IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 6, 2019, accepted November 16, 2019, date of publication November 25, 2019,
date of current version December 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2955387

A Systematic Review of Intelligence Video Surveillance: Trends, Techniques, Frameworks, and Datasets

GURUH FAJAR SHIDIK[1], EDI NOERSASONGKO[1], ADHITYA NUGRAHA[1], PULUNG NURTANTIO ANDONO[1], JUMANTO JUMANTO[2], AND EDI JAYA KUSUMA[1]

[1]Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia
[2]Faculty of Humanities, Universitas Dian Nuswantoro, Semarang 50131, Indonesia

Corresponding author: Guruh Fajar Shidik (guruh.fajar@research.dinus.ac.id)

This work was supported in part by the Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jendral Penguatan Riset, Teknologi, dan Pendidikan Tinggi, Indonesia, under Grant 026/L6/AK/SP2H.1/PENELITIAN/2019 and Grant 060/A38.04/UDN-09/V/2019.

ABSTRACT Video surveillance systems obtain a great interest as application-oriented studies that have been growing rapidly in the past decade. The most recent studies attempt to integrate computer vision, image processing, and artificial intelligence capabilities into video surveillance applications. Although there are so many achievements in the acquisition of datasets, methods, and frameworks published, there are not many papers that can provide a comprehensive picture of the current state of video surveillance system research. This paper provides a comprehensive and systematic review on the literature from various video surveillance system studies published from 2010 through 2019. Within a selected study extraction process, 220 journal-based publications were identified and analyzed to illustrate the research trends, datasets, methods, and frameworks used in the field of video surveillance, to provide an in-depth explanation about research trends that many topics raised by researchers as a focus in their researches, to provide references on public datasets that are often used by researchers as a comparison and a means of developing a test method, and to give accounts on the improvement and integration of network infrastructure design to meet the demand for multimedia data. In the end of this paper, several opportunities and challenges related to researches in the video surveillance system are mentioned.

INDEX TERMS Artificial intelligence, cloud video surveillance, intelligent video surveillance, video surveillance framework.

I. INTRODUCTION

The practical needs of the surveillance system have been growing rapidly in the past decade. This growth rates are also supported by the development of high definition technology. Surveillance systems have earned a huge attention as application-oriented researches which are integrating the ability of computer vision, machine learning, and image processing. The major purpose of this approach is to automatically interpret the scene in a video, to observe, and to predict the interaction of an object in the scene based on information that has been gathered from camera sensors.

The initial utilization of the surveillance system was constructed from tube cameras which were distributed to monitor or to relay the factory activities in the 1940s. The conventional video surveillance system was often named Close-Circuit Television (CCTV), which was fragile and expensive, since the cameras were provided and installed by a security team to inspect the occurrence in the scene from a video display. The appearance of digitalized CCTV and advanced computation system has increased the expansion of a (semi-) self-loading system which is known as the 2nd generation of surveillance system. This new system has received advantages from the previous advancement in digital video communication, such as effective data compression, steady transmission, and bandwidth decrement [1]. The progress of this new generation is to enhance the efficiency of CCTV monitoring. The issues that often come up in the surveillance system are the detection and tracking algorithms which are required for behavioral analysis. To solve these issues,

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

the automatic video surveillance system integrates a real-time live computer vision and an intelligence approach. The implementation of this approach can, therefore, help the system to give real-time alert and forensic examination results for the security staffs because it is supported by an advanced video analysis technique. The implementation of advanced use of video surveillance system in recent decades has led to various domains, such as crime prevention [2], [3]; elder treatment [4]; accident detection [5], [6] traffic oversee and control [5], [7]–[11]; counting moving object such as pedestrians or vehicles [12]–[14]; human activity understanding [15]–[22]; motion detection [23]–[27], activity analysis [16], [28]–[30], identifying, tracking, and classifying vehicles, human, and any object of interest [3], [31]–[37]. The availability, development, and price of processors and sensors have also led to the utilization of this system in term of supporting indoor and outdoor neighborhoods such as shopping mall, airport, train station, and parking lots [16], [20], [38]–[40]. The study of video surveillance system is often associated with another multidisciplinary area, such as image or pattern analysis and recognition, signal processing, distributed system, and communication.

Currently, lots of studies have proposed the implementation of network infrastructure, and hardware and software components which support and provide efficient and effective video surveillance services to customers. There are several datasets, processing methods, and proposed frameworks which have been published. However, the currently proposed studies of intelligence system of video surveillance have provided an unclear explanation due to their complex implementation.

Due to various publications in numerous journals, conducting a direct comparison may face difficulties. The main purposes of this study are, therefore, as follows:

- To colligate and classify the studies related to the surveillance system.
- To provide and propose a concept of framework that can be integrated and classified based on the corresponding articles.
- To present recommendations for the development of surveillance system researches upon the completion of this review.

This study is conducted as follows: the methodology of the study is presented in Section 2. The outcomes and answers to the research questions are then discussed in Section 3. Finally, the study is summarized in the last section.

II. METHODOLOGY

A. REVIEW METHOD

Systematic Literature Review (SLR) is described as an activity to identify, specify, and analyze all researches or publications in a specific area to present the answer to each research question (RQ). The study proposed by Kitchenham [41] provides accounts on the original guideline to conduct the structure of this SLR. Other parameters such as reviewing techniques, styles, and several figures that appear in this part have also been encouraged by other researchers [42], [43].

TABLE 1. Summary of PICOC.

Population	Video surveillance, video surveillance application, video surveillance system, intelligent video surveillance, cloud video surveillance
Intervention	Detection methods or technique, behavior and activity analysis, dataset, models, classification, network infrastructure, framework
Comparison	n/a
Outcomes	Successful object detecting and tracking methods for video surveillance system, proposed improvement methods in video surveillance system, architecture for successful video surveillance system
Context	Studies in factory and academia, small and huge datasets

There are three stages that are performed in this SLR, i.e. planning, conducting, and reporting. Each performed stage consists of substages which have specific purposes. For the initial stage of planning strategy, the necessity of SLR is identified. The objectives for conducting the SLR are explained in the introduction section of this study. Then, the existing SLRs which are discussing video surveillance system are recollected and reviewed.

The next substages in the planning stage are the developments of review protocol. This review protocol is developed and assessed to direct the review process and to reduce the probability of research bias. The purpose of this process is to generate the next stage parameter such as the RQ, searching strategy, research selection, data extraction, quality evaluation, and data synthesis.

The review protocol is assessed and iteratively developed, while the conducting process and reporting stage of the review result are being carried out.

B. RESEARCH QUESTION

The important part of the systematic review is determining research questions (RQs). A research question (RQ) is defined at the beginning to direct the review process to stay focused in which in turn will also encourage the whole process of systematic review methodology, in which:

- the search stage must identify the main study that addresses the research question.
- the data-extraction stage must extract the data items needed to answer the RQs.
- the data-analysis stage must synthesize the data, thus the RQs can be answered.

RQs are obtained based on consideration of the Population, Intervention, Comparison, Result, and Context (PICOC) criteria [41]. Table 1 illustrates the PICOC structure associated with the RQs of this study.

Based on Table 1, the research questions for this literature review can be constructed, as displayed in Table 2.

From the main research, RQ1 up to RQ4 provide a resume or synopsis of a specific area of study in video surveillance system field. Video surveillance datasets, methods, and

TABLE 2. Research question and literature review.

ID	Research Question	Motivation
RQ1	Which publication has the most significant effort in a video surveillance system?	Identify the most significant publication in the video surveillance system field
RQ2	Who are the most active and influential researchers in the video surveillance system field?	Identify the most active and influential researchers that gave the most contribution to a research area of video surveillance system field
RQ3	What domains are implementing a surveillance system?	Identify trends for application video surveillance
RQ4	What kind of topic is popular among the researchers in the intelligent video surveillance field?	Identify research topics and trends in video surveillance system field
RQ5	What kind of datasets are the most used for intelligent video surveillance?	Identify datasets commonly used in intelligent video surveillance
RQ6	What kind of machine learning method is used most often for video surveillance system?	Identify the common machine learning methods for visual surveillance and object analysis on the video surveillance system
RQ7	What kind of network architecture improvement is used for video surveillance system?	Identify the most influential network architecture in video surveillance system
RQ8	What kind of framework is proposed for cloud video surveillance?	Identify the most significant framework used in intelligence video surveillance

frameworks to answer RQ5 to RQ8 are extracted. Then, the video surveillance datasets, methods and frameworks are observed to decide whether they give a significant effect for each question or not. The main research is RQ5 to RQ8 and the remaining questions which are RQ1 till RQ4 contribute to assessing the contextual information of the primary research.

The main goal of this SLR is to identify the trends of surveillance system research, the datasets used in conducting the research, the machine learning (ML) methods, and the frameworks used in the video surveillance system.

C. SEARCH STRATEGY

A searching process is considered as a set of activities such as selecting online libraries, determining the search query, performing a pilot search, purifying the search keyword, and regaining an initial table of main studies from digital libraries which match the keyword. A proper collection of databases has to be selected to improve the discovery possibility of relevant articles before the searching process begins. The well-known literature databases in this area especially in the surveillance system are searched to provide a comprehensive set of possibly available researches. A comprehensive perspective is important for a wide coverage of the literature. Below is the list of digital libraries which were used as sources for searches:

- Association for Computer Machinery (ACM) Digital Library which can be accessed in dl.acm.org.
- Institute of Electrical and Electronics Engineers (IEEE) Xplore digital library which can be accessed in ieeexplore.ieee.org.

TABLE 3. Inclusion and exclusion criteria.

Inclusion Criteria	(1)	For research which have both types of publication (the conference proceeding and journal versions), only the journal version that will be included.
	(2)	Studies discussing modeling, techniques, performance measurement in the area of intelligent video surveillance.
	(3)	Large and small scale datasets used in academic and industry studies.
	(4)	For multiple published journals in the same interest, only the most recent and exhaustive one will be included.
Exclusion Criteria	(1)	Studies without a robust validation or experimental results of intelligent video surveillance will be excluded.
	(2)	Studies discussing methods, datasets, frameworks in a context other than intelligent video surveillance
	(3)	Studies not written in English

- Science Direct which can be opened in sciencedirect.com.
- Springer Link which can be observed in link.springer.com.

The search query was constructed based on the following steps:

1. Identify the searching terms from PICOC table, especially information in Population and Intervention field.
2. Determine the searching terms from constructed Research Questions (RQs).
3. Identify the related titles, abstracts, and keywords of the studies.
4. Identify the synonyms, antonyms, and alternative orthography of search terms.
5. Construct an advanced search query of identified search terms using Boolean ANDs and ORs.

The following search query was eventually used:

(cloud OR intelligent OR application) AND (video) AND (surveillance) AND (system∗ OR services∗ OR detection∗ OR methods∗ OR models∗ OR techniques∗ OR dataset)

The arrangement of the query has been done, but the initial query is stored. This is because the arrangement of the query will significantly improve the existence of irrelevant studies. After that, the query was adjusted to match the certain criterion of each database. The publication databases were searched by title, abstract, and keyword. The search was done in the year range of 2010 to 2019. This year range was the limit of the search carried out in this SLR. Moreover, this SLR only involved journal papers, excluding conference proceedings. The search was limited to journal publications in English.

D. STUDY SELECTION

The research selection within inclusion and exclusion criteria were purposed to choose the main studies as described in Table 3.

Mendeley (http://mendeley.com) is an application used to manage and store the searching results. There are two steps to conduct the selection process of study, which are:

TABLE 4. Data extraction properties mapped to research questions.

Property	Research Questions
Researchers and Publications	RQ1, RQ2
Research Trends and Topics	RQ3, RQ4
Intelligent Video Surveillance Datasets	RQ5
Intelligent Video Surveillance Methods	RQ6, RQ7, RQ8
Intelligent Video Surveillance Frameworks	RQ8

1. Exclude the primary research based on its title and abstract.
2. Exclude the primary research by considering its full text.

The literature reviews and other researches which do not relate with the experiment results are excluded. The similar discussion of the study within the area of video surveillance system is also included in the studies.

The selected list of studies from the first stage has consisted of 220 full-text primary publications. The analysis process was considered as the selection process criteria (inclusion and exclusion), the assessment of the primary research quality, and the relevancy of the primary studies with the RQs and research similarity. Resembling researches conducted by the same authors in primary journals were removed from the list. There are 220 primary research after the exclusion of studies based on the full-text selection. The comprehensive list of selected researches is presented in Table 6 of this paper.

E. DATA EXTRACTION

In this review, the selected main studies will be extracted to collect the information. This collection is used to construct the answers regarding the research questions (RQs). Data extraction forms are designed to gather evidences from the main researches, which are needed to answer the RQs. The parameter is identified based on the RQs and the analysis process to run. Five parameters that are used to answer the RQs are described in Table 4 where data extraction will be done iteratively.

