
Human Behavior Recognition in Surveillance Systems

Abstract

This paper is a literature review that focuses on detecting and analyzing the human behavior and its application in surveillance systems. Surveillance systems play a very important role in tracking and monitoring human behavior generally, this is one of the reasons why it has recently become a major interesting research topic. Human effort is not very effective in monitoring human behavior via surveillance systems as human operators miss a lot of information in different video frames.

Sometimes, it is difficult to differentiate a normal behavior from an abnormal behavior as what is normal in a context may be abnormal in another context. Ordinary surveillance systems on their own cannot classify human behavior into normal or abnormal behavior and for this reason, they require machine learning techniques. With machine learning techniques, surveillance systems can detect different objects, find anomalies in the monitor screen automatically and alert the necessary authority.

We aim in this paper to write a literature review article aimed to provide a foundation of knowledge on human behavior recognition and establish similarities and differences between previous studies by analyzing major formative work in the field of surveillance systems.

Introduction

Human action detection, motion tracking, scene modelling, and behavior understanding (human activity recognition and discovery of activity patterns) have received a lot of attention in the computer vision and machine learning communities (Popoola & Wang, 2012) due to the uprising of several fields that these tools can be implemented efficiently in.

Researchers have developed a large interest in the areas of computer vision and machine learning communities to analyze the problem of human behavior in video clips. This is because in the last decade, systems used to consist of a low-level/mid-level computer vision system (to detect and analyze moving objects such as humans or cars) and a higher level interpretation module (to categorize motion into behaviors, for example, a pointing gesture or a car turning right). However, there have been few efforts to understand human behaviors that have vast extent in time, particularly when they involve serious interactions between people (Oliver, Rosario, & Pentland, 2000). Therefore, our main goal of this paper is to focus on detecting and analyzing the human behavior and its application in surveillance systems.

Video surveillance systems have become more important in today's life, especially with the developments of computer science, communication technology, and internet engineering. Governments install video surveillance systems that work 24/7 in major government buildings and important facilities (e.g. airports, borders, roads) to fight against crime and keep an eye on places where incidents such as theft and accidents are likely to occur (Mishra & Saroha, 2016). The increasing need of security and safety in different areas has pushed video surveillance systems to emerge from the first generation "dumb camera" surveillance that needs a human

eye to evaluate its images to the second generation “computer-linked camera” surveillance systems that evaluate its own video images and reduce the human factor in surveillance (Surette, 2005). It is difficult to monitor simultaneous events in surveillance displays due to many limitations, therefore efforts are being made to design intelligent surveillance systems that are capable of distinguishing between what is a normal or an abnormal behavior within the context, because a normal behavior in one context may be abnormal in another (Kale, Jawale, Dhende, Abak, & Kosht, 2019).

Intelligent surveillance systems can detect different objects and find anomalies in the monitor screen automatically. As a result, it provides security personnel with the quickest and best way to alert and provide useful information (Zhao et al., 2014). These intelligent surveillance systems use various machine learning techniques and methods in order to understand human actions/behaviors. However, understanding of human actions using machine learning is a complex, diverse, and challenging area that has received much attention within the past ten years. Applications have included video surveillance and human-computer interfaces (Popoola & Wang, 2012) in addition to many more. Previous surveys and the number of conferences, papers, and good surveys on detecting or tracking human motion and gestures as well as anomaly detection and behavior recognition or analysis show that it is a highly researched area.

Video surveillance generally involves acquiring and processing visual data from a scene, in order to detect target(s) along time and space so as to recognition suspicious behaviors in awkward situations and if possible, generate alarms. It begins with change detection and motion data capture for moving targets (using tracking or non-tracking methods), to enable successive high-level event analysis (Popoola & Wang, 2012).

As a result, we aim in this paper to write a literature review article aimed to provide a foundation of knowledge on human behavior recognition and establish similarities and differences between previous studies by analyzing major formative work in the field of surveillance systems in order to draw a conclusion on the best practices of recognizing human behavior in surveillance systems and the best machine learning method used.

