Multimedia Tools and Applications
https://doi.org/10.1007/s11042-019-07793-w

Big data analytics for video surveillance

Badri Narayan Subudhi[1] · Deepak Kumar Rout[2] · Ashish Ghosh[3]

Received: 1 September 2018 / Revised: 19 March 2019 / Accepted: 15 May 2019
Published online: 05 June 2019

Abstract
This article addresses the usage and scope of Big Data Analytics in video surveillance and its potential application areas. The current age of technology provides the users, ample opportunity to generate data at every instant of time. Thus in general, a tremendous amount of data is generated every instant throughout the world. Among them, amount of video data generated is having a major share. Education, healthcare, tours and travels, food and culture, geographical exploration, agriculture, safety and security, entertainment etc., are the key areas where a tremendous amount of video data is generated every day. A major share among it are taken by the daily used surveillance data captured from the security purpose camera and are recorded everyday. Storage, retrieval, processing, and analysis of such gigantic data require some specific platform. Big Data Analytics is such a platform, which eases this analysis task. The aim of this article is to investigate the current trends in video surveillance and its applications using Big Data Analytics. It also aims to focus on the research opportunities for visual surveillance in Big Data frameworks. We have reported here the state-of-the-art surveillance schemes for four different imaging modalities: conventional video scene, remotely sensed video, medical diagnostics, and underwater surveillance. Several works were reported in this research field over recent years and are categorized based on the challenges solved by the researchers. A list of tools used for video surveillance using Big Data framework is presented. Finally, research gaps in this domain are discussed.

Keywords Video surveillance · Big Data · Data Science · Big Data Analytics for video

✉ Ashish Ghosh
ash@isical.ac.in

[1] Department of Electrical Engineering, Indian Institute of Technology Jammu, Nagrota, Jammu and Kashmir, India

[2] Department of Electronics and Communication Engineering, National Institute of Technology Goa, Farmagudi, Ponda, Goa, India

[3] Center for Soft Computing Research, Indian Statistical Institute, 203 B. T. Road, Kolkata, West Bengal 700108, India

1 Introduction

A smart surveillance system design is one of the most active research areas in computer vision. In a city, on a daily basis thousands of cameras are collecting a huge amount of video data for different surveillance purposes. Recently, the demand for intelligent visual surveillance has been raised due to the growing importance of public security at different places. It needs the development of computer vision based detection and tracking algorithms for detecting frauds, suspected movements, crime suspected vehicles, etc. An important task in computer vision based surveillance systems is detection followed by tracking of the moving objects. The process of moving object detection can be defined as identifying the object in a video which is changing its position with respect to the field of view of a scene. The process of tracking may be defined as identifying the path/ trajectory of an object in a given video sequence [1].

In regular life, we generally deal with different surveillance based devices that are constantly generating video data. In 2014, estimates put worldwide surveillance data generation at a staggering 3ZB. It is expected that by 2018, due to the development of high-quality visual lenses each surveillance camera is supposed to generate 100 GB of video data every month. At the same time, it is expected that the amount of such video data is supposed to grow at a 40% higher rate. This shows that data grows exponentially [2, 3]. Analysis of surveillance tasks needs large computational environment under different complex scenarios such as occlusion, camouflage, illumination variation, etc. Hence video data tends to “Big Data” as its volume, velocity, and variety is very high to ingest, store, analyze, and process [4]. Hence performing surveillance task over such data is expected to be on Big Data framework [3].

Big Data is a conglomeration of the booming volume of heterogeneous data sets, which is so huge and intricate that processing it becomes difficult, using the existing database management tools. For a surveillance camera the challenges include capturing, storing, and detection and tracking of a moving object from the field of coverage by the camera. Analysis of a single large dataset provides additional information related to data, as compared to the analysis of separate smaller sets with the same total amount of data (like collection over several cameras covering the view over a wide area of coverage). Big Data analysis for surveillance is the application of different intelligent mechanisms for storing, processing, and analyzing the huge amount of surveillance video data that was previously not utilized because of unavailability of modern and sophisticated data handling tools. [5].

The major challenge for this task is that 2.5 quintillion bytes of data are generated every day. Around 90% of the data in the world has been created in the last two years alone. Thus we can predict the size of data going to be generated in the coming years. In this study, we would like to concentrate on the analysis of a massive collection of digital videos which come from different heterogeneous sources in the domain of video analysis, and medical video analysis. As already mentioned, managing and analysis of massive data is a challenging field in recent years. In this article, “Big Data analysis” issues in the domain of digital video processing are addressed. With the improvement of video acquisition tools, variety, veracity, and volume of different videos are increasing day-by-day. To handle this huge amount of videos and extract information from them, the necessity of a new platform (i.e., new algorithms, software, and mechanism) is felt.

Form the above discussion, it may be noted that surveillance in Big Data framework is always a challenging task. To the best of the authors’ knowledge, in the literature, there is no article, which reported different state-of-the-art techniques on video surveillance algorithms under Big Data framework. The main contribution of this paper is to investigate the current trends of surveillance in video analysis using Big Data Analytics and to exploit the scope of future usage. It also aims to focus on the research opportunities for video analysis in Big Data platform. In this article we have reported surveillance schemes for four different imaging modalities: conventional videos, remotely sensed videos, medical videos, and underwater videos.

The rest of the article is organized as follows: Section-2 discusses the evolution of Big Data concept and its requirement. The evolution and future growth of big video streams are discussed in Section 3. Section 4 discusses the importance of Big video Data and its use in different surveillance applications. Big Data in cloud computing is briefly discussed in Section-5. Section-6 describes technology evolution in Big Data analysis and different programming schemes. The role and scope of deep learning for Big video Data Analytics is discussed in Section 7. Section 8 investigates the future scope of research in Big Data for video surveillance applications. The paper is concluded in Section-9.

2 Small data to Big Data

The Big Data Analytics is the process of investigating a large amount of heterogeneous data to identify and extract hidden patterns and the correlations that exist among the data, so as to make a decision making system more efficient. The advantage of Big Data Analytics over conventional analytics and business intelligence solutions is that it can analyze huge volume of data, whereas the latter two fail to do so effectively. Let us consider a case of handling a billion rows of data with hundreds of millions of data combinations in multiple data stores and abundant formats. Then some high-performance computing resources are necessary to process or mine them and this enters the domain of Big Data Analytics [5]. In 2005, Roger Magoulas from O'Reilly media introduced the concept of Big Data for first time in order to define a large amount of data which the then existing data handling approaches fail to handle and process because of huge complexity and size. However, the evolution of Big Data as a research and scientific topic existed since the 1970s. In the recent time the concept of Big Data and its implications has been serving the computational world with different perspectives in many fields [6]. One can view the concept of Big Data Analytics as having the following characteristics [7]:

Volume: The first characteristic of Big Data is 'volume'. It represents the quantity of data that is generated; however, it also depends on the value and potential of the data that is under consideration. Statisticians expect that the overall data worldwide will increase by 44 times by 2020 as compared to 2009. By then the amount of data would be around 35ZB from 0.8ZB in 2009 which can be seen in Fig. 1.

Variety: It represents the various categories of data that has to be handled. The knowledge of the type of data helps the data analysts, data managers, and data scientists, use the data efficiently as well as effectively. Categorization of data can be based on different formats, types, and structures of the data. It could be text, numerical, images, audio, video, sequences, time series, social media data, spatial data, etc. Above that data also can be classified as structured, semi-structured or unstructured data. The data can be static or streaming too.

Velocity: The next characteristic of Big Data is 'velocity' which means the speed with which data is being generated and processed in order to achieve the demands and challenges raised in the growth and development of infrastructure. This demands online data analytics as the late decision may lead to missing opportunities. Let us consider the following issues where decisions need to be made on the data stream. In case of e-promotions, based on the current location, the purchase history, what a person likes and

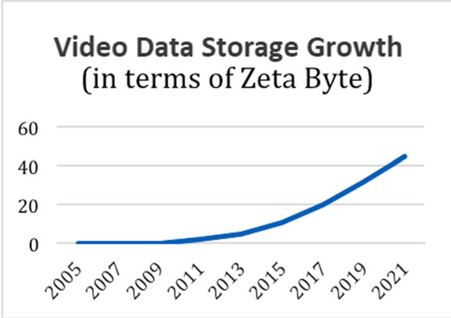


Fig. 1 Increasing volume of data (Source: Cisco VNI [8])

availability of the product, promotional offers can be sent from the store to the customer. Here, Big Data analysis could help the end user to learn about different offers on the type of products he/she may be interested in. This saves a lot of time of the customer in taking the decision of whether to buy the product or not. Another instance can be cited in the healthcare monitoring sector. Here, the sensors monitoring the activities of the human body conditions, alert the concerned persons in case of any abnormality in health measures. The health consultant of the specific person can directly instruct for a proper medication.

There are some more important properties of Big Data like variability, veracity, viscosity, and virality. They are combined with the volume, variety, and velocity and sometimes referred to as the 7 V's of Big Data.

Variability: One of the major challenges is handling huge amount of data which keeps on varying. Variability here refers to the inconsistency of data with respect to time, which affects the mechanism of effective data management. Variability is mostly confused with the term variety. Let us assume that there is a stationery shop that sells pen of five different brands. That is variety. Let us consider a case where we go to that shop and bought three pens with the same brand but with three different colors. That is called as variability. Variability is a quite popular term in the context of sentiment analyses. Variability means changing. It may happen that a word can have a different meaning at different instances in a single tweet. Hence for proper analysis of sentiment, algorithms need to be able to understand the use of the word for different contexts.

Veracity: Mainly relates to the quality and origin of the data to determine whether it is trustworthy, conflicting or impure. The quality of data that is captured varies to a great extent with source. Thus it is very important to consider the veracity of the data source in order to accurately analyse the data. Data usually comes in large volume with high speed. Incorrect data may cause problems for consumers. Therefore, the company needs to ensure that the data is correct. In case of automated decision-making systems, it needs to ensure that the data and the analyses are correct.

Viscosity: It measures the resistance to flow in the data. It may refer to the inertia when navigating through a data collection. It may be noted that due to the variety of data sources, the velocity of data flows may be reduced and the complexity of the processing may be increased. Viscosity can be controlled by improved streaming, agile integration bus, and complex event processing.

Virality: Virality refers to data transmission through the network. It measures the speed at which the data can spread through networks. Time is the major quantity which needs to be taken care of with the rate of spread.

Besides the 7 V's of Big Data, complexity is another closely associated term. The management of data becomes too complex, mainly when a huge amount of data is being generated or captured from multiple sources. In order to extract the information that the data under consideration is expected to convey, different tasks are carried out, such as establishing links between data, and correlating them. This is, therefore, termed as the 'complexity' of Big Data [9].

The key essence of Big Data is its capability to handle huge amount of data of various types generated from multiple sources effectively. A well-defined Big Data with effective usage can bring huge improvement to an organization's business, in terms of improvement in manufacturing of product to sales. Efficient management of Big Data can make companies achieve a better understanding of their products, demands of the customers, feedback from different sale points as well as customers, competitors, etc. which will lead to growth of their business [9, 10].

The dimensionality of usage is a very crucial issue during the analysis of Big Data. These are the aspects on which the users want to interact with their data. These contain frequency, totality, exploration, and dependency [11]. Frequency refers to "how frequently user wants the data to be analyzed to get better accuracy and in-time business intelligence". The desire of the user to process and analyze all the available data is called as totality. Exploration is application of analytic methods by the user, where the entire process of managing the data is defined in response to the nature of the query. Maintaining a proper balance of investment between the existing methodologies, and the adoption of new techniques by the user is known as dependency.

3 Big video data, its evolution and future growth

The size of data is increasing day by day. Every minute, lots of video data is getting generated and adding up to the existing data. This increases the volume of data, the velocity of data, and the variety of data gathered; thus leading to Big Data. The CAGR (Compound Annual Growth Rate) prediction done by Cisco VNITM is expecting a 31% CAGR in overall video data generation.

Figure 2, depicts the video data generation in different segments such as business and consumer. The CAGR for the consumer data is expected to be 23% whereas for the business data it is around 18%. The CAGR prediction for the Asia Pacific, North America, Western Europe, Central and Eastern Europe, Latin America, Middle East and Africa are 22%, 19%, 20%, 27%, 21% and 41% respectively. Thus prediction infers that the growth rate will be too high in the Middle East and African region.

The overall CAGR would be around 22%, which is quite high. The tremendous amount of data will come up with plenty of challenges. Thus the future of Big Data Analytics is quite open and bright. A survey by Cisco has measured that the growth of video data is 72,521 PB per month during 2015, whereas it is predicted to cross 194,374 PB by the end of 2020. This increase in data leads to increase in the dimension of research opportunities. This makes the issues of storage, handling, processing and analysis of Big Data for video much more challenging.