F. QUALITY ASSESSMENT OF STUDY AND DATA SYNTHESIS

The quality assessment is to be utilized as an interpretation guideline for the synthesis findings and for determining the potential conclusion outlines. The data synthesis purpose is to collect the information from selected researches to answer the RQs that were previously described. This review will extract the quantitative and qualitative data.

Various strategies are used to synthesize the data from selected studies which are extracted to answer various types of RQs. In general, a narrative synthesis method is used where the data is reconstructed in such a way that is relevant to the research questions. Furthermore, several visualization tools which contain charts and tables are used to improve the explanation on the distribution method for video surveillance system and its data accuracy.

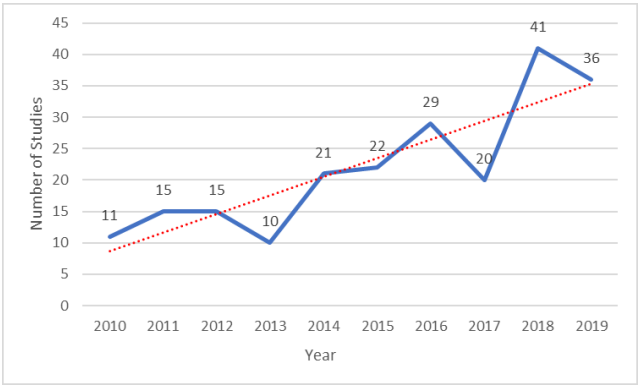

FIGURE 1. Distribution of selected studies over the years.

G. THREATS TO VALIDITY

This review goal is to analyze researches on video surveillance systems based on statistical trend calculation, methods, datasets, and several machine learning techniques. This research does not discuss the existence of bias in the selection of studies. The search might exclude some filtered video paper systems which result from several conferences or journals because the review is not based on a manual reading of all the contents of the published papers.

III. RESEARCH RESULT

A. SIGNIFICANCE OF JOURNAL PUBLICATION

In this systematic literature review, 220 primary studies were selected based on the procedure in Table 3. These primary studies were organized according to their publication years. Thus, the distribution of researches by years can show and illustrate the interests of researchers in the video surveillance area.

A short explanation of the study interests in the surveillance system over a few years is presented in Figure 1. Since 2010, the research interest in surveillance system has increased, which denotes that this research field is still relevant today. Although in 2013 and 2017 the number of publications decreased, the overall research trend in video surveillance system has increased quite significantly.

Figure 2 displays the top 20 publications related to video surveillance system. As aforementioned, the conference process is excluded from this graph.

The information about the publications such as Scimago Journal Rank (SJR) value, Q-level categories (Q1-Q4), and the number of publications from selected primary studies are shown in Table 5. Journal publications are ordered alphabetically according to their journal names.

B. MOST ACTIVE AND INFLUENTIAL RESEARCHERS

During this stage, several researchers were identified for their very active roles and much contribution a lot to the field of video surveillance system research. The distribution of the most active researchers based on the number of studies in

TABLE 5. Information related to the selected journals.

No	Journal Publications	SJR	Q-Category	Number of Studies
1	ACM Transactions on Embedded Computing Systems	0,33	Q2	1
2	ACM Transactions on Multimedia Computing, Communications, and Applications	0,57	Q1	4
3	Artificial Intelligence Review	1,06	Q1	1
4	Automation in Construction	1,41	Q1	1
5	Biosystems Engineering	0,83	Q1	1
6	Cluster Computing	0,34	Q2	4
7	Computer Vision and Image Understanding	0,77	Q1	6
8	Computers and Electrical Engineering	0,44	Q2	3
9	Computers and Electronics in Agriculture	0,95	Q1	1
10	Computers and Mathematics with Applications	1	Q1	2
11	Energy and Buildings	1,93	Q1	1
12	Environmental Monitoring and Assessment	0,62	Q2	1
13	Eurasip Journal on Advances in Signal Processing	0,32	Q2	1
14	Eurasip Journal on Embedded Systems	0,14	Q4	1
15	Eurasip Journal on Image and Video Processing	0,34	Q2	4
16	European Journal of Wildlife Research	0,58	Q2	1
17	Expert Systems with Applications	1,19	Q1	11
18	Fire Safety Journal	1,4	Q1	2
19	Forensic Science International	0,88	Q1	1
20	Future Generation Computer Systems	0,84	Q1	1
21	IEEE Access	1	Q1	8
22	IEEE Geoscience and Remote Sensing Letters	1,52	Q1	1
23	IEEE Journal on Emerging and Selected Topics in Circuits and Systems	0,72	Q1	1
24	IEEE Sensors Journal	0,73	Q1	1
25	IEEE Signal Processing Letters	0,79	Q1	1
26	IEEE Transactions on Big Data	0,93	Q1	1
27	IEEE Transactions on Circuits and Systems for Video Technology	0,98	Q1	11
28	IEEE Transactions on Cloud Computing	0,92	Q1	1
29	IEEE Transactions on Consumer Electronics	0,45	Q1	4
30	IEEE Transactions on Cybernetics	3,55	Q1	1
31	IEEE Transactions on Image Processing	1,81	Q1	5
32	IEEE Transactions on Industrial Electronics	2,4	Q1	1
33	IEEE Transactions on Industrial Informatics	1,68	Q1	3
34	IEEE Transactions on Information Forensics and Security	1,36	Q1	2
35	IEEE Transactions on Intelligent Transportation Systems	1,41	Q1	1
36	IEEE Transactions on Multimedia	1,22	Q1	5
37	IEEE Transactions on Pattern Analysis and Machine Intelligence	3,76	Q1	1
38	IEEE Transactions on Systems, Man, and Cybernetics: Systems	2,15	Q1	1
39	IEEE Transactions on Vehicular Technology	1,09	Q1	1
40	Image and Vision Computing	0,63	Q1	1
41	Information Fusion	2,24	Q1	2
42	Information Systems Frontiers	0,8	Q1	1
43	International Journal of Computer Vision	3,6	Q1	1
44	International Journal of Information Security	0,33	Q2	3
45	Journal of Big Data	1,12	Q1	1
46	Journal of Electronic Imaging	0,24	Q3	1
47	Journal of Grid Computing	0,68	Q1	1
48	Journal of Industrial Information Integration	1,57	Q1	1
49	Journal of Intelligent Information Systems	0,42	Q2	1
50	Journal of Mathematical Imaging and Vision	0,67	Q1	1
51	Journal of Medical Systems	0,57	Q2	1
52	Journal of Network and Computer Applications	0,9	Q1	3
53	Journal of Parallel and Distributed Computing	0,42	Q2	1
54	Journal of Real-Time Image Processing	0,37	Q2	3
55	Journal of Supercomputing	0,39	Q2	1
56	Journal of Visual Communication and Image Representation	0,51	Q1	10
57	Machine Vision and Applications	0,38	Q2	4
58	Medical Engineering and Physics	0,66	Q2	1
59	Mobile Networks and Applications	0,43	Q2	1
60	Multimedia Systems	0,31	Q2	6
61	Multimedia Tools and Applications	0,34	Q1	38
62	Neural Computing and Applications	0,64	Q2	1
63	Neurocomputing	0,61	Q1	8
64	Optik	0,4	Q2	4
65	Pattern Analysis and Applications	0,36	Q2	2
66	Pattern Recognition	1,36	Q1	3

TABLE 5. *(Continued.)* **Information related to the selected journals.**

67	Pattern Recognition Letters	0,66	Q1	4
68	Personal and Ubiquitous Computing	0,46	Q1	1
69	Progress in Nuclear Energy	1,12	Q1	1
70	Safety Science	1,29	Q1	2
71	Signal Processing	0,91	Q1	1
72	Signal Processing: Image Communication	0,56	Q2	2
73	Signal, Image and Video Processing	0,5	Q2	1
74	Soft Computing	0,62	Q2	1
75	SpringerPlus	0,43	Q1	1
76	Telecommunication Systems	0,29	Q2	1
77	Transportation Research Part C	2,61	Q1	1
78	Visual Computer	0,38	Q2	2
79	Wireless Networks	0,4	Q2	1
80	Wireless Personal Communications	0,25	Q3	2

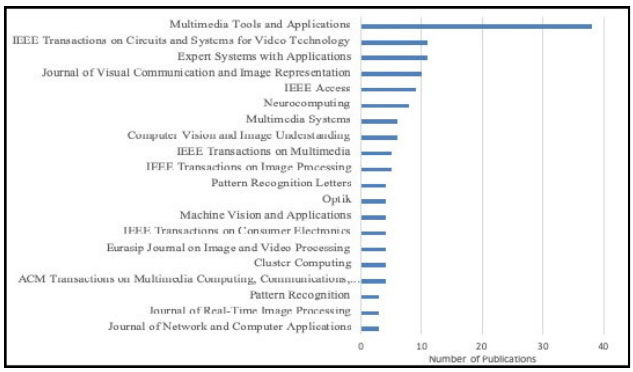

FIGURE 2. Top 20 of journal publications and distribution of selected studies.

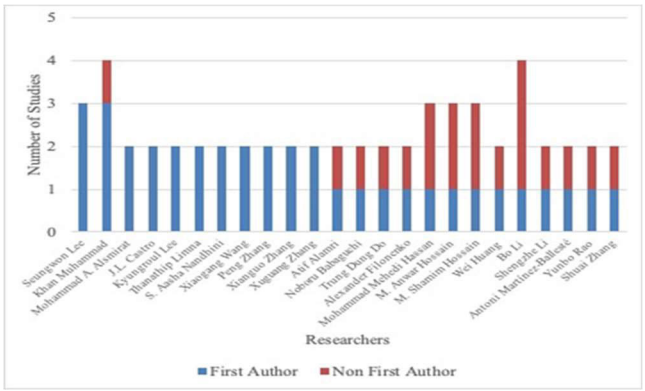

FIGURE 3. Influential researchers and number of studies.

the field of surveillance systems is shown in Figure 3. Some of these researchers are either as the first authors or not the first authors in their publications, i.e. Seungwon Lee, Khan Muhammad, Mohammad A. Alsmirat, J.L. Castro, Kyungroul Lee,, M. Anwar Hossain, Thanathip Limna, S. Aasha Nandhini, Xiaogang Wang, Peng Zhang, Xianguo Zhang, Xuguang Zhang, Atif Alamri, Noboru Babaguchi, Trung Dung Do, Alexander Filonenko, Mohammad Mehedi Hassan, Wei Huang, Bo Li, Shengzhe Li, M. Shamim Hossain, Antoni Martínez-Ballesté, Yunbo Rao and Shuai Zhang.

TABLE 6. References of application surveillance system.

No.	Classification Criteria	References
(a)	Protection and Privacy	[4], [16], [20], [39], [40], [44]–[53]
(b)	Object Detection and Tracking	[17], [22], [24]–[26], [31], [34], [54]–[100]
(c)	Activity recognition and behavior understanding	[30], [60], [101]–[137]
(d)	Traffic controlling and monitoring	[5], [7]–[11], [138]–[142]
(e)	Disaster relief and monitoring	[6], [26], [132], [143]–[148]

C. APPLICATION AREAS IDENTIFIED

Currently, the implementation of the surveillance system is used in various fields such as crime prevention, shopping, airport monitoring, elderly care, and monitoring and controlling important infrastructures and highways. Different types of video data are captured and streamed by cameras from remote areas to the server.

The characteristics of video surveillance systems are providing services that can operate without human control or within autonomous systems, detect, analyze and anticipate events or behaviors of objects, and then then give a warning of any unexpected activities.

a. Protection and/or privacy criteria explains issues on the implementation of surveillance system in public and private areas issues, elder treatment, ATM configuration, and specific security monitoring (museum, train station, airport, bank, etc).
b. Object detection and tracking category consists of all researches that focus on the process of detecting, tracking, identifying, and monitoring activities such as pedestrian tracking, identification of people, independent navigation, observing moving objects, and monitoring complex environments.

c. Category of activity recognition and behavior understanding has discussed the issues of motion segmentation, object classification, and identification in the area (face recognition, pedestrian identification, etc.), object and activity understanding, counting objects, and classification of moving vehicles.
d. Traffic controlling and monitoring category discusses several issues, such as vehicle flow rate estimation, prevention of traffic accidents, automated speed detection, and traffic flow density.
e. Disaster and accident monitoring describes the preventive applications, such as fire detection, flood detection, and the safety properties of construction workers.

D. RESEARCH TOPIC IN THE INTELLIGENCE VIDEO SURVEILLANCE

The analysis results of selected main studies show that the majority of recent video surveillance system studies focuses on several topics:

a. Visual Surveillance, which discusses image processing or computer vision techniques that are recently used for video surveillance;
b. Intelligent surveillance systems, integration of the different computer vision and image processing algorithms, which analyzes object, activity and behavior recognition, and is alert for any unplanned event for a completed surveillance system.
c. Distribution, communication, and system design, which discusses how infrastructure, network design, and operating protocols are compiled into a huge system to reflect the future practical needs of CCTV installations.

