



A Review On State Of The Art Abnormal Activity Recognition Approaches

Devang Jani

Research Scholar, Ganpat University, Ganpat Vidyanagar, Mehsana. 384012,

State : Gujarat, Country: India.

devangjani645@gmail.com

Dr. Anand Mankodia

Associate Professor, EC, Ganpat University - U. V. Patel College of Engineering, Ganpat Vidyanagar, Mehsana. 384012,

State : Gujarat, Country: India

ABSTRACT

In last few decades, technological revolution has accelerated the deployment of large scale surveillance systems on almost all public places such as malls, hospitals, airports, railways, bus stations, roads, etc. These intelligent surveillance systems can play crucial role in governance of situations, collective security and safety, mitigating as well as prevention of adversaries. With gradual increase in multi camera surveillance systems enclosing multi angle views of same as well as different scenes has increased complexity of monitoring the systems by manual inspection. Abnormalities also known as anomalies or outliers are inevitable part of the existence and presumed to be rare in occurrence. Manual monitoring of such abnormalities is susceptible to errors and limited by human capabilities such as inattention and tiresome. Hence in the field of computer vision, automated abnormal activity recognition (AAR) from surveillance systems is emerging research area. The intent of this research is to shed a light on recent innovations and developments that have made a mark in abnormal activity recognition (AAR) involving deep learning. This paper also includes conventional categorization of anomalies based on different perspectives which can provide better understanding to young researchers. Though recent developments still poses many real time challenges in automatic abnormal activity recognition, some of them are enclosed in this paper.

Key words: Surveillance Systems, Adversaries, Abnormalities, Deep Learning, Anomalies, Activity Recognition

1. INTRODUCTION

Over last few decades, there is a growing hype for autonomous intelligent surveillance systems in the field of computer vision with vast application potential. Technological revolution has propelled the development of large scale surveillance systems deployed on various private/public places. The modern urban world is equipped with infrastructure which can play vital role in

governance, security, safety, risk management, prediction of events and prevention of adversaries etc. Mounted cameras collect data 24x7 continuously. In a way, surveillance systems helps to better understand environmental dynamics. Just as existence comprises both certain events based on known variables or patterns and uncertain events based on unknown variables or patterns, to function effectively intelligent systems should be able to grasp both the dimensions of existence.

Autonomy of such systems rely on ability to function in uncertainty. The behavior of events that occur in real time if can be confide to well established reference frames than those events would be called as norms or normal behavior. Events that occur but could not be correlated to established frame of reference could be termed as abnormal events or anomalous events. For example road accidents, vehicle on pedestrian crossing, theft, robbery, falling, riots, etc. There exists no single definition to accurately describe an anomaly. Due to underlying frame of reference out which events are normalized, the idea of abnormal events is ambiguous. Surveillance videos as sequence of images may or may not have recorded such events or behavior which can be labeled as abnormal or anomalous. Understanding and recognition of such abnormalities can help in dealing with harmful situations with its wide range of applications. Hence there is huge demand for abnormal activity recognition in the field of computer vision.

Abnormal activity recognition from surveillance system data poses many real time challenges vary from the defining the abnormal behavior. Abnormal activity recognition is a subpart of behavior understanding in computer vision. Commonly, abnormalities are studied based on popular ground assumptions. Three attributes are taken for consideration which are frequency, characteristics and semantics. For example, abnormalities are rare in occurrence. So majority of considered data must fall into non-abnormal category. Secondly, abnormalities are assumed to demonstrate different characteristics than well defined normal behavior. And at last, abnormalities should amount to a different but meaningful behavior. Abnormalities are relative in nature. They are always defined in relation with a reference frame a. k. a. context

[1]. Due to this nature, it is very difficult to come up with generalized framework which can recognize all abnormal behaviors. Generic single solution for all abnormal behavior still remains a far fetched idea in computer vision. Hence most of current research revolves around context specific recognition of abnormal activity. Due to this ambiguous nature of anomaly, comparison of different state of art approaches is very difficult and are done from the

perspective of data. Contextual choice of feature representation plays significant role in determining performance of abnormal activity recognition.