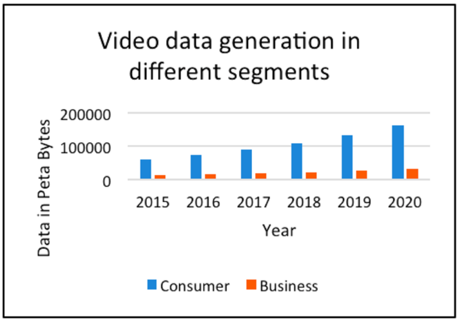


Fig. 2 Video data generation by consumers and business (Source: Cisco VNI [8])

4 Current trends and their applications

In the following sections, we briefly review Big Data Analytics techniques for structured and unstructured data; specifically the case of Big video Data. The role of Big Data in video processing and analysis is discussed in detail and specific attention is given to the application of Big Data concepts in the scene visualization, remote sensing, healthcare, and underwater.

4.1 Role in conventional video surveillance applications

Video processing and analysis is used for many computer vision tasks. Currently, video processing is an emerging field of research [12]. Easy availability of digital devices and cheap sensors has increased the demand of surveillance systems also. But the case is not limited to recording of surveillance video only. Starting from surveillance cameras, GoPros, Drop cams, cell phones and even old-fashioned camcorders, were able to record video at unprecedented scale. Currently, YouTube sees 100 h of new content added every minute and may contain much information embedded in different frames of the video. Video analytic otherwise known as video content analysis (VCA), involves a number of approaches to oversee, analyze, and extract useful information from video streams. Surveillance is also a part of VCA [13]. The major contributors to the exponential growth of computerized video analysis are the closed-circuit television (CCTV) cameras and all the video-sharing websites. One of the most important challenges for the VCA is the sheer size of video data [14], which is seen as an opportunity for the Big Data platform. The prime focus of video analytics in recent years has been on the designing of sophisticated indoor and outdoor surveillance infrastructures as well as enhanced security systems. Video analytics can efficiently and effectively perform various surveillance tasks such as detection of suspicious movements, identification of objects removed or left unattended in a scene, detection of loitering in a prohibited area, detecting attempts of camera tampering, etc. [15]. The amount of video data generated from different resources are increasing day by day and thus the amount of total video data generated on monthly basis is expected to cross million Petabytes (PB) in near future. Handling of such huge data is a big challenge. In case of content-based searching, instant video retrieval would be too difficult. Processing such huge data to extract the requisite information would be a much difficult task, because of the infrastructural limitations. The state-of-the-art algorithms designed so far are not efficient to handle such large amount of video data at a time. Thus lots of research opportunities are there in the area of Big Data Analytics for video analysis [16]. Few of the issues are highlighted in the subsequent part in this section [17].

Text detection in surveillance Text detection from the video is a quite interesting and a challenging area of research in video analysis. It has different applications including video indexing [13], assistive technology for the visually impaired [18], automatic localization for businesses [19], robotic navigation [20], etc. Texts in a video occur in two ways: artificially or scene text. The artificial text occurs in a video as it is added by the source user so as to enhance the content or to better represent the contents of it. Scene text occurs in a video as it appears in different parts of the scene naturally e.g., text on T-shirts or road signs. Identification of texts in the video in Big Data environment plays an important role in the field of video analytics. Ayed et al. [21] have proposed a MapReduce based algorithm which can detect texts from the natural scene videos in Big Data framework. Texture-based features are used for this process. Fetching a video from an online or offline source and extracting frames from it followed by detection of the text region is computationally challenging. To handle this complexity, the Hadoop Distributed File System (HDFS) [8] has been used and analyzed and classification of text region is carried out by MapReduce [6, 21] model.

The problem of text detection in outdoor and complex background images is a challenging task. Recently deep learning based approaches are gaining huge popularity. Optiz et al. [22] have proposed a strategy to recognize text from outdoor image or video data by an amalgamation of local ternary pattern (LTP), maximally stable external region (MSER) and deep convolutional neural network (CNN). The performance is further improvised by Turki et al. [23], that utilized edge enhancement by filtering out the complex background, detection of MSER candidate, filtering of non-text regions, and classification of character candidates using SVM and CNN, and elimination of false positive by grouping of words. The use of MSER helps in identifying the text regions in the scene. However, Selmi et al. [24] presented a system that employ deep learning for the automatic detection of text regions (basically license plates) in a scene and recognize the characters. Although the performance of Selmi et al. [24] is satisfactory for good resolution visual data, but it involves lots of pre-processing tasks prior to the application of the CNN. Furthermore, the performance is poor in case of low resolution data. In order to address this, Shivakumara et al. [25] have proposed a method which employs an amalgamation of CNN and recurrent neural network (RNN) to recognize the text, few results of which can be seen in Fig. 3.

A detailed review on the development of text and document analysis and the role and future of deep learning based approaches has been carried out by Vincent and MarcOgier [26], which concludes that a balancing between the conventional and deep learning based approaches will sustain in future. The Big Data aspect and future of text based urban data management system has been thoroughly discussed by Babar et al. [27].

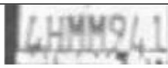
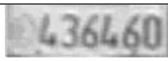
Text image			
Recognized text	"WUL96"	"BJY320"	"BHP3141"
Text image			
Recognized text	"HWE949"	"4HMM941"	"436460"

Fig. 3 Text recognition results of Shivakumara et al. [25]

Human activity detection in surveillance Analysis of human activities in video surveillance is also a challenging task [28]. The growth of online and offline video contents demands now the automatic indexing of multimedia content for ease of search and retrieval. This can be achieved by the use of metadata, soundtrack, transcripts, and the visual content of the video [14]. Use of social signal processing was investigated by different researchers too where the nonverbal cues, such as facial expressions, posture of body, gestures, vocal characteristics, etc. are explored to analyze the behavior [29]. To analyze the behavior of workers in China metro construction, Guo et al. [30] developed a Big Data framework. The authors established a behavioral risk knowledge base and from the recorded video, the worker's unsafe behaviors are collected. The Big Data cloud platform is used for storing data by the distributed file manager. One of the potential usages of video analytics is the analysis of the buying behavior of groups of people and people in a group [31]. For instance, in case of a family that shops as a unit, only one pays the bills or communicate with the cash counter of the store. In this way, the traditional systems miss collecting data on the buying patterns of other members of the family. At situations like this, video analytics can assist the store managers, to understand the size, demographics, and the buying behavior of individual members of the family.

Multi view bag of words (MVBoW) representation, and graph model based dictionary learning for human action recognition scheme has been proposed by Gao et al. [32], in order to recognize human action under various challenging situations. A detailed review on the object tracking using visual data has been presented in [33]. Recently Shao et al. [34] have proposed a scheme for processing of big surveillance video data, in order to monitor the behavior and activities in sensitive geographical areas. Their scheme uses the spatio-temporal information of the particular geographical area under the coverage of the smart camera and claims to detect any abnormality in the behavior of people and issue pre-alarm before any mishap. A smart surveillance mechanism that uses fog computing is being developed by Liu et al. [35] for tracking of objects in varying lighting conditions. A correlation filter based tracker is also employed to tackle the illumination variation issue, whereas fog computing is used to speed up the processing. Gao et al. [36] have explored the feature learning, unsupervised cross domain learning, and supervised cross domain learning to generalize the action recognition process. They have also introduced a multi-view and multi-modality human action recognition dataset. Ray and Chakraborty [37] have used a 3-dimensioanl Gabor filter and minimum spanning tree scheme to detect and track object of interest in a dynamic background scenario. In order to handle the issue of occlusion they have employed Kalman filter and for the data association task liner assignment algorithm is used.

Restricted video filtering for surveillance Increase in the web-based video hosting sites, increased the challenges of the telecom operators and service providers. Detection and filtering of videos which are banned from distribution or having restriction of viewing and distributing is a big challenge around the world. A huge number of databases are there containing billions of videos. Detecting a restricted video is too much challenging. Jansohn et al. [38] investigated a technique where an amalgamation of conventional key-frame-based methods and statistical analysis of MPEG-4 motion vectors are used for pornography detection. Behrad et al. [39] also proposed a scheme where the connected components in consecutive frames are used for feature extraction. Combination with color histogram and entropy of motion vectors are explored for obscene video recognition. The analysis of billions of videos of variable length and resolution is extremely challenging [40]. The YouTube partners, who have created a channel, can upload videos on their channel. The type of video and its content analysis is an emerging area of research [41]. It is observed that the content of a video may be violating the ethics and code of conduct of the YouTube. In case YouTube wants to restrict or delete those videos, it has to know the content of it. Another challenging area is the cloud radio access network (CRAN) research [42]. Sheng et al. [43], discussed about the strategies and architecture for the delivery of video data in the heterogeneous cloud radio access network. A Big Data based video distribution scheme has been proposed by Ruiz et al. [44] for content-based data distribution.

Crowd tracking The common weakness of most of the video processing systems is their incapability to address densely crowded scenes. With the increase of the density of moving objects in the scene, the performance in terms of surveillance gets degraded significantly. A change of the field of view, and density of people, the ambiguous appearance of body parts (e.g. some parts of one object in the scene may be similar to another nearby object), adds up to the complexity. This inability to deal with crowded scenes represents a significant problem. Such a system has different applications in real life scenario including, crowd management, behavior analysis, public space design, virtual environment, intelligent environment, etc. [14]. A logistic and infrastructural coherence is required to handle such issues by the help of Big Data [45].

A combination of motion and appearance saliency cue is used recently [46] for the detection of objects and the hierarchical representation of targets from groups to individuals in crowded scenes. The said scheme is designed in Big Data Analytic framework which allows simultaneous micro (individualistic) and macro (holistic) levels of inference on visual information. A detailed discussion on the physics inspired schemes, direction of research for crowd management in Big Data framework, its data capture, software development with crowd behavior understanding are discussed in [47]. Block matching, and social force model are employed by Kajo et al. [48] for crowd tracking. Huang et al. [49] have proposed a scheme for counting the crowd using deep learning approach which uses body structure and convolutional neural network. A support vector machine classifier is also used for counting number of people in a dense crowd, where Shami et al. [50], used SURF [51] feature. The head detection results of Shami et al. [50] are shown in Fig. 4. In order to track people with abnormal behavior in a crowded scene, a convolutional neural network based crowd motion analysis scheme has been proposed by Ravanbakhsh et al. [52]. Li [53] has developed a deep learning scheme to understand behavior of crowd using a spatio-temporal framework. A deep convolutional neural residual network based scheme has been proposed by Mandal et al. [54] to recognize the behavior of crowd.

Fig. 4 Results of the head detector employed by Shami et al. [50]

Multi-camera network based surveillance Designing of a smart video surveillance system is one of the proactive areas of research in computer vision. On daily basis thousands of cameras are collecting a large volume of video data in a moderate sized city, in the recent time. In order to extract relevant information from these heaps of video data, researchers and data experts are suggested to develop algorithms and intelligent systems. The capability of a camera to capture the view of a scene is finite and limited by scene structures. Thus, a multi-camera network needs to be used in order to overcome the limitations of a single camera with respect to the field of view. A multi-camera system can be employed for the surveillance of a wide area, like tracking a vehicle traveling through the road network of a city or investigating the global activities happening in a large railway station. Traditional computer vision challenges in tracking and recognition may be found to be robust to pose, illumination, occlusion, and clutter. However, tracking targets over a wide area as they appear and disappear is still a potential challenge. So the possible options in this regard is a decentralized or distributed camera network (both distributed and decentralized may be categorized under collaborative and distributed sensing schemes) [55, 56].

A video coverage monitoring framework is proposed for identification of problematic camera sensor and thereafter reduces the data volume generation by eliminating faulty sensors. It analyzes the projection of a set of 3D volume to check if it satisfies a number of constraints predicting the success of specified tasks [57].

Distributed and collaborative sensing is one of the emerging trends in recent days, which deals with the development of a mechanism that provides ways to the constituent sensors to interpret each other's observations and measurements. The implementation of such a system largely depends on the availability of global reference information of all the sensors, in the absence of which, the sensors effectively get isolated, and fail to incorporate the relevance of the observations of other sensors [55, 56]. Chen et al. [58] developed an object tracking model over a multiple-camera network in conjunction with Big Data framework. The authors have used Histogram oriented gradient with SVM to detect pedestrian from a particular camera view. The trajectory of the object in the multiple cameras is obtained by geometric projection to optimize the trajectory. Blat et al. [59] recently, proposed an integration of multiple data source for improving and monitoring the generated data, which lead to an analysis of semantic data in a better way.