Literature Review

The applications of security surveillance in content based video retrieval and annotation, human computer interaction, human fall detection, video summarization, robotics, etc., has made human behavior recognition an important area of research in computer vision. Video surveillance systems are employed in the monitoring and analyzing of human behavior and activities. The main aim of intelligent surveillance system is to detect suspicious human or object behavior (otherwise known as abnormal behavior) in certain scenes and provide real time notification to the relevant person (Kale, Jawale, Dhende, Abak, & Kosht, 2019). J (2015) defines an abnormal behavior to be one that deviates from usual known behavior. (Kale and Patil, 2016), states that human behavior recognition “is a systematic approach to understand and analyze the movement of people in a camera captured content”.

The analysis of human behavior with the use of surveillance systems involves a pre-processing phase that detects the foreground and the background, tracking, feature extraction and then a machine learning method for classifying behavior into normal behavior or abnormal behavior (Farzad & Asli, 2015). Machine learning, as a branch of artificial intelligence, enables machines

to perform given tasks skillfully with intelligent software (Mohammed, Khan, & Bashier, 2016).

Supervised learning: Supervised learning is basically concerned with labeled data. Hence, it is easy to apply its algorithms to pattern classification and data regression problems. It is used to discover mapping rules for predicting outputs of unknown inputs (Wang & Sng, 2015).

Unsupervised learning: Unlike supervised learning, unsupervised learning is concerned with unlabeled data. Unsupervised learning can be used to handle intra-class variation as it does not require labels of data. This learning algorithm deals with a form of network learning that gives correct response without any interference of an external behavior (Mishra & Saroha, 2016).

Semi-supervised Learning: Semi-supervised learning is a combination of both supervised learning and unsupervised learning. (He & He, 2018) proposed an end-to-end Unsupervised Multi-Object Detection framework for video surveillance, where a neural model learns to detect objects from each video frame by minimizing the image reconstruction error.

Clustering Algorithms: Clustering refers to the process of data gathering with identical features into groups. These groups are otherwise known as clusters (Al-Dhamari, Sudirman, & Mahmood, 2017). Abnormal detection with the aid of clusters is made possible using a clustering algorithm to analyze and then classify the data. (Lei, Jiang, Wu, & Du, 2016) further classified clustering algorithms into hierarchical clustering, partitioning clustering, model-based clustering, density-based clustering, grid-based clustering, etc.

Classification: A classification technique is used to manage or group of different objects individually. Object classification detect object and provide information that is needed for video analytics. It gives more detailed description of events or behaviors of objects. Object classification is useful for detection accuracy and creates high level metadata that describes another data in the video sequence (Mishra & Saroha, 2016).

Candamo, Shreve, Goldgof, Sapper, and Kasturi (2010) presented a survey research on human behavior recognition algorithms for surveillance in transit scenes with image processing techniques. The methods of detection Candamo et al. (2010) researched includes slow movement (individual), fighting (for multiple person) individual-vehicle recognition and object recognition. It was stated in their research that algorithms for human behavior recognition can aid in averting certain occurrences.

Popoola and Wang (2012) in their review on video based abnormal behavior recognition pointed out some methods of detecting abnormal human behavior using machine learning techniques.

Methodology

Selecting Research Papers

In this paper, we review research works in the area of detecting and analyzing human behavior together with its application in surveillance systems for a period of nineteen years (19) years, i.e., only research work done between 2000 and 2019 were selected. Information gathered was done by studying the abstract of different research works including journals, conferences materials, academic articles and surveys that have to do with - "human behavior", "intelligent

surveillance systems”, “abnormal behavior detection”, “machine learning in surveillance systems” and then extracting necessary information for our research purpose. Forty (40) research works were reviewed and are all documented in English language including journals from IEEE, Engineering and Technology International Journal of Computer and Information Engineering, Journal of Engineering and Applied Sciences, etc.

Result

Abnormal behaviors can sometimes be difficult to detect in video surveillance by a human observer (Sultani, Chen, & Shah, 2018), since there are no clear distinctions of abnormal behaviors from normal behaviors in certain circumstances especially in emergency situations like traffic accidents, crimes and terrorist attacks (Shao, Cai, & Wang, 2018). In other words, the abnormal behavior to be detected is determined by the defined normal behavior (Eisa & Moreira, 2017). Walking, running or sitting down are common normal behaviors that we exhibit as humans in our daily life (Chaaroui, Padilla-Lopez, Ferrandez-Pastor, Nieto-Hidalgo, & Florez-Revuelta, 2014). However, abnormal behaviors may include abandoned object detection, suspicious human behavior, and intruder detection (Kotkar & Sucharita, 2017).