The first type of works (Visual Surveillance of moving object and behavior) is to detect, identify, track specific objects from the sequences of frame image, and to understand and explain the object behaviors. There are a number of publications and researches in the visual surveillance field dealing with implementation of algorithms to perform traditional techniques of image processing such as motion detection [24], [26], [57], [58], object detection [17], [25], [31], [57], [59]– [71], [99], object or event classification [16], [20], [34], [35], [54], [66], [144], [145], [149]–[158], object tracking [2], [7], [15], [32], [35], [56], [70], [98], [99], [134], [136], [139], [159]–[168], behavior and activity recognition [60], [101], [102], [111], [134], [137] and database or system operations [153], [169]–[172]. The aim of the researches is to build intelligent visual surveillance that can be used to take over the traditional passive video surveillance or monitoring methods using human operators that have already been proven ineffective for human to monitor a number of cameras at one time.

The second type of works (Integration of different algorithms for complete intelligent surveillance system) aims to present an interpretation on video scenes and to identify and predict the activities and interactions of the noticed objects based on the information produced by cameras and sensors automatically [3], [9], [30], [132], [173]–[175]. Along with the increment of surveillance camera system usage number, the task of monitoring multiple video streams by security teams becomes very tough due to the limits of human capability. Therefore, implementation of semi-automated intelligence video analysis paradigms is needed. This implementation will assist the security teams in video scene analysis and understanding. Furthermore, it also provides timely warnings to alert security personnels. With the development in sensory technology, surveillance cameras, and sound recording systems, it is possible to provide the intelligent video surveillance system integrated with the real-time monitoring of moving objects within a specific environment [54], [98], [111], [176]–[178].

The third type of works (Distribution, communication, and system design) classifies purposed communication and system design, network infrastructures of a video surveillance system. Numerous advancements in surveillance camera technology are affecting today's network performances. The higher resolution of surveillance video cameras is, the more systems are being added to network infrastructures. As a result, outdated network infrastructures are struggling to keep up with these demands [31], [179], [180]. The implementation of video surveillance system is complicated and impractical in many rapid situations. This problem appears because the installation of network infrastructure is both time-consuming and highly costly in operation. There is a number of researches dealing with network infrastructure upgrades, such as [27], [33], [143], [180]–[192]. The network infrastructure must be reliable, and the robust performance of surveillance video is not compromised.

E. DATASET USED FOR VIDEO SURVEILLANCE SYSTEM

A dataset is a collection of data in a certain amount that is used for certain machine learning purposes. It is a difficult job to collect a video surveillance dataset for some activities such as crowd, pedestrian, traffic, etc. Nowadays, several researchers and organizations have shared their video datasets for open use, hence facilitating other researchers in verifying the effectiveness of their methods and contributing even more to the field of video surveillance.

Some videos result contains monitoring activities for specific purposes such as crowd, traffic, face, pedestrian, etc. Most of the videos can be used for the purposes of tracking objects, counting moving objects and density estimation, objects of motion segmentation, objects of behavior analysis and objects of abnormal event detection. Table 7 shows some of these public datasets.

F. MACHINE LEARNING METHOD USED IN VIDEO SURVEILLANCE SYSTEM

The record of video data from surveillance cameras which have been installed in residential areas, industrial factory, government institutions, public areas (such as malls, downtown, etc.) and other places, generates a large number of video datasets. Advances in technology and high-performance

TABLE 7. References for publication dataset used.

Name of Public Dataset	References
AVSS	[59], [70], [164], [193], [194]
BEHAVE	[195]
Caltech	[74]
CANTATA	[196]
CAVIAR	[2], [39], [59], [164], [168], [195], [197]
CUHK Crowd Dataset	[67], [70], [155], [198]
HDA+	[155]
OTCBVS	[7], [165], [168], [199]
PETS	[7], [8], [66], [67], [74], [103], [151], [164], [178], [193], [194], [196], [12], [200]–[202], [13], [14], [20], [28], [45], [59], [61][203]
SPEVI	[197]
Town Centre	[28], [66], [74]
TRECVid	[202], [204]
UCF Dataset	[45], [162]
UCSD	[108]
UMN	[45], [108], [196], [205]
Viper	[155], [178], [201]
Virat Dataset	[162], [192], [206]
ViSOR	[60], [178], [207]
XJTU	[70]

computers make it easier for researchers to apply the concept of machine learning that can automatically extract features in varied data to distinguish objects from one another.

Since 2010, nineteen machine learning methods have been proposed and applied as solutions for automatic object detection, tracking, segmentation, classification, and behavior analysis for visual surveillance. A summary of the methods used in visual surveillance and object analysis is shown in Figure 4.

G. THE MOST INFLUENTIAL NETWORK RCHITECTURE

There are several approaches in the development of video surveillance system, starting with conventional analog systems, digital-analog video recording systems (DVRs), network analog DVR systems, network video encoder systems, network IP camera systems, and cloud-based video systems [208], [58], [135], [163], [209], [210].

An intelligent video system design aims to combine the an integration of hardware systems, management, and video processing for end users. Complete solutions need to provide for integration of network infrastructure design and installation to keep up with the demands of multimedia data produced by surveillance peripheral. Some of advanced system designs are based on several areas such as wireless solution [210]–[216] cloud infrastructure solutionn [183], [186], [187], [217]–[220], sensor improvement [4], [108], [124], [212], protocol enhancement [111], [159], [181],

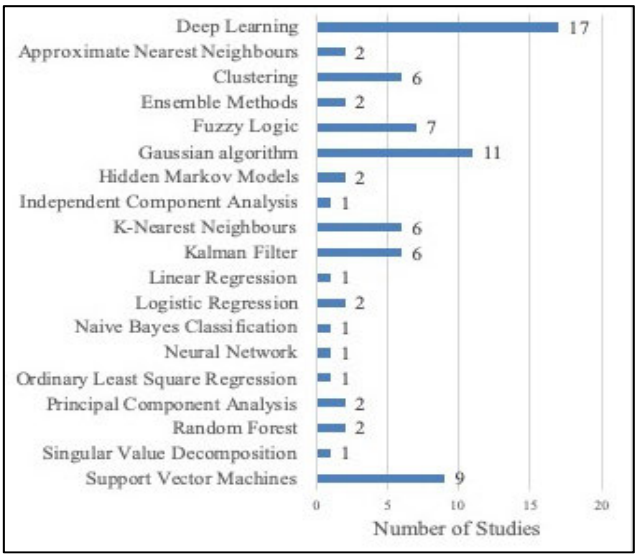

FIGURE 4. Machine learning method used in surveillance system.

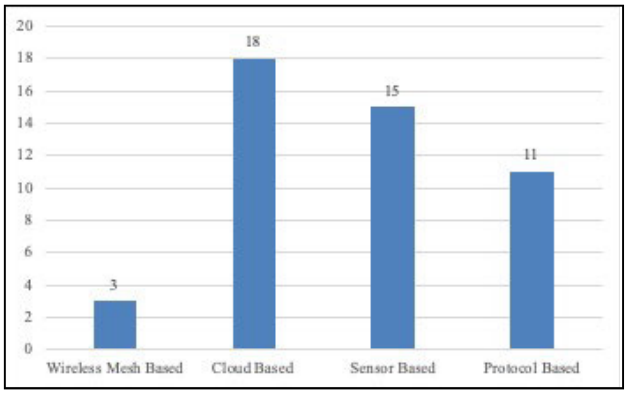

FIGURE 5. Research on network infrastructure design in video surveillance system.

[188], [221], [222]. Figure 5 shows numbers of research focus on network infrastructure improvement.

H. INTELLIGENT VIDEO SURVEILLANCE FRAMEWORK

Five proposed frameworks which are most cited and are also influencing to the development of intelligent video surveillance field are as follows:

1) Kyungroul Lee et al.'s Framework (2012)

Kyungroul Lee et al. [177] proposed a secure framework that implemented a prototype sample of server and the implemented client module based on the architecture that had been introduced in his research to be able to adopt heterogeneous video networks and protocols for surveillance management. This paper briefly explained that most of the problems in video surveillance were due to the lack of connection and interoperatability of network cameras which made the effort to integrate video surveillance systems into a global large scale more difficult and expensive.

Through an in-depth research, Kyungroul Lee et al.'s framework as seen in Figure 6 [177], introduced a solution

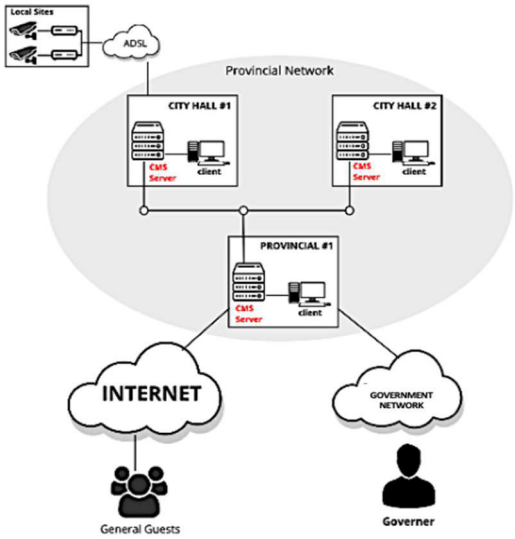

FIGURE 6. Kyungroul framework.

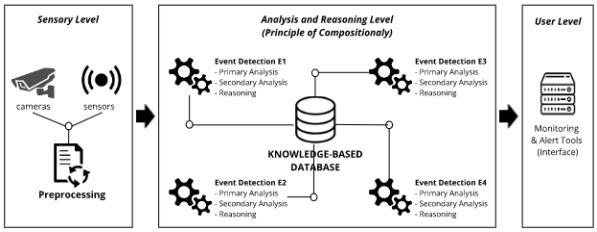

FIGURE 7. iSurveillance framework.

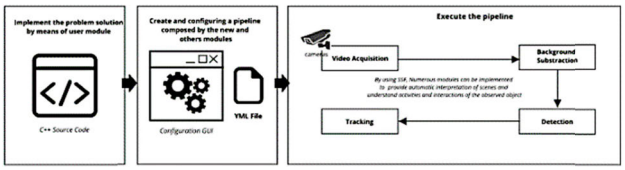

FIGURE 8. Smart surveillance framework.

that might be used in developing a detailed architecture of video surveillance systems to adopt heterogeneous connection protocols and video formats from video equipment. However, network cameras that were physically connected to public networks were vulnerable to cybercrime. Strict access control policies were, therefore, needed, especially for large-scale video surveillance networks.

The cryptography approach is integrated to be used as user authentication as well as for securing entire images and partial privacy masking. A private key mechanism based on the Kerberos protocol has been chosen to secure images from exposure and eavesdropping by the attacker. Through the concept of restoring the privacy mask, the key mechanism allows the system to mask information based on the level of privilege that is given and restored when viewed by privileged users. Based on the results of comparisons achieved through researches, the framework shows reasonable performance and reduced consumption of resources, and resolves existing problems.

This proposed surveillance framework has been deployed for local government to distribute visualization of the transportation system and global disaster surveillance system.

2) Mei Kuan Lim et al.'s Framework (2014)

Mei Kuan Lim et al. [196] proposed a new framework that was capable to identify several events, in various regions-of-interests (ROI) of a video scene, at certain times as seen in Figure 7. This framework was called an intelligent framework for multiple event detection in surveillance videos or was referred to as iSurveillance. The strength of this framework was that it could adapt and apply knowledge-based architecture to the video surveillance domain for a larger and integrated analysis of real-world surveillance scenarios.

Knowledge-based understanding of the environment is constructed by modulating supervisory problems into a set of variables consisting of ROI, class (for instance human, vehicle), attributes (such as speed or locality), and a set of ideas or rules associated with each attribute. The combination of several variables employed, enables a broader and integrated understanding of the complex patterns of activities in the scene.

The experimental results have shown that the proposed method provides comprehensive detection of various events both in public and real-time datasets. This shows how the system provides effectiveness and endurance performance in providing the flexibility of detecting various events in different ROI.

3) Antonio C. Nazare Jr and William Robson Schwartz's Framework (2016)

Antonio C. Nazare Jr and William Robson Schwartz's [223] proposed a new framework as seen in Figure 8, called the Smart Surveillance Framework (SSF) which was used to easily integrate various computer vision algorithms into a functional surveillance system, so as to facilitate the process of video analysis.

The purpose of this framework is to provide an environment that allows researchers to implement and evaluate their algorithms in relation to the monitoring system in an integrated manner. SSF is a tool built to provide a set of functions to help researchers develop functional surveillance systems, as well as create new algorithms to solve problems related to video surveillance.