Big Data based surveillance architecture There exist two perspectives to video analytics, as proposed by Gandomi and Haider [4] and they are server based and edge based analytics. The video obtained by each camera is routed back to a dedicated centralized server which carries out the analysis in case of the server-based architecture. The video is compressed by reducing the frame rates and/or the image resolution in order to cope with the limitation of bandwidth. The advantages of such analytics are like it provides the economies of scale and facilitates easier maintenance, whereas, their limitation is that they are less accurate because of the loss of information during compression. On the other hand, the analytics are applied at the edge of the system in case of edge-based architecture, where the video analytics is carried out locally on the crude data captured by the camera. The advantage of such an infrastructure is that the entire content of the video stream is available for the analysis, making it a more effective content analysis; however, these are more expensive to maintain and have a lower processing capability compared to the server-based systems.

4.2 Role of surveillance in remotely sensed video analysis

Remotely sensed (RS) video analysis is a promising application domain of digital video processing. Remote sensing [60, 61] can be defined as a process of acquisition of physical data of an object from a distance without direct-contact. There are vast applications of remote sensing in the earth observations like an early prediction of environmental change, land-usage monitoring, maintaining forest ecosystem, and supervising agricultural growth and so on. The term "Remote Sensing" was first used by Evelyn Pruitt, a scientist working for the U.S. Navy's Office of Naval Research. The year 1972 is the most important in the history of remote sensing in terms of the earth observation. In that year, Landsat 1, the first earth-orbiting satellite was launched for monitoring the earth surface in a systematic manner.

Many mechanisms have been proposed to take care of vehicle tracking in case of visible imagery sequences collected by sensors on multiple aircrafts or low-rate video [62–70]. The surface object tracking capabilities are also seen in synthetic aperture radar sensors [64], which can be used for the tracking of ships. Vehicle tracking schemes are also developed using a combination of airborne spectral and polarimetry imaging [65], to illustrate the potential usage of remote sensing platforms. For the tracking of ships in ocean, satellite imagery [66] along with other data sources [63] have been used. A target tracking scheme is proposed by Meng et al. [69], where the target is modeled by spectral and spatial features. The high dimensionality of remotely sensed video data is also difficult to handle. To analyze this imagery, dimensionality reduction is an alternative way, which in turn may lead to the loss of information. On the contrary, the necessity of efficient mechanism to handle higher dimensional data (without reducing dimensionality) is also considered in the literature [69]. An amalgamation of Bhattacharyya distance [71], histogram intersection, and pixel count similarity is used for the target matching process. Detection of change in geographical attributes of earth surface by the help of Gibbs Markov random filed based model and fuzzy c-means clustering algorithm has been proposed by Subudhi et al. [72]. Rathore et al. [73] proposed a real-time Big Data Analytics architecture where remote sensing data acquired by satellite were sent to the base station. Hadoop based architecture is used for filtration, load balancing, and parallel processing. Cavallaro et al. [74] proposed the use of Support Vector Machine (SVM) to contribute to smart data analysis processes with scalable and parallel processing techniques. The authors intended the scheme mainly for land cover identification. The information about the measured signal and the expected signal are utilized in the detection framework [74] that evaluates the effects of moving objects suffering from the SAR focusing process. Chi et al. [75] have made two case studies for remote sensing data in Big Data framework. In the first case, the Big Data framework is used to detect marine oil spills. In the second case, content-based information retrieval is performed to extract information from a large database of remote sensing images.

In remote sensing, energy radiated from the earth surface is measured by sensors mounted on the aircraft or satellite platform. In turn, these measurements are used for generating images in digital format. Sensor data is mainly used to view the physical entities like water, bare soil, and vegetation using electromagnetic radiation emitted or reflected from these. A major foundation of remotely sensed images lies in its spectral, spatial, and radiometric behaviors. Depending on the variety in spectral, spatial and radiometric characteristics, many kinds of remote sensing data are generated from different types of sensor i.e., Landsat Multi-Spectral Scanner (MSS), Thematic Mapper (TM), Satellite Probatoired' Observation de la Terre

(SPOT), Advanced Very High-Resolution Radiometer (AVHRR). RS Big Data challenges include the following: (i) difficulties in managing the massive RS data including the extract information from a large amount of data and (ii) handling RS data in the cluster and cloud-based parallel systems [75].

The geometrical diversity of massive RS data is exponentially increased with the growth in the amount of video data. Due to the fact that the videos come from different data centers, which are far away from each other. The videos could be of a different format, different length and their way of interpretation could be different. Better analysis of earth systems demands that all geometrically related RS videos are necessarily accessed in parallel. Though storing and managing of these are related, diverse nature of RS videos for a particular application is more challenging [75]. In real life, due to computational overhead, it is a big challenge to load massive higher dimensional RS video data in a single computer system [75]. To overcome this limitation, the video can be analyzed intensively in cluster-based parallel computer systems. NASA took the first initiative for global RS data processing in NEX system on a Beowulf cluster with 16 computers [70]. Instead of loading data into the local memory of parallel computer systems, data can be accessed directly from cloud platform also. Cloud computing provides software, infrastructure, and platform as service (i.e., SaaS, IaaS, and PaaS, respectively). For example, Amazon EC2 provides IaaS; whereas Google AppEngine and Microsoft Azure offer PasS. Hadoop [76] also provides an important infrastructure for the cloud. The open source environment of Hadoop in MapReduce framework helps to analyze RS Big Data. Hadoop-GIS system [76] also provides a platform for large-scale data processing.

4.3 Role in healthcare and diagnosis

The usage of Big Data in the field of healthcare and telemedicine is an emerging trend these days. Starting from the symptom-based health diagnosis to medical transcription, from open surgery to tele-robotic surgery, from medicine prescription to telemedicine applications, Big Data Analytics is taking up its usage. It is also very much obvious that without using the Big Data Analytics methodologies, handling and analyzing a huge amount of data generated would be extremely difficult.

A light field (LF) based 3D telemedicine system has been developed by Xiang et al. [77]. They have used light field video data formatting, which leads to the Big LFV Data analysis. In order to handle the issues of storage and processing of the huge data, they extended the multi-view video coding or MVC, to LF – MVC or light field MVC. A Big Data framework has been developed by them which integrates the LF video with the conventional telemedicine framework, which leads to an efficient system for healthcare units.

The advent of video data acquisition methodologies has opened a wide avenue in understanding human physiology with a wide range of precisions. But with this advancement in modalities, there comes a burden of storing and processing of the data generated by them. A single human scan using modalities like MRI, DT-MRI, CT, FMRI etc., can generate megabytes of data. This scales exponentially when the precision in imaging increases (e.g. a single stack of electron-microscopy images of the cell is more than a gigabyte). When such data is accumulated for a year, it scales to terabytes and according to a statistical survey, every five years such data doubles. Thus from the above, we can understand the 3 V’s (which are the major characteristics of any Big Data problem) are very much inherent in medical imaging informatics also. Processing and analysis of such huge variety of data arriving at such a high velocity is a major issue in medical paradigm and this has motivated medical researchers to adopt the Big Data platform for its processing. Analysis of large-scale medical imaging data not only provides information about the precise information on the physiological condition of individuals but provide gross information on the risk of epidemiology in the community. The derived information can provide the health status of the community and help in designing policy for the health sector, insurance sector etc. [78].

Another instance can be considered where a video recorder needs to be installed to record the behavior of a patient, instead of employing a nursing staff round the clock for health monitoring. The video data can be analyzed and the patient's health monitoring can be carried out. If a single such recorder is producing 2GB of video data per day, and if it is assumed that around 1 Lakh such cases are there throughout the Globe, then the amount of data generated would be around 100000GB per day. Handling of such huge data is a real challenge, most importantly detection of the abnormal behavior of the patient under observation is the key task. Normal movements of the patients are not of any interest to the doctors, any abnormal event or movement only needs to be addressed on urgent basis. Thus any abnormal behavior only is needed to be stored and remaining time, even if the camera records the video, it is not required to store the data. Thus the data storage problem can be taken care of up to certain extent. There comes the use of Big Data Analytics for video analysis.

Big Data Analytics can provide a handful of tools to exploit the latent information in the images taken from different perspectives. This can be from the assessment of structural information about the organ to explore its physiological state to its cross-correlation analysis at the genomic level. Big Data Analytics have been used in studies of risk factor studies, genotype phenotype-association, and causality analysis.

The heterogeneity in modality and lack of unified protocol makes the processing of the data difficult [77]. This hinders the process of viewing the physiological condition of an organ at the community level. To circumvent this problem there has been a global drive for archiving the images at a common repository and the result of which is cardiac atlas project and cancer imaging archives. Recently, Bansal et al. [79] made an extensive review on the recent advances of Big Data for disease surveillance, monitoring medical adverse events and tracking of the patient sentiments and mobility.

With the availability of these diverse images from different sources at a common place, there comes the issue of processing then to extract useful information. Analysis of medical images involves the following steps: localization of the region of interest, segmentation, characterization, information integration from different modality, model reconstruction, and physiological parameter estimation [80]. Processing as a whole becomes challenging due to the huge size of the 3D stacks, diverse modality and drastic change in perspective of viewing the organ. This demands the design of features and algorithms which is invariant to the change in perspective, diversity of the modality and explores the inter-relationship of the slices for processing the 3D stacks of images. The integration of intra-domain information is already a challenging issue in medical imaging domain. This becomes more complex when we have a hierarchy of inter-domain information for fusion. Churning this information to understanding the causality and precisely predict the future state of the disease, demands scaling/design of the new methodologies.

Angioplasty video analysis is a quite popular methodology used for detection of blood flow blockage in the blood vessels, basically in the heart. Tremendous data is generated every hour around the globe. Analysis of such large video data using Big Data Analytics methodologies is gaining popularity.

Medical video analysis has been a growing field since its inception. With the advent of modalities, there has been a growing need for designing integrated automated tool for analysis of the organ under investigation. Eberhardt et al. [81] diagnosed lung lesion using multi-modal bronchoscopy images. Upcoming of repositories like the cancer imaging archive, cardiac atlas project, virtual skeleton database etc., [82, 83] motivated researcher to understand the causality of different disease conditions and develop tools for 3D modeling of the diseased organ for better visualization and precise intervention during surgery. This kind of work was carried out for brain tumor localization by Menze et al. [84]. Young and Frangi [85] designed an integrated model for cardiac physiologies useful for understanding disease and planning intervention. Such multi-modality modeling helped in the study of different cardiovascular diseases. One such study by Manolis et al. [86] reported the applicability of such models in assessing the cardiac damage due to hypertension. Researchers are now moving towards integrating cross-domain information along with the imaging information for understanding the causality of diseases at the genomic level. Liu et al. [87] designed one such tool for automated diagnosis of Glaucoma. They integrated the screening data, fundus imaging information and genomic data for prediction of glaucoma using multiple kernel learning. Vallieres et al. [88] designed a radio genomics model for joint FDG-PET and MRI texture features for prediction of lung metastases in soft-tissue sarcomas of the extremities. There are many such studies in the literature which have helped researchers for better understanding and precise design of their intervention policy.

Imaging informatics is also getting its popularity in recent days for predicting and monitoring infectious disease like Ebola. Detection and prediction of this kind of diseases need a huge volume of medical data for processing. Luo et al. [78] provided a study on medical informatics integrated with cloud platforms on Big Data platform for predicting and monitoring infectious diseases. Recently, Lee et al. [89] made an extensive review of challenges, procedures, and advancement of medical image data in Big Data framework. The usage of Big Data Analytics in health care sector has been elaborated in [90].

4.4 Role in underwater surveillance

Underwater surveillance is recently getting lots of attention because of the potential challenges involved in it. It is useful for several applications including underwater surveillance of ocean, underwater wreck detection, automated underwater vehicle (AUV) design and navigation, submarine navigation, preservation of underwater animals, coastal surveillance, ship-wreck detection, etc. The underwater imaging model depends on many factors such as salinity of the water, the composition of the water, oxygen content, amount of pollutant, size of all the intrinsic particles, haze, etc. Thus to have accurate detection of objects in such case, the surveillance algorithm should be self-sufficient to enhance or restore the input data quality by some de-hazing mechanism. Then it should be capable of detecting or tracking the moving objects accurately. Thus ships and cargos can be installed with such systems; hence any case of an intruder trying to get into the ship or cargo can be detected. Most of the existing state-of-the-art literature focused on the underwater image processing system development which follows enhancement of the underwater image/video and thereby detection of moving objects and tracking of the same. Some of the works are reported in existing literature which used local motion in the scene. Few works are done with global motion in the scene; however, most of the work focused on the availability of a priori foreground information. The problem becomes more challenging under poor visibility as the source of illumination is not able to illuminate properly or viscosity of the water may not allow so. The poor visibility may be due to underwater degradation such as degradation of chromaticity, improper visibility of the scene due to poor lighting etc. When the scene is not visible properly, then detection of an object of interest becomes very challenging. Again haze in a water body is an important parameter which needs to be considered. Erroneous object detection may lead to false tracking too.

