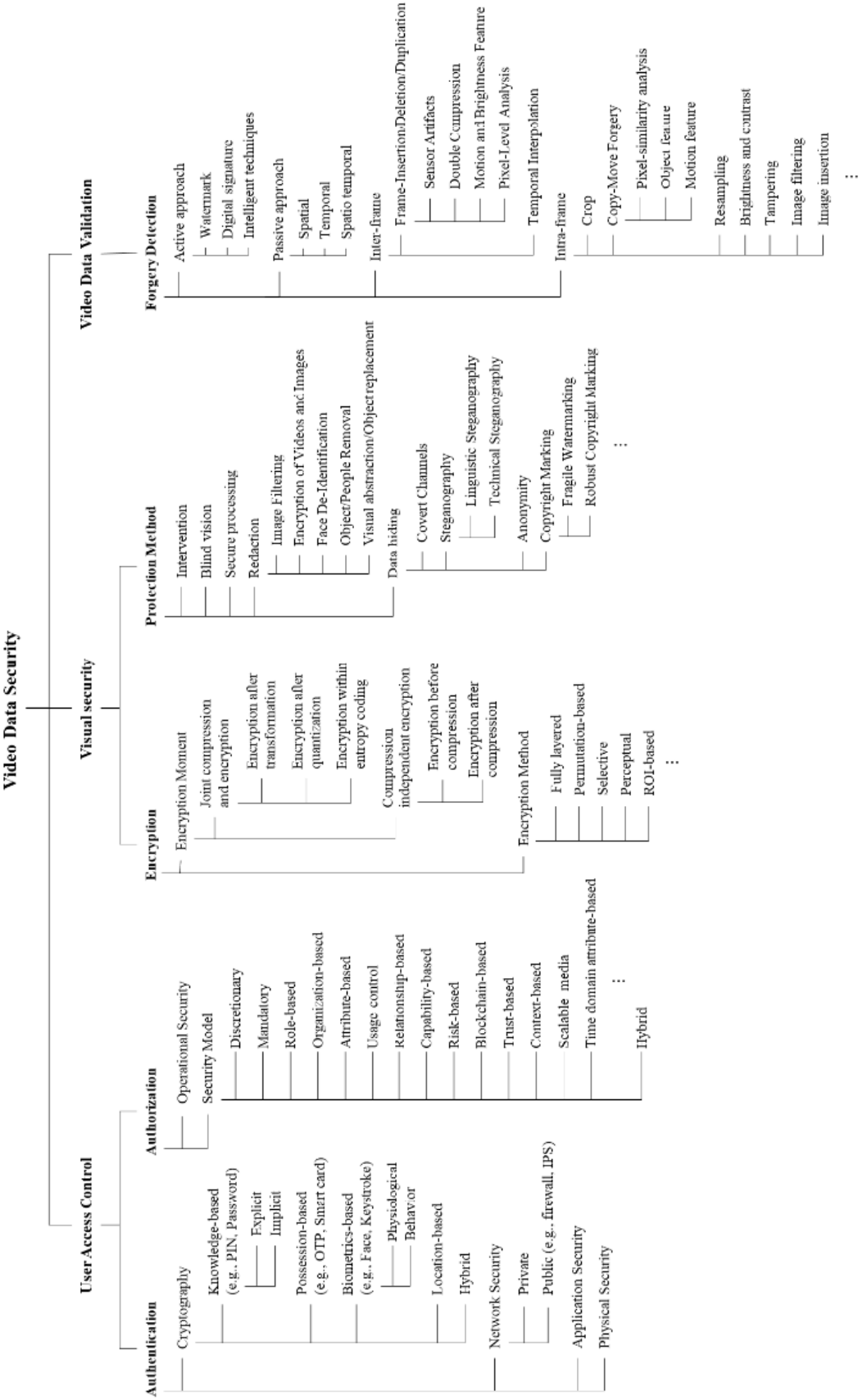

FIGURE 2. Taxonomy of the video data security.

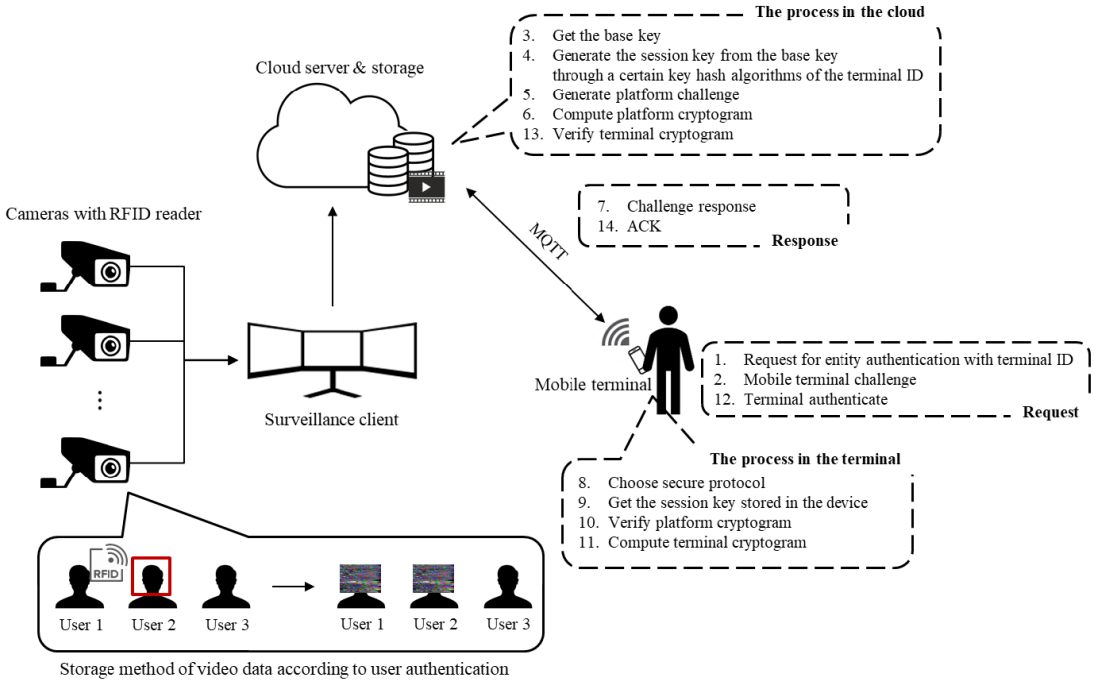

FIGURE 3. The model of user access control in video data security.

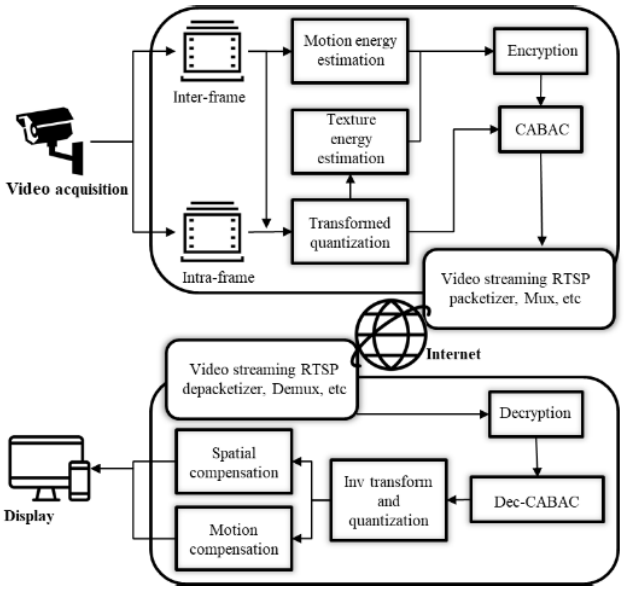

FIGURE 4. Overview of the video data encryption [P5].