Dimensions of Abnormal Human Behavior

We can classify abnormal human behavior recognition into two dimensions – individual abnormal behavior recognition and crowd abnormal behavior recognition (Dhiman & Vishwakarma, 2018).

- **Individual Abnormal Behavior Recognition:** In most cases, according to Dhiman & Vishwakarma (2018), the individual abnormal human behavior recognition is used in Ambient Assisted Living (AAL). AAL is a group of concepts, services, and products that combine new technologies with social environment in order to create better life quality specially for older people by using, for example, wearable sensors and health-monitoring devices that transmit data such as blood pressure to a person’s health care facility (e.g. hospital), updating the person’s medical records and providing early alerts to changes in health status. These surveillance systems are able to detect abnormal behavior in terms of health problems (Eisa & Moreira, 2017) for old people or health deterioration and life-threatening situations (Paula Lago & Roncancio, 2017). They are usually used in smart homes where wearable sensors or non-wearable sensors are employed. The wearable sensors fixed on the individual’s clothes or body and the non-wearable sensors are position in certain parts of the house (Debes, et al., 2016).

Meng, Miao, and Leung (2017) presented an Online Daily Habit Modeling and Anomaly Detection (ODHMAD) model with a list of defined activities to be detected and their various sensors. It provides online analysis of old person’s sensory behavior data for daily behavior detection.

- **Crowd Abnormal Behavior Recognition:** The presence of intelligent surveillance systems in the public is important for the safety of the general public (Zhao, Zhou, & Wei, 2012). It is difficult to detect abnormal behavior in a crowded environment due to the complexity of the image. The detection of abnormal crowd behavior depends solely on the motion characteristics, interactions, and spatial distribution of a crowd (Marsden, McGuinness,

Little, & O'Connor, 2016).

How it Works: Detecting Abnormal Behaviors

In order to detect behaviors (normal and abnormal), visual data need to be obtained from the surveillance system for the main purpose of recognizing suspicious behaviors and alert the authority (J, 2015). It starts by tracking targets (with the use of either tracking or non-tracking methods) when there is a change in motion for analysis. Detecting abnormal behavior involves the search of unexpected problem in a data pattern (Kotkar & Sucharita, 2017).

According to Al-Dhamari et al. (2017), sensors are major components for surveillance systems. They aid in sighting targets that exhibit abnormal behaviors. Visible Light Camera (VLC), Thermal Camera (TC), Forward Looking Infrared (FLIR) camera are examples of sensors used in surveillance.

Feature extraction is done to detect area of interest of human movement (Xu et al., 2018), after which the extracted features are then trained or tested with model and finally classified into normal or abnormal behavior.

Machine Learning Techniques

Some machine learning techniques that can be used to detect and classify abnormal human behaviors include:

- Support vector machine (Xu, Li, Fang, & Zhang, 2018) - Balasubramanian, Archana, Elangovan, and Abhaikumar (2010) proposed a system to detect abnormal behaviors by extracting skeleton features and inputting them into support vector machine to process and then classify the behaviors into normal or abnormal behaviors.
- Fuzzy clustering technique (Chen, Tian, Zeng, & Huang, Detecting Abnormal Behaviors in Surveillance Videos Based on Fuzzy Clustering and Multiple Auto-Encoders, 2015): From figure 3.3 below, Chen et al. (2015), presented a framework consisting of fuzzy clustering to distribute some trajectories and Auto-Encoders to classify behaviors.

The Auto-Encoder acquires similar structures of normal video sequences and does estimation of normal patterns.

- Convolutional Neural Network (CNN) - Tay, Connie, Ong, Goh, and Teh (2019) in their work presented a CNN-based method to detect abnormal behavior. The CNN learns the characteristics of different abnormal behaviors on its own and analyses the abnormal behavior of an individual, more than one individual and even crowds.
- Kang, Shin, and Shin (2010) proposed the use of Hierarchical Markov Model to identify human behavior states and an algorithm based on data acquired from state-change sensors to identify abnormal behavior in a smart home. There is an improved detection accuracy with this model as it was used to predict human behavior while cooking a meal.
- Recurrent Neural Networks (RNN) - Arifoglu and Bouchacia (2017) proposed a method for detecting anomaly in old people with dementia (a disease in which a person experiences a decline in mental ability) using recurrent neural networks. In their proposed method, they extract sensor based features after a dataset has been

segmented into frames and the RRN used to recognize already trained activities and encode behavioral patterns.