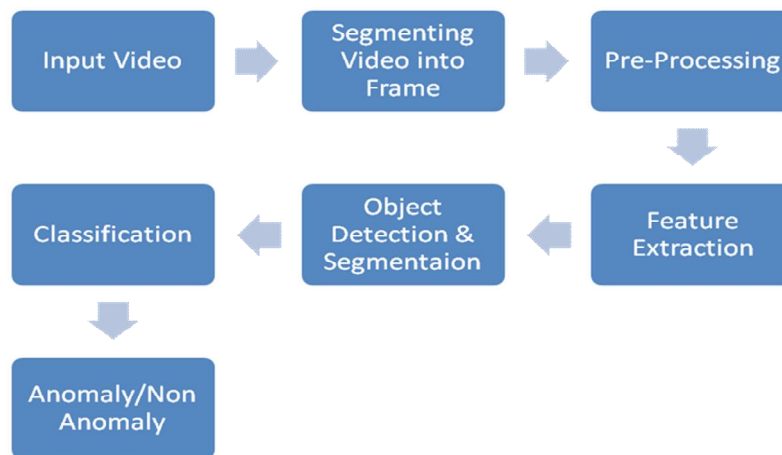


Figure 1: General Approach For Abnormal Activity Recognition

1.1 Classification of Abnormalities

Traditional Classification: Traditionally abnormalities are of three types. *Point Abnormalities* refer to distinctive difference from usual distribution of data. For example, vehicle moving when the signal is red. *Contextual Abnormalities* refer to same event can be normal and abnormal based on context. For example, fast moving vehicle in slow moving traffic which is abnormal but same event can be normal if there is less traffic. *Collective Abnormalities* occur when engaging entities in event correspond to normal behavior when considered individually, but show deviated behavior when grouped. For example, group of people suddenly leaving the area.

Classification Based on Coverage: Based on coverage in surveillance frames abnormalities can be classified as follows. *Local Abnormalities:* Local abnormalities are detected in specific region of frame which are commonly assumed to be temporal persistent. *Non Local Abnormalities:* Also known as global anomalies which can be found in entire frame.

1.2 Learning Approaches

Based on magnitude of guidance required for a model to learn, there are three types of learning approaches popularly known as (I) *Supervised Learning*, (II) *Semi Supervised Learning* and (III) *Unsupervised Learning*. In *supervised learning*, model learns about normal and abnormal behavior

by known labeled data. Supervised learning works well for limited environment where events are known but can be less efficient to learn unknown (unlabeled) normal and abnormal events. In *Unsupervised learning*, normal and abnormal behavior are grouped based on statistical methods and inter relationships. In *Semi supervised learning* combination of known labeled data as well as statistical methods are used to learn about events nature.

1.3 Abnormal Activity Recognition

In general abnormal activity recognition follows below steps as shown in Fig. 1. At first, surveillance videos are segmented into image frames and pre-processing is done to make data clean and uniform. Then from segmented data feature vectors are extracted. These feature vectors are dependent on application domain. These extracted features are then fed to object detection algorithms to identify object of interest for behavior modeling. From training data model learns are normal and abnormal object behavior which can be classified using binary or multi class classifier.

1.4 Related Work

In last two decades, a great amount of research has been done in this area. Table 1 summarizes recent research work that is done in abnormal activity recognition during last few years. We have primarily focused on implementation approach, context of application and novelty of work for our survey.