The existing challenges in literature can thus be focused in resolving two types of complexity in detection and tracking of underwater video [91]. The first type of complexity is related to the scene. The scene complexity includes the degraded and obscured videos because of underwater turbulence. The second type of complexity is related to the object itself. The object complexity may be due to improper illumination, object deformation, change in viewing angle, occluded objects, non-static background etc.

Minami et al. [92], proposed an algorithm for fish tracking with visual surveying. It is basically a shape based tracker which is allowed to evolve through a genetic algorithm (GA). Use of searching scheme like GA needs a high-performance computing application for real-time application. Foresti et al. [93], developed an algorithm for underwater pipeline detection. It works well with small occlusion of pipelines by sand, mud, seaweed etc. Sehgal et al. [94], proposed an algorithm called TOUCH (Track Object Underwater based on Color Hue) for tracking underwater objects. They have proposed a modification to the 4-connectivity of pixels and Hue is used as the feature to distinguish the object of interest and track it. An algorithm for underwater fish tracking for moving cameras based on deformable multiple kernels has been developed by Chuang et al. [95], and few results of the same scheme are shown in Fig. 5. It performs well at low illumination underwater videos and can track single as well as multiple objects.

A frame-differencing algorithm has been proposed in [96], which takes five consecutive frames and detects the underwater moving objects using a set of logical operators. Mondal et al. [97] proposed a tracking-by-detection algorithm to detect and track camouflaged objects from underwater and conventional video scenes. The authors have used an integrated feature space for better separability of the camouflaged object and the background. A probabilistic neural network (PNN) is modified and used for detection of the objects. Figure 6 presents the tracking results obtained by Mondal et al. [97] in case of occlusion. Rout et al. [98] have developed a Wronskian framework based underwater object detection scheme which explored the background modeling capability of Gaussian mixture model.

One of the major challenges in underwater data is the study of water ecology. A massive amount of data may be needed for this where the sound signal coming from the underwater animal and object-specific videos were recorded for different applications. Storing and analysis of this vast amount of video and sound signal need a special attention. A case study of large-scale video processing for studying underwater ecology is reported in [99] where a high-performance computing facility is developed for underwater fish monitoring in the context of the Fish4Knowledge project. One effective pipeline architecture for underwater Big Data Analytics is proposed by Alharbi et al. [100], which uses pipelining to speed up data processing and transmission. Lebart et al. [101] have presented an algorithm for automatic indexing of underwater videos and illustrated scheme for benchmarking them. A detailed

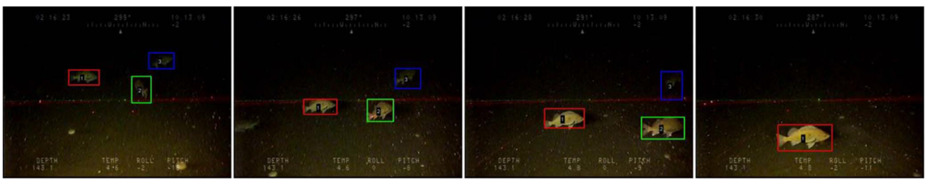


Fig. 5 Tracking results of Chuang et al. [95]

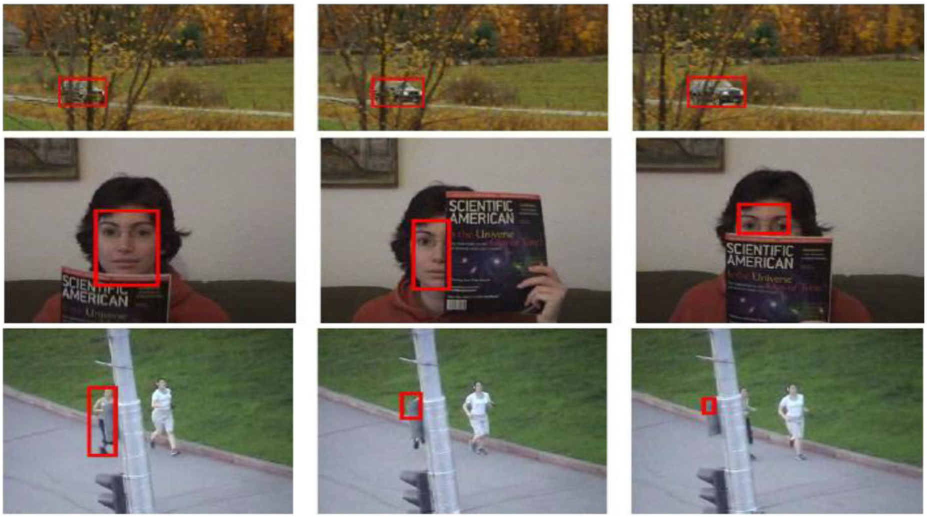


Fig. 6 Tracking results of Mondal et al. [97]

survey on the subsea object tracking is also carried out by Trucco et al. [102], with emphasis on the imaging model, underwater object motion analysis, methods to tackle the subsea tracking challenges. Xiang et al. [103] have discussed the use of fuzzy-logic in tracking and controlling of underwater vehicles. Trends and future scopes with respect to soft computing approaches are also reported in it.

The major contributions, strengths and weaknesses of the state-of-the art techniques are summarized in Table 1; whereas Table 2 focuses on the key contribution of the survey papers in the relevant research areas.

5 Big Data and cloud computing

The technological growth is exponentially adding up to the amount of data generated. Analysis of such gigantic amount of data requires a well capable computing infrastructure, which can handle an enormous amount of data and should be capable of extracting information from it. It should be able to analyze the statistical information. The progress in the development of Big Data handling tools like media server [104], Hadoop [9] and R [105] platforms, have given a better scope for analysis of big multimedia data. Cloud computing can provide the infrastructural facility for Big Data analytics. Cloud computing is a well-established powerful technology which can perform very large scale complex computing. The modern-day information communication technology has got a paradigm shift towards cloud processing. The cloud has advantages of parallel processing, virtualization of resources, data security, and scalable data storage capability. Above all, any kind of data service can be integrated with the cloud, which eases the access of requisite information by the end user. Although the integration of cloud computing for Big Data Analytics is taking up its pace, still it is in the infancy. A lot many issues are yet to be resolved in order to efficiently integrate the two powerful technologies for future data analysis. Few of these issues are like data security, data accessibility, usability etc.

Abaker et al. [106] have investigated the relationship between Big Data and cloud computing, Big Data storage systems, and Hadoop technology. Trends in Big Data Analytics

Table 1 Comparative analysis of some key surveillance schemes

Year	Author	Methodology	Strength	Limitation
1999	Minami et al. [92]	Genetic algorithm based fish tracking scheme	Able to track fish in scene dynamism	In a school of fish the performance is not satisfactory
2000	Foresti et al. [93]	Tracking of underwater pipe-line and cable structures based on color compensation and artificial neural network	Robust to decolorization, accuracy of detection is higher as geometrical reasoning is used to suppress false alarms	Under dense haze the scheme fails terribly
2004	Ezaki et al. [18]	Text extraction scheme using Sobel edge detection, Otsu binarization, connected component extraction and rule based filtering	Better accuracy in case of large size text	Accuracy in case of small characters is not satisfactory and not reliable for practical use
2004	Sehgal et al. [94]	Color based tracking of underwater vehicles	Works well in case of shallow underwater environment	Fails in case of real underwater scenario with dense haze, poor lighting and decolorization
2006	Saracoglu et al. [19]	Multiple hypothesis based text recognition strategy is implemented with the help of statistical language model	Text localization and recognition is satisfactory	In case of challenging outdoor videos, accuracy is not satisfactory
2008	Fleuret et al. [55]	A probabilistic occupancy map based multi-camera object tracking scheme	It can accurately track upto six people inspite of occlusion and illumination variation	Not suitable for dense crowd
2009	Carrano [63]	Ultra-scale tracking of multiple vehicles using path dynamics, mover map and object features in low-frame rate videos	Track multiple vehicles and efficiently perform in case of go-stop-go case.	Not suitable for dense traffic condition
2010	Palaniappan et al. [65]	A combination of appearance model, saliency, and motion prediction scheme for low-frame-rate object tracking	Work effectively for low-frame-rate wide area video	Fails in case of dense traffic scenario
2010	Xiao et al. [66]	Behavioral model of the vehicle, vertex and pairwise edge matching used in relation to graph matching framework for traffic surveillance	Works good at low resolution input video with dense traffic condition	Not suitable for non-stationary camera setup and dynamic background condition
2011	Subudhi et al. [1]	Detection of moving objects based on spatio-temporal change information and tracking by centroid localization	The processing speed and accuracy of detection are good in case of stationary camera set up	Performance is not good in case of non-stationary camera based video
2012	Meng et al. [69]	Object detection and tracking by spectral and spatial features based target modeling, and histogram based matching	Works satisfactorily in case of high-resolution multi-spectral satellite video	Small objects, dense traffic and low-resolution input affect the performance of the scheme

Table 1 (continued)

Year	Author	Methodology	Strength	Limitation
2014	Opitz et al. [22]	LTP, MSER and deep CNN based end to end text recognition	Detects text regions and recognize characters satisfactorily, works fine with images of wild and outdoor environment	In case of complex background the efficacy of the scheme is not encouraging
2015	Ayed et al. [21]	MapReduce programming model for text extraction using sub-image level analysis	Texts can be extracted with lesser computation time irrespective of different languages, fonts, size, color and orientation	Lots of false detections because of text like regions in image
2015	Gou et al. [30]	Construction worker's behavior observation using Big Data Analytics	Practically validated and the unsafe behavior of workers were identified satisfactorily	Designed for a specific construction site, thus the scheme may not be suitable for different area
2015	Zitouni et al. [31]	An amalgamation of the object saliency and motion is used to characterize individuals and groups in crowded scenes	Able to detect people with higher accuracy and delineate people within group	In case of dense crowd scene, performance of the scheme is not good
2015	Chen et al. [58]	HOG and SVM based detection and spatio-temporal local context for the tracking of pedestrian	Satisfactory outcome in case of road-traffic monitoring	This strategy cannot work for tracking of objects with arbitrary motion
2015	Gao et al. [32]	Multi view bag of words (MVBoW) representation, and graph model based dictionary learning for human action recognition	Efficiently recognize human action, under various challenging situations	Effectiveness is not known in case of crowded scene
2017	Turki et al. [23]	Detection of text region by MSER and recognition of characters by CNN	Able to provide satisfactory performance in case of wild as well as complex background image or video data	Involve a number of filtering operations after the MSER based region selection
2017	Selmi et al. [24]	Used deep learning for detection of text region (basically the license plate) and recognition of the text characters	Works well in case of good resolution images, can handle uneven illumination	A lots of pre-processing tasks are involved before the recognition step, the performance is not encouraging in case of low resolution images
2017	Chuang et al. [95]	Deformable multiple kernel based scheme for underwater fish tracking with stationary and non stationary camera set up	With great accuracy tracks the random movement of fish, even in case of challenging illumination	In case of high deformation, and sharp change in the trajectory the scheme fails
2017	Mondal et al. [97]	Color, shape and texture feature based probabilistic neural network framework with fuzzy energy based active contour	Efficiently tracks objects in outdoor as well as underwater environment	Not suitable for multiple object tracking

Table 1 (continued)