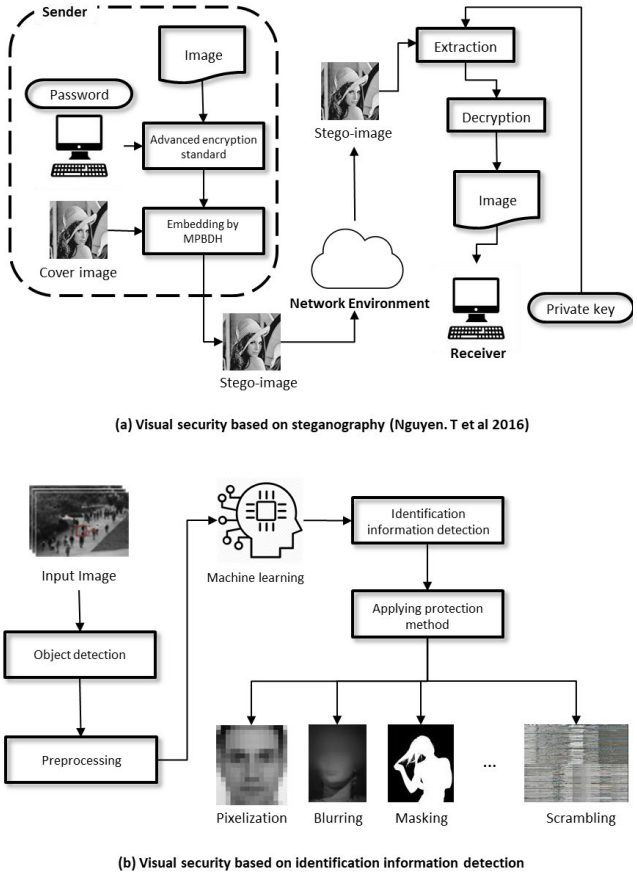

FIGURE 5. Visual security models in video data security.

data security as shown in Fig.3. Unfortunately, we did not find any research specifically on intelligent authorization schemes for video data security. Accordingly, this research area has to be investigated as an open challenge in order to establish a secure video data security environment. This challenge is further discussed in Section III-A.2 and IV.

1) AUTHENTICATION

Jinsu. K *et al.* [P1] proposed access control based on video surveillance combined with a face recognition system through machine learning and radio frequency identification (RFID). The reason for devising two-factor authentication technology in this study is due to inaccurate recognition

TABLE 3. Research area of intelligent user access control in video data security.

			Recognition P1	Context-awareness P2
Authentication	Cryptography	Knowledge-based	✓	
		Possession-based		✓
		Biometrics-based	✓	
		Location-based		
		Hybrid	✓	
	Network security	Private		
		Public		
	Application security			
	Physical security			

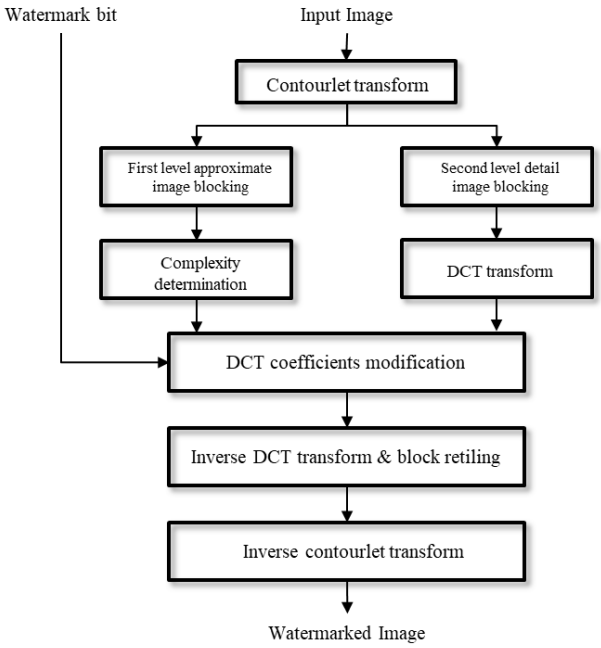

FIGURE 6. Block diagram for describing embedding watermarks scheme [P13].

rates and the issue of the RFID tag being violated by a harsh environment. Although face recognition technology has achieved more accurate recognition rates than ever before thanks to machine learning and deep learning, 100% recognition accuracy is impossible due to elements that interfere with recognition such as video quality, weather, poses, facial expressions, and hairdos. In addition, RFID-based authentication enables authentication even in harsh environments with low resolution.

Shen. J. *et al.* [P2] proposed a situation-aware mobile web middleware that provides convenient access to video taken in public places. This system also promises security and protects privacy through mobile terminals in public places, and it inspired the development of a context-aware mobile web middleware system based on near-field communication (NFC) that is expected to tackle the privacy issue. The present study addresses two security challenges by introducing the use of architecture in multiple contexts, unlike the prior study. The first is that middleware with context-awareness, privacy-awareness, and a security scheme makes it safe to access surveillance cameras placed in public infrastructure based on a lightweight web runtime on mobile devices. It also provides an interoperable and platform-independent solution. The second is that the researchers' proposal for accessing cameras is more comfortable and secure, which leads to a reduction in the system's awareness responsibility using NFC and the service layer of the NFC stack.

2) THE CHALLENGES OF USER ACCESS CONTROL

As shown in Table 3, we represented the research areas of those studies by in-depth analysis of the selected studies. In addition, we analyzed the challenges for each study. However, there are still open challenges to be addressed. Hence, we mention and explore challenges that were not presented in the abovementioned study. As the use of video surveillance increases in public infrastructures, such as smart cities and airports, access control to determine which users can access video content is becoming increasingly important. To tackle this problem, we have to consider the purpose of video surveillance, which is typically used to record specific events, crimes, or accidents [26]. Video surveillance has some disadvantages for access control such as minimizing the exposure of object information and quickly determining a specific situation. Anomaly detection can be one of the approaches to access control.

Anomaly detection is that an object's abnormal behavior captured on video can be judged stochastically, depending on whether the event setting recognizes the behavior as normal or abnormal [27]. Anomaly detection may be effective in

TABLE 4. Research area of intelligent visual security in video data security.

			A					R						C
			P3	P4	P6	P10	P12	P7	P8	P9	P11	P13	P14	P5
Encryption	Moments	Joint compression and encryption												✓
		Compression independent encryption		✓				✓		✓				
	Method	Fully layered												✓
		Permutation-based	✓		✓									
		Selective	✓	✓	✓			✓						
		Perceptual												
		ROI-based						✓	✓	✓				
Protection Method		Intervention												
		Blind vision								✓			✓	
		Secure processing												
		Redaction	✓	✓	✓			✓	✓	✓		✓	✓	✓
		Data hiding				✓	✓				✓			

* **A**: Adaptive technology, **R**: Recognition, **C**: Context-awareness

privacy protection by excluding scenes that are not detected in the pre-set event as well as by applying methods such as the blurring and masking of objects exhibiting normal behavior. This effect seems to be represented for visual security, but it is also related to access control in that it extracts only the information necessary for the user and conducts security for unnecessary information. However, there are some challenges in anomaly detection that related studies have explored for a long time. First, most anomaly detection is dependent on machine-learning technology for training to identify abnormal behavior, but there is not enough anomaly video data for effective training [28]. Second, it is difficult to define anomalies when two or more objects interact. For example, it may be difficult to distinguish between a soccer player who is trying to steal a ball from another player and a robber who is taking something from another person. This is because computers are trained using features extracted from frames. Third, although object detection has been significantly studied, detected objects cannot always be completely identified. An object should be identified to detect abnormal behavior, but detection itself is difficult. This is because, if the target object is mostly rigid, it should handle all possible variations through tractable computations [29]. In this section, we proposed anomaly detection as a countermeasure to point out and cope with the challenges that access control in the video domain will address in the future. Moreover, we mentioned that anomaly detection is closely related to access control, it can also be a countermeasure to visual security. Although the proposed countermeasure could be a response to other challenges, we believe it is more closely related to access control. Therefore, Section III-B.3 does not make any additional mention of anomaly detection as a countermeasure to visual security.

B. VISUAL SECURITY

Visual security is a technology to protect the confidentiality of video data from unauthorized users. Fig. 4 shows the latest trend in video data encryption. Video data encryption can be further classified into two methods: performing compression and encryption simultaneously and performing compression and encryption separately. When compression and encryption occur simultaneously, encryption is conducted after transformation, after quantization, and within entropy coding. When compression and encryption are performed independently, they can be subdivided into pre-compression encryption and post-compression encryption [30]. Video encryption types include fully layered encryption, permutation-based encryption, selective encryption, and perceptual encryption [31]. Fully layered encryption encrypts all bytes in the entire moving picture expert group stream using standard encryption schemes such as DES or AES, with the entire content compressed first. Permutation-based encryption scrambles or encrypts visual data using different permutation algorithms. Algorithms for impairment visual perception include pure,

TABLE 5. Research area of intelligent video data validation in video data security.