This framework aims to bring about some improvements that provide scalability and flexibility, which allows users (researchers) to only focus on solving the problems that they are interested in without creating the infrastructure for each case to be handled. SSF provides a mechanism for data representation, storage, communication, and parallelism processes. It also provides features as tools for performing scene understanding, scalability, real-time operation, distributed multi-sensor environment, and communication control.

This work proposes a new framework that enables researchers to carry out further development of computer vision methods and surveillance applications. The framework

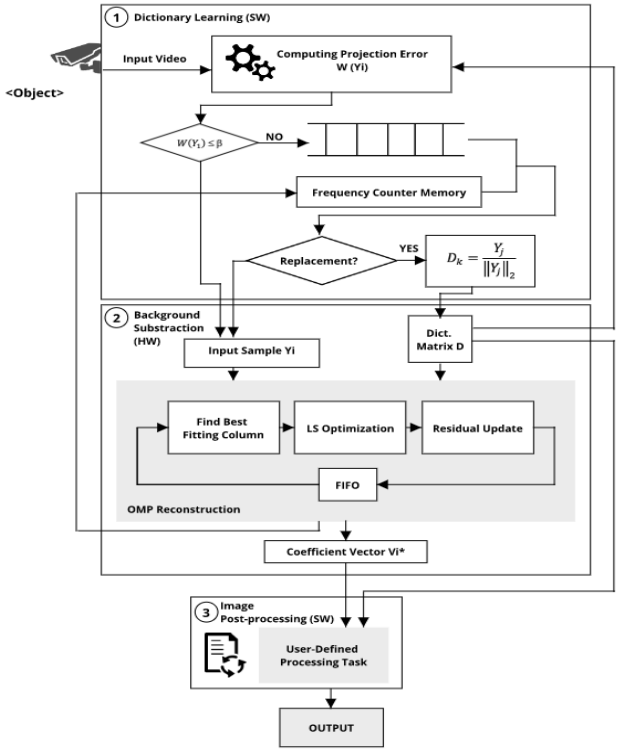

FIGURE 9. RISE framework.

component makes it possible to create various types of projects, from simple prototypes to complex applications, making it a supporting tool for the development of algorithms and visual surveillance solutions.

In real world implementation, the aim of this framework is to monitor the gates and to recognize faces of people entering the building. By making this framework publicly available, it is possible for researchers to be able to share datasets, methods and system designs that can later be used by other researchers. This also makes it possible for other researchers to be able to compare the performance of the methods or algorithms being studied with the methods available by other researchers. Therefore, SSF might also contribute as an accurate validation of the computer vision algorithms related to the surveillance system.

4) Bita Darvish Rouhani et al.'s Framework (2017)

Bita Darvish Rouhani et al. [224] proposed RISE, a new automatic computing framework approach for real-time background reduction as seen in Figure 9, which was implemented in intelligent video surveillance on Field-Programmable Gate Array (FPGA). RISE consisted of three blocks of object detection/tracking application systems, i.e. the dictionary learning units, the background reduction units, and the user-defined post-processing units.

The main step in the background reduction algorithms is to effectively study the dictionary matrix or reference models that represent background scenes in the video stream. It is important to construct dynamic streaming algorithms/tools to handle disturbances, such as changes in illumination,

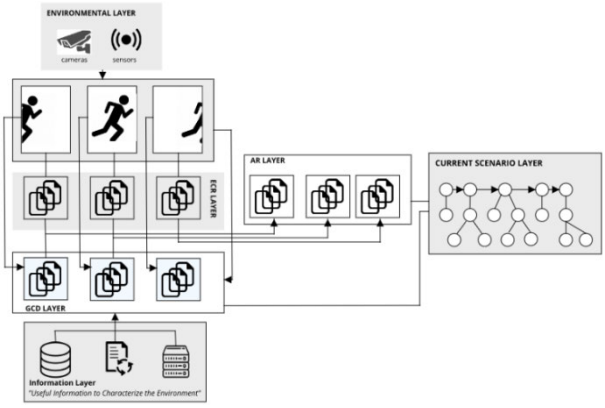

FIGURE 10. EDCAR framework.

the presence of shadows, repetitive movements in background scenes in real-world settings that may appear in input data.

RISE is designed with a new, adaptive, and streaming-based methodology that studies/updates the corresponding dictionary matrices of background pixels in real-time monitoring of persistent and temporary objects in an uncontrolled environment. RISE works by processing input data in the form of video frame streams taken by surveillance cameras and then highlighting the foreground content in each frame to further reduce the object detection/tracking workload so that it meets the provisions/constraints of related resources. To provide the best representation of input data, RISE calculates video stream sketches by implementing a series of dynamic dictionaries. Data sketches are then used in the RISE framework to perform user-specific learning applications such as object tracking/localization.

RISE supports the API, so designers can use it for rapid prototyping and deployment of different video surveillance applications using FPGA-based system-on-chip platforms.

5) Loredana Caruccio et al.'s Framework (2019)

Loredana Caruccio et al. [102] published a video surveillance framework called Elements and Descriptors of Context and Action Representations (EDCAR) as seen in Figure 10. EDCAR applied video-based occurrence recognition to interpret information in video frame sequences to define, recognize, and isolate several events, then provide a warning for any unplanned activities.

By using the EDCAR framework, it is possible to understand current events and to recognize abnormal situations through stored knowledge by analyzing relevant elements, modeling objects or actors, contexts, and scenarios that must be captured that allow to be detected by using ECR (Elements of Context Representations), GCD (General Context Descriptors), and AR (Action Representations); and to define the composition of actions and the sequence of events.

Defining the target scenarios through the EDCAR framework, the steps are to reduce the complexity of events to be simple initiating elements and context descriptors, to take action according to the actual data, to give reasons for events and provide information lost in events, and to make conclusions based on data and information used.

In this paper, Intelligent Video SurveillanT (IVIST) is also introduced. IVIST is a prototype system with EDCAR knowledge that provides automatic support to implement related inference procedures such as object tracking and detection to trigger alarms.

IVIST system has been investigated by using several target scenarios such as ''steal baggage'', ''crowd activity'', ''unattended baggage'', and ''fighting''. From investigation result, each scenario can precisely be detected using IVIST system integrated with EDCAR Framework.

The results of the study show that the EDCAR framework is able to represent and characterize several types of events or information in terms of ECR, AR, and GCD, which are still in good performance, although the framework still faces some detection errors.

IV. CONCLUSION AND FUTURE WORK

The purpose of this systematic literature review is to identify and analyze research interest development, datasets, models, and proposed frameworks used in video surveillance system, which provides the answers to the 8 research questions (RQs) that have been given earlier. There are 220 video surveillance system studies published between January 2010 and August 2019 that have gone through a screening process based on the pre-designed inclusion and exclusion criteria. Based on the results of distribution studies during 2010-2019, the field of research on video surveillance system is still very relevant today.

We have shown that the research topics most currently analyzed based on selected primary studies reveal that video surveillance system research focuses on three topics and trends, i.e. the visual surveillance method, intelligent and integrated video surveillance, distribution, and communication and system design for video surveillance. In addition, we have also discussed publicly on the available datasets that were built for testing the methodology proposed by the authors.

The states-of-the-art of the method that is widely applied in every problem-solving visual supervision is focused on three main tasks, i.e. detection, tracking, and recognition of activities or understanding behaviors. Nineteen machine learning methods have been proposed and applied as solutions for visual surveillance. Of the nineteen methods, the five methods most widely applied in visual surveillance systems are Deep Learning, Gaussian, Support Vector Machine (SVM), Fuzzy Logic, and Nearest Neighbor.

In addition to visual surveillance that applies computer vision and image processing algorithms, there are a number of studies that focus on improving and integrating network infrastructure designs to meet the demand for multimedia data generated by surveillance devices. Some advances in system design are based on approaches such as wireless solutions, cloud infrastructure solutions, sensor enhancements, and protocol enhancements.

This paper has also identified five proposed frameworks that have an important role in the video surveillance system field. In their research, Kyungroul Lee et al. has introduced a framework that manages to adopt heterogeneous video networks and protocols in supervisory management. Meanwhile, Mei Kuan Lim et al. have proposed iSurveillance that is able to detect several events, in various areas of interest (ROI) of a scene, at certain times by adapting and applying knowledge-based architecture to the video surveillance domain for a broader and integrated analysis of real-world surveillance scenarios. Antonio C. Nazare Jr and William Robson Schwartz have proposed SSF as a tool that provides an environment and a set of functions that enable researchers to implement and evaluate their algorithms in relation to an integrated monitoring system. Bita Darvish Rouhani et al. have also proposed RISE, which is designed by applying a background subtraction algorithm, a new adaptive methodology, based on streaming that studies/updates the appropriate dictionary matrix of background pixels for real-time monitoring of persistent and transient objects in an uncontrolled environment. Loredana Caruccio et al. have proposed EDCAR which implements video-based event recognition to interpret activities or behaviors in video sequences to detect, recognize, isolate certain events, and then warn of any unplanned activities.

There are many limitations and open challenges for researches in the field of video surveillance. The problem with motion detection in dynamic scenes is a difficult task to deal with illumination and weather changes, and shadow detection. A fast and accurate method is still needed to apply segmentation techniques to improve process performance. Tracking many people or groups of people is difficult due to crowded environments, poor lighting, noisy images, and the presence of camera movements.

REFERENCES

[1] B. Cheng, J. Yang, S. Wang, and J. Chen, ''Adaptive video transmission control system based on reinforcement learning approach over heterogeneous networks,'' *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 3, pp. 1104–1113, Jul. 2015.

[2] R. Xu, Y. Guan, and Y. Huang, ''Multiple human detection and tracking based on head detection for real-time video surveillance,'' *Multimedia Tools Appl.*, vol. 74, no. 3, pp. 729–742, Feb. 2015.

[3] H.-M. Moon, S.-H. Chae, D. Moon, Y. Chung, and S. B. Pan, ''Intelligent video surveillance system using two-factor human information,'' *Telecommun. Syst.*, vol. 52, no. 4, pp. 2249–2257, 2013.

[4] W.-Y. Shieh and J.-C. Huang, ''Falling-incident detection and throughput enhancement in a multi-camera video-surveillance system,'' *Med. Eng. Phys.*, vol. 34, no. 7, pp. 954–963, 2012.

[5] V. C. M. Vishnu, M. Rajalakshmi, and R. Nedunchezhian, ''Intelligent traffic video surveillance and accident detection system with dynamic traffic signal control,'' *Cluster Comput.*, vol. 21, no. 1, pp. 135–147, 2018.

[6] A. Filonenko, D. C. Hernández, and K.-H. Jo, ''Fast smoke detection for video surveillance using CUDA,'' *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 725–733, Feb. 2018.

[7] J. Zhu, Y. Lao, and Y. F. Zheng, ''Object tracking in structured environments for video surveillance applications,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 2, pp. 223–235, Feb. 2010.

[8] H. Cheng and J. Hwang, ''Integrated video object tracking with applications in trajectory-based event detection,'' *J. Vis. Commun. Image Represent.*, vol. 22, no. 7, pp. 673–685, 2011.

[9] A. Finogeev, A. Finogeev, L. Fionova, A. Lyapin, and K. A. Lychagin, ''Intelligent monitoring system for smart road environment,'' *J. Ind. Inf. Integr.*, vol. 15, pp. 15–20, Sep. 2019.

[10] C. Ma, D. Liu, X. Peng, L. Li, and F. Wu, ''Traffic surveillance video coding with libraries of vehicles and background,'' *J. Vis. Commun. Image Represent.*, vol. 60, pp. 426–440, Apr. 2019.