Table 1: Related Research Work In Abnormal Activity Recognition and Localization

Paper	Description	Context	Comments
Kelathodi Kumaran Santhosh , Debi Prosad Dogra, et.al (2019) [1]	A concise multi perspective critical analysis of recent research innovations was conducted	Road Traffic Scenarios	Aptly represents importance of context and dire need for long duration of relevant data in anomaly detection and tracking.
Amira Ben Mabrouk et.al (2017) [2]	Level based intelligent video surveillance system (IVVS) was presented. Merits and Demerits of different feature extraction as well as learning methods were discussed. Comparison between popularly used in abnormal activity recognition was presented with respective context.	Scene Density and Activity Type	Gives better understanding about choice of feature extraction method, learning model, contextual data and the trade off that comes along with it.
Nian Chi Tay , Tee Connie et.al (2019) [3]	CNN based abnormal behavior detection on 5 benchmark data-sets was demonstrated. Two parts where first part was binary classification into normal and abnormal events and then secondly 6 types of events were classified with high accuracy.	Events such as kicking, punching, pushing, hand shaking, pointing, hugging in indoor as well as outdoor environment	It shows how learning rate should not be too high for overshooting and should not be too low for over-fitting prevention and low convergence.
M. R. Sumalatha, P. Lakshmi Harika et.al (2019) [4]	Occasionally moving object and abnormal human behavior recognition in crowded area was presented. The proposed system used GMM for background modelling and HOG , Optical flow for abnormal activity detection.	Abnormalities in crowded areas	The proposed model achieved 96% in real time performance for PETS, ABODA, UMN data-sets
Jun Wang, Limin Xia (2018) [5]	The proposed method used two SDAEs to learn appearance and motion feature. Vision words and agglomerative information bottleneck approach was used to optimize extracted features. Finally sparse representation was used to detect abnormal events.	Non crowded events such as walk, follow, split, ignore, chase, fight grabbing newspaper, fighting, harassing, fainting etc	The proposed approach gives three benefits (I) Doesn't need lot of samples to learn and can work with small datasets, (ii) improved dense trajectory (iii) low computational cost. The proposed system achieved 0.982 AUC performance on BEHAVE and BOSS datasets
Kai Zhou, Tingting Wu et.al (2019) [6]	A new framework was introduced which was combination of skeleton spatio temporal correspondent CNN (ST-CNN) and attention based LSTM (ATT-LSTM) for abnormal human behavior recognition. 3D depth information of object skeleton recorded in 2D was used for recognition.	Eight types of human actions in indoor environment was considered where pushing, kicking , boxing was considered as abnormal where others were normal.	The proposed system achieved highest 84.2% accuracy on new constructed own dataset called "Top1" and 83.3% on SBU public dataset.
Rashmika Nawaratne, Damminda Alahakoon et.al (2019) [7]	Incremental Spatio-Temporal Learner (ISTL) for real time abnormal activity recognition and localization from surveillance videos was proposed. Active	UCSD Pedestrian dataset and CUHK Avenue dataset consisting outdoor events such as pedestrian on sidewalk, group walk,	Highlights limitation of most current researches that evolving nature of normal as well as abnormal events is unaddressed. Most research works are highly dependent on known