			Adaptive technology					Recognition		
			P17	P18	P19	P20	P21	P15	P16	P22
Forgery Detection	Active	Watermark	✓	✓	✓	✓	✓	✓		
		Digital signature								
		Intelligent techniques							✓	✓
	Passive	Spatial		✓	✓	✓	✓	✓	✓	✓
		Temporal								
		Spatio temporal	✓							
	Inter-frame	Frame-insertion /deletion/duplication	✓							
		Temporal interpolation	✓							
	Intra-frame	Crop			✓		✓		✓	
		Copy-move forgery						✓		
		Resampling				✓	✓			
		Brightness and contrast							✓	
		Tampering			✓					✓
		Image insertion			✓					
		Image filtering		✓	✓	✓	✓		✓	

zig-zag, Huffman code word, compression logic-based random, and correlation preserving. Selective encryption is useful for real-time encryption because it selectively encrypts the bytes within a video frame by reducing computational complexity. In perceptual encryption, the quality of aural/visual data is only partially degraded by encryption, which makes it suitable for commercial video systems and on-demand encryption.

The visual security method is similar to visual privacy protection in that it protects visually identifiable privacy from unauthorized users in the context of video data security. Different types of visual security to protect the information in images include intervention, blind vision, secure processing, redaction, and data hiding [32]. Fig. 5 shows a representative method for visual security. (a) is steganography for image or frame, and (b) is an intelligent visual security method using machine learning. The intervention creates an anti-capture domain to prevent capturing scenes that feature private information from unauthorized users, and it is conducted by intervening with physical devices. Blind vision refers to the processing of images or videos as anonymous, and it can be applied using secure multi-party computation (SMC) technology. The aim of secure processing is to anonymously process video or images in order to protect individuals' privacy using technologies other than SMC. Redaction is a method of protecting a subject's privacy by modifying sensitive areas of an image or video through techniques such as image filtering, video and image encryption, face de-identification, object/people removal, and visual abstraction/object replacement. Finally, data hiding is employed to protect privacy by overwriting a fake image over the original image through approaches such as covert channels, steganography, anonymity, and copyright marking. The subsections in this study focus mainly on redaction and data hiding because these methods can be more easily combined with intelligent security technologies.

1) ENCRYPTION

Fu. Y. *et al.* [P3] considered an encryption method based on reversible data hiding (RDH) to protect the confidentiality of multimedia data in a cloud environment. They stated that reversible data hiding in the encrypted image (RDHEI) can be applied to image processing to ensure privacy and secure cloud computing based on the RDH method, which utilizes the compressibility of natural images. The proposed scheme was able to achieve satisfactory rate-distortion performance and a reasonable embedding rate. In this scheme, source images are masked by applying stream encryption. Then, in order to make room for the data, the most significant

bits (MSB) layer of the embeddable block is adaptively compressed according to the MSB's frequency of occurrence. This enables the embedding of additional data in the MSB layer.

Song. Y. *et al.* [P4] argued that the effectiveness of encryption for real-time video data security has not been fully considered. Accordingly, the researchers proposed a chaotic selective encryption scheme (CSES) based on context-adaptive binary arithmetic coding (CABAC) of H.264 / advanced video coding (H.264 / AVC). Sensitive objects are identified based on CABAC's syntax element analysis to meet the requirements of real-time encryption. Then, two encryption methods are presented in CSES, combined with two chaos-based keystream generators (KSGs) for efficient encryption of identified objects. However, as various binarization methods affect performance in high-efficiency video coding (HEVC), it is necessary to continuously research the characteristics of the syntax element selected in HEVC's CABAC in order to design an efficient encryption scheme.

Thiyagarajan. K. *et al.* [P5] pointed out that HEVC, which is the latest video coding standard used for video compression, can make encryption less efficient when applied to energy-constrained environments such as the internet of multimedia things (IoMT). Although HEVC is capable of encrypting certain syntax elements in a video stream and reduce computational complexity by decreasing the overhead of bitrates, it is not suitable for energy-constrained systems because it increases the overall computational cost due to integrated encryption and compression. In order to tackle this issue, the study discussed here proposed an energy-aware HEVC encryption scheme that can be applied in an energy-constrained environment with reference to related literature. The proposed scheme identifies syntax elements that will be encrypted according to the structure, texture, and motion in each frame of a video. As a result, the study's researchers demonstrated that the proposed scheme can be effectively applied to energy-constrained environments by reducing the encryption overhead.

Amani. H. R. *et al.* [P6] argued that traditional encryption techniques (e.g., DES, IDES, AES, DNA-based encryption) used for digital image security are not sufficiently secure. However, the researchers suggested that the chaos system is appropriate because it features a variety of important characteristics for encryption such as high sensitivity to initial conditions, certainty, and ergodicity. In particular, the study in question proposed a scheme using a DNA sequence and hyper-chaotic dynamics in the adaptive encoding of color images. The researchers also analyzed and cited many studies related to the hyper-chaos system and DNA encoding in order to describe their proposed scheme in detail. The main challenge with this model is that pixel logic arrays are removed using Arnold's cat map. Moreover, scrambled images are combined with Chen's chaotic system, DNA sequences, and the proposed adaptive method to increase the complexity of the encryption algorithm.

Although most cloud services provide visual security for privacy, Guo. J. *et al.* [P10] pointed out that these services tend to focus on parameter estimation and anomaly detection in an encrypted video bitstream. To address this problem, the study investigated encrypted bitstream focusing on parameter estimation and anomaly detection. The effectiveness and feasibility of conventional anomaly detection schemes usually depend on the available pixel values or bitstream parameters, but the anomaly detection of privacy is excluded. Since input data and/or parameters are unavailable, it is challenging to design a practical privacy-preserving anomaly detection scheme. Therefore, the researchers proposed a method combining macroblock sizes, macroblock partitions, and motion vector difference magnitude to omit video decryption, full compression/decompression, and an interactive protocol because the video process can become quite complex if an encryption tool is used. The proposed scheme is compatible with different video encryption methods, and it was implemented in a parallel structure to accelerate the running time.

Jiang. R. *et al.* [P12] pointed to scrambling as a prominent technique for preserving privacy because it is relatively harder to decrypt than full encryption. However, the researchers also highlighted an issue with scrambling; namely, that chaotic signal processing approaches are required for face recognition in scrambling domains. This is because face models turn into chaotic signals after scrambling. The study discussed here focused on the challenges of face recognition, using existing data-based face recognition algorithms to alleviate them. Moreover, the study proposed a new ensemble approach featuring many kernels random discriminant analysis (MK-RDA) to detect differential patterns from chaotic signals. The researchers also combined the salience-aware strategy proposed ensemble technique to pattern the protruding areas. First, the training set is transferred to the training procedure and the offline procedure learns its semantic salience map. Then, the database is scrambled, and the feature space becomes reconstructed based on semantic salience by multiplying salient features. After that, random sampling is applied to select the features sparsely and construct the kernel that is allowed. Discriminant analysis is used to learn a kernel subspace for each kernel. Finally, the scrambled dataset is tested, the input dataset is projected into each kernel subspace, and the distance to each training sample is computed.

The key challenge of selective visual privacy protection is identifying whether a particular individual belongs to a privacy group. Ying. L [P14] proposed a visual privacy protection system that best matches the abovementioned information using iris patterns to address privacy problems caused by the widespread use of surveillance cameras. The study also proposed an anonymous subject identi?cation procedure to protect personal information from unidentified individuals. This approach guarantees privacy by determining whether there is a clear match with an iris probe signal or not.

2) PROTECTION METHOD

Hosam. O. *et al.* [P7] have referred to steganography as a technique that can be used to safely hide data including images and videos. This study, based on Wu and Tsai [33], proposed a PVD (pixel value differencing) method for embedding security data in digital images. The PVD steganography algorithm embeds data in an image based on pixel adjacency differences. The image is divided into 3×3 blocks, and the difference between the pixel's minimum and maximum values is adjusted to distribute the security data featured in an image. Secure data is embedded in the image's content area through edge and intensity transitions, and a textured image provides a higher embedding size than a normal image.

