Anomaly event detection based on people trajectories for surveillance videos

Anonymous Author(s)
Affiliation
Address
email

Abstract

In this work, we propose a novel approach to detecting anomalous events in videos based on people movements, which are represented through time as trajectories. Given a video scenario, we collect trajectories of normal behavior using people pose estimation and employ a multi-tracking data association heuristic to smooth trajectories. We propose two distinct approaches to describe the trajectories, based on Convolutional Neural Network and Recurrent Neural Network. We use these models to describe all trajectories where anomalies are those that differ much from all normal trajectory. Experimental results show that our model is comparable with state-of-art methods and also support the idea of using trajectories to find other type of useful information, helping to understand people behavior, for instance the existence of rare trajectories.

1 Introduction

Abnormal event detection for video surveillance refers to the problem of finding patterns in sequences that do not conform to expected events [10]. It is a challenging problem because the definition of anomaly is subjective to the particular scene context, giving origin to a large number of possibilities. For instance, someone running at a marathon is a normal event, while someone running during a regular working day might be due to an emergency, an anomalous event. Therefore, the difficulty of anomaly recognition is related to the semantics that are observed in the scene.

Due to the success of Deep Neural Networks (DNN), researchers started to employ them to solve the anomaly recognition problem [12]. For instance, CNN-based approaches describe anomalies by creating models that combine optical flow and texture information from spatiotemporal regions [22]. Models that use AE or Convolutional AE (CAE) [20] aim at describing events in non-supervised fashion. Thus, anomalies are representations that differ from normal (i.e., an anomaly occurs when the AE is not able to perform a satisfactory reconstruction). Similar to AE, GAN-based approaches learn the normal behavior using a generative model [18], in which anomalies are recognized by the discriminator since the generator built an anomaly representation based in normal situations. Furthermore, RNN models usually appear accompanied with DNN, specially for movement data [6]. The idea is to combine the recurrent information of what is considered normal and create a representation of it. Nevertheless, most of these models depend on the camera position. Thus, these models learn specific patterns of the camera view which cannot be transferred to other views without retraining. Similarly to handcrafted features [9], these techniques also extract texture (appearance) and movement (flow) information. On the other hand, in our model, the source of information for anomaly representation is different. Specifically, our model extracts information from trajectories. An important difference with these models is the fact that our model is not affected by large color intensity changes.