- Deep convolutional framework - Ko and Sim (2018) in their research proposed a unified framework which is based on deep convolutional framework to detect anomaly in images with RGB color scheme. The aim of the unified framework was to enhance recognition speed.

Applications of Surveillance Systems

Video surveillance systems have wide range of applications like traffic monitoring and human activity understanding (Harish, Swati, Sonali, & Sangmesh, 2015), hospital for monitoring patient, in industry and process plant to monitor the activity of the production line (Jadhav, Suryawanshi, & Jadhav, 2017). Rai, Husain, Maity, & Yadav (2018), says that governments generally use surveillance systems during investigations and for the protection the masses.

- Traffic Monitoring: Surveillance systems are found on very busy roads and junctures in major highways (Rai et al., 2018) mostly to monitor pedestrians (Kotkar & Sucharita, 2017), track vehicles, estimate vehicle speed and monitor highway activities. Video clips are collected from the surveillance and processed frame by frame with image processing algorithms. Results obtained from the detection and tracking results are used for vehicle counts and vehicle speed estimation (Park, Kim, Lee, Park, & Suh, 2017).
- Crime Prevention: Surveillance systems have been positioned majorly in urban regions to eliminate crime and reduce insecurity in the areas. Data obtained from the surveillance systems in the form of recorded voice and video clips are used as evidence and also to investigate committed crimes (Shao, Cai, & Wang, 2018). They are very useful in crime prevention and forensic.
- Hospitals: Surveillance systems are required to monitor psychiatric patients (Hsu, Chuang, Huang, Teng, & Lin, 2018), mange issues that include the abduction of infants (Dasic, Dasic, & Crvenkovic, 2017), detect disorders and diseases (Candas, et al., 2014).
- Industry: In industries, surveillance systems can be used to predict human error.

Drawbacks of the Video Surveillance in Abnormality Detection

- For crowd analysis Lv, Liu, Xu, & Zhou (2018) reports that there is difficulty in identifying new movements which results in emergencies.
- Surveillance systems can be complex sometimes. They do not inter-link local patterns (Bouindour, Snoussi, Mohamad, Tazi, & Wang, 2019).
- Occlusion problem: This refers to inter-object interference mostly in crowded scene(s). It produces poor quality of already available surveillance video clips (Charara, Jarkass, Sokhn, Mugellini, & Khaled, 2012).
- It is difficult to abstract vigorous descriptors and classify algorithms, which results in false alarms (Bouindour, Snoussi, Mohamad, Tazi, & Wang, 2019).
- The cost of storage of surveillance systems largely increases because video data increases in size (Shao, Cai, & Wang, 2018).

Conclusion

In this paper, we did a literature review on the existing knowledge of human behavior recognition and its applications in the field of surveillance systems. We introduced the most significant machine learning techniques and methods used in detecting and classifying abnormal human behaviors.

Obviously, based on our study in the previous sections, we can see that human behavior detection algorithms face several challenges such as:

- Labeled data for training and validation of models, employed by abnormal detection technique (e.g. the support vector machine technique), is not always available which causes a significant issue.
- The meaning of abnormality differs from an application to another. Also, applying the same technique into different fields is not always possible.
- Video clips that have complex scenes take more time to process while extracting features and abnormal events.

Having these challenges and many more shows that solving the abnormal behavior detection issue is not straightforward. Using a certain method in one environment/application does not mean that it will work well in other applications. Therefore, there is a need to always examine the performance of machine learning abnormal detection methods in order to improve its work within the wanted context.

Applying these methods in surveillance systems was also among the topics that we discussed throughout our paper. We firstly introduced surveillance systems first generations (dumb camera) that needs a human eye to evaluate its images. Then, we focus on the second generation “intelligent surveillance systems” that are capable of distinguishing between what is a normal or an abnormal behavior and can detect different objects automatically. As we mentioned in previous sections, the detection of abnormal behaviors using surveillance systems is used in several applications such as traffic monitoring and preventing crimes. Data obtained from intelligent surveillance systems experience pre-processing phase that detects the foreground and the background, feature extraction, and then one of the mentioned machine learning method is applied for classifying behaviors.

At the end, it is of note that more work should be devoted in order to build up real-time intelligence surveillance frameworks that can particularly overcome the scalability of video analysis especially in the real videos that contain a lot of moving objects, activities, and a crowd.