Duan. X. *et al.* [P8] stated that although multimedia information (such as text and images) can be quickly shared over a network, this poses a number of security risks. Thus, steganography is a basic algorithm suitable for hiding information that can reveal more complex statistical features using the simplest algorithms such as LSB to HUGO, SUNIWARD, and WOW. Recently, steganalysis features have become more complex and high-dimensional in order to compete with more advanced adaptive steganography. To this end, recent steganography-related studies have analyzed the high-order statistical characteristics of steganalysis based on complex correlation modeling in the image domain. Accordingly, in the research described here, a neural network was introduced to hide image information. Moreover, a WGAN-GP model was implemented and used to conduct experiments.

Traditionally, data is hidden by passing the payload based on a quantized discrete cosine transform (QDCT). However, QDCT coefficients have the disadvantage of causing intra-frame and inter-frame distortion drift because they expose video frames' texture and motion characteristics. In order to avoid visual artifacts and to maintain an adequate bitrate, frame distortions should be considered. With this in mind, Chen. Y. *et al.* [P9] proposed an adaptive data hiding scheme based on cost assignment and the syndrome-trellis code (STC) in which a cost assignment function is constructed based on video sequence features such as texture characteristics, motion properties, and frame position. Specifically, the proposed scheme considered intra-frame changes and intra-frame distortion drift, for which frames' texture and motion changes can be measured. Finally, they used STC to embed data by minimizing overall distortion.

This study employed the OpenFace approach proposed in the previous study [34] to recognize personal information and explain bottom-up ecosystem configuration focused on real-time video analysis for visual privacy. Wang. J. *et al.* [P11] attempted to apply high-performance techniques such as high processing speed and high accuracy to DeepFace and FaceNet based on OpenFace. OpenFace can be combined with interframe tracking to implement RTFace, a privacy protection system that provides low privacy leak rates. RTFace is a mechanism for denaturing video streams that selectively blurs human faces based on specific policies. This approach enables the management of privacy during real-time video analysis while providing a secure method for handling retrospective policy exceptions. However, long-standing challenges remain. For instance, there are issues in recognizing human faces when disease, aging, or facial alteration are present. Faces can be transformed through makeup, facial expressions, and styling. Moreover, it is impossible to recognize a large number of faces at the same time in the real-time environment targeted by this study.

Brkic. K. *et al.* [P13] argued that the observer in public places where surveillance takes place may need to selectively access objects via particular scenes or authority regarding a particular event. To this end, the study proposed a computer vision-based automatic de-identification pipeline. This scheme is able to protect an object's identifying information by obfuscating appearance while maintaining the naturalness and utility of unidentified data. This method is unable to obtain accurate identifying information (e.g., hair, skin color, clothing, etc); however, shapes can be distinguished using the blur technique. The study's research was conducted in a video surveillance environment using the improved GrabCut algorithm to only extract pedestrians. Then, objects can be rendered in a different style by employing the neural art algorithm, which uses the responses of a deep neural network that alter the appearance of the segmented object.

3) THE CHALLENGES OF VISUAL SECURITY

The use of video surveillance in public infrastructure is steadily increasing, increasing the amount of data collected. To manage collected video data, various methods such as compression and disposal have been proposed. However, it isn't easy to securely manage vast amounts of collected data. As shown in Table 4, we analyzed the studies that deal with intelligent security technologies on video data, but open challenges remain. In this study, we analyzed not only video but also studies on intelligent visual security focusing on images, but the number of studies that only considered video is rare. Like steganography, it has a relevance to visual security, but I can't answer some of the items in Table 4 (i.e., encryption moment, encryption method) because it also includes studies that are relatively less relevant to encryption.

When considering the purpose of video surveillance, particular video scenes required by users are likely to contain specific events or accidents, and users typically expect these scenes to be compact. Synopsis technology can reduce the burden of visual security processing by representing an original video scene compactly. Furthermore, this approach can bypass security issues associated with privacy and access control when combined with anomaly detection because only a video's detected objects will be represented compactly. Although video synopsis is an effective technology for representing objects and storing video data, it may not be suitable for complex situations. This is because synopsis heavily depends on the pre-processing results of foreground segmentation and multiple object tracking, and pre-processing techniques usually result in poor performance in crowded

scenes [35]. In other words, this implies that object collision [36] and the optimization process [37] need to be improved.

While the foregoing was referring to a capacitive challenge, the challenge covered below is caused by video technology development. As video technology develops, the video information collected clearly identifies a lot of personal information (e.g., race, gender). Although video surveillance in public places is generally accepted because there is no expectation of privacy, rapid improvements in surveillance systems have made it possible for video surveillance to capture information that exceeds existing expectations [38]. Accordingly, to increase the efficiency of visual security, we should prefer partial encryption of critical areas rather than overall encryption of the video. However, there are still challenges of object detection and identification. Although many studies have been conducted on object detection, prior research has not figured out how to identify sensitive information through high-performance methods. There are several challenges to improving the performance of object detection, including:

- **Dual priorities:** In order to detect an object, we should classify the image object in addition to locating the object.
- **Real-time detection speed:** Quick decisions are often needed depending on the situation. It is necessary to accurately classify and localize the objects that should be detected and process them in real time.
- **Multiple spatial scales and aspect ratios:** When many objects exist, they cannot be expressed individually. Thus, we should be able to represent the objects of interest with the appropriate size and aspect ratio.
- **Limited data:** There are currently fewer annotations available for object detection compared to classes for object images.
- **Class imbalance:** Class imbalance is a hindrance to classifying objects. At this time, everything except the main object in the image can be recognized as the background.

C. VIDEO DATA VALIDATION

Video data validation is a technology to protect integrity through verification of forgery and modulation of video data. Fig. 6 is the latest trend to verify the integrity of video data. There are inter-frame-based and intra-frame-based methods for video forgery. Depending on the forgery method, the passive-based approach or the active-based approach [39] should be selected for the video forgery detection method. In the case of the active approach, digital video undergoes some pre-processing like watermark insertion, digital signatures, etc., which would degrade the video's performance to a certain extent at the time of creating the video. If the video is forged, then recovery of the watermark or signature would not be possible, and its identification will result in tampering detection. The active approach cannot be used in scenarios with no pre- processing insertions (i.e., watermark or signature absent). On the other hand, the passive approach works on the basic assumption that video contains naturally occurring properties or inherent fingerprints unique to it due to different video imaging devices and their characteristics. If the video is not forged, then the given video's underlying statistical correlations remain consistent after some post-processing operations. Before and after video acquisition, identifying the difference, which may be a malicious or non-malicious alteration, is the main task of various detection methodologies. In most cases, passive forensics can be converted to a problem of pattern recognition.

Inter-frame forgery affects video frame sequences and includes specific types of forgery such as frame removal, insertion, and copying [40]. Techniques for detecting forgeries include sensor artifacts, double compression, motion and brightness features, and pixel-level analysis. Intra-frame forgery refers to the modification of content—such as copy-paste and upscale-crop—in a video's individual frames. Techniques for copy-move forgery detection include pixel-similarity analysis and object and motion features. These forgery methods induce spatial, temporal, and spatio-temporal tempering [41]. In the subsections below, we present our analysis of studies incorporating intelligent security in video data encryption and video forgery modulation detection.

Chakravarthy. S. *et al.* [P15] claimed that images containing sensitive information should be watermarked to detect ownership verification and tampering. In particular, related works focused mainly on robustness and imperceptibility, which are critical factors in watermarking technology. However, a few issues remain. For instance, a complication arises when an intruder is prevented from obtaining a watermark signal from a watermarked image because this can expose future point-to-point correspondences. The proposed scheme for tamper detection combines various methods such as integer wavelet transform (IWT), singular value decomposition (SVD), and Fibonacci Lucas transform (FLT) to improve the security of a watermarking scheme where there is fewer data to exchange before each transaction. Moreover, this model is optimized through heuristics and metaheuristics and employs the sine cosine algorithm (SCA) to achieve a trade-off between robustness and imperceptibility. The researchers also built an artificial neural network (ANN) to assume attack types and show image quality metrics using peak signal-to-noise ratio (PSNR), structured similarity (SSIM), and correlation coefficient.

As a method for detecting face manipulation, Dang. L. M *et al.* [P16] proposed a framework to detect deep learning-based manipulated face images, diverging from existing studies that rely on traditional manuals. Although some problems exist with this method—such as the lack of training data and poor performance—they can be solved by employing ensemble techniques that increase the frequency of minority layers. Such techniques include undersampling and oversampling. The proposed manipulated face (MANFA) is a convolutional neural network (CNN)

model for identifying manipulated face images. This model extracts abstract features from the manipulation regions and eliminates dataset imbalances through adaptive boosting (AdaBoost) and extreme gradient boosting (XGBoost).