[11] R. Mohammadi and R. Javidan, "An adaptive type-2 fuzzy traffic engineering method for video surveillance systems over software defined networks," *Multimedia Tools Appl.*, vol. 76, no. 22, pp. 23627–23642, 2017.
[12] G. Castañón, M. Elgharib, V. Saligrama, and P.-M. Jodoin, "Retrieval in long-surveillance videos using user-described motion and object attributes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 12, pp. 2313–2337, Dec. 2016.
[13] C. R. del-Blanco, F. Jaureguizar, and N. García, "An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications," *IEEE Trans. Consum. Electron.*, vol. 58, no. 3, pp. 857–862, Aug. 2012.
[14] D. Conte, P. Foggia, G. Percannella, F. Tufano, and M. Vento, "A method for counting moving people in video surveillance videos," *EURASIP J. Adv. Signal Process.*, vol. 2010, Dec. 2010, Art. no. 231240.
[15] S. Vishwakarma and A. Agrawal, "A survey on activity recognition and behavior understanding in video surveillance," *Vis. Comput.*, vol. 29, no. 10, pp. 983–1009, Oct. 2013.
[16] D. Karthikeswaran, N. Sengottaiyan, and S. Anbukaruppusamy, "Video surveillance system against anti-terrorism by using adaptive linear activity classification (ALAC) technique," *J. Med. Syst.*, vol. 43, no. 8, p. 256, 2019.
[17] T. Sikandar, K. H. Ghazali, and M. F. Rabbi, "ATM crime detection using image processing integrated video surveillance: A systematic review," *Multimedia Syst.*, vol. 25, no. 3, pp. 229–251, 2018.
[18] W. Bouachir, R. Gouiaa, B. Li, and R. Noumeir, "Intelligent video surveillance for real-time detection of suicide attempts," *Pattern Recognit. Lett.*, vol. 110, pp. 1–7, Jul. 2018.
[19] X. Zhang, Q. Yu, and H. Yu, "Physics inspired methods for crowd video surveillance and analysis: A survey," *IEEE Access*, vol. 6, pp. 66816–66830, 2018.
[20] T. T. Zin, P. Tin, H. Hama, and T. Toriu, "Unattended object intelligent analyzer for consumer video surveillance," *IEEE Trans. Consum. Electron.*, vol. 57, no. 2, pp. 549–557, May 2011.
[21] P. C. Ribeiro, R. Audigier, and Q. C. Pham, "RIMOC, a feature to discriminate unstructured motions: Application to violence detection for video-surveillance," *Comput. Vis. Image Understand.*, vol. 144, pp. 121–143, Mar. 2016.
[22] P. Marti-Puig, S. Rodríguez, J. F. De Paz, R. Reig-Bolaño, M. P. Rubio, and J. Bajo, "Stereo video surveillance multi-agent system: New solutions for human motion analysis," *J. Math. Imag. Vis.*, vol. 42, nos. 2–3, pp. 176–195, 2012.
[23] M. Chen, S. Gonzalez, H. Cao, Y. Zhang, and S. T. Vuong, "Enabling low bit-rate and reliable video surveillance over practical wireless sensor network," *J. Supercomput.*, vol. 65, no. 1, pp. 287–300, 2013.
[24] T. Zhang and P. J. Jin, "A longitudinal scanline based vehicle trajectory reconstruction method for high-angle traffic video," *Transp. Res. C, Emerg. Technol.*, vol. 103, pp. 104–128, Jun. 2019.
[25] A. M. Pinto, P. G. Costa, M. V. Correia, and A. P. Moreira, "Enhancing dynamic videos for surveillance and robotic applications: The robust bilateral and temporal filter," *Signal Process. Image Commun.*, vol. 29, no. 1, pp. 80–95, 2014.
[26] F. Yuan, "An integrated fire detection and suppression system based on widely available video surveillance," *Mach. Vis. Appl.*, vol. 21, no. 6, pp. 941–948, 2010.
[27] M. Wang, B. Cheng, and C. Yuen, "Joint coding-transmission optimization for a video surveillance system with multiple cameras," *IEEE Trans. Multimedia*, vol. 20, no. 3, pp. 620–633, Mar. 2018.
[28] H. Hattori, N. Lee, V. N. Boddeti, F. Beainy, K. M. Kitani, and T. Kanade, "Synthesizing a scene-specific pedestrian detector and pose estimator for static video surveillance," *Int. J. Comput. Vis.*, vol. 126, no. 9, pp. 1027–1044, 2018.
[29] W. Huang, H. Ding, and G. Chen, "A novel deep multi-channel residual networks-based metric learning method for moving human localization in video surveillance," *Signal Process.*, vol. 142, pp. 104–113, Jan. 2018.
[30] X. Wang, "Intelligent multi-camera video surveillance: A review," *Pattern Recognit. Lett.*, vol. 34, no. 1, pp. 3–19, 2013.
[31] V. Tsakanikas and T. Dagiuklas, "Video surveillance systems-current status and future trends," *Comput. Elect. Eng.*, vol. 70, pp. 736–753, Aug. 2018.
[32] N. Liu, H. Wu, and L. Lin, "Hierarchical ensemble of background models for PTZ-based video surveillance," *IEEE Trans. Cybern.*, vol. 45, no. 1, pp. 89–102, Jan. 2015.
[33] M. Lotfi, S. A. Motamedi, and S. Sharifian, "Time-based feedback-control framework for real-time video surveillance systems with utilization control," *J. Real-Time Image Process.*, vol. 16, no. 4, pp. 1301–1316, 2019.
[34] A. S. Murugan, K. S. Devi, A. Sivaranjani, and P. Srinivasan, "A study on various methods used for video summarization and moving object detection for video surveillance applications," *Multimedia Tools Appl.*, vol. 77, no. 18, pp. 23273–23290, 2018.
[35] L. Zhao, Z. He, W. Cao, and D. Zhao, "Real-time moving object segmentation and classification from HEVC compressed surveillance video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1346–1357, Jun. 2018.
[36] S. Zhang, D. Cheng, Y. Gong, D. Shi, X. Qiu, Y. Xia, and Y. Zhang, "Pedestrian search in surveillance videos by learning discriminative deep features," *Neurocomputing*, vol. 283, pp. 120–128, Mar. 2018.
[37] Y. Nam, S. Rho, and J. H. Park, "Intelligent video surveillance system: 3-tier context-aware surveillance system with metadata," *Multimed. Tools Appl.*, vol. 57, no. 2, pp. 315–334, 2012.
[38] K. Sitara and B. M. Mehtre, "Automated camera sabotage detection for enhancing video surveillance systems," *Multimed. Tools Appl.*, vol. 78, no. 5, pp. 5819–5841, 2019.
[39] R. Arroyo, J. J. Yebes, L. M. Bergasa, I. G. Daza, and J. Almazán, "Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7991–8005, 2015.
[40] M. Abu-Lebdeh, F. Belqasmi, and R. Glitho, "An architecture for QoS-enabled mobile video surveillance applications in a 4G EPC and M2M environment," *IEEE Access*, vol. 4, pp. 4082–4093, 2016.
[41] S. Keele, "Guidelines for performing systematic literature reviews in software engineering, version 2.3," EBSE, Goyang-si, South Korea, EBSE Tech. Rep., 2007.
[42] B. Kitchenham, R. Pretorius, D. Budgen, O. P. Brereton, M. Turner, M. Niazi, and S. Linkman, "Systematic literature reviews in software engineering—A tertiary study," *Inf. Softw. Technol.*, vol. 52, no. 8, pp. 792–805, 2010.
[43] J. Wen, S. Li, Z. Lin, Y. Hu, and C. Huang, "Systematic literature review of machine learning based software development effort estimation models," *Inf. Softw. Technol.*, vol. 54, no. 1, pp. 41–59, 2012.
[44] A. Martínez-Ballesté, H. Rashwan, D. Puig, and A. Solanas, "Design and implementation of a secure and trustworthy platform for privacy-aware video surveillance," *Int. J. Inf. Secur.*, vol. 17, no. 3, pp. 279–290, 2018.
[45] H. Sohn, W. De Neve, and Y. M. Ro, "Privacy protection in video surveillance systems: Analysis of subband-adaptive scrambling in JPEG XR," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 2, pp. 170–177, Feb. 2011.
[46] N. Babaguchi, A. Cavallaro, R. Chellappa, F. Dufaux, and L. Wang, "Guest editorial: Special issue on intelligent video surveillance for public security and personal privacy," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 10, pp. 1559–1561, Oct. 2013.
[47] Y. Nakashima, N. Babaguchi, and J. Fan, "Intended human object detection for automatically protecting privacy in mobile video surveillance," *Multimedia Syst.*, vol. 18, no. 2, pp. 157–173, 2012.
[48] J. L. Castro, M. Delgado, J. Medina, and M. D. Ruiz-Lozano, "Intelligent surveillance system with integration of heterogeneous information for intrusion detection," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11182–11192, 2011.
[49] L. Du, W. Zhang, H. Fu, W. Ren, and X. Zhang, "An efficient privacy protection scheme for data security in video surveillance," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 347–362, Feb. 2019.
[50] H. Fradi and J.-L. Dugelay, "Towards crowd density-aware video surveillance applications," *Inf. Fusion*, vol. 24, pp. 3–15, Jul. 2015.
[51] H. A. Rashwan, A. Solanas, D. Puig, and A. Martínez-Ballesté, "Understanding trust in privacy-aware video surveillance systems," *Int. J. Inf. Secur.*, vol. 15, no. 3, pp. 225–234, 2016.
[52] D. Frejlichowski, K. Gościewska, P. Forczmański, and R. Hofman, "Application of foreground object patterns analysis for event detection in an innovative video surveillance system," *Pattern Anal. Appl.*, vol. 18, no. 3, pp. 473–484, 2015.
[53] C. A. F. Jorge, J. M. de Seixas, E. A. B. Silva, A. C. A. Mól, J. G. T. Monteiro, G. R. Lira, and P. H. S. Silva, "Assessment of radiation dose received by nuclear plant personnel through a video-based surveillance system," *Prog. Nuclear Energy*, vol. 103, pp. 51–59, Mar. 2018.

[54] S.-C. Huang, "An advanced motion detection algorithm with video quality analysis for video surveillance systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 1, pp. 1–14, Jan. 2011.

[55] X. Song, L. Sun, J. Lei, D. Tao, G. Yuan, and M. Song, "Event-based large scale surveillance video summarization," *Neurocomputing*, vol. 187, pp. 66–74, Apr. 2016.

[56] P. Zhang, T. Thomas, and S. Emmanuel, "Privacy enabled video surveillance using a two state Markov tracking algorithm," *Multimedia Syst.*, vol. 18, no. 2, pp. 175–199, 2012.

[57] E. Kermani and D. Asemani, "A robust adaptive algorithm of moving object detection for video surveillance," *EURASIP J. Image Video Process.*, vol. 2014, p. 27, Dec. 2014.

[58] A. Verdant, P. Villard, A. Dupret, and H. Mathias, "Three Novell analog-domain algorithms for motion detection in video surveillance," *EURASIP J. Image Video Process.*, vol. 2011, Dec. 2011, Art. no. 698914.

[59] A. V. Yoganand, A. C. Kavida, and D. Rukmanidevi, "An efficient PCA based pose and occlusion invariant face recognition system for video surveillance," *Cluster Comput.*, vol. 22, pp. 11443–11456, Sep. 2019.

[60] H. Zhang, C. Luo, Q. Wang, M. Kitchin, A. Parmley, J. Monge-Alvarez, and P. Casaseca-de-la-Higuera, "A novel infrared video surveillance system using deep learning based techniques," *Multimed. Tools Appl.*, vol. 77, no. 20, pp. 26657–26676, 2018.

[61] K. Goyal and J. Singhai, "Review of background subtraction methods using Gaussian mixture model for video surveillance systems," *Artif. Intell. Rev.*, vol. 50, no. 2, pp. 241–259, 2018.

[62] L. Kong and R. Dai, "Efficient video encoding for automatic video analysis in distributed wireless surveillance systems," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 14, no. 3, 2018, Art. no. 72.

[63] Y. Fang, G. Ding, Y. Yuan, W. Lin, and H. Liu, "Robustness analysis of pedestrian detectors for surveillance," *IEEE Access*, vol. 6, pp. 28890–28902, 2018.

[64] X. Li, M. Ye, Y. Liu, F. Zhang, D. Liu, and S. Tang, "Accurate object detection using memory-based models in surveillance scenes," *Pattern Recognit.*, vol. 67, pp. 73–84, Jul. 2017.

[65] S. Lee, N. Kim, K. Jeong, K. Park, and J. Paik, "Moving object detection using unstable camera for video surveillance systems," *Optik*, vol. 126, no. 20, pp. 2436–2441, 2015.

[66] D. Zeng and M. Zhu, "Multiscale fully convolutional network for foreground object detection in infrared videos," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 617–621, Apr. 2018.

[67] H. Wang, P. Wang, and X. Qian, "MPNET: An end-to-end deep neural network for object detection in surveillance video," *IEEE Access*, vol. 6, pp. 30296–30308, 2018.

[68] W. Liu, S. Liao, and W. Hu, "Perceiving motion from dynamic memory for vehicle detection in surveillance videos," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.

[69] A. J. Tom and S. N. George, "Video completion and simultaneous moving object detection for extreme surveillance environments," *IEEE Signal Process. Lett.*, vol. 26, no. 4, pp. 577–581, Apr. 2019.

[70] S. Zhang, C. Wang, S. C. Chan, X. Wei, and C. H. Ho, "New object detection, tracking, and recognition approaches for video surveillance over camera network," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2679–2691, May 2015.

[71] S. Kim, S. Kwak, and B. C. Ko, "Fast pedestrian detection in surveillance video based on soft target training of shallow random forest," *IEEE Access*, vol. 7, pp. 12415–12426, 2019.

