

A SURVEY ON ABNORMAL HUMAN OBJECT DETECTION IN VIDEO SURVEILLANCE SYSTEM

Vishnu Priya.P¹, Rajeswari.R²

¹ M. Phil Research Scholar, Department of Computer Applications, Bharathiar University, Coimbatore.

² Associate Professor, Department of Computer Applications, Bharathiar University, Coimbatore.

ABSTRACT

The anomaly detection is obtained using the smart video surveillance system. This survey mainly forces on the detecting the different human actions i.e. abnormal activities. At present, one of the biggest problem is to manage and store the videos manually and so, smart video surveillance plays a vital role in detecting actions and make alert through SMS and GSM. Sometimes it is necessary to store the videos based on the human actions and it also provides storage space. It is made possible by identifying the object from the video, and extract the features and classify what type of action it was. This survey deals about the different anomaly detections techniques that was used in past ten decades and given a brief explanation with references.

Keywords: Anomaly Detection, Model Creation, Feature Extraction, Classification, Dataset

1. INTRODUCTION

Smart Video Surveillance System (SVSS) has more scope in the digital world, which is widely used for activity recognition. Video surveillance is used in many applications such as surveillance environments, healthcare system, security purposes and traffic controls. SVSS's are available in many public places to monitor the activities against the public properties, and also used for many private/personal uses, for example in house, parking and garden. This system is used mainly for safety and security purposes and most important in health care facilities to monitor the elderly person's health. SVSS performs various tasks such as motion detection, human detection and counting, alarm activation, Short Service Message (SMS) notification through Global System for Mobile (GSM), along with techniques for optimising storage and advanced management features.

A large number of the security system works just inside a specific territory limit, for instance, Closed Circuit Television (CCTV), as a person need to see camera footage from control room. The current security systems against robbery are costly as a certain amount of cash must be paid to administration supplier. A lot of man power is needed to keep track of unusual human activities in the videos. In this paper, the video based technologies for human anomaly detection and techniques will be extensively reviewed and discussed. The sections are classified based on the process involved in surveillance system. The second section describes the literature survey on smart video surveillance system survey for past ten years. The third section describes the overview and importance of the background modelling and foreground extraction. The fourth section describes feature extraction which is considered as most important. The fifth section describes classification techniques that are used to classify the features extracted. The sixth section gives an overview of various datasets that are been used so far in human action recognition.

2. LITERATURE SURVEY

In this paper, an extensive literature survey is carried about various existing smart video surveillance systems and also about the other techniques involved in SVSS. Generally the human activity recognition is classified into several categories based on the functions. Figure 1. shows the video surveillance process. The first level is capturing the videos and pre-processing i.e. making the videos available to work efficiently with the help of operations such as background modelling and foreground extraction. Background modelling plays an important