Feng. C. *et al.* [P17] pointed to the existence of numerous digital watermarking studies focused on detecting image tampering; however, these studies require special techniques or equipment to detect tampering. In contrast, digital forensics explores the intrinsic features of a particular media rather than using prebuilt ones. The study discussed here explored frame deletion forgery detection. Detection methods are usually classified into side effects detection by frame deletion in compressed images and detection using the difference between frame deletion point (FDP) and the reference frame. However, the methods presented in existing studies do not apply to image sequences with variable motion intensities. Moreover, the effects of interference frames have not been considered. Therefore, to alleviate this problem, Feng. C. *et al.* developed an assistant tool based on frame motion residuals to identify the FDP.

Watermarking, a technique for preventing digital image manipulation, often causes deterioration of the host image. To tackle this challenge, Rangel-Espinoza. K. *et al.* [P18] proposed a scheme based on removable visible watermarking systems. The researchers considered the transparency and visibility of visible watermarks and the quality of images after removing visible watermarks. In other words, their study proposed a removable visible watermarking system based on dual watermark technology and blind removal. The cosine discrete transform (DCT) was used to generate watermarks taking into account the texture and luminance characteristics of watermarks and host images. In addition, QIM-DM (Quantization Index Modulation-Dither Modulation) technology is used to prevent illegal watermark removal. In sum, the proposed scheme can obtain a high-quality host image when the correct key is used; however, the use of the wrong key severely distorts the image. Similar studies have proposed adaptive watermarking for detecting forgery including region-adaptive semi-fragile dual watermarking [P19], image-adaptive watermarking [P20], and adaptive blind image watermarking [P21].

Copy-move forgery refers to the pasting of an area of a particular image onto another area of the same image. This type of forgery is not easy to detect because it is performed on the same image and image properties such as noise components and colored text in the paste area are compatible. Accordingly, Bi. X. *et al.* [P22] proposed adaptive over-segmentation and feature point matching (ASFPM) by integrating both features. Research on copy-move forgery detection has been conducted in the past few years, and the main issue revolves around the block-based function and key point-based function for detecting forgery. Block-based algorithms divide the host image into blocks and extract block functions, while key point-based algorithms ensure geometric transformations' robustness and selectively extract matching key point functions. The proposed adaptive over-segmentation and feature point matching (ASFPM) works by integrating both features. The proposed scheme generates multiple scales by dividing the patch into blocks through the block-based function and applying the scale-invariant feature transform to extract the function points from all the patches. An adaptive patch-matching algorithm is proposed to find a match that represents a suspected forged region at each scale. Then, the key point-based feature is used to detect and generate suspicious scales.

1) THE CHALLENGES OF VIDEO DATA VALIDATYION

As shown in Table 5, we analyzed studies related to video data validation. However, challenges, including technical challenges for video data validation, remain. Although this section analyzes video data validation studies, including images, the method to verify video forgery is not the same as the method to verify video forgery. In particular, a subtle difference exists between the semantic of a digital data and the authenticity of digital evidence, some methods are only used to detect simple distortion but cannot be used in detecting complex ones [42]. In other words, compressed videos are a serious challenge to their method because noise correlation is not a reliable feature in a compressed video. The task of noise residue extraction still is a complex one.

In the aspect of technique, there are three major challenges in watermarking for video forgery verification [43]. First, there are many non-hostile video processings, which are likely to alter the watermark signal. These are photometric attacks, spatial resynchronization, temporal resynchronization, and video editing that gather all attacks that modify the pixel values of a frame caused by extensive video processing. Second, resilience to collusion is much more critical in the context of the video. Collusion refers to a set of malicious users who merge other watermark data to create unwatermarked data. Collusion type is classified as a case where the same watermark contains different data or if the same data contains different watermarks. Third, real-time is often a requirement for digital video watermarking. When considering real-time, the watermarking algorithm designed by Philips Research could be often considered as a reference.

IV. OPEN CHALLENGES OF VIDEO DATA SECURITY IN PUBLIC INFRASTRUCTURE

In the present study, we have analyzed research on intelligent video data security and also discussed concerns about remaining challenges in order to suggest future research directions. In particular, recent video surveillance has exposed many threats as both the size of the software stack and the openness of the network infrastructure increase [4]. Accordingly, the American Civil Liberties Union [44] has emphasized specific issues of surveillance in public infrastructure (ACLU 2020) as follows:

- **There is not enough proof regarding the effects of video surveillance:** The recent increase in safety concerns has justified the strengthening of video surveillance. However, for some crimes, video surveillance has

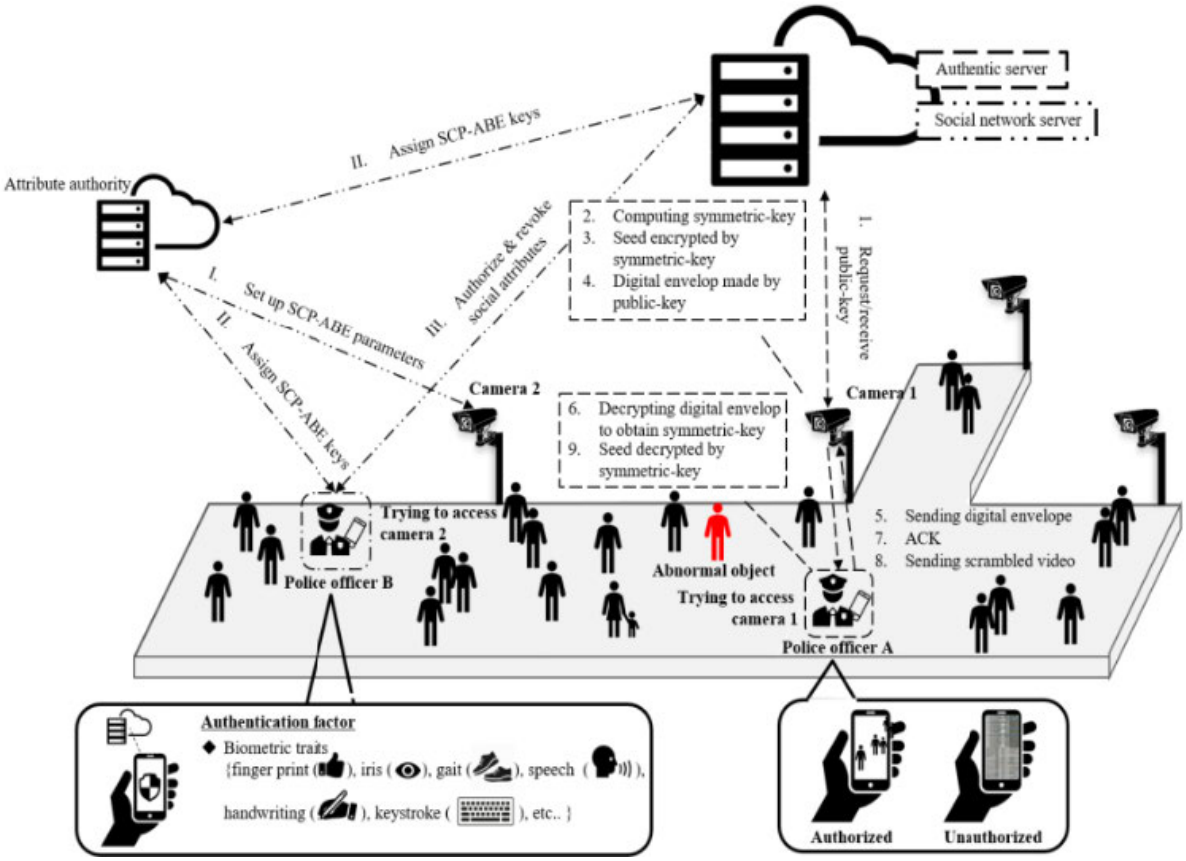

FIGURE 7. User access control approaches for video data security in public infrastructure.

not functioned properly, and the indiscriminate expansion of such systems could improperly waste state budgets. On the basis of this allegation, some sociologists have said that it is difficult to characterize the correlation between video surveillance and crimes in public places. Moreover, technical experts in the U.S. government have pointed out that video screen monitoring is typically allotted a timeframe of under 20 minutes.