[72] S.-C. Cheng, K.-F. Hsiao, C.-K. Yang, P.-F. Hsiao, and W.-H. Yu, "A novel unsupervised 3D skeleton detection in RGB-D images for video surveillance," *Multimedia Tools Appl.*, pp. 1–29, Jul. 2018.

[73] A. R. A. Mateus, C. Grilo, and M. Santos-Reis, "Surveying drainage culvert use by carnivores: Sampling design and cost–benefit analyzes of track-pads vs. video-surveillance methods," *Environ. Monit. Assessment*, vol. 181, nos. 1–4, pp. 101–109, 2011.

[74] S. Lee, N. Kim, and J. Paik, "Adaptively partitioned block-based contrast enhancement and its application to low light-level video surveillance," *Springerplus*, vol. 4, no. 1, 2015, Art. no. 431.

[75] P. Cocca, F. Marciano, and M. Alberti, "Video surveillance systems to enhance occupational safety: A case study," *Saf. Sci.*, vol. 84, pp. 140–148, Apr. 2016.

[76] J. Meng, A. Wu, and W.-S. Zheng, "Deep asymmetric video-based person re-identification," *Pattern Recognit.*, vol. 93, pp. 430–441, Sep. 2019.

[77] X. Yin, W. Li, Y. Liu, B. Wang, and M. Zhang, "FPGA-based real time video stitching method for video surveillance," *Optik*, vol. 126, no. 21, pp. 2804–2808, 2015.

[78] J. Sun, J. Shao, and C. He, "Abnormal event detection for video surveillance using deep one-class learning," *Multimedia Tools Appl.*, vol. 78, no. 3, pp. 3633–3647, 2019.

[79] B. Li, L. Liu, M. Shen, Y. Sun, and M. Lu, "Group-housed pig detection in video surveillance of overhead views using multi-feature template matching," *Biosyst. Eng.*, vol. 181, pp. 28–39, May 2019.

[80] S. C. Lee and R. Nevatia, "Hierarchical abnormal event detection by real time and semi-real time multi-tasking video surveillance system," *Mach. Vis. Appl.*, vol. 25, no. 1, pp. 133–143, 2014.

[81] D. Wu, S.-J. Zheng, X.-P. Zhang, C.-A. Yuan, F. Cheng, Y. Zhao, Y.-J. Lin, Z.-Q. Zhao, Y.-L. Jiang, and D.-S. Huang, "Deep learning-based methods for person re-identification: A comprehensive review," *Neurocomputing*, vol. 337, pp. 354–371, Apr. 2019.

[82] S. A. Nandhini, S. Radha, and R. Kishore, "Efficient compressed sensing based object detection system for video surveillance application in WMSN," *Multimedia Tools Appl.*, vol. 77, no. 2, pp. 1905–1925, 2018.

[83] H. Woo, Y. M. Jung, J.-G. Kim, and J. K. Seo, "Environmentally robust motion detection for video surveillance," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2838–2848, Nov. 2010.

[84] Y. S. Lee and S.-H. Kim, "A proposal of novel design on the WAVE MAC algorithm with low-latency for seamless video surveillance in V2X environment," *Multimedia Tools Appl.*, pp. 1–5, Jan. 2019.

[85] D.-Y. Huang, C.-H. Chen, T.-Y. Chen, W.-C. Hu, and B.-C. Chen, "Rapid detection of camera tampering and abnormal disturbance for video surveillance system," *J. Vis. Commun. Image Represent.*, vol. 25, no. 8, pp. 1865–1877, 2014.

[86] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T. M. Rose, and W. An, "Detecting non-hardhat-use by a deep learning method from far-field surveillance videos," *Autom. Construct.*, vol. 85, pp. 1–9, Jan. 2018.

[87] C.-H. Yeh, C.-Y. Lin, K. Muchtar, H.-E. Lai, and M.-T. Sun, "Three-pronged compensation and hysteresis thresholding for moving object detection in real-time video surveillance," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4945–4955, Jun. 2017.

[88] J. Zou, Q. Zhao, W. Yang, and F. Wang, "Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation," *Energy Buildings*, vol. 152, pp. 385–398, Oct. 2017.

[89] D. Wang, J. Tang, W. Zhu, H. Li, J. Xin, and D. He, "Dairy goat detection based on Faster R-CNN from surveillance video," *Comput. Electron. Agricult.*, vol. 154, pp. 443–449, Nov. 2018.

[90] H.-M. Hu, X. Zhang, W. Zhang, and B. Li, "Joint global-local information pedestrian detection algorithm for outdoor video surveillance," *J. Vis. Commun. Image Represent.*, vol. 26, pp. 168–181, Jan. 2015.

[91] A. Castillo, S. Tabik, F. Pérez, R. Olmos, and F. Herrera, "Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos with deep learning," *Neurocomputing*, vol. 330, pp. 151–161, Feb. 2019.

[92] P. Russo, E. Gualdi-Russo, A. Pellegrinelli, J. Balboni, and A. Furini, "A new approach to obtain metric data from video surveillance: Preliminary evaluation of a low-cost stereo-photogrammetric system," *Forensic Sci. Int.*, vol. 271, pp. 59–67, Feb. 2017.

[93] P. Yang, L. Chen, J. Tian, and X. Xu, "Dust particle detection in surveillance video using salient visual descriptors," *Comput. Elect. Eng.*, vol. 62, pp. 224–231, Aug. 2017.

[94] J. Son, S. Kim, and K. Sohn, "Fast illumination-robust foreground detection using hierarchical distribution map for real-time video surveillance system," *Expert Syst. Appl.*, vol. 66, pp. 32–41, Dec. 2016.

[95] V. Mutneja and S. Singh, "GPU accelerated face detection from low resolution surveillance videos using motion and skin color segmentation," *Optik*, vol. 157, pp. 1155–1165, Mar. 2018.

[96] J. Jeong, S. Kwon, M.-P. Hong, J. Kwak, and T. Shon, "Adversarial attack-based security vulnerability verification using deep learning library for multimedia video surveillance," *Multimed. Tools Appl.*, pp. 1–15, Jan. 2019.

[97] A. B. Hamida, M. Koubaa, H. Nicolas, and C. B. Amar, "Video surveillance system based on a scalable application-oriented architecture," *Multimedia Tools Appl.*, vol. 75, no. 24, pp. 17187–17213, 2016.

[98] N. Cooharojananone, S. Kasamwattanarote, R. Lipikorn, and S. Satoh, "Automated real-time video surveillance summarization framework," *J. Real-Time Image Process.*, vol. 10, no. 3, pp. 513–532, 2015.

[99] N.-W. Lo, H. T. Chang, and J.-Y. Chang, "Caged mice mating behavior detection in surveillance videos," *J. Vis. Commun. Image Represent.*, vol. 25, no. 5, pp. 755–762, 2014.

[100] D. Y. Kim, B.-N. Vo, B.-T. Vo, and M. Jeon, "A labeled random finite set online multi-object tracker for video data," *Pattern Recognit.*, vol. 90, pp. 377–389, Jun. 2019.
[101] A. K. S. Kushwaha and R. Srivastava, "Multiview human activity recognition system based on spatiotemporal template for video surveillance system," *J. Electron. Imag.*, vol. 24, no. 5, 2015, Art. no. 051004.
[102] L. Caruccio, G. Polese, G. Tortora, and D. Iannone, "EDCAR: A knowledge representation framework to enhance automatic video surveillance," *Expert Syst. Appl.*, vol. 131, pp. 190–207, Oct. 2019.
[103] H. Kim and H.-J. Lee, "A low-power surveillance video coding system with early background subtraction and adaptive frame memory compression," *IEEE Trans. Consum. Electron.*, vol. 63, no. 4, pp. 359–367, Nov. 2017.
[104] J. Xiong, X. Long, R. Shi, M. Wang, J. Yang, Y. Li, and G. Gui, "Background error propagation model based RDO for coding surveillance and conference videos," in *Proc. Int. Conf. Adv. Hybrid Inf. Process.*, vol. 279, 2019, pp. 504–513.
[105] K. Muhammad, T. Hussain, and S. W. Baik, "Efficient CNN based summarization of surveillance videos for resource-constrained devices," *Pattern Recognit. Lett.*, to be published.
[106] W. Lin, M.-T. Sun, R. Poovendran, and Z. Zhang, "Group event detection with a varying number of group members for video surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 8, pp. 1057–1067, Aug. 2010.
[107] X. Hu, S. Peng, L. Wang, Z. Yang, and Z. Li, "Surveillance video face recognition with single sample per person based on 3D modeling and blurring," *Neurocomputing*, vol. 235, pp. 46–58, Apr. 2017.
[108] H. Kim, S. Lee, Y. Kim, S. Lee, D. Lee, J. Ju, and H. Myung, "Weighted joint-based human behavior recognition algorithm using only depth information for low-cost intelligent video-surveillance system," *Expert Syst. Appl.*, vol. 45, pp. 131–141, Mar. 2016.
[109] S. Zhang, H. Zhou, B. Zhang, Z. Han, and Y. Guo, "'Signal, image and video processing' special issue: Semantic representations for social behavior analysis in video surveillance systems," *Signal, Image Video Process.*, vol. 8, no. 1, pp. 73–74, 2014.
[110] J. Fang, W. Qian, Z. Zhao, Y. Yao, and Z. Wen, "Adaptively feature learning for effective power defense," *J. Vis. Commun. Image Represent.*, vol. 60, pp. 33–37, Apr. 2019.
[111] H. Kavalionak, C. Gennaro, G. Amato, C. Vairo, C. Perciante, C. Meghini, and F. Falchi, "Distributed video surveillance using smart cameras," *J. Grid Comput.*, vol. 17, no. 1, pp. 59–77, 2019.
[112] S. Li, V. H. Nguyen, M. Ma, C.-B. Jin, T. D. Do, and H. Kim, "A simplified nonlinear regression method for human height estimation in video surveillance," *EURASIP J. Image Video Process.*, vol. 2015, no. 1, p. 35, 2015.
[113] N. Zhang, C. Wu, Y. Wu, and N. N. Xiong, "An improved target tracking algorithm and its application in intelligent video surveillance system," *Multimed. Tools Appl.*, pp. 1–19, Dec. 2018.
[114] F. F. E. Guraya and F. A. Cheikh, "Neural networks based visual attention model for surveillance videos," *Neurocomputing*, vol. 149, pp. 1348–1359, Feb. 2015.
[115] W. N. I. Al-Obaydy and S. A. Suandi, "Open-set single-sample face recognition in video surveillance using fuzzy ARTMAP," *Neural Comput. Appl.*, pp. 1–8, Aug. 2018.
[116] X. Ding, H. Lei, and Y. Rao, "Sparse codes fusion for context enhancement of night video surveillance," *Multimed. Tools Appl.*, vol. 75, no. 18, pp. 11221–11239, 2016.
[117] P. Napoletano, G. Boccignone, and F. Tisato, "Attentive monitoring of multiple video streams driven by a Bayesian foraging strategy," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3266–3281, Nov. 2015.
[118] S. A. Nandhini and S. Radha, "Efficient compressed sensing-based security approach for video surveillance application in wireless multimedia sensor networks," *Comput. Elect. Eng.*, vol. 60, pp. 175–192, May 2017.
[119] X. Cao, S. Gao, L. Chen, and Y. Wang, "Ship recognition method combined with image segmentation and deep learning feature extraction in video surveillance," *Multimedia Tools Appl.*, pp. 1–16, Jan. 2019.
[120] P. V. W. Radtke, E. Granger, R. Sabourin, and D. O. Gorodnichy, "Skew-sensitive Boolean combination for adaptive ensembles—An application to face recognition in video surveillance," *Inf. Fusion*, vol. 20, no. 1, pp. 31–48, 2014.
[121] N. Cardoso, P. Rodrigues, J. Vale, P. Garcia, P. Cardoso, J. Monteiro, J. Cabral, J. Mendes, M. Ekpanyapong, and A. Tavares, "A generative-oriented model-driven design environment for customizable video surveillance systems," *EURASIP J. Embedded Syst.*, vol. 2012, p. 7, Dec. 2012.
[122] Y. Rao, "Automatic vehicle recognition in multiple cameras for video surveillance," *Vis. Comput.*, vol. 31, no. 3, pp. 271–280, 2015.
[123] W. Chen, Y. Tian, Y. Wang, and T. Huang, "Fixed-point Gaussian mixture model for analysis-friendly surveillance video coding," *Comput. Vis. Image Understand.*, vol. 142, pp. 65–79, Jan. 2016.
[124] S. M. M. Rahman, M. A. Hossain, H. Mouftah, A. El Saddik, and E. Okamoto, "Chaos-cryptography based privacy preservation technique for video surveillance," *Multimedia Syst.*, vol. 18, no. 2, pp. 145–155, 2012.
[125] M. S. Sayed and J. G. R. Delva, "An efficient intensity correction algorithm for high definition video surveillance applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 11, pp. 1622–1630, Nov. 2011.
[126] A. Czyzewski and K. Lisowski, "Employing flowgraphs for forward route reconstruction in video surveillance system," *J. Intell. Inf. Syst.*, vol. 43, no. 3, pp. 521–535, 2014.
[127] C. Pham, A. Makhoul, and R. Saadi, "Risk-based adaptive scheduling in randomly deployed video sensor networks for critical surveillance applications," *J. Netw. Comput. Appl.*, vol. 34, no. 2, pp. 783–795, 2011.
[128] Y. Zhou and Q. Zhang, "An efficient FMO-based error-resilient scheme for video surveillance coding," *Multimed. Tools Appl.*, vol. 66, no. 3, pp. 339–359, 2013.
[129] P. Szwed, P. Skrzynski, and W. Chmiel, "Risk assessment for a video surveillance system based on fuzzy cognitive maps," *Multimed. Tools Appl.*, vol. 75, no. 17, pp. 10667–10690, 2016.
[130] X. Yu, J. Cheng, S. Wu, and W. Song, "A framework of timestamp replantation for panorama video surveillance," *Multimedia Tools Appl.*, vol. 75, no. 17, pp. 10357–10381, 2016.
[131] A. Unterweger, K. V. Ryckegem, D. Engel, and A. Uhl, "Building a post-compression region-of-interest encryption framework for existing video surveillance systems: Challenges, obstacles and practical concerns," *Multimedia Syst.*, vol. 22, no. 5, pp. 617–639, 2016.
[132] E. Cermeño, A. Pérez, and J. A. Sigüenza, "Intelligent video surveillance beyond robust background modeling," *Expert Syst. Appl.*, vol. 91, pp. 138–149, Jan. 2017.
[133] F. Tung, J. S. Zelek, and D. A. Clausi, "Goal-based trajectory analysis for unusual behaviour detection in intelligent surveillance," *Image Vis. Comput.*, vol. 29, no. 4, pp. 230–240, Mar. 2011.
[134] O. Guzhva, H. Ardö, A. Herlin, M. Nilsson, K. Åström, and C. Bergsten, "Feasibility study for the implementation of an automatic system for the detection of social interactions in the waiting area of automatic milking stations by using a video surveillance system," *Comput. Electron. Agricult.*, vol. 127, pp. 506–509, Sep. 2016.
[135] R. Gula, J. Theuerkauf, S. Rouys, and A. Legault, "An audio/video surveillance system for wildlife," *Eur. J. Wildlife Res.*, vol. 56, no. 5, pp. 803–807, 2010.
[136] J. Ren, M. Xu, J. Orwell, and G. A. Jones, "Multi-camera video surveillance for real-time analysis and reconstruction of soccer games," *Mach. Vis. Appl.*, vol. 21, no. 6, pp. 855–863, 2010.
[137] G. Sreenu and M. A. S. Durai, "Intelligent video surveillance: A review through deep learning techniques for crowd analysis," *J. Big Data*, vol. 6, no. 1, 2019, Art. no. 48.
[138] Z. Zhang, C. Trivedi, and X. Liu, "Automated detection of grade-crossing-trespassing near misses based on computer vision analysis of surveillance video data," *Saf. Sci.*, vol. 110, pp. 276–285, Dec. 2018.
[139] B. Tu, W. Kuang, Y. Shang, D. He, and L. Zhao, "A multi-view object tracking using triplet model," *J. Vis. Commun. Image Represent.*, vol. 60, pp. 64–68, Apr. 2019.
[140] E. Belyaev, A. Vinel, A. Surak, M. Gabbouj, M. Jonsson, and K. Egiazarian, "Robust vehicle-to-infrastructure video transmission for road surveillance applications," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 2991–3003, Jul. 2015.
[141] J. Xiao, R. Hu, L. Liao, Y. Chen, Z. Wang, and Z. Xiong, "Knowledge-based coding of objects for multisource surveillance video data," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1691–1706, Sep. 2016.
[142] J. L. Castro, M. Delgado, J. Medina, and M. D. Ruiz-Lozano, "An expert fuzzy system for predicting object collisions. Its application for avoiding pedestrian accidents," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 486–494, 2011.
[143] R.-I. Chang, T.-C. Wang, C.-H. Wang, J.-C. Liu, and J.-M. Ho, "Effective distributed service architecture for ubiquitous video surveillance," *Inf. Syst. Frontiers*, vol. 14, no. 3, pp. 499–515, 2012.