	<p>unsupervised deep learning approach with fuzzy aggregation to leverage evolutionary learning behaviors was realized. Two thresholds anomaly threshold and temporal thresholds were used along with sparse reconstruction error for model evaluation.</p>	<p>loitering, littering, walking on grass, walking towards camera, abandoned objects, skateboarding, cycling, vehicle movement, etc.</p>	<p>knowledge about normality and abnormality. The proposed approach comes close to evolutionary human cognition. Human observation was used for validation of detected anomalies. The proposed work was conducted on 36 Core CPU with 128GB RAM and Dual NVIDIA Quadro of 24GB GPU workstation for training, where as testing was done on 4-Core CPU with 2.6Ghz and 24GB RAM with NVIDIA GeForce GTX 970M PC. The proposed work achieved 91.1/8.9 AUC/EER performance on considered data.</p>
<p>Radu Tudor Ionescu, Fahad Shahbaz Khan et.al (2019) [8]</p>	<p>Novel approach where abnormal behavior recognition formulated as one versus rest binary classification problem. Object centric CAEs to learn about target appearance and motion was used. Then k- means clustering was used to detect types of normal events. Then supervised binary classification was used separate each normality from different clusters where events will be relatively normal and abnormal to each other. Negative highest classification score was considered was detecting anomaly.</p>	<p>Four benchmark datasets Avenue, Shanghai Tech, UCSD and UMN was considered. Public place events were observed and tested for model.</p>	<p>After clustering normalities, each event can act as normal and rest as abnormal events to that particular normal event in one versus rest classifier. It is helpful for training models where data has significantly low abnormal events. It constructs dummy abnormal events based on context of training. The proposed work showed better performance for the data-sets considered with compared state of art approaches. It was experimented with NVIDIA Titan Xp GPU with 12 GB RAM.</p>
<p>Ullah W., Ullah A. et.al (2020) [9]</p>	<p>A new framework based on deep CNN features and multi layer bi-directional LSTM (BD-LSTM) network for abnormal activity recognition was proposed.</p>	<p>UCF Crime and UCFCrime2Local data sets were used for experimentation covering fighting, explosion, abuse, accidents, burglary, arrest, robbery, assault, stealing, vandalism etc.</p>	<p>The proposed method showed improved results with 85% , 89.05% AUC on both data-sets. Framework could process video of one second length in half second. The proposed method is adaptable in evolving nature of normal and abnormal events. The proposed system has potential application in real time surveillance and monitoring. The experiment was conducted in Tensor Flow with GeForce Titan X Graphics environment.</p>
<p>S. Jothi Shri, S. Jothilakshmi (2019) [10]</p>	<p>Deep CNN based anomaly recognition and localization approach was presented where upon detection of anomalies alert sounds were fired. GSM module was used to send message to the controller of fired area of anomaly.</p>	<p>Crowd events scenes such as riots, fighting, protest, running, fire etc were enclosed for anomaly detection.</p>	<p>VGG16 and baseline architecture was used to perform training and testing of data. Alert sound system can help prevent harmful effects of anomalies and can help better govern the situation.</p>
<p>Damla Arifoglu, Abdelhamid Bouchachia (2017)</p>	<p>Abnormal human behavior recognition for old people with dementia was studied. Vanilla</p>	<p>Popular data-set from three households A, B, C was collected by Van Kasteren</p>	<p>The proposed method can be very helpful in detecting early signs of dementia by abnormal activity</p>

[11]	RNN, LSTM and GRU frameworks were implemented and compared against SVM, Naive Bayes, HMM, HSMM, CRF approaches.	was used for experimentation. Indoor single person based routine activities such as leave house, use toilet, brushing teeth, preparing dinner, eating, getting snack, take shower, prepare breakfast, prepare dinner, etc were covered.	recognition on physical repetitive movements of subject. The proposed work had considered three feature representations from sensor readings, that were binary, point change and last fired. The proposed work was conducted on Keras Deep Learning Library and Theanos. The results showed different best network for different feature representation methods. Overall LSTM performed slightly better as compared all state of art methods on all of datasets.
Waqas Sultani, Chen Chen et.al (2019) [12]	Anomaly detection using deep multiple instance ranking through weakly labeled data was proposed. New large scale dataset of untrimmed real world surveillance videos was introduced. The proposed work classifies anomaly detection as regression problem.	Public datasets such as UCSD Peds1, Peds2, UMN, Avenue, Subway Exit, Subway Entrance, BOSS covering various indoor and outdoor surveillance places events were considered for experiment. Newly introduced dataset of untrimmed videos covering 13 types of anomalies such as abuse, arrest, arson, assault, accident, burglary, Explosion, fighting, robbery, shooting, stealing, shoplifting, vandalism was also included.	The proposed method performed significantly better as compared binary SVM classifier, dictionary based approach and deep auto encoder based approach on all considered data-sets with reduced false alarm rates.

Kelathodi Kumaran Santhosh et.al presented concise multi perspective survey of research innovations in anomaly detection and tracking and highlighted how anomaly detection task is context specific in nature and what may be normal for one context can be abnormal for another. [1]

Amira Ben Mabrouk et.al covered comprehensive analysis of various intelligent systems based on feature representation and modeling highlighting current short comings in abnormal activity recognition. It gives insights into which feature extraction method will be robust against what circumstances and will be prone to errors in other circumstances. Similarly, comparison between models enhances understanding of limitations with respect to context and data. [2]

Nian Chi Tay, Tee Connie et.al implemented convolutional neural network based abnormal human behavior recognition on 5 different public benchmark datasets. A CNN with 6 layers in which 3 convolution layers, 2 fully connected layers and 1 softmax layer was configured. For feature extraction, edge, corner, text information was used. ReLU layer was used for faster training. The proposed work demonstrated importance of choosing learning rate for training model and its impact of accuracy [3]