- **Video data prone to abuse:** As mentioned earlier, video data now contains more information than we previously thought. This information can be abused in five different ways: criminal abuse, institutional abuse, abuse for personal purposes, discriminatory targeting, and voyeurism.
- **The lack of limits or controls on camera use:** The rapid development of camera-related technologies has made it possible to have access to many devices with desired performance. However, there are no clear boundaries defining cameras' functions depending on location. How should we map the performance and function of specific locations and cameras?

Although the challenges mentioned above have been fully considered in Section III of the present study, this section analyzes the inherent challenges of video data security that should be considered when video surveillance is used in public infrastructure locations. First, we should consider the absence of intelligent access control research in the video domain (mentioned earlier in Section III). When searching for relevant literature using our proposed search strategy, we located studies on general access control for multimedia and domain security. These studies did not address intelligent security technology. However, we devised a method of intelligent access control in public infrastructure with reference to [P1, P2], as Fig. 7 shows. The referenced study [45] systematically reviewed research related to biometrics-based one-factor and two-factor authentication, classifying and analyzing classic research methods. Mondal and Bours *et al.* [46] proposed a keystroke, an intelligent authentication method based on the pairwise user coupling technique. Ma *et al.* [47] proposed a scalable media access control (SMAC) system that takes into account the complex and diverse social relationships of social network users to protect their personal information in large-scale sharing systems. Additionally, Yang *et al.* [48] proposed a time-domain attribute-based access control (TAAC) as a way to safely share the contents of a large-scale sharing system. Based on [P2] and a combination of the above studies, Fig. 7 presents an intelligent access control method that can safely share large-scale video data in public infrastructure. The overview of the scenario is as follows.

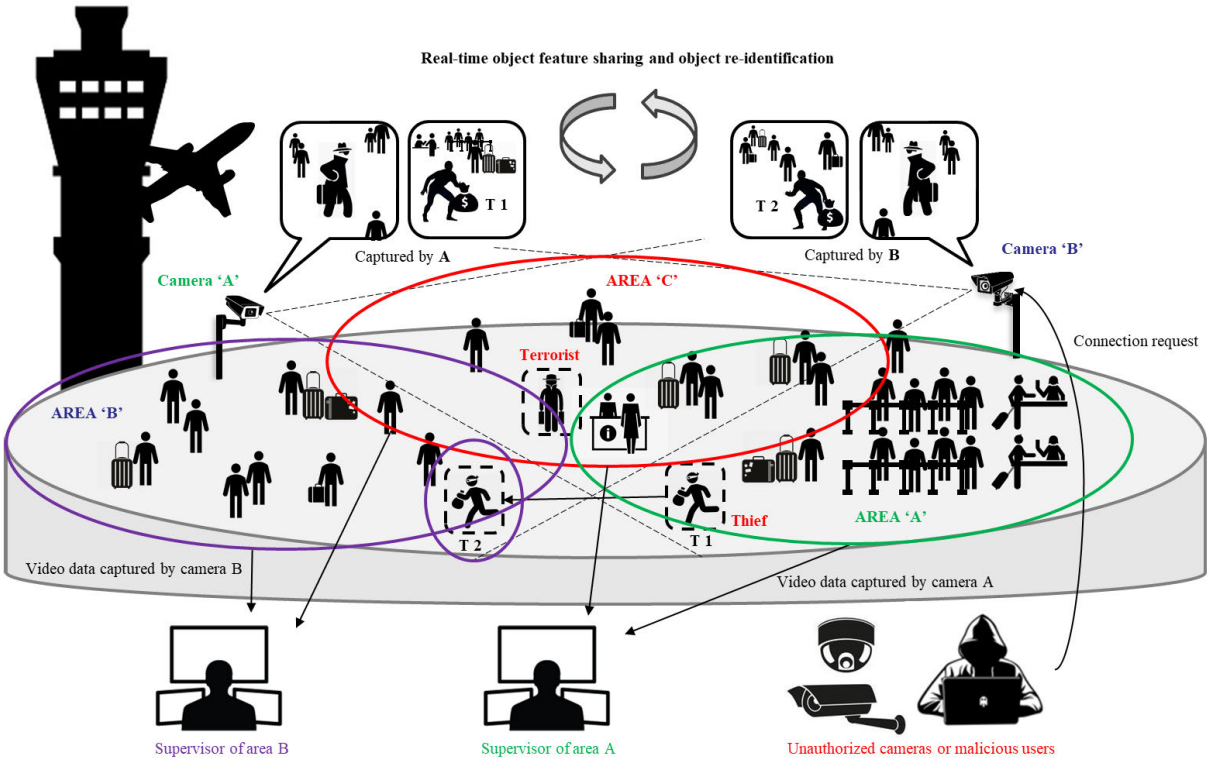

FIGURE 8. Open challenges of video data security in public infrastructure.

1. In order to catch an "Abnormal object" such as sexual harassment in public infrastructure, "Police officer A," a video information requester, sends a request to view video information to "Camera 1," a surveillance camera.
2. "Camera 1"delivers a request to view video information from "Police officer A" to the cloud system.
3. The cloud system generates an encryption key to use for video information encryption, puts it in a digital envelope, and delivers it to the IoT devices of "Camera 1" and "Police officer A."
4. Camera 1" opens the electronic envelope, decrypts the video information with the encryption key received from the cloud system, and extracts the synopsis video of the "Abnormal object."
5. "Camera 1" allows "Police officer A" to perform authentication for viewing video information to the cloud system through an IoT device and encrypts and delivers the extracted synopsis video.
6. "Police officer A" attempts to decrypt the synopsis image through biometric authentication factors such as iris, fingerprint, and voice of IoT devices.
7. In the case of "Police officer A" equipped with an appropriate authentication factor, it is possible to track the "Abnormal object" by viewing the video information.
8. Even if the "Abnormal object" flees with "Camera 2," the "Police officer B" can continue to grasp the escape route through the above authentication process.

In addition to intelligent user access control technology, a cloud system capable of storing a large amount of video information, synopsis technology is required to meet the above detailed analysis of existing studies and further research on evaluation and performance should be conducted.

Fig. 8 is also one of the examples to illustrate the challenges of video data security in public infrastructure. Public infrastructure has increased the demand for intelligent video surveillance systems that integrate surveillance devices through the web to continuously create safe public environments [49]. However, the source such as manufacturer, operation system of video surveillance system devices can change. Even if demand comes from the same source, it can exhibit heterogeneity with previous devices owing to devices' evolution or changes to improve performance. In other words, it can be difficult to share information between existing devices and newly distributed devices due to this heterogeneity. To address this challenge, heterogeneous devices require specialized architecture whereby a centralized management server (CMS) and the client handle various types of media encoding and connection protocols [50]. Furthermore, various standards have been established to tackle interoperability challenges, while emerging research on interoperability has been conducted [51].

However, further research should be done on interoperability security to solve a number of challenges. For example, as Fig. 8 shows, sharing information between cameras is a limit to tracking the criminal with one camera, so it is necessary to be able to identify specific hazards (e.g., terrorists,

TABLE 6. Selected Studies.

Code	Author	Title	Year	Publication
P1	Jinsu. K and Namje. P	Lightweight knowledge-based authentication model for intelligent closed-circuit television in mobile personal computing	2019	Personal and Ubiquitous Computing
P2	Shen. J et al.	A context-aware mobile web middleware for service of surveillance video with privacy	2015	Multimedia Tools and Applications
P3	Fu. Y et al.	Effective reversible data hiding in encrypted image with adaptive encoding strategy	2019	Information Sciences
P4	Song. Y et al.	Efficient protection using chaos for Context-Adaptive Binary Arithmetic Coding in H. 264/Advanced Video Coding	2019	Multimedia Tools and Applications
P5	Thiyagarajan. K et al.	Energy aware encryption for securing video transmission in Internet of Multimedia Things	2018	IEEE Transactions on Circuits and Systems for Video Technology
P6	Amani. H et al.	A new approach in adaptive encryption algorithm for color images based on DNA sequence operation and hyper-chaotic system	2019	Multimedia Tools and Applications
P7	Guo. J et al.	Efficient Privacy-Preserving Anomaly Detection and Localization in Bitstream Video	2019	IEEE Transactions on Circuits and Systems for Video Technology
P8	Jiang. R et al.	Face recognition in the scrambled domain via salience-aware ensembles of many kernels.	2016	IEEE Transactions on Information Forensics and Security
P9	Ying. L et al.	Anonymous subject identification and privacy information management in video surveillance	2017	International Journal of Information Security
P10	Hosam et al.	Adaptive block-based pixel value differencing steganography	2016	Security and Communication Networks
P11	Duan et al.	A coverless steganography method based on generative adversarial network	2020	EURASIP Journal on Image and Video Processing
P12	Chen. Y et al.	Adaptive Video Data Hiding through Cost Assignment and STCs	2019	IEEE Transactions on Dependable and Secure Computing
P13	Wang. J et al.	Enabling live video analytics with a scalable and privacy-aware framework	2018	ACM Transactions on Multimedia Computing, Communications
P14	Brkic. K et al.	Protecting the privacy of humans in video sequences using a computer vision-based de-identification pipeline	2017	Expert Systems with Applications