[144] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional neural networks based fire detection in surveillance videos," *IEEE Access*, vol. 6, pp. 18174–18183, 2018.
[145] K. Muhammad, J. Ahmad, and S. W. Baik, "Early fire detection using convolutional neural networks during surveillance for effective disaster management," *Neurocomputing*, vol. 288, pp. 30–42, May 2018.
[146] N. Alamgir, K. Nguyen, V. Chandran, and W. Boles, "Combining multi-channel color space with local binary co-occurrence feature descriptors for accurate smoke detection from surveillance videos," *Fire Saf. J.*, vol. 102, pp. 1–10, Dec. 2018.
[147] S. C. Kong, D. Jin, S. Li, and H. Kim, "Fast fire flame detection in surveillance video using logistic regression and temporal smoothing," *Fire Saf. J.*, vol. 79, pp. 37–43, Jan. 2016.
[148] M. Hashemzadeh and A. Zademehdi, "Fire detection for video surveillance applications using ICA K-medoids-based color model and efficient spatio-temporal visual features," *Expert Syst. Appl.*, vol. 130, pp. 60–78, Sep. 2019.
[149] S. Lee, N. Kim, K. Jeong, I. Paek, H. Hong, and J. Paik, "Multiple moving object segmentation using motion orientation histogram in adaptively partitioned blocks for high-resolution video surveillance systems," *Optik*, vol. 126, no. 19, pp. 2063–2069, 2015.
[150] Y. Kong, J. Huang, S. Huang, Z. Wei, and S. Wang, "Learning spatiotemporal representations for human fall detection in surveillance video," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 215–230, Feb. 2019.
[151] A. B. Mabrouk and E. Zagrouba, "Abnormal behavior recognition for intelligent video surveillance systems: A review," *Expert Syst. Appl.*, vol. 91, pp. 480–491, Jan. 2018.
[152] K. Quast and A. Kaup, "AUTO GMM-SAMT: An automatic object tracking system for video surveillance in traffic scenarios," *EURASIP J. Image Video Process.*, vol. 2011, Dec. 2011, Art. no. 814285.
[153] M. Frikha, E. Fendri, and M. Hammami, "People search based on attributes description provided by an eyewitness for video surveillance applications," *Multimedia Tools Appl.*, vol. 78, no. 2, pp. 2045–2072, 2019.
[154] P. Kumar, A. Pande, and A. Mittal, "Efficient compression and network adaptive video coding for distributed video surveillance," *Multimedia Tools Appl.*, vol. 56, no. 2, pp. 365–384, 2012.
[155] T. D. Do, V. H. Nguyen, and H. Kim, "Real-time and robust multiple-view gender classification using gait features in video surveillance," *Pattern Anal. Appl.*, pp. 1–15, Feb. 2019.
[156] Y. Lu, K. Boukharouba, J. Boonært, A. Fleury, and S. Lecœuche, "Application of an incremental SVM algorithm for on-line human recognition from video surveillance using texture and color features," *Neurocomputing*, vol. 126, pp. 132–140, Feb. 2014.
[157] X. Zhang, T. Huang, Y. Tian, and W. Gao, "Background-modeling-based adaptive prediction for surveillance video coding," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 769–784, Feb. 2014.
[158] X. Zhang, Y. Tian, T. Huang, S. Dong, and W. Gao, "Optimizing the hierarchical prediction and coding in HEVC for surveillance and conference videos with background modeling," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4511–4526, Oct. 2014.
[159] S. C. Chan, S. Zhang, J.-F. Wu, H.-J. Tan, J. Q. Ni, and Y. S. Hung, "On the hardware/software design and implementation of a high definition multiview video surveillance system," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, no. 2, pp. 248–262, Jun. 2013.
[160] A. Maalouf, M.-C. Larabi, and D. Nicholson, "Offline quality monitoring for legal evidence images in video-surveillance applications," *Multimedia Tools Appl.*, vol. 73, no. 1, pp. 189–218, 2014.
[161] Z. Jia, J. Zhao, H. Wang, Z. Xiong, and A. Finn, "A two-step face hallucination approach for video surveillance applications," *Multimedia Tools Appl.*, vol. 74, no. 6, pp. 1845–1862, 2015.
[162] A. Walha, A. Wali, and A. M. Alimi, "Video stabilization with moving object detecting and tracking for aerial video surveillance," *Multimedia Tools Appl.*, vol. 74, no. 17, pp. 6745–6767, 2015.
[163] J. S. Kim, D. H. Yeom, and Y. H. Joo, "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems," *IEEE Trans. Consum. Electron.*, vol. 57, no. 3, pp. 1165–1170, Aug. 2011.
[164] P. Zhang, L. Wang, W. Huang, L. Xie, and G. Chen, "Multiple pedestrian tracking based on couple-states Markov chain with semantic topic learning for video surveillance," *Soft Comput.*, vol. 19, no. 1, pp. 85–97, 2015.
[165] G. S. Walia and R. Kapoor, "Robust object tracking based upon adaptive multi-cue integration for video surveillance," *Multimedia Tools Appl.*, vol. 75, no. 23, pp. 15821–15847, 2016.
[166] I. Elafi, M. Jedra, and N. Zahid, "Unsupervised detection and tracking of moving objects for video surveillance applications," *Pattern Recognit. Lett.*, vol. 84, pp. 70–77, Dec. 2016.
[167] K. S. Ray and S. Chakraborty, "Object detection by spatio-temporal analysis and tracking of the detected objects in a video with variable background," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 662–674, Jan. 2019.
[168] J. Guo, P. Zheng, and J. Huang, "An efficient motion detection and tracking scheme for encrypted surveillance videos," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 13, no. 4, 2017, Art. no. 61.
[169] L. Zhang, D. V. Kalashnikov, S. Mehrotra, and R. Vaisenberg, "Context-based person identification framework for smart video surveillance," *Mach. Vis. Appl.*, vol. 25, no. 7, pp. 1711–1725, 2014.
[170] R. S. Feris, B. Siddiquie, J. Petterson, Y. Zhai, A. Datta, L. M. Brown, and S. Pankanti, "Large-scale vehicle detection, indexing, and search in urban surveillance videos," *IEEE Trans. Multimedia*, vol. 14, no. 1, pp. 28–42, Feb. 2012.
[171] Y. Luo, S.-C. S. Cheung, R. Lazzeretti, T. Pignata, and M. Barni, "Anonymous subject identification and privacy information management in video surveillance," *Int. J. Inf. Secur.*, vol. 17, no. 3, pp. 261–278, 2018.
[172] L. Lo Presti, S. Sclaroff, and M. La Cascia, "Path modeling and retrieval in distributed video surveillance databases," *IEEE Trans. Multimedia*, vol. 14, no. 2, pp. 346–360, Apr. 2012.
[173] C.-T. Fan, Y.-K. Wang, and C.-R. Huang, "Heterogeneous information fusion and visualization for a large-scale intelligent video surveillance system," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 4, pp. 593–604, Apr. 2017.
[174] Z. Shao, J. Cai, and Z. Wang, "Smart monitoring cameras driven intelligent processing to big surveillance video data," *IEEE Trans. Big Data*, vol. 4, no. 1, pp. 105–116, Mar. 2018.
[175] D. Samaiya and K. K. Gupta, "Intelligent video surveillance for real time energy savings in smart buildings using HEVC compressed domain features," *Multimedia Tools Appl.*, vol. 77, no. 21, pp. 29059–29076, 2018.
[176] M. Valera, S. A. Velastin, A. Ellis, and J. Ferryman, "Communication mechanisms and middleware for distributed video surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 12, pp. 1795–1809, Dec. 2011.
[177] 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, 2012.
[178] R. Vezzani and R. Cucchiara, "Video surveillance online repository (ViSOR): An integrated framework," *Multimedia Tools Appl.*, vol. 50, no. 2, pp. 359–380, Nov. 2010.
[179] L. Tian, H. Wang, Y. Zhou, and C. Peng, "Video big data in smart city: Background construction and optimization for surveillance video processing," *Future Gener. Comput. Syst.*, vol. 86, pp. 1371–1382, Sep. 2018.
[180] T. Limna and P. Tandayya, "A flexible and scalable component-based system architecture for video surveillance as a service, running on infrastructure as a service," *Multimedia Tools Appl.*, vol. 75, no. 4, pp. 1765–1791, 2016.
[181] G. D'Angelo and S. Rampone, "A NAT traversal mechanism for cloud video surveillance applications using WebSocket," *Multimedia Tools Appl.*, vol. 77, no. 19, pp. 25861–25888, 2018.
[182] X. Ma, L. T. Yang, Y. Xiang, W. K. Zeng, D. Zou, and H. Jin, "Fully reversible privacy region protection for cloud video surveillance," *IEEE Trans. Cloud Comput.*, vol. 5, no. 3, pp. 510–522, Jul./Sep. 2017.
[183] M. A. Hossain and B. Song, "Efficient resource management for cloud-enabled video surveillance over next generation network," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 806–821, 2016.
[184] B. Song, M. M. Hassan, Y. Tian, M. S. Hossain, and A. Alamri, "Remote display solution for video surveillance in multimedia cloud," *Multimedia Tools Appl.*, vol. 75, no. 21, pp. 13375–13396, 2016.
[185] M. A. Alsmirat, I. Obaidat, Y. Jararweh, and M. Al-Saleh, "A security framework for cloud-based video surveillance system," *Multimedia Tools Appl.*, vol. 76, no. 21, pp. 22787–22802, 2017.
[186] M. M. Hassan, M. A. Hossain, M. Abdullah-Al-Wadud, T. Al-Mudaihesh, S. Alyahya, and A. Alghamdi, "A scalable and elastic cloud-assisted publish\subscribe model for IPTV video surveillance system," *Cluster Comput.*, vol. 18, no. 4, pp. 1539–1548, 2015.

[187] Y. Xiong, S. Wan, J. She, M. Wu, Y. He, and K. Jiang, "An energy-optimization-based method of task scheduling for a cloud video surveillance center," *J. Netw. Comput. Appl.*, vol. 59, pp. 63–73, Jan. 2016.
[188] T. Sultana and K. A. Wahid, "Choice of application layer protocols for next generation video surveillance using Internet of video things," *IEEE Access*, vol. 7, pp. 41607–41624, 2019.
[189] D. Kirubha and K. Ramar, "MCCA scheduling for enhancing QoS based video streaming for video surveillance applications," *Cluster Comput.*, pp. 1–11, Mar. 2018.
[190] M. Nasir, K. Muhammad, J. Lloret, A. K. Sangaiah, and M. Sajjad, "Fog computing enabled cost-effective distributed summarization of surveillance videos for smart cities," *J. Parallel Distrib. Comput.*, vol. 126, pp. 161–170, Apr. 2019.
[191] B.-H. Chen, L.-F. Shi, and X. Ke, "A robust moving object detection in multi-scenario big data for video surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 4, pp. 982–995, Apr. 2019.
[192] T. Limna and P. Tandayya, "Workload scheduling for Nokkhum video surveillance as a service," *Multimedia Tools Appl.*, vol. 77, no. 1, pp. 1363–1389, 2018.
[193] D. Ilias, M. C. El Mezouar, N. Taleb, and M. Elbahri, "An edge-based method for effective abandoned luggage detection in complex surveillance videos," *Comput. Vis. Image Understand.*, vol. 158, pp. 141–151, May 2017.
[194] A. Filonenko, A. Filonenko, and K.-H. Jo, "Unattended object identification for intelligent surveillance systems using sequence of dual background difference," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2247–2255, Dec. 2016.
[195] K. Kardas and N. K. Cicekli, "SVAS: Surveillance video analysis system," *Expert Syst. Appl.*, vol. 89, pp. 343–361, Dec. 2017.
[196] M. K. Lim, S. Tang, and C. S. Chan, "ISurveillance: Intelligent framework for multiple events detection in surveillance videos," *Expert Syst. Appl.*, vol. 41, no. 10, pp. 4704–4715, 2014.
[197] Z. Wei, Z. Yan, Y. Wu, and R. H. Deng, "Trustworthy authentication on scalable surveillance video with background model support," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 12, no. 4s, 2016, Art. no. 64.
[198] X. Wang, M. Wang, and W. Li, "Scene-specific pedestrian detection for static video surveillance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 2, pp. 361–374, Feb. 2014.
[199] A. Torabi, G. Massé, and G.-A. Bilodeau, "An iterative integrated framework for thermal–visible image registration, sensor fusion, and people tracking for video surveillance applications," *Comput. Vis. Image Understand.*, vol. 116, no. 2, pp. 210–221, 2012.
[200] D. Li, L. Xu, and E. D. Goodman, "Illumination-robust foreground detection in a video surveillance system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 10, pp. 1637–1650, Oct. 2013.
[201] C. Poppe, G. Martens, P. De Potter, and R. Van de Walle, "Semantic Web technologies for video surveillance metadata," *Multimedia Tools Appl.*, vol. 56, no. 3, pp. 439–467, 2012.
[202] P. L. Venetianer and H. Deng, "Performance evaluation of an intelligent video surveillance system—A case study," *Comput. Vis. Image Understand.*, vol. 114, no. 11, pp. 1292–1302, 2010.
[203] K. Lin, S.-C. Chen, C.-S. Chen, D.-T. Lin, and Y.-P. Hung, "Abandoned object detection via temporal consistency modeling and back-tracing verification for visual surveillance," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 7, pp. 1359–1370, Jul. 2015.
[204] G. Wang, B. Li, Y. Zhang, and J. Yang, "Background modeling and referencing for moving cameras-captured surveillance video coding in HEVC," *IEEE Trans. Multimedia*, vol. 20, no. 11, pp. 2921–2934, Nov. 2018.
[205] J. Wang, W. Fu, H. Lu, and S. Ma, "Bilayer sparse topic model for scene analysis in imbalanced surveillance videos," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5198–5208, Dec. 2014.
[206] P. K. Atrey, B. Trehan, and M. K. Saini, "Watch me from distance (WMD): A privacy-preserving long-distance video surveillance system," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 15, no. 2, 2019, Art. no. 37.
[207] X. Jin and S. Goto, "Encoder adaptable difference detection for low power video compression in surveillance system," *J. Image Commun.*, vol. 26, no. 3, pp. 130–142, 2011.
[208] H. Liu, S. Chen, and N. Kubota, "Intelligent video systems and analytics: A survey," *IEEE Trans. Ind. Informat.*, vol. 9, no. 3, pp. 1222–1233, Aug. 2013.
[209] S. J. Y. Go, C. A. M. Festin, and W. M. Tan, "An SDN-based framework for improving the performance of underprovisioned IP video surveillance networks," *J. Netw. Comput. Appl.*, vol. 132, pp. 49–74, Apr. 2019.
[210] H. Yoon, Y. Jung, and S. Lee, "An image sequence transmission method in wireless video surveillance systems," *Wireless Pers. Commun.*, vol. 82, no. 3, pp. 1225–1238, 2015.
[211] M. A. Alsmirat, Y. Jararweh, I. Obaidat, and B. B. Gupta, "Automated wireless video surveillance: An evaluation framework," *J. Real-Time Image Process.*, vol. 13, no. 3, pp. 527–546, 2017.
[212] I. M. de Diego, I. S. Román, J. C. Montero, C. Conde, and E. Cabello, "Scalable and flexible wireless distributed architecture for intelligent video surveillance systems," *Multimedia Tools Appl.*, vol. 78, no. 13, pp. 17437–17459, 2019.
[213] N. Kumar, J.-H. Lee, and J. J. P. C. Rodrigues, "Intelligent mobile video surveillance system as a Bayesian coalition game in vehicular sensor networks: Learning automata approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1148–1161, Jun. 2015.
[214] P. Hu, W. L. Tan, R. Wishart, M. Portmann, and J. Indulska, "MeshVision: An adaptive wireless mesh network video surveillance system," *Multimedia Syst.*, vol. 16, nos. 4–5, pp. 243–254, 2010.
[215] S. C. Lubobya, M. E.-H. Dlodlo, G. de Jager, and A. Zulu, "Mesh IP video surveillance systems model design and performance evaluation," *Wireless Pers. Commun.*, vol. 100, no. 2, pp. 227–240, 2018.
[216] K.-W. Lim, Y. Seo, W.-S. Jung, Y.-B. Ko, and S. Park, "Design and implementation of adaptive WLAN mesh networks for video surveillance," *Wireless Netw.*, vol. 19, no. 7, pp. 1511–1524, 2013.
[217] S. Zhizhuo, T. Yu-An, and L. Yuanzhang, "An energy-efficient storage for video surveillance," *Multimedia Tools Appl.*, vol. 73, no. 1, pp. 151–167, 2014.
[218] Y. Xiao, Z. Changyou, X. Yuan, Z. Hongfei, L. Yuanzhang, and T. Yu-An, "An extra-parity energy saving data layout for video surveillance," *Multimedia Tools Appl.*, vol. 77, no. 4, pp. 4563–4583, 2018.
[219] B. N. Subudhi, D. K. Rout, and A. Ghosh, "Big data analytics for video surveillance," *Multimedia Tools Appl.*, vol. 78, no. 18, pp. 26129–26162, 2019.
[220] A. Alamri, M. S. Hossain, A. Almogren, M. M. Hassan, K. Alnafjan, M. Zakariah, L. Seyam, and A. Alghamdi, "QoS-adaptive service configuration framework for cloud-assisted video surveillance systems," *Multimed. Tools Appl.*, vol. 75, no. 21, pp. 13333–13348, 2016.
[221] M. S. Hossain, "QoS-aware service composition for distributed video surveillance," *Multimedia Tools Appl.*, vol. 73, no. 1, pp. 169–188, 2014.
[222] W. Zhang, P. Duan, X. Xie, F. Xia, Q. Lu, X. Liu, and J. Zhou, "QoS4IVSaaS: A QoS management framework for intelligent video surveillance as a service," *Pers. Ubiquitous Comput.*, vol. 20, no. 5, pp. 795–808, 2016.
[223] A. C. Nazare, Jr., and W. R. Schwartz, "A scalable and flexible framework for smart video surveillance," *Comput. Vis. Image Understand.*, vol. 144, pp. 258–275, Mar. 2016.
[224] B. D. Rouhani, A. Mirhoseini, and F. Koushanfar, "RISE: An automated framework for real-time intelligent video surveillance on FPGA," *ACM Trans. Embedded Comput. Syst.*, vol. 16, no. 5s, 2017, Art. no. 158.

GURUH FAJAR SHIDIK was born in West Borneo, Indonesia, in February 1987. He has received the Bachelor of Computer Science degree from Universitas Dian Nuswantoro (UDINUS), Semarang, Indonesia, in 2009, the Master of Computer Science degree from Technical Malaysia Melaka University (UTeM), in 2011, the Ph.D. degree from Universitas Gadjah Mada (UGM), Indonesia, in 2016. He is currently a Lecturer with UDINUS. His areas of interests include cloud computing, wireless communication, machine learning, and artificial intelligence.

EDI NOERSASONGKO was born in Semarang, Central Java, Indonesia, in June 1955. He is currently a Senior Professor of artificial intelligence and technopreneurship. He also has a position as the President or Rector of Universitas Dian Nuswantoro, Semarang, Indonesia. His area of interests includes artificial intelligence and technopreneurship.

ADHITYA NUGRAHA was born in Palangkaraya, Indonesia, in March 1987. He received the Bachelor of Computer Science from Universitas Dian Nuswantoro (UDINUS), Semarang, Indonesia, in 2010, and the Master of Computer Science from Technical Malaysia Melaka University (UTeM), in 2012. He is currently a Lecturer with UDINUS. His research interests are computer networks and artificial intelligence.

PULUNG NURTANTIO ANDONO was born in Central Java, Indonesia, in September 1982. He received the Bachelor of Engineering from Universitas Trisakti, Jakarta, in 2006, and the Master of Computer Science from Universitas Dian Nuswantoro (UDINUS), in 2009, and the Ph.D. degree from the Institut Teknologi Sepuluh November (ITS), Surabaya, in 2016. He is currently a Lecturer with UDINUS. His area of interest includes 3D image reconstruction and computer vision.

JUMANTO JUMANTO was born in Pati, Central Java, Indonesia, in August 1968. He received the Ph.D. degree in linguistics (Pragmatics) from Universitas Indonesia, Jakarta, Indonesia, in 2006. His research interests include Linguistics, pragmatics, semiotics, discourse analysis, and verbal as well as non-verbal communication, including cross-cultural understanding and misunderstanding.

EDI JAYA KUSUMA was born in Central Java, Indonesia, in May 1997. He received the Bachelor of Computer Science from Universitas Dian Nuswantoro (UDINUS), Semarang, Indonesia, in 2018, where he is currently pursuing the master's degree. His research interest is in machine learning, artificial intelligence, image processing, and security.