Year	Author	Methodology	Strength	Limitation
		detection scheme for camouflage object tracking		
2018	Huang et al. [49]	Deep convolutional neural network for crowd counting	Performs satisfactorily in case of crowded scene	Performance of the scheme in case of dense crowd is not known
2018	Mandal et al. [54]	Crowd behavior recognition using subclass discriminant analysis and deep residual network	Performs well with low crowd density	Performance under high density crowd is not demonstrated
2018	Ravanbakhsh et al. [52]	Convolutional neural network for crowd analysis and detection of abnormal event	Perform well in case of low density crowd image	The performance is not clear under medium and high density crowd image
2018	Li [53]	Deep learning approach in spatio-temporal framework for crowd behavior analysis	Effectively detects the spatial and temporal motion dynamics, effective for path forecasting, classification of the behavior of crowd	The performance in a highly dense crowd scene is not clear
2018	Shao et al. [34]	Intelligent camera system for the processing of big surveillance video data	Able to detect and activate the alarming system in case of any abnormal behavior in the crowd, start retrieving data in case of detection of any alarming situation	Not suitable for dense crowd situation
2018	Liu et al. [35]	Multi position detection using alternate template (MPAT) based fog computing strategy is used for object tracking in varying lighting situation	Can handle occlusion, illumination variation and slow and fast movement of object in public places	Real time deployment of fog computing strategy is still a challenge
2018	Pradhan et al. [116]	Deep convolutional neural network for prediction of the intensity of cyclonic storm	The scheme is able to predict cyclone with high accuracy	The effectiveness in real situation is not verified
2018	Shami et al. [50]	SURF feature and SVM based classifier to detect presence of people, use of regression to determine average head size and determination of crowd density	Works very well in case of medium and high density crowd scene	The processing time is a key issue, because of the complex architecture
2018	Kajo et al. [48]	Block based matching, particle advection, and social force model for crowd motion analysis	Effectively predict the direction of motion of pedestrian using the social force model. Perform well in low density crowded scene	The performance is not clear under high density crowd scene
2018	Pan et al. [33]	Visual attention feature (VAF) based outdoor tracking scheme	Handles efficiently any variation of scale, velocity, and acceleration of object, illumination variation, occlusion etc.	The performance in complex situations is not satisfactory

Table 1 (continued)

Year	Author	Methodology	Strength	Limitation
2018	Shamsolmoali et al. [120]	Recovery of resolution by a deep learning approach	The scheme is quite useful to handle artifacts occurred due to compression and denoising	Performance on real time data is not investigated. The processing time could be a limitation for real time implementation
2018	Xie et al. [121]	Machine learning approach for detection of abnormal behavior in crowd video	Performs well in case of sudden agitation, and abnormal behavior of a group of people or individual in a crowd	Speed, real-time immunity, and robustness of the scheme needs improvement
2018	Xu et al. [122]	Identification of any abnormal behavior of crowd using dual channel convolutional neural network and crowd feature engineering	The scheme is efficient in case of identification of abnormal behavior in online crowd data	The efficacy of the scheme in case of high density crowd video is not investigated
2018	Gao et al. [36]	Feature learning, unsupervised cross domain learning, and supervised cross domain learning are explored to generalize the action recognition process	Introduced a multi-view and multi-modality human action recognition dataset, effective scheme for cross-domain learning	The situation of human action recognition under crowded scene is not considered for evaluation of scheme
2018	Shivakumara et al. [25]	An amalgamation of CNN and RNN including Bi-directional Long Short Term Memory (BLSTM) for text character recognition	Works well with degraded images as well as of poor resolution	The scheme fails to handle text images which are dis-oriented
2019	Ray et al. [37]	3D Gabor filter, and minimum spanning tree based spatio-temporal analysis for object tracking in outdoor environment	Works well with stationary as well as non-stationary camera set up, no initialization is required, can handle occlusion	Performance is not satisfactory in case of dense crowd, sever illumination variation, and complex background

Table 2 Some notable survey papers and their key contributions

Year	Author	Survey on	Key Contributions
1999	Meggitt et al. [91]	Underwater tracking	Methodologies for installation and recovery of sea sentinel array systems
2003	Lebart et al. [101]	Underwater video indexing	Automatic indexing of the underwater survey videos on the basis of their inherent characteristics and methodologies for the benchmarking of those videos
2006	Trucco et al. [102]	Underwater surveillance	Motion analysis issues in underwater environment, imaging models used and methodologies employed for subsea object tracking
2008	Zhan et al. [14]	Crowd analysis	Methods to handle crowd analysis issues and their usability to computer vision domain
2008	Ko [28]	Outdoor surveillance	Development of embedded systems for homeland security by the amalgamation of motion, behavior and biometrics
2011	Hu et al. [13]	Video indexing and retrieval	Broader prospectus on analysis of video data including indexing, retrieval, segmentation, annotation, shot boundary detection, key-frame extraction, mining, etc
2011	Murtaza et al. [56]	Multi-camera tracking	Distributed and decentralized multiple stationary camera based tracking methods
2012	Labrinidis et al. [5]	Big Data	Controversies and issues with Big Data Analytics and ways to debunk them
2012	Heidemann et al. [17]	Online social networks	The role of huge data generated by the social media platforms and their impact on the people over the globe
2013	Yadav et al. [3]	Big Data	Comprehensive review of the schemes and tools available to handle Big Data
2013	Louridas et al. [105]	Big Data	Provides an overview of the tools, libraries, and infrastructure available for Big Data Analytics
2013	Cristani et al. [29]	Human behavior surveillance	A socio-technical prospective of existing computer vision schemes are discussed and the scope for future developments are suggested
2014	Fan et al. [6]	Big Data	Focuses on the features of Big Data and their effect on the methods and computing infrastructures
2014	Kambatla et al. [107]	Big Data	Hardware infrastructure as well as software tools available to handle Big Data and scope for research in the domain
2014	Lyon [152]	Big Data for surveillance	Existing tools, speed of processing, extend of control of the data, risk associated with the massive data, strategies available to handle Big Data, and the ethical issues involved in Big surveillance video data
2014	Chen et al. [45]	Big Data	Challenges and opportunities in Big Data, with strategies for data deluge issues
2014	Leeson et al. [46]	Crowd analysis	Detailed discussion on the stand of current technology, gap in logistics and infrastructure facility, ways of data collection, processing and analysis.
2015	Gandomi et al. [4]	Big Data	Survey on the analytics of unstructured Big multimedia Data and need for new tools to handle them
2015	Sheng et al. [43]	Video delivery architecture	Development of architecture for networking, strategies for transmission, challenges incurred for the designing of such systems and ways to evaluate the schemes
2015	Belle et al. [83]	Big Data in health care	Use of signal, image, and genomics data generated in health care sector for the better medical data analytics applications
2015	Kehoe et al. [110]	Cloud robotics	Descriptive study on the Big Data, cloud computing, collective robot learning, and human computation strategies for cloud robotics

Table 2 (continued)

Year	Author	Survey on	Key Contributions
2015	Abaker et al. [106]	Big Data and cloud computing	The strengths of cloud computing environment and techniques and tools available for the analysis of Big Data in cloud platform are analyzed
2015	Ma et al. [126]	Big Data for remote sensing	Remote sensing Big Data, its properties, features, computing mechanisms, and the data-intensive computing issue incurred in handling such massive data
2015	Assunção et al. [108]	Big Data and cloud computing	Handling the Big Data in cloud and the gap in technology for cloud based analytics
2016	Fang et al. [90]	Big Data in health care	Big Data in computational health informatics, issues, opportunities and possible step by step (pipeline strategy) approach to handle it
2016	Chi et al. [75]	Big Data for remote sensing	Challenges and opportunities in remote sensing Big Data Analytics along with a couple of case studies
2016	Lou et al. [78]	Big Data in health care	Use of Big Data for public health informatics, bioinformatics, imaging informatics and clinical informatics applications, challenges and scopes in healthcare data analytics
2016	Ghosh [123]	Big Data	Description on Big Data, issues and future scopes in the development of tools and computing facilities
2016	Verma et al. [124]	Big multimedia data analytics	Analytics of social media multimedia data based on the content, and challenges in handling them. It also discusses the probable applications of the outcome.
2016	Bansal et al. [79]	Disease	Developments in the area of surveillance of infectious diseases and ways in which Big health Data can be utilized for the better health care of public
2017	Pan et al. [111]	Outdoor surveillance	The advantages and limitations of existing methodologies for outdoor surveillance are discussed in detail and possible research directions
2017	Khan et al. [109]	Big Data and cloud computing	Discusses the data analytics tools and mathematical approaches available for the Big Data Analytics. Presented the issues and challenges of cloud computing platform to handle Big Data
2017	Lee et al. [89]	Big Data in health care	Characteristics of medical Big Data, their special attributes and methods for efficient analysis of those data
2018	Xiang et al. [103]	Underwater vehicle control	Controlling of underwater vehicles using fuzzy logic and soft computing approaches, their usages and future scope
2018	Pouyanfar et al. [16]	Big multimedia data	Tools and infrastructure available to handle the Big multimedia data, the challenges still persisting, and the gap in technology to handle them are discussed, along with descriptive suggestions for scope of research in the domain
2018	Zhang et al. [47]	Crowd analysis	Physics-inspired algorithms used for crowd scene analysis and the future scope in the domain
2018	Pouyanfar et al. [113]	Deep learning	Multimedia data analysis methods using deep learning approach are discussed. Black-box models, unsupervised learning, online learning issues are elaborated with possible research directions for the development of futuristic technologies
2018	Sindagi et al. [114]	Crowd analysis	A comprehensive study on the schemes available using convolutional neural network for crowd analysis is presented and recent datasets available for evaluation purpose
2018	Zhang et al. [115]	Deep learning for Big Data	Deep learning models for the analysis of Big Data are presented and the existing challenges in the implementation of deep learning schemes to handle Big Data are discussed

Table 2 (continued)

Year	Author	Survey on	Key Contributions
2019	Helbing [117]	Big Data and deep learning	Ethical, societal, legal, and economical issues in the implementation of Big Data, deep learning, artificial intelligence, and manipulative technologies
2019	Vincent [26]	Deep learning for Big text data	Various aspects of conventional and deep learning based document analysis schemes are analyzed and the authors have given equal importance to both kind of schemes, and concluded that the emerging trend of deep learning based schemes although have a brighter prospectus, but conventional schemes will also sustain in future
2019	Babar et al. [27]	Urban data management system and role of Big Data	Analyzed various requirements of the data generated in urban area, and discussed the schemes, and available tools to handle the big urban data consisting of all possible formats of signal. It also emphasized the usability of Hadoop in handling the data management system
2016	Dubuisson, and Gonzales [141]	Dataset for visual tracking	Provided an extensive list of datasets available for the visual tracking applications; also listed out the issues and challenges involved in each dataset

have been thoroughly discussed in [106–108]. Different strategies and environments for Big Data analysis on the cloud have been discussed in [108]. Khan et al. [109] have discussed the cloud-based Big Data handling and analysis techniques for the scholarly resources. Data security is a major concern because the cloud storage is not hundred percent secure to attacks and if the data is too personal, then the risk is also high. Data accessibility is not possible without the internet connection. Thus if the end user is in an area where internet access is not available, then data cannot be accessed. Big Data does not use the local storage devices; rather it uses the cloud-based distributed storage technique.

Kehoe et al. [110], have analyzed the research trends in cloud robotics and automation. The amount of video data generated by the cloud-connected robots is the key concern for the efficient analysis of the robotic activities. Big Data Analytics provides the end user, the most refined information to take a decision, whereas the cloud computing provides an underlying infrastructure which clean, integrate and represent the raw data in a suitable manner for data analysis. It then provides the basic platform required for the analysis of the refined data and brings it to a suitable form for data visualization and interpretation, which gives the end user most accurate and requisite information to take the decision. Recently Pan et al. [111] have proposed a visual attention feature (VAF) based visual tracking scheme using cloud platform.

6 Technology evolutions in Big Data analysis

There are many technologies and platforms discovered by companies for their purposes. Many new technologies are also being built and research is going on in many companies to get better and better technologies to serve their needs and handle Big Data.

6.1 Programming with Hadoop

In recent years, with the emergence of low cost new technologies the process of data capture, data storage, and data analysis have improved to a great extent. Capturing data from multiple sources and types (blogs, social media feeds, audio and video files) is now very much simple for organizations. Optimal storage of data and its processing have enriched and emerged dramatically. Technologies like MapReduce, and in-memory computing provide highly optimized capabilities for different business purposes [21].

Distributed computing can be carried out in a framework called Hadoop, which provides its own distributed file system, known as the Hadoop Distributed File System (HDFS) and open source libraries for the MapReduce software. This framework is developed in such a way that it is capable of handling a few computing nodes as well as to thousands of machines, each offering local computation and storage. Besides all these advantages, the most important strength of Hadoop is that it is designed to run on servers or personal computers, and exhibits a high degree of tolerance to hardware failure. The backbone of HDFS is a fault-tolerant (storing multiple copies) storage system that can store large volume of data, and withstand storage failure. The clusters of Hadoop are built with sophisticated but low-cost computing devices. In case of the failure of one computing device, the cluster continues to carry-on the operation without losing data. In such situations, the load of the failed device or node is re-distributed among the remaining machines in the cluster. HDFS takes care of the storage issue by breaking the files into small blocks and stores them in the cluster across multiple nodes. In the recent time, a number of Hadoop vendors are providing customized versions of Hadoop in the market [9].

6.2 Programming with R

The programming language 'R' [105] is actually developed from its predecessor 'S' [105] language, which was developed by Jhon Chambers and his colleagues at Bell Labs during the 1970s for programming with large datasets. The basic motivation of developing such a language was to provide a programming environment for the quantitative computation and visualization of large data. It aims to provide an exclusive and robust structure for the manipulation, statistical analysis, and visualization of large data. The inclusion of the graphics handling capability to the 'S' leads to the "S-Plus", variant which was made available for commercial usage by Bell Labs.