TABLE 6. *(Continued.)* **Selected Studies.**

P15	Chakravarthy. S et al.	An optimized hierarchical encryption technique for tamper recognition	2019	Multimedia Tools and Applications
P16	Dang. L et al.	Face image manipulation detection based on a convolutional neural network	2019	Expert Systems with Applications
P17	Feng. C et al.	Motion-adaptive frame deletion detection for digital video forensics	2016	IEEE Transactions on Circuits and Systems for Video Technology
P18	Rangel-Espinoza. K et al.	Adaptive removable visible watermarking technique using dual watermarking for digital color images	2018	Multimedia Tools and Applications
P19	Shi. H et al.	A region-adaptive semi-fragile dual watermarking scheme	2016	Multimedia Tools and Applications
P20	Bhinder. P et al.	An improved robust image adaptive watermarking with two watermarks using statistical decoder	2020	Multimedia Tools and Applications
P21	Fazlali. H. R et al.	Adaptive blind image watermarking using edge pixel concentration	2017	Multimedia Tools and Applications
P22	Bi. X et al.	Multi-scale feature extraction and adaptive matching for copy-move forgery detection	2018	Multimedia Tools and Applications

thieves) on other cameras through quick sharing. However, interoperability should also be considered for situations other than these special circumstances. For instance, information about normal people from zone R should not be shared in zone P by Camera A. This would concern not only visual security issues but also access control issues. In this regard, large numbers of cameras are often deployed in public infrastructure, meaning that there are vast amounts of information that can be shared. Thus, security challenges such as external access or connections with unauthorized cameras should be considered. The main contents of scenario are as follows.

Abnormal behavior detection and tracking

1. Theft crime occurred in the area controlled only by ''Camera A.''
2. ''Camera A'' detects and tracks an abnormal object (the thief), extracts the synopsis video at the same time, and delivers it to ''Supervisor A.''
3. ''Camera A'' continuously tracks the thief(T2) from the occurrence of the incident (T1) to determine the path of movement and transmits the thief's features to ''Camera B,'' which collects video information on the path the thief travelled.
4. ''Camera B'' identifies and verifies the object based on the thief's features delivered from ''Camera A.''
5. ''Camera B'' detects and tracks thieves, extracts synopsis video in real-time, and delivers them to ''Supervisor B.''

Risk recognition and re-identification

1. ''Camera A'' recognizes risk factors (e.g., terrorists, wanted offenders, persons in possession of abnormal objects) that exist in ''AREA C.''
2. ''Camera A'' detects risk and extracts a video synopsis that contains the risk factors and delivered to ''Supervisor A.''
3. ''Camera A'' extracts the features of the detected risk and shares it in real-time with the camera (Camera B) that controls ''AREA C'' in common.
4. ''Camera B'' identifies and verifies the object based on the risk factor features received.
5. ''Camera B'' extracts a video synopsis that is containing detected risk factor and delivers it to ''Supervisor B.''

In addition to the aforementioned events, the proposed scenario considers countermeasures for complex threats, including network and physical abnormal access security. Identification, and recognition technology for risk factors, re-identification technology that can identify the same entity, tracking technology for continuous observation of objects, abnormal behavior detection technology that can identify anomaly things, real-time sharing technology for real-time context information sharing, access control technology that allows only authenticated and authorized users to view the video, and video synopsis technology capable of compacting large amounts of video information are required to make up the secure public infrastructure. As a future study, we will try

to create a safe public infrastructure environment by applying intelligent video data security technologies to the actual public infrastructure environment (i.e., smart city).

V. CONCLUSION

The purpose of this study was to analyze the current status of video data security technology combined with intelligent technology and to classify the research area dealing with video data security. To this end, we attempted to cite reliable literature by adopting the SLR methodology because feasibility is critical for survey-based studies such as this one. We also designed a video data security taxonomy and analyzed the literature that was collected literature via the proposed search strategy. This taxonomy was obtained by examining the diverse research areas of video data security based on the proposed process rather than through the results of analyzing the extracted literature. Based on our knowledge, this study is the first to consider the diverse aspects and challenges of video data security and to propose a specific taxonomy. Moreover, almost no prior literature has analyzed the trends of video data security technologies combined with intelligent technology by locating and analyzing relevant literature through a search strategy. In other words, this study's findings serve to clarify the research area of video data security. Furthermore, we have clearly discussed the open challenges of implementing video technology in public infrastructure. In regard to future research, more detailed taxonomies could be constructed with reference to this study. Finally, we also hope that further studies will be done to explore unaddressed video data security research areas.

APPENDIX

See Table 6.