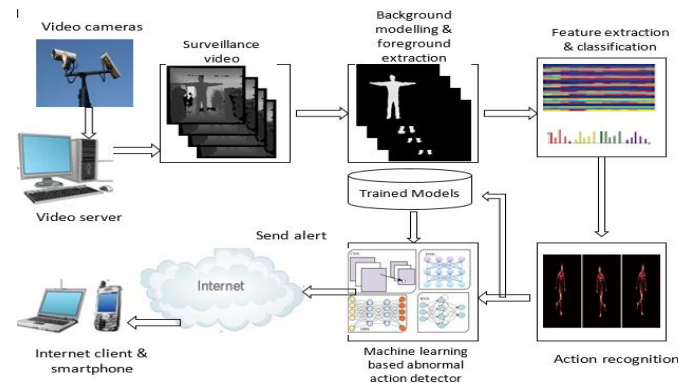


Figure-1: The overview of a general system for human activity recognition

Table-1: Comparison of various video surveillance systems

Year, Author(s)	Topic	Techniques used	Dataset used	Accuracy/time complexity
2011, Oluwatoyin et al.	Video-Based Abnormal Human Behaviour Recognition -A Review	Mean- Scale Invariant Features Transform (SIFT)	Benchmark datasets like UMN, UCSD PETS, CAVIAR	Reduce time complexity
2013, Li et al.	Anomaly Detection and Localization in Crowded Scenes	MDT- Mixture of Dynamic Textures	UCSD Anomaly UMN dataset Subway, U-turn	99.5% 99.5% 90.8% 95.2%
2014, Lglesias et al.	Analysis of network traffic features for anomaly detection	IPFIX measurements	NSL-KDD	80%
2015, Arshad et al.	Matlab Based Intelligent Video Surveillance System	Processed Histogram based face recognition algorithm	ORL database	---
2016, Kai Wang et al.	Research on healthy anomaly detection model based on deep learning	Multivariate Gaussain Distribution Conventional Neural Network-CNN	DEAP dataset	0.2 threshold value
2017, Aaron Chadha et al.	Video classification with CNN	Conventional Neural Network (CNN) - Long Short Term Memory (LSTM)	UCF HMDB	77.2% 48.0%
2017, Sun et al.	Automatic temporal segment detection vis bilateral LSTM-RNN	Support Vector Machine -SVM	FABO	95.30%
2018, Bo-Hao Chen et al.	A robust moving object detection in multi scenario bigdata for video surveillance	Short-term and life-time repeat consumption Model (SLRC)	Benchmark dataset CD net LASIESTA	Computational time 2.21%
2019, Divya Thakur et al.	An optimized CNN based real world anomaly detection in surveillance videos	CNN MSER-CNN	UCF Crime dataset	98.36% 95.06%
2019, Waqas Sultani et al.	Real world anomaly detection in surveillance video	C3D	UCF Crime dataset	94.0%

role in the SVSS, it is used in order to detect foreground objects like shadow removal, differencing incoming frames with the models. Next level is feature extraction to extract features based on movement of the object, gestures, colours, poses, and body motions. Based on the set of features extracted, classification of the activities is done by creating models. The classification of activities is carried out mostly by training a model. Classification techniques involve training and testing to classify the video whether it is the normal or abnormal. The last level is detecting the anomaly behaviour and then giving an alert by sending an SMS only when an unusual activity is detected.

A summary of the video surveillance system available in the literature spread over past ten years (2010-2020) is presented in Table 1. Some of the anomaly behaviours are abuse, road accident, burglary, fighting and assault. There are numerous research efforts for various applications based on human activity recognition [1], [2], [16], abnormal activities [3], vehicle detection [4], anomaly detection [5], [19], [21], object detection and tracking [7], [18], traffic controls [6], healthy anomaly detection [10].

3. BACKGROUND MODELLING AND FOREGROUND EXTRACTION

The aim of background modelling is to remove intrusion from the background and extract the foreground moving object, from the background in the sequence of video frames. Background modelling methods can be divided into two groups parametric and nonparametric methods given in table 2. Background modelling method are categorised as Pixel-based methods, Region-based methods and Hybrid method.

Table-2: Background modelling technique

Category	Pixel-based methods	Region-based methods	Hybrid method
Parametric	GMM, Oliver, AGMM	Russell, Heikkila	Huang, Tsai, Cristani
Non-parametric	Vibe, SACON, SOBS, Schick, Codebook, PBAS	KDE, Seki, Liu	Chen, Toyama

The most commonly used background modelling techniques are Gaussian mixture model (GMM), Mixture of Gaussian (MoG), Kernel Density Estimator (KDE), Visual Background Extractor (ViBe), Codebook (CB) [14]. GMM is one of the most classic methods and widely used for background modelling in cases of dynamic scenes, illumination changes, moving objects and shadows. MoG approach is more suitable for outdoor environment. The MoG method has low rate of complexity, it can handle multi-modal distributions along with its robustness. The purpose of MOG is to solve the shadow effect by extending the color video sequences. The advantage in adaptive mixture of Gaussian is lightning conditions are better in case of changes. The pixel value that is within a scaling factor of background standard deviation is considered as part of background. This can be determined by comparing the pixel value with Gaussian component tracking. Illumination will be a major drawback in MoG. ViBe model is a pixel based techniques mainly used for quick background subtraction in moving target object based on probability. The CB is the popular and fast background modelling technique to detect moving object in real-time video.