'R' language was developed by Ross Ithaka and Robert Gentleman at the University of Auckland during mid of the 1990s. 'R' is currently distributed under the GNU open software license. Applications of 'R' uses a specific package designed for the specific purpose. Thousands of such packages available are mostly written by the users. It is an integrated software, featured with data manipulation, computing and graphics handling capabilities. It has an efficient data handling and storage capabilities along with the features of 'S' programming language. It has a wide range of statistical and graphical tools, which includes linear as well as nonlinear modeling, time series analysis, classification, conventional statistical tests, clustering, etc. It gets extended by functions and packages. The advantages of 'R' is that it can be linked with 'C', 'C++', JAVA,. NET, Python and Fortran code in case of computationally complex tasks, and the 'R' objects directly can be altered as per requirement. It has advantages of object-oriented programming capabilities also.

6.3 Media server

Media Server [104] is one of the most powerful video analytics tools available in the current days. It is capable of recognizing and reading text in multimedia data. It can identify people, logo, any living or non-living objects. It can detect the shot boundaries as well as record the times at which these scene changes occurred in a video data. It also can detect, if a video is having audio in it or not. In case of a video having audio in it, this tool can sense the audio signals and convert them into text. Thus for content based classification of big multimedia data, media server is a very efficient tool. Moreover, media server can make the visual surveillance process monitoring and analysis task very efficient by automatically extracting useful information from the big surveillance video data. The key steps involved in the operation of a media server are: ingest, analyze, and encode. The ingest step deals with feeding multimedia data to the system, analysis step takes care of all the analysis tasks such as: key-frame extraction from a video, detection of a specified object or objects in a scene, recognize the objects, characters, patterns in a scene, classification of multimedia data, speech to text conversion, etc. In the encode step, the multimedia data is encoded using some source coding strategies to fit the data as per the specific requirement or application. Media server also can be utilized to identify duplicate files, content based similar files, etc. Customization of the tool is also possible with the help of the Lua scripting language [112]. Thus it can be observed that Media server is quite efficient in handling big multimedia data.

7 Deep learning for Big video Data analysis

Deep learning has been one of the most discussed approaches in the recent years. It deals with hierarchical feature learning, and characterized by many-layered artificial neural networks. The most popular deep learning schemes are based on multi-layered perceptron networks (MPN), convolutional neural networks (CNN), recurrent neural networks (RNN) and their variants. The use of deep learning to handle big video data has attracted many computer vision and data science researchers [113–117].

Ahmad et al. [118] have proposed a data retrieval approach to compress the huge amount of deep features to binary codes by the help of Fourier decomposition, which reduces the computational cost during the retrieval task. In order to estimate the real-time crowd density, Li et al. [119] have used CNN along with pixel level information for high density crowd and CNN with texture level information for the low density crowd. Deep CNN based scheme has been proposed by Shamsolmoali et al. [120], to increase the resolution of video data for different surveillance applications. Deep learning is also employed in crowd scene analysis and abnormal behavior detection [121, 122]. Figure 7 shows few results of abnormal behavior detection by Xie et al. [121] for crowd scene surveillance.

8 Challenges and future scope

The amount of video data generated per day is increasing exponentially and it is predicted that the amount of surveillance video data will go up to 859 Petabytes, as per the prediction by a UK based information management company, IHS markit. Handling huge amount of data requires a proper integration of Big Data and cloud computing framework. The most demanding field of Big Data Analytics for video data analysis is the social networking sites and video content analysis. Video-based search engine optimization is getting popular. Video lectures are in high demand nowadays, but the problem in getting a particular video is still a challenging issue.

Video retrieval and annotation have tremendous scopes for research. Different challenges include the problem of storage optimization for a huge amount of video data, segmentation of video based on the user-defined criteria, search time optimization, data security in the cloud, integration of heterogeneous information based video data, indexing and processing of video data etc. Scalability, reliability, and efficacy of the strategies employed for handling video image analysis require to be explored. Cross-platform strategies need to be developed for the users' convenience. The amount of video data which contains some information is very less as compared to the amount of video data generated. Thus identifying the information containing data is a big challenge. The case of surveillance video data can be considered. The CCTV cameras are recording round the clock, thus producing a large amount of video data which is stored unnecessarily, increasing the storage complexity. If the video data is stored only when some important event is occurring then the storage problem can be reduced to a large extent. Machine vision and artificial intelligence methodologies can be used for accomplishing this [123–126].

Surveillance is a task of detecting an object or multiple objects in a scene and tracking their where about in the subsequent time instants. In order to keep track of the location of the object

Fig. 7 Normal and abnormal behavior of the population and their social interaction force flow diagram presented by Xie et al. [121]

or objects, the tracker should have knowledge of the uniqueness of the objects. The attributes which uniquely describe an object in a video are nothing but the visual features. Li et al. [127] have discussed in detail about the commonly used visual features for surveillance applications. The feature extraction schemes used for human activity recognition are discussed by Ali et al. [128]. Besides the commonly used features, KAZE [129], SHIFT [130]], SURF [51], features and different variants of them are some of the popular features used for visual surveillance applications in the recent time. Hore et al. [131] have elaborated about different features used for scene characterization. Multi-cue approaches are also quite popular in the literature [132]. Deep learning approaches are also getting popularity in the recent time for the representation of the object of concern in surveillance applications [133].

The use of neural networks [134], artificial intelligence [135] and deep learning [136] has been increasing in the recent time because of their capability to learn from examples and use the learning to take effective decision in subsequent time instances. Thus new approaches can be developed using deep learning strategies so as to handle Big video surveillance data. Besides the learning based schemes, conventional approaches using statistical features [137], Markov random field based image models [138–140] have proven effective in detecting moving objects in video for surveillance applications. Hence the amalgamation of the deep learning based approaches and conventional modeling based techniques can be encouraging for futuristic surveillance applications.

A huge number of datasets are available [141] for the purpose of testing and evaluating the performance of the schemes and strategies developed for video surveillance applications. Some of the established datasets are like MOT dataset [142], MILtrack dataset [143], CAVIAR dataset [144], TRECVID dataset [145], UCSD dataset [146], CDNET dataset [147], VOT dataset [148] etc. UWCD dataset [149], F4K dataset [150], and ReefVid dataset [151] are some of the underwater datasets available for research usages. Recently Gao et al. [36] have also contributed a dataset for cross-domain action recognition task.

Video surveillance becomes more challenging for distributed and collaborative sensing systems. The individual sensors get isolated in case of unavailability of pairwise or global reference information. In order to tackle such situations, manually systems that are consisting of few numbers of sensors can be fed with reference information. However, systems deployed with large number of sensors require the design of automated approaches [152, 153].

9 Conclusions

This article extensively presents the development of Big Data Analysis techniques for video surveillance. It explores different research opportunities which exist in the video data storage, handling and analysis domain. Use of Big Data analysis techniques in different sectors of information processing in general and in video image analysis, in particular, has been discussed. The challenges involved in the present day computing infrastructure have been discussed and different scopes for research and development have been identified. The growth of video data compels the improvement of the video processing technologies to efficiently handle the huge amount of data. Inversely, it is also very true that video technologies need advancement to handle such huge amount of data. With the exponential growth of video data, the efficiency and data handling capabilities of video analysis algorithms must be improvised with exponential pace, else, the excellent methodologies of the present day will certainly fail in near future. Thus a lot of challenges are there and the scope of Big Data for video analysis is going to be one of the most eventful research areas in the coming years.

References

1. Subudhi BN, Nanda PK, Ghosh A (2011) A change information based fast algorithm for video object detection and tracking. IEEE Trans on Cir and Syst for Vid Tech 21(7):993–1004
2. Kwon O, Lee N, Shin B (2014) Data quality management, data usage experience and acquisition intention of big data analytics. Int. J. of Info. Man. 34(3):387–394
3. Yadav C, Wang S, Kumar M (2013) Algorithm and approaches to handle large data- a survey. Int J of Comp Sci and Net 2(3):1–5
4. Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. Int J of Info Man 35:137–144
5. Labrinidis A, Jagadish HV (2012) Challenges and opportunities with big data. Proc of the VLDB Endowment 5(12):2032–2033
6. Fan J, Han F, Liu H (2014) Challenges of big data analysis. Nat Sci Rev 1(2):293–314
7. Cohen J, Dolan B, Dunlap M, Hellerstein JM, Welton C (2009) Mad skills: new analysis practices for big data. Very large databases conf, pp 1–6
8. Whitepaper: Cisco VNI forecast and methodology, 2015-2020, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html. Accessed 18 Jan 2018
9. Lin J, Ryaboy D (2014) Scaling big data mining infrastructure: the twitter experience. SIGKDD Explorations 14(2):6–19
10. Dean, J (2014) Big data, data mining, and machine learning. Wiley
11. Herodotou H, Lim H, Luo G, Borisov N, Dong L, Cetin FB, Babu S (2011) Starfish: a self-tuning system for big data analytics. 5th biennial Conf. On Inno. Data. Syst Res:261–272
12. Ghosh A, Subudhi BN, Ghosh S (2012) Object detection from videos captured by moving camera by fuzzy edge incorporated Markov random field and local histogram matching. IEEE Trans. on Cir. and Syst. for Vid. Tech. 22(8):1127–1135
13. Hu W, Xie N, Li L, Zeng X, Maybank S (2011) A survey on visual content-based video indexing and retrieval. IEEE Trans on Syst Man, and Cyb, Part C: Appl and Rev 41(6):797–819
14. Zhan B, Monekosso D, Remagnino P, Velastin S, Xu L-Q (2008) Crowd analysis: a survey. Mac Vis and Appl 19(5–6):345–357
15. Subudhi BN, Nanda PK, Ghosh A (2011) Entropy based region selection for moving object detection. Patt Recog Lett 32(15):2097–2108
16. Pouyanfar S, Yang Y, Chen SC, Shyu ML, Iyengar SS (2018) Multimedia big data analytics: a survey. ACM Comp. Sur. 51(1):10.1–10.34
17. Heidemann KMJ, Probst F (2012) Online social networks: a survey of a global phenomenon. Comput Netw 56(18):3866–3878
18. Ezaki, N., Bulacu, M., Schomaker, L (2004) Text detection from natural scene images: towards a system for visually impaired persons. 17th Int. Conf. On Patt. Recog., 2, 683–686
19. Saracoglu A, Alatan AA (2006) Automatic video text localization and recognition. IEEE 14th Sig Proc and Com Appl:1–4
20. Lin W, Jia S, Yang F, Takase K (2004) Topological navigation of mobile robot using ID tag and WEB camera. Int Conf on Intel Mech and Auto:644–649
21. Ayed AB, Halima MB, Alimi AM (2015) MapReduce based text detection in big data natural scene videos. Procedia Comp Sci 53:216–223
22. Opitz M, Diem M, Fiel S, Kleber F, Sablatnig R (2014) End-to-end text recognition using local ternary patterns, MSER and deep convolutional nets, 11th IAPR Int. Wor on Doc Ana Sys:186–190
23. Turki H, Ben Halima M, Alimi AM (2017) Text detection based on MSER and CNN features, 14th IAPR Int. Conf. On doc. Ana. And Recog., 949–954
24. Selmi Z, Ben Halima M, Alimi AM (2017) Deep learning system for automatic license plate detection and recognition, 14th IAPR Int. Conf. On doc. Ana. And Recog., 1132–1138
25. Shivakumara P, Tang D, Asadzadehkaljahi M, Lu T, Pal U, Hossein Anisi M (2018) CNN-RNN based method for license plate recognition. CAAI Trans on Intel Tech 3(3):169–175
26. Vincent N, Ogier JM (2019) Shall deep learning be the mandatory future of document analysis problems? Pat Recog 86:281–289
27. Babar M, Arif F, Jan MA, Tan Z, Khan F (2019) Urban data management system: towards big data analytics for internet of things based smart urban environment using customized Hadoop. Fut. Gen. Comp. Sys. 96:398–409
28. Ko T (2008) A survey on behavior analysis in video surveillance for homeland security applications. 37th IEEE App Im Pat Rec Work:1–8