REFERENCES

[1] K. Low and U. Sheikh, "Review on human re-identification with multiple cameras," *J. Telecommun., Electron. Comput. Eng.*, vol. 8, no. 9, pp. 89–95, 2016.
[2] A. Shifa, M. N. Asghar, M. Fleury, and M. S. Afgan, "Ontology based intelligent security framework for smart video surveillance," in *Proc. Future Technol. Conf.* Cham, Switzerland: Springer, 2018, pp. 118–126.
[3] A. Malathi and S. S. Baboo, "Evolving data mining algorithms on the prevailing crime trend-an intelligent crime prediction model," *Int. J. Sci. Eng. Res.*, vol. 2, no. 6, pp. 1–6, 2011.
[4] T. Winkler and B. Rinner, "Privacy and security in video surveillance," in *Intelligent Multimedia Surveillance*. Berlin, Germany: Springer, 2013, pp. 37–66.
[5] K. He, J. Chen, Y. Zhang, R. Du, Y. Xiang, M. M. Hassan, and A. Alelaiwi, "Secure independent-update concise-expression access control for video on demand in cloud," *Inf. Sci.*, vol. 387, pp. 75–89, May 2017.
[6] F. K. Tabash, M. Izharuddin, and M. I. Tabash, "Encryption techniques for H.264/AVC videos: A literature review," *J. Inf. Secur. Appl.*, vol. 45, pp. 20–34, Apr. 2019.
[7] S. K. and B. M. Mehtre, "Detection of inter-frame forgeries in digital videos," *Forensic Sci. Int.*, vol. 289, pp. 186–206, Aug. 2018.
[8] N. Ruchaud and J.-L. Dugelay, "Privacy protecting, intelligibility preserving video surveillance," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2016, pp. 1–6.
[9] G. V. Glass and M. L. Smith, "Meta-analysis of research on class size and achievement," *Educ. Eval. Policy Anal.*, vol. 1, no. 1, pp. 2–16, Jan. 1979.
[10] R. Mallett, J. Hagen-Zanker, R. Slater, and M. Duvendack, "The benefits and challenges of using systematic reviews in international development research," *J. Develop. Effectiveness*, vol. 4, no. 3, pp. 445–455, Sep. 2012.
[11] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering, version 2.3," Keele Univ., Univ. Durham, Durham, U.K., Tech. Rep., 2007.
[12] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering—A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, Jan. 2009.
[13] K. S. Khan, R. Kunz, J. Kleijnen, and G. Antes, "Five steps to conducting a systematic review," *JRSM*, vol. 96, no. 3, pp. 118–121, Mar. 2003.
[14] H. Small, "On the shoulders of Robert Merton: Towards a normative theory of citation," *Scientometrics*, vol. 60, no. 1, pp. 71–79, 2004.
[15] R. K. Merton, "The matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property," *Isis*, vol. 79, no. 4, pp. 606–623, Dec. 1988.
[16] N. Dey, A. E. Hassanien, C. Bhatt, A. Ashour, and S. C. Satapathy, *Internet of Things and Big Data Analytics Toward Next-Generation Intelligence*. Berlin, Germany: Springer, 2018.
[17] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE Access*, vol. 6, pp. 32328–32338, 2018.
[18] J. Wen, M. Severa, W. Zeng, M. H. Luttrell, and W. Jin, "A format-compliant configurable encryption framework for access control of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 545–557, Jun. 2002.
[19] K. Kraus, O. Martikainen, and R. Reda, "Security management process or video surveillance systems in heterogeneous communication networks," in *Proc. 1st IFIP Wireless Days*, Nov. 2008, pp. 1–5.
[20] M. A. Hamoudy, M. H. Qutqut, and F. Almasalha, "Video security in Internet of Things: An overview," *Int. J. Comput. Sci. Netw. Secur.*, vol. 17, no. 8, p. 199, 2017.
[21] X. Ma, W. K. Zeng, L. T. Yang, D. Zou, and H. Jin, "Lossless ROI privacy protection of H.264/AVC compressed surveillance videos," *IEEE Trans. Emerg. Topics Comput.*, vol. 4, no. 3, pp. 349–362, Jul. 2016.
[22] S. Tm and K. Ramesh, "Reviewing the effectivity factor in existing techniques of image forensics," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 7, no. 6, p. 3558, Dec. 2017.
[23] P. Remagnino, S. A. Celastin, G. L. Foresti, and M. Trivedi, "Novel concepts and challenges for the next generation of video surveillance systems," *Vis. Appl.*, vol. 18, pp. 135–137, May 2007.
[24] S. W. Shah and S. S. Kanhere, "Recent trends in user authentication—A survey," *IEEE Access*, vol. 7, pp. 112505–112519, 2019.
[25] A. Majumder, S. Namasudra, and S. Nath, "Taxonomy and classification of access control models for cloud environments," in *Continued Rise Cloud*. London, U.K.: Springer, 2014, pp. 23–53.
[26] C. Norris, *A Review of the Increased Use of CCTV and Video-Surveillance for Crime Prevention Purposes in Europe*. Brussels, Belgium: European Parliament, 2009.
[27] J. R. Medel and A. Savakis, "Anomaly detection in video using predictive convolutional long short-term memory networks," 2016, *arXiv:1612.00390*. [Online]. Available: http://arxiv.org/abs/1612.00390
[28] X. Mo, V. Monga, R. Bala, and Z. Fan, "Adaptive sparse representations for video anomaly detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 4, pp. 631–645, Apr. 2014.
[29] X. Wang, M. Yang, S. Zhu, and Y. Lin, "Regionlets for generic object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 10, pp. 2071–2084, Oct. 2015.
[30] F. Liu and H. Koenig, "A survey of video encryption algorithms," *Comput. Secur.*, vol. 29, no. 1, pp. 3–15, Feb. 2010.
[31] J. Shah and D. V. Saxena, "Video encryption: A survey," 2011, *arXiv:1104.0800*. [Online]. Available: http://arxiv.org/abs/1104.0800
[32] J. R. Padilla-López, A. A. Chaaraoui, and F. Flórez-Revuelta, "Visual privacy protection methods: A survey," *Expert Syst. Appl.*, vol. 42, no. 9, pp. 4177–4195, Jun. 2015.
[33] D.-C. Wu and W.-H. Tsai, "A steganographic method for images by pixel-value differencing," *Pattern Recognit. Lett.*, vol. 24, nos. 9–10, pp. 1613–1626, Jun. 2003.
[34] J. Wang, B. Amos, A. Das, P. Pillai, N. Sadeh, and M. Satyanarayanan, "A scalable and privacy-aware IoT service for live video analytics," in *Proc. 8th ACM Multimedia Syst. Conf.*, Jun. 2017, pp. 38–49.
[35] X. Li, Z. Wang, and X. Lu, "Video synopsis in complex situations," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 3798–3812, Aug. 2018.
[36] X. Li, Z. Wang, and X. Lu, "Surveillance video synopsis via scaling down objects," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 740–755, Feb. 2016.

[37] Z. Zhang, Y. Nie, H. Sun, Q. Zhang, Q. Lai, G. Li, and M. Xiao, "Multi-view video synopsis via simultaneous object-shifting and view-switching optimization," *IEEE Trans. Image Process.*, vol. 29, pp. 971–985, 2020.
[38] Senior, Andrew, "Privacy protection in a video surveillance system," in *Protecting Privacy Video Surveillance*. London, U.K.: Springer, 2009, pp. 35–47.
[39] K. Sowmya and H. Chennamma, "A survey on video forgery detection," *Int. J. Comput. Eng. Appl.*, vol. 9, no. 2, pp. 17–27, 2015.
[40] R. D. Singh and N. Aggarwal, "Video content authentication techniques: A comprehensive survey," *Multimedia Syst.*, vol. 24, no. 2, pp. 211–240, Mar. 2018.
[41] R. Sawant and M. Sabnis, "A review of video forgery and its detection," *J. Comput. Eng.*, vol. 20, pp. 1–4, 2018.
[42] A. W. A. Wahab, M. A. Bagiwa, M. Y. I. Idris, S. Khan, Z. Razak, and M. R. K. Ariffin, "Muhammad, passive video forgery detection techniques: A survey," in *Proc. 10th Int. Conf. Inf. Assurance Secur.*, Nov. 2014, pp. 29–34.
[43] G. Doërr and J.-L. Dugelay, *Video Watermarking: Overview and Challenges*. Boca Raton, FL, USA: CRC Press, 2003.
[44] (2020). *ACLU*. [Online]. Available: https://www.aclu.org/other/whats-wrong-public-video-surveillance
[45] S. Dargan and M. Kumar, "A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities," *Expert Syst. Appl.*, vol. 143, Apr. 2020, Art. no. 113114.
[46] S. Mondal and P. Bours, "Person identification by keystroke dynamics using pairwise user coupling," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 6, pp. 1319–1329, Jun. 2017.
[47] C. Ma, Z. Yan, and C. W. Chen, "Scalable access control for privacy-aware media sharing," *IEEE Trans. Multimedia*, vol. 21, no. 1, pp. 173–183, Jan. 2019.
[48] K. Yang, Z. Liu, X. Jia, and X. S. Shen, "Time-domain attribute-based access control for cloud-based video content sharing: A cryptographic approach," *IEEE Trans. Multimedia*, vol. 18, no. 5, pp. 940–950, May 2016.
[49] (2020). *Ezine Articles*. [Online]. Available: https://ezinearticles.com/?How-Effective-are-CCTV-Security-Systems-at-Reducing-Crime?&id=149735
[50] K. Lee, K. Yim, and M. A. Mikki, "A secure framework of the surveillance video network integrating heterogeneous video formats and protocols," *Comput. Math. Appl.*, vol. 63, no. 2, pp. 525–535, Jan. 2012.
[51] J. Koo, S.-R. Oh, and Y.-G. Kim, "Device identification interoperability in heterogeneous IoT platforms," *Sensors*, vol. 19, no. 6, p. 1433, Mar. 2019.

JIN-YONG YU received the B.E. degree in computer science from the Academic Credit Bank System, Chung-Ang University, Seoul, South Korea, in 2017. He is currently pursuing the Ph.D. degree with the Department of Computer and Information Security, and Convergence Engineering for Intelligent Drone, Sejong University. His current research interests include the Internet of Things security and video data security based on artificial intelligence.

YUJUN KIM received the B.E. degree in computer and information security from Sejong University, where he is currently pursuing the master's degree with the Department of Computer and Information Security, and Convergence Engineering for Intelligent Drone. His research interests include computer vision, deep learning, and information security.

YOUNG-GAB KIM (Member, IEEE) received the B.S. degree in biotechnology and genetic engineering and minored in computer science and engineering and the M.S. and Ph.D. degrees in computer science and engineering from Korea University, Seoul, South Korea, in 2001, 2003, and 2006, respectively. He was an Assistant Professor with the School of Information Technology, Catholic University of Daegu. He is currently an Associate Professor with the Department of Computer and Information Security, and Convergence Engineering for Intelligent Drone, Sejong University. He has published more than 180 research articles in the field of computer science and information security. His current research interests include the Internet of Things (IoT) security, big data security, network security, home network, security risk analysis, and security engineering. As a Korean ISO/IEC JTC 1 member, he is contributing in developing data exchange standards.