M. R. Sumalatha, P. Lakshmi Harika et.al experimented with PETS, ABODA, UMN public datasets to mine real time abnormal behaviors in crowded places and experimental results showed 96% accuracy. The proposed experiment was conducted with GMM, HOG and Optical flow models. [4]

Jun Wang, Limin Xia presented new low cost method for abnormal activity detection where two SDAEs were used to learn about motion and appearance feature automatically. Vision words and agglomerative information bottleneck approach were used for reducing feature dimensions and compression. And abnormal activity detection was carried out using sparse reconstruction method. The proposed system showed 0.982 AUC efficiency on BEHAVE, BOSS public datasets. [5]

Kai Zhou, Tingting Wu et.al proposed new framework using skeleton based spatio temporal features which are immune to illumination, scale, occlusion trained on ST-CNN and attention based ATT-LSTM network for abnormal human behavior recognition. The new framework showed improvement in performance on public dataset and own skeleton based data-set “Top1” for indoor abnormalities. [6]

Rashmika Nawaratne, Daminda Alahakoon et.al did impressive work and proposed novel approach for real time

anomaly detection and localization. The proposed approach was known as Incremental Spatio Temporal Learner (ISTL) which was inspired by evolutionary human cognition by overcoming limitations of most current research works that is consideration of evolving nature of normal and abnormal behavior over time. The proposed work included three steps : First spatio-temporal autoencoder was used to learn about target appearance and motion, then CNN was used for learning spatial anomalies and Conv LSTM was used to learn about temporal anomalies as second step, and finally in third step active unsupervised deep learning with fuzzy aggregation was used to leverage evolving behavior of the model with respect to normal and abnormal events with marginal human observation for anomaly validation. The proposed work showed improved performance as compared to state or art. [7]

Radu Tudor Ionescu, Fahad Shahbaz Khan et.al proposed to formulate abnormal activity detection problem as one versus rest classification problem. In the cases of lack of data with sufficient abnormal events, relative contextual abnormality property can be helpful in training the model which was presented in proposed work. K-Means clustering to learn about types of normal events and then for each normality one versus rest binary classification was used to learn about abnormal events. Highest classification score was considered for labeling the behavior. Samples having negative score value labeled as abnormal. The proposed work showed better performance for benchmark data-sets consisting public place events with respect to other approaches. [8]

Ullah W., Ullah A. et.al demonstrated new multi layer bi-directional LSTM framework (BD-LSTM) for leveraging real time abnormal activity recognition. Two RNNs stacked together one for backward and one for forward pass was configured. The proposed method overcomes short coming of learning solely based on available training data. The results showed improved performance as compared to SVM, MIL, Optical flow, TSN, weakly supervised etc approaches on UCF Crime and UCFCrime2Local data sets. [9]

S. Jothi Shri, S. Jothilakshmi implemented deep CNN based anomaly recognition and localization along with GSM module to generate alert sounds and warn respective controllers. It was mainly focussed on public crowd events such as protests, running, fighting, firing etc. The proposed method achieved close to 90% accuracy. [10]

Damla Arifoglu, Abdelhamid Bouchachia proposed detecting early anomalous behavior in patients suffering from dementia. The proposed work implemented with three variants of RNN namely Vanilla, LSTM and GRU. Comparative results of experiments showed promising results for application in smart homes for patients with dementia. Overall LSTM proved to better as compared to other methods. [11]

Waqas Sultani, Chen Chen et.al thought of anomaly detection as regression problem and implemented solution based on deep multiple instance ranking through weakly labeled data. Newly collected large scale data sets covering 13 types of anomalous behavior was used for training along with publicly available data sets. The proposed work demonstrated increased accuracy in anomaly detection as compared to binary SVM, dictionary approach and deep auto-encoder based approach. [12]

2. CHALLENGES

Despite many notable works in last decade for abnormal activity recognition, anomaly detection task still has many real time challenges. Defining normal and abnormal behavior which can be context neutral is very difficult. The clear distinction between normal and abnormal activity is not well established, leading to more false alarms [1]. Multi contextual anomalies can overlap with each other in real time surveillance scenario and would be difficult to separate from each other which can undermine efficiency of model. There is still limited availability of large scale wide range anomaly data for training and validation. Notion of normal and abnormal behavior is not static, it evolve over time. Accuracy and robustness of anomaly detection methods are limited by feature representations which are robust to scene variations. Real time anomaly detection is limited by high computational and infrastructure cost. Optical illusions can distort reality of the same events for abnormal behavior recognition.