4. FEATURE EXTRACTION

Feature extraction is a process of dividing the raw data to a set of manageable groups by reducing the dimensions. It is the name for methods that select and/or combine variables into features, effectively detects the data redundancy and give accurate and complete dataset. Feature extraction is used to identify key features in the data from the old data set to the new one with valid key points. Feature extraction algorithms are used to detect features such as shapes, edges, or motion in a digital image or video. The feature extraction is the most required step in model creation. There are many feature detection techniques used in different fields such as robot navigation, pattern recognition, image and video detection. A basic survey on feature detection techniques Scale Invariant Features Transform (SIFT), Speed-up Robust Features (SURF), KAZE, Accelerated-KAZE (AKAZE), Oriented FAST and Rotated BRIEF (ORB), and Binary Robust Invariant Scalable Keypoints

(BRISK) are discussed in [17]. Some of the other feature extraction techniques are MSER, Features from accelerated segment test (FAST), Fast Retin keypoints (FREAK), Harris corner and hybrid methods.

Table-3: Comparison between feature extraction descriptors and time complexity

Types of descriptors	Feature type	Feature Detection Description	Time in ms.
FAST	Corner	No scale, rotation	0.0008
SURF	Blob	Scale and rotation changes	0.0110
KAZE	Blob	Scale and rotation	0.0059
BRISK	Corner	Scale, rotation, affine variants	0.0006
MSER	Region with uniform intensity	Scale and rotation	0.3824
ORB	Corner	Rotation only	0.0008
SIFT	Corner	Scale, rotation, affine variants	0.0068

5. CLASSIFICATION

The classification is also an important step used in SSVS. These techniques include SVM, LSTM, Conventional Neural Network (CNN), Gaussians Mixture Model (GMM), Re-current Neural Network (RNN). Hierarchical Dirichlet Process (HDP), Hidden Markov Model (HMM) and Convolutional neural networks (CNN), a popular technique, could be used in many fields such as image and video classification, natural language processing, pedestrian detection, generic visual recognition, face recognition, and image recognition.

Table-4: Learning of classification

methods and types techniques

Learning methods	Classification techniques	Reference
Supervised	HMM, Matrox Imaging Library (MIL), SVM, CNN, LSTM	15, 19, 21
Un-supervised	Latent Dirichlet allocation (LDA), HDP, GMM, DataBase-SCAN, Fisher kernel method, Practice swam optimization	20
Hybrid	HDP+HMM GAN+LSTM CNN+LSTM	16

In video classification the first step is extracting the frames from the given set of videos and it is the simplest way to deal with video data. The anomaly classification is categorized based on learning methods [4] and these methods are given in Table 4. Anomaly detection systems are used to overcome many difficulties in surveillance system and some of them explained [19]. The cheapest visual sensors with large quality of video

data is also increased. In [21] they have considered normal and anomalous video as bags and video segments as instance, which automatically study a deep anomaly ranking model that predicts high anomaly score for anomalous video segments. Due to the use of only machine learning techniques, the accuracy is very low and need to be improve by classifiers. 3D-CNN architecture for video classification is explained in [16], where they have used Motion Vector (MV) extraction and selective RGB decoding speed with lightweight CNN processing. CNN based video classification methods base their training on full frame video decoding and optical flow estimation.

6. DATASET

The dataset plays a vital role in constructing intelligence surveillance system. It is important to extract features from the given set of data. One challenge for supervised machine learning method is that they require sufficient labelled data for the model to learn, but it's required to collect datasets based abnormal events. And so benchmark datasets are used to analyse the performance of any action recognition method The detection of anomalous events in crowded scenes [5] can be evaluated in a few datasets. University of the California, San Diego (UCSD) contains videos of two pedestrian ped1 and ped2, which have various limitations in terms of size, saliency of the anomalies and evaluation criteria The University of Minnesota (UMN) dataset contains three escape scenes [5]. The Subway dataset consists of two sequences recorded from the entrance and exit [5]. The U-turn dataset with one video sequence recorded by a static camera overlooking the traffic at a road intersection [5]. A standard benchmark videos from data sets in CD-net and LASIESTA are used [18] for mulita scenario surveillance videos. Using CNN techniques [16] for UCF-101 dataset used to detect regular activities like walk, jump, play etc.