29. Cristani M, Raghavendra R, Bue AD, Murino V (2013) Human behavior analysis in video surveillance: a social signal processing perspective. Neurocomputing. 100:86–97
30. Guo S, Luo H, Yong L (2015) A big data-based workers behavior observation in China metro construction. Procedia Eng 123:190–197
31. Zitouni MS, Dias J, Al-Mualla M, Bhaskar H (2015) Hierarchical crowd detection and representation for big data analytics in visual surveillance. IEEE Int Conf on Syst, Man, and Cyb:1827–1832
32. Gao Z, Zhang H, Xu GP, Xue YB, Hauptmann AG (2015) Multi-view discriminative and structured dictionary learning with group sparsity for human action recognition. Sig Pro 112:83–97
33. Pan Z, Liu S, Fu W (2017) A review of visual moving target tracking. Multi. Tools and App. 76(16): 16989–17018
34. Shao Z, Cai J, Wang Z (2018) Smart monitoring cameras driven intelligent processing to big surveillance video data. IEEE Trans on Big Data 4(1):105–116
35. Liu G, Liu S, Muhammad K, Sangaiah AK, Doctor F (2018) Object tracking in vary lighting conditions for fog based intelligent surveillance of public spaces. IEEE Access 6:29283–29296
36. Gao Z, Han TT, Zhu L, Zhang H, Wang Y (2018) Exploring the cross-domain action recognition problem by deep feature learning and cross-domain learning. IEEE Access 6:68989–69008
37. Ray KS, Chakraborty S (2019) Object detection by spatio-temporal analysis and tracking of the detected objects in a video with variable background. J of Vis Com and Im Rep 58:662–674
38. Jansohn C, Ulges A, Breuel TM (2009) Detecting pornographic video content by combining image features with motion information. In: 17th ACM Int. Conf. On multimedia, pp 601–604
39. Behrad A, Salehpour M, Ghaderian M, Saiedi M, Nasrollah Barati M (2012) Content-based obscene video recognition by combining 3D spatiotemporal and motion-based features. EURASIP J on Image and Vid Proc 23: 1–17
40. Zhu T, Phipps D, Pridgen A, Crandall JR, Wallach DS (2013) The velocity of censorship: high-fidelity detection of microblog post deletions. 22nd USENIX Conf. On security, 227–240
41. Cheng X, Mehrdad F, Ma X, Zhang C, Liu J (2014) Understanding the YouTube partners and their data: measurement and analysis. China Com 11(12):26–34
42. Wu J, Zhang Z, Hong Y, Wen Y (2015) Cloud radio access network (C-RAN): a primer. IEEE Netw 29(1):35–41
43. Sheng M, Han W, Huang C, Li J, Cui S (2015) Video delivery in heterogenous crans: architectures and strategies. IEEE Wireless Com 22(3):14–21
44. Ruiz M, Germán M, Contreras LM, Velasco L (2016) Big data-backed video distribution in the telecom cloud. Comp Com 84:1–11
45. Chen CLP, Zhang C-Y (2014) Data-intensive applications, challenges, techniques and technologies: a survey on big data. Info. Sci. 275:314–347
46. Leeson A, Pablo A, Ghosh S (2014) Understanding how big data and crowd movements will shape the cities of tomorrow. Euro Trans Conf:1–12
47. Zhang X, Yu Q, Yu H (2018) Physics inspired methods for crowd video surveillance and analysis: a survey. IEEE Access 6:66816–66830
48. Kajo I, Kamel N, Malik AS (2018) An adaptive block-based matching algorithm for crowd motion sequences. Multi Tools and App 77(2):1783–1809
49. Huang S, Li X, Zhang Z, Wu F, Gao S, Ji R, Han J (2018) Body structure aware deep crowd counting. IEEE Trans. on Im. Pro. 27(3):1049–1059
50. Shami M, Maqbool S, Sajid H, Ayaz Y, Cheung SCS (2018) People counting in dense crowd images using sparse head detections. IEEE trans. On Cir. And sys. For vid. In: Tech
51. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Comp Vis and Im Und 110(3):346–359
52. Ravanbakhsh M, Nabi M, Mousavi H, Sangineto E, Sebe N (2018) March) Plug-and-play cnn for crowd motion analysis: an application in abnormal event detection. IEEE Win Conf on App of Com Vis:1689–1698
53. Li Y (2018) A deep spatiotemporal perspective for understanding crowd behavior. IEEE Trans on Multi 20(12):3289–3297
54. Mandal B, Fajtl J, Argyriou V, Monekosso D, Remagnino P (2018) Deep residual network with subclass discriminant analysis for crowd behavior recognition. 25th IEEE Int. Conf. On Im. Pro., 938–942
55. Fleuret F, Berclaz J, Lengagne R, Fua P (2008) Multicamera people tracking with a probabilistic occupancy map. IEEE Trans on Patt Anal and Mach Intel 30(2):267–282
56. Murtaza T, Cavallaro A (2011) Distributed and decentralized multicamera tracking. IEEE Sig Proc Magazine 28(3):46–58
57. Gundecha P, Liu H (2012) Mining social media: a brief introduction. Tutorials in Operations Research 1(4):1–17
58. Chen Z, Liao W, Xu B, Liu H, Li Q, Li H, Xiao C, Zhang H, Li Y, Bao W, Yang D (2015) Object tracking over a multiple-camera network. IEEE Int Conf on Multi Big Data:276–279

59. Blat J, Evans A, Kim H, Imre E, Polok L, Ila V, Nikolaidis N, Zemcˇık P, Tefas A, Smrzˇ P, Hilton A, Pitas I (2016) Big data analysis for media production. Proc. of the IEEE. 104(11):2085–2113
60. Richards JA, Jia X (2006) Remote sensing digital image analysis: an introduction. Springer-Verlag, Berlin
61. Campbell JB, Wynne RH (2011) Introduction to remote sensing. The Guilford Press, New York
62. Lenhart D, Hinz S, Leitloff J, Stilla U (2008) Automatic traffic monitoring based on aerial image sequences. Patt Recog and Image Anal 18:400–405
63. Carrano C (2009) Ultra-scale vehicle tracking in low spatial resolution and low frame-rate overhead video. Proc. of SPIE. 7445, LLNL-CONF-413376
64. Presnar M, Raisanen A, Pogorzala D, Kerekes J, Rice A (2010) Dynamic scene generation, multimodal sensor design, and target tracking demonstration for hyperspectral/polarimetric performance-driven sensing. Proc of SPIE 7672:76720T
65. Palaniappan K, Bunyak F, Kumar P, Ersoy I, Jaeger S, Ganguli K, Haridas A, Fraser J, Rao R, Seetharaman G (2010) Efficient feature extraction and likelihood fusion for vehicle tracking in low frame rate airborne video. 13th Conf. On info. FUSION (FUSION), 1–8
66. Xiao J, Cheng H, Sawhney H, Han F (2010) Vehicle detection and tracking in wide field-of-view aerial video. IEEE Conf on Comp Vis and Patt Recog (CVPR) 679–684(2010)
67. Palaniappan K, Rao R, Seetharaman G (2011) Wide-area persistent airborne video: architecture and challenges. Distributed video sensor networks. B. Bhanu et al. Springer, London, pp 349–371
68. Skyland N (2012) Big data: what is NASA doing with big data today. Open. Gov. open-access article
69. Meng L, Kerekes JP (2012) Object tracking using high resolution satellite imagery. IEEE J. of Sel. Top. in Appl. Earth Observ. and Remote Sens. 5(1):146–152
70. Oliveira SF, Furlinger K, Kranzlmuller D (2012) Trends in computation, communication and storage and the consequences for data-intensive science. IEEE 14th Int. Conf. On high Perfor. Comp. And com. & IEEE 9th Int. Conf. On embed. Software and Syst. (HPCC-ICESS), 572–579
71. Bhattacharyya A (1943) On a measure of divergence between two statistical populations defined by their probability distributions. Bull of the Cal Math Soc 35:99–109
72. Subudhi BN, Bovolo F, Ghosh A, Bruzzone L (2014) Spatio-contextual fuzzy clustering with Markov random field model for change detection in remotely sensed images. Opt & Las Tech 57:284–292
73. Rathore MMU, Paul A, Ahmad A, Chen BW, Huang B, Ji W (2015) Real-time big data analytical architecture for remote sensing application. IEEE J. of Sel. Top. in Appl. Earth Observ. and Remote Sens. 8(10):4610–4621
74. Cavallaro G, Riedel M, Richerzhagen M, Benediktsson JA, Plaza A (2015) On understanding big data impacts in remotely sensed image classification using support vector machine methods. IEEE J of Sel Top in Appl Earth Observ and Remote Sens 8(10):4634–4646
75. Chi M, Plaza A, Benediktsson JA, Sun Z, Shen J, Zhu Y (2016) Big data for remote sensing: challenges and opportunities. Proc of the IEEE 104(11):2207–2219
76. Dean J, Ghemawat S (2008) MapReduce: simplified data processing on large clusters. ACM Com 51(1):107–113
77. Xiang W, Wang G, Pickering M, Zhang Y (2016) Big video data for light-field-based 3D telemedicine. IEEE Netw 30(3):30–38
78. Luo J, Wu M, Gopukumar D, Zhao Y (2016) Big data application in biomedical research and health care: a literature review. Biomed Infor Insights 8(1–10)
79. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C (2016) Big data for infectious disease surveillance and modeling. The J of Infect Diseases 214(4):S375–S379
80. Rangayyan RM (2004) Biomedical image analysis. CRC Press
81. Eberhardt R, Anantham D, Ernst A, Feller-Kopman D, Herth F (2007) Multimodality bronchoscopic diagnosis of peripheral lung lesions. Am J of Resp and Critical Care Med 176(1):36–41
82. Suinesiaputra A, Brett C, Pau M-G, Abram Y (2015) Big heart data: advancing health informatics through data sharing in cardiovascular imaging. IEEE J of Biomed and Health Infor 19(4):1283–1290
83. Belle, A., Thiagarajan, R., Reza Soroushmehr, S. M., Navidi, F., Beard, D. A., Najarian, K.: Big data analytics in healthcare. Biomed Res Int 2015(370194), 1–16 (2015).
84. Menze BH, Bjoern H, Leemput KV, Lashkari D, Weber M, Ayache N, Golland P (2010) A generative model for brain tumor segmentation in multi-modal images. Med Image Comp and Comp-Assist Inter–MICCAI:151–159
85. Young AA, Alejandro FF (2009) Computational cardiac atlases: from patient to population and back. Exp Physio 94(5):578–596
86. Manolis AJ, Eftichia C, Ioanna Z (2015) Modern diagnostic approach for the assessment of cardiac damage in hypertension: 3D, CT and MRI. Ass of Preclin Org Dam in Hyp:25–37
87. Liu J, Zhang Z, Wong DW, Xu Y, Yin F, Cheng J, Tan NM (2013) Automatic glaucoma diagnosis through medical imaging informatics. J of the American Med Infor Asso 20(6):1021–1027