3. FUTURE SCOPE

Most existing work have made great progress with help of deep learning yet there is still room for research innovations. Current infrastructures are not yet affordable and suitable for real time anomaly detection as it requires very intensive computation and multi layer deep learning due to which real time processing gets delayed. It is observed that most existing works are done on known training data to learn about normal behavior and anything deviating from that considered as abnormal. Very less research is conducted based on evolving nature of normal and abnormal behavior to cover unknown normal and abnormal events. Very little work has been done with respect to anomaly prediction and prevention which can helpful for automatic handling of such events.

4. CONCLUSION

In this paper, what is anomaly and how its defined along with its classification was covered. How anomalies are context sensitive and how it affects choice of feature and limits generalization of abnormal activity recognition methods was discussed. A brief survey of various state of art abnormal activity recognition methodologies are studied. How deep learning helps in feature extraction and classification of events was explored. Despite significant research from last decade, there are still many challenges in

abnormal activity recognition such as scarcity of benchmark data sets, ambiguity of anomaly definition, real time analysis etc. Generic method for abnormal activity recognition is still a novel research area.

REFERENCES

- [1] Kelathodi Kumaran Santhosh , Debi Prosad Dogra, Partha Pratim Roy, Anomaly Detection in Road Traffic Using Visual Surveillance: A Survey, Preprint, 2019.
- [2] Amira Ben Mabrouk, Ezzeddine Zagrouba, Abnormal behavior recognition for intelligent video surveillance systems : a review, *Expert Systems With Applications* (2017), doi: 10.1016/j.eswa.2017.09.029.
- [3] Nian Chi Tay , Tee Connie et.al , A Robust Abnormal Behavior Detection Method Using Convolutional Neural Network, Springer Nature Singapore Pte Ltd., 1-11, 2019.
- [4] M. R. Sumalatha, P. Lakshmi Harika et.al , Detection of Static and Dynamic Abnormal Activities in Crowded Areas Using Hybrid Clustering , Springer Nature Switzerland AG, 1-10,2019.
- [5] Jun Wang, Limin Xia, Abnormal behavior detection in videos using deep learning, Springer Science+Business Media, LLC, part of Springer Nature, 1-11,2018.
- [6] Kai Zhou, Tingting Wu et.al, Skeleton Based Abnormal Behavior Recognition Using Spatio-Temporal Convolution and Attention-Based LSTM, International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2019) , 1-9, 2019.
- [7] Rashmika Nawaratne, Daminda Alahakoon et.al , Spatiotemporal Anomaly Detection using Deep Learning for Real-time Video Surveillance , TII-18-2964.R2, IEEE, 1-10, 2019. DOI 10.1109/TII.2019.2938527.
- [8] R. T. Ionescu, F. S. Khan, M. Georgescu and L. Shao, "Object-Centric Auto-Encoders and Dummy Anomalies for Abnormal Event Detection in Video," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 7834-7843, doi: 10.1109/CVPR.2019.00803.
- [9] Ullah, W., Ullah, A., Haq, I.U. et al. CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks. *Multimed Tools Appl* (2020). <https://doi.org/10.1007/s11042-020-09406-3>.
- [10] S. Jothi Shri, S. Jothilakshmi, Anomaly Detection in Video Events using Deep Learning, *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-9 July, 1-4, 2019.
- [11] Damla Arifoglu, Abdelhamid Bouchachia, Activity Recognition and Abnormal Behaviour Detection with Recurrent Neural Networks , The 14th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2017), 1-8, 2017.
- [12] Waqas Sultani, Chen Chen et.al, Real-world Anomaly Detection in Surveillance Videos, 1-10, 2019.