The most common dataset that have been used in anomaly detections in surveillance videos is University of Central Florida (UCF) Crime dataset [16, 19, 21]. It consists of long untrimmed surveillance videos which cover 13 real world anomalies, including Abuse, Arrest, Arson Assault, Road Accident, Burglary, Explosion, Fighting, Robbery, Shooting, Stealing, Shoplifting, and Vandalism. This dataset is used because the data's are mostly related to public safety. There is a need for real-life datasets to see the effectiveness of anomaly detection techniques

7. CONCLUSION

In this survey paper we have discussed about the different video surveillance system for past ten decades. We have discussed about the background model and also up-to-date feature extraction techniques. This survey also included classification techniques and describes the models as normal or abnormal. Finally we have discussed about the common benchmark datasets that are been commonly used and also some real-time dataset are also used according to the requirement. This survey helps to know about the techniques used in SVSS.

REFERENCES

- [1] Yannick Benezeth, Pierre-Marc Jodoin, "Comparative Study Of Background Subtraction Algorithms," *Journal Of Electronic Imaging*, 2010.
- [2] C. Lakshmi Devasena, R. Revathi, "Video Surveillance Systems – A Survey," *International Journal Of Computer Science Issues (IJCSI)*, 2011.
- [3] Oluwatoyin P. Popoola And Kejun Wang, "Video-Based Abnormal Human Behavior Recognition—A Review," *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, 2011.
- [4] A. A. Shafie, M. H. Ali, "Smart Video Surveillance System For Vehicle Detection And Traffic Flow Control," *Journal Of Engineering Science And Technology*, 2011.
- [5] Weixin Li, Vijay Mahadevan, "Anomaly Detection And Localization In Crowded Scenes," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 2013.
- [6] Félix Iglesias, Tanja Zseby, "Analysis Of Network Traffic Features For Anomaly Detection," 2014.
- [7] K. Kalirajan And M. Sudha, "Moving Object Detection For Video Surveillance," *The Scientific World Journal*, 2015.

- [8] Arun Hampapur, Lisa Brown, "Smart Surveillance: Applications, Technologies And Implications," IEEE Conference Paper, 2014.
- [9] Arshad Y. Maniyar, Ganesh K. Andurkar, "Matlab Based Intelligent Video Surveillance System," International Journal Of Advance Engineering And Research Development, 2015.
- [10] Kaiwang, Youjinzhao, "Research On Healthy Anomaly Detection Model Based On Deep Learning From Multiple Time-Series Physiological Signals," 2016.
- [11] Alessandro Pieropan, "Feature Descriptors For Tracking By Detection: A Benchmark," arXiv, 2016.
- [12] David J. Hill, Barbara S. Minsker, "Real-Time Bayesian Anomaly Detection For Environmental Sensor Data," Office Of Naval Research Under Grant
- [13] Zhuang Miao, Shan Zou, "Intelligent Video Surveillance System Based On Moving Object Detection And Tracking," International Conference On Information Engineering And Communications Technology (IECT) 2016.
- [14] Yong Xu, Jixiang Dong, "Background Modeling Methods In Video Analysis: A Review And Comparative Evaluation," CAAI Transactions On Intelligence Technology, 2016.
- [15] Bo Sun, Siming Cao, "Automatic Temporal Segment Detection Via Bilateral Long Shortterm Memory Recurrent Neural Networks", Journal Of Electric Imaging, 2017.
- [16] Aaron Chadha, Alhabib Abbas, "Video Classification With Cnns: Using The Codec As A Spatio-Temporal Activity Sensor," IEEE Transactions On Circuits And Systems For Video Technology, 2017.
- [17] Shaharyar Ahmed Khan, Zahra Saleem "A Comparative Analysis Of SIFT, SURF, KAZE, AKAZE, ORB, And BRISK," International Conference On Computing, Mathematics And Engineering Technologies, 2018.
- [18] Bo-Hao Chen, Ling-Feng Shi, Student And Xiao Ke, "A Robust Moving Object Detection In Multi-Scenario Big Data For Video Surveillance," IEEE Transactions On Circuits And Systems For Video Technology, 2018.
- [19] Divya Thakur, Rajdeep Kaur, "An Optimized CNN Based Real World Anomaly Detection In Surveillance Videos," International Journal Of Innovative Technology And Exploring Engineering (IJITEE), 2019.
- [20] Kavitha B R, Ramya G, "Performance Comparison Of Various Feature Descriptors In Object Category Detection Application Using SVM Classifier," 2019.
- [21] Waqas Sultani, Chen Chen, "Real-World Anomaly Detection In Surveillance Videos," 2019.