88. Vallieres M, Freeman C, Skamene S, Issam El N (2015) A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys in Med and Bio 60(14):5471–5496
89. Lee CH, Yoon H (2017) Medical big data: promise and challenges. Kidney Res and Clin Pract 36(1):3–11
90. Fang R, Pouyanfar S, Yang Y, Chen SC, Iyengar SS (2016) Computational health informatics in the big data age: a survey. ACM Comp Sur 49(1):12.1–12.36
91. Meggitt DJ, Roderick DK, Cooke KP (1999) Advanced technologies for undersea surveillance of modern threats: riding the crest into the 21st century. OCEANS '99 MTS/IEEE 1:289–294
92. Minami M, Agbanhan J, Asakura T (1999) Manipulator visual servoing and tracking of fish using genetic algorithm. Ind Rob 26(4):278–289
93. Foresti GL, Gentili S (2000) A vision based system for object detection in underwater images. Int J of Patt Recog and Art Intel 14(2):167–188
94. Sehgal A, Kadarusman J, Fife LD (2004) TOUCH: a robotic vision system for underwater object tracking. IEEE Conf on Robo, Auto and Mech 1:455–460
95. Chuang MC, Hwang JN, Ye JH, Huang SC, Williams K (2017) Underwater fish tracking for moving cameras based on deformable multiple kernels. IEEE Trans on Syst, Man, and Cyb: Syst 7(9):2467–2477
96. Rout DK, Bhat PG, Veerakumar T, Subudhi BN, Chaudhury S (2017) A novel five-frame difference scheme for local change detection in underwater video. 4th IEEE Int. Conf. On Im. Info. Proc., 1–6
97. Mondal A, Ghosh S, Ghosh A (2017) Partially camouflaged object tracking using modified probabilistic neural network and fuzzy energy based active contour. Int J of Comp Vis 122(1):116–148
98. Rout DK, Subudhi BN, Veerakumar T, Chaudhury S (2018) Spatio-contextual Gaussian mixture model for local change detection in underwater video. Exp Sys With Appl 97:117–136
99. Palazzo S, Spampinato C, Giordano D (2014) Large scale data processing in ecology: a case study on long-term underwater video monitoring. 22nd Euromicro Int. Conf. On Paral., Distri., and net.-based proc., Torino, 312–316
100. Alharbi A, Reda AA, Hesham A, Sanguthevar R, Jun H (2014) Efficient pipeline architectures for underwater big data analytic. IEEE Int Sym on Sig Proc and Info Tech pp:161–166
101. Lebart K, Smith C, Trucco E, Lane DM (2003) Automatic indexing of underwater survey video: algorithm and benchmarking method. IEEE J. of Ocean. Eng. 28(4):673–686
102. Trucco E, Plakas K (2006) Video tracking: a concise survey. IEEE J of Ocean Eng 31(2):520–529
103. Xiang X, Yu C, Lapierre L, Zhang J, Zhang Q (2018) Survey on fuzzy-logic-based guidance and control of marine surface vehicles and underwater vehicles. Int J Fuzzy Syst 20:572–586
104. Hewlett Packard Enterprise website, Big Data solutions, available from: https://www.hpe.com/us/en/solutions/big-data.html. Accessed 1 Mar 2019
105. Louridas P, Ebert C (2013) Embedded analytics and statistics for big data. IEEE Softw 30(6):33–39
106. Abaker I, Hashem T, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU (2015) The rise of "big data" on cloud computing: review and open research issues. Info Syst 47:98–115
107. Kambatla K, Kollias G, Kumar V, Grama A (2014) Trends in big data analytics. J. of Paral. and Distri. Comp. 74(7):2561–2573
108. Assunção MD, Calheiros RN, Bianchi S, Netto MAS, Buyya R (2015) Big data computing and clouds: trends and future directions. J of Paral and Distri Comp 79–80:3–15
109. Khan S, Shakil KA, Alam M (2017) Big data computing using cloud-based technologies: challenges and future perspectives, networks of the future: architectures, technologies, and implementations, editors: Mahmoud Elkhodr, Qusay Hassan, Seyed Shahrestani, Chapman and Hall/CRC
110. Kehoe B, Patil S, Abbeel P, Goldberg K (2015) A survey of research on cloud robotics and automation. IEEE Trans on Auto Sci and Eng 12(2):398–409
111. Pan Z, Liu S, Sangaiah AK, Muhammad K (2018) Visual attention feature (VAF): a novel strategy for visual tracking based on cloud platform in intelligent surveillance systems. J of Par and Dist Comp 120:182–194
112. Lua software tool available at: https://www.lua.org/. Accessed 1 Mar 2019
113. Pouyanfar S, Sadiq S, Yan Y, Tian H, Tao Y, Reyes MP, Shyu ML, Chen SC, Iyengar SS (2018) A survey on deep learning: algorithms, techniques, and applications. ACM Comp. Sur. 51(5):92.1–92.36
114. Sindagi VA, Patel VM (2018) A survey of recent advances in cnn-based single image crowd counting and density estimation. Pat Rec Let 107:3–16
115. Zhang Q, Yang LT, Chen Z, Li P (2018) A survey on deep learning for big data. Info Fus 42:146–157
116. Pradhan R, Aygun RS, Maskey M, Ramachandran R, Cecil DJ (2018) Tropical cyclone intensity estimation using a deep convolutional neural network. IEEE Trans on Im Pro 27(2):692–702
117. Helbing D (2019) Societal, economic, ethical and legal challenges of the digital revolution: from big data to deep learning, artificial intelligence, and manipulative technologies. Tow Dig Enl 47-72
118. Ahmad J, Muhammad K, Lloret J, Baik SW (2018) Efficient conversion of deep features to compact binary codes using Fourier decomposition for multimedia big data. IEEE Trans on Ind Info 14(7):3205–3215

119. Li B, Han X, Wu D (2018) Real-time crowd density estimation based on convolutional neural networks. Int Conf on Intel Trans, Big Data & Smart City:690–694
120. Shamsolmoali P, Zareapoor M, Jain DK, Jain VK, Yang J (2018) Deep convolution network for surveillance records super-resolution. Multi. Tools and App.:1–15
121. Xie S, Zhang X, Cai J (2018) Video crowd detection and abnormal behavior model detection based on machine learning method. Neu Comp and App:1–10
122. Xu Y, Lu L, Xu Z, He J, Zhou J, Zhang C (2018) Dual-channel CNN for efficient abnormal behavior identification through crowd feature engineering. Mac Vis and App:1–14
123. Ghosh A (2016) Big Data and its Utility Consulting Ahead 10(1):52–69
124. Verma JP, Agrawal S, Patel B, Patel A (2016) Big data analytics: challenges and applications for text, audio, video, and social media data. Int J on Soft Comp, Art Intel and Appl 5(1):41–51
125. Ghosh A, Seiffert U, Jain L (2007) Evolutionary computation in bioinformatics. J of Intel and Fuzzy Syst 18(7):25–26
126. Ma Y, Wu H, Wang L, Huang B, Ranjan R, Zomaya A, Jie W (2015) Remote sensing big data computing: challenges and opportunities. Fut Gen Comp Sys 51:47–60
127. Li Y, Wang S, Tian Q, Ding X (2015) A survey of recent advances in visual feature detection. Neurocomputing 149:736–751
128. Ali HH, Moftah HM, Youssif AA (2018) Depth-based human activity recognition: a comparative perspective study on feature extraction. Fut Comp and Info J 3(1):51–67
129. Alcantarilla PF, Bartoli A, Davison AJ (2012) KAZE features. Euro Conf on Comp Vis:214–227
130. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int. J. of Comp. Vis. 60(2):91–110
131. Hore S, Chatterjee S, Chakraborty S, Shaw RK (2018) Analysis of different feature description algorithm in object recognition. Comp. Vis.: con., meth., tools, and Appl., IGI Global, 601–635
132. Sadeghian A, Alahi A, Savarese S (2017) Tracking the untrackable: learning to track multiple cues with long-term dependencies. IEEE Int Conf on Comp Vis:300–311
133. Li P, Wang D, Wang L, Lu H (2018) Deep visual tracking: review and experimental comparison. Pat. Recog. 76:323–338
134. Subudhi BN, Ghosh S, Ghosh A (2015) Application of Gibbs–Markov random field and Hopfield-type neural networks for detecting moving objects from video sequences captured by static camera. Soft Comp 19(10):2769–2781
135. Ghosh A, Chakraborty D, Law A (2018) Artificial intelligence in internet of things. CAAI Trans. on Intel. Tech. 3(4):208–218
136. Chakraborty D, Narayanan V, Ghosh A (2019) Integration of deep feature extraction and ensemble learning for outlier detection. Pat. Recog. 89:161–171
137. Subudhi BN, Ghosh S, Shiu SC, Ghosh A (2016) Statistical feature bag based background subtraction for local change detection. Info Sci 366:31–47
138. Subudhi BN, Ghosh S, Cho SB, Ghosh A (2016) Integration of fuzzy Markov random field and local information for separation of moving objects and shadows. Info. Sci. 331:15–31
139. Subudhi BN, Ghosh S, Nanda PK, Ghosh A (2017) Moving object detection using spatio-temporal multilayer compound Markov random field and histogram thresholding based change detection. Multi Tools and Appl 76(11):13511–13543
140. Subudhi BN, Ghosh S, Ghosh A (2017) Moving object detection using multi-layer Markov random field model. Pat Recog and Big Data:687–711
141. Dubuisson S, Gonzales C (2016) A survey of datasets for visual tracking. Mac. Vis. and Appl. 27(1):23–52
142. Multiple object tracking benchmark: http://motchallenge.net. Accessed 2 Mar 2019
143. MILtrack dataset: https://bbabenko.github.io/miltrack.html. Accessed 28 May 2019
144. CAVIAR test case scenarios: http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/. Accessed 2 Mar 2019
145. TRECVID homepage: http://www-nlpir.nist.gov/projects/trecvid/. Accessed 2 Mar 2019
146. UCSD pedestrian database (2009) http://www.svcl.ucsd.edu/projects/peoplecnt/index.htm. Accessed 2 Mar 2019
147. CDNET dataset: http://changedetection.net/. Accessed 2 Mar 2019
148. VOT Challenge dataset: http://www.votchallenge.net/. Accessed 2 Mar 2019
149. UWCD dataset: http://underwaterchangedetection.eu/. Accessed 2 Mar 2019
150. F4K dataset: http://f4k.dieei.unict.it/datasets/bkg_modeling/. Accessed 2 Mar 2019
151. ReefVid dataset: http://www.reefvid.org/. AccessNational Institute of Technology Goaed 2 Mar 2019
152. Lyon D (2014) Surveillance, Snowden, and Big Data: Capacities, consequences, critique. J of Big Data & Soc 1(2):1–12
153. Ruhe MHO, Dalaff C, Kuhne RD (2003) Traffic monitoring and traffic flow measurement by remote sensing systems. IEEE Intel Transport Syst 1:760–764

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Badri Narayan Subudhi : received B. E degree in Electronics and Telecommunication from Biju Patnaik University of Technology, Orissa, India, in 2004, and the M.Tech. in Electronics and System Communication from National Institute of Technology, Rourkela, India, in 2009. He obtained his PhD from Jadavpur University in year 2014 with the financial support from Council of Scientific and Industrial Research (CSIR) scheme at Machine Intelligence Unit, Indian Statistical Institute, Kolkata, India. He was a visiting scientist at University of Trento, Italy during Aug. 2010 to Feb 2011. Currently he is serving as an Assistant Professor at Indian Institute of Technology Jammu from 2017 onwards. Prior to this He served as an Assistant professor at National Institute of Technology Goa, India during 2014–2017. His research interests include Video Processing, Image Processing, Machine Learning, Pattern Recognition, and Remote Sensing Image Analysis.

Deepak Kumar Rout received B.E. degree in Electronics and Telecommunication and M.Tech. degree in Electronics and Communication engineering from Biju Patnaik University of Technology, Rourkela, India, in 2005 and 2009 respectively. He is currently a Research Scholar with the department of Electronics and Communication Engineering, National Institute of Technology Goa, India. Prior to this, he worked with C. V. Raman College of Engineering, Bhubaneswar, India at a capacity of Assistant Professor for 10 years. He is a member of IEEE since 2007 and TC affiliate member of IVMSP group of IEEE signal processing society since 2015. His current research interests include underwater image and video processing.

Ashish Ghosh is Head and Professor of the Machine Intelligence Unit at Indian Statistical Institute, Kolkata, India. He received the B.E. degree in Electronics and Telecommunication from the Jadavpur University, Kolkata, in 1987, and the M.Tech. and Ph.D. degrees in Computer Science from the Indian Statistical Institute, Kolkata, in 1989 and 1993, respectively. He received the prestigious and most coveted *Young Scientists* award in Engineering Sciences from the Indian National Science Academy in 1995; and in Computer Science from the Indian Science Congress Association in 1992. He has been selected as an *Associate* of the Indian Academy of Sciences, Bangalore, in 1997. He visited the Osaka Prefecture University, Japan, with a Post-doctoral fellowship during October 1995 to March 1997, and Hannan University, Japan as a *Visiting Faculty* during September to October, 1997 and September to October, 2004. He has also visited Hannan University, Japan, as *Visiting Professor* with a fellowship from Japan Society for Promotion of Sciences (JSPS) during February to April, 2005. During May 1999 he was at the Institute of Automation, Chinese Academy of Sciences, Beijing, with CIMPA (France) fellowship. He was at the German National Research Center for Information Technology, Germany, with a German Government (DFG) Fellowship during January to April, 2000, and at Aachen University, Germany in September 2010 with an European Commission Fellowship. During October to December, 2003 he was a *Visiting Professor* at the University of California, Los Angeles, and during December 2006 to January 2007 he was at the Computer Science Department of Yonsei University, South Korea. His visits to University of Trento and University of Palermo (Italy) during May to June 2004, March to April 2006, May to June 2007, 2008, 2009 and 2010 were in connection with collaborative international projects. He also visited various Universities/Academic Institutes and delivered lectures in different countries including Poland and The Netherlands. His research interests include *Pattern Recognition and Machine Learning, Data Mining, Image Analysis, Remotely Sensed Image Analysis, Video Image Analysis, Soft Computing, Fuzzy Sets and Uncertainty Analysis, Neural Networks, Evolutionary Computation and Bioinformatics*. He has already published more than 120 research papers in internationally reputed journals and referred conferences, has edited 8 books and is acting as a member of the editorial board of various international journals. He is a member of the founding team that established a *National Center for Soft Computing Research* at the Indian Statistical Institute, Kolkata, in 2004, with funding from the Department of Science and Technology (DST), Government of India, and at present is the In-Charge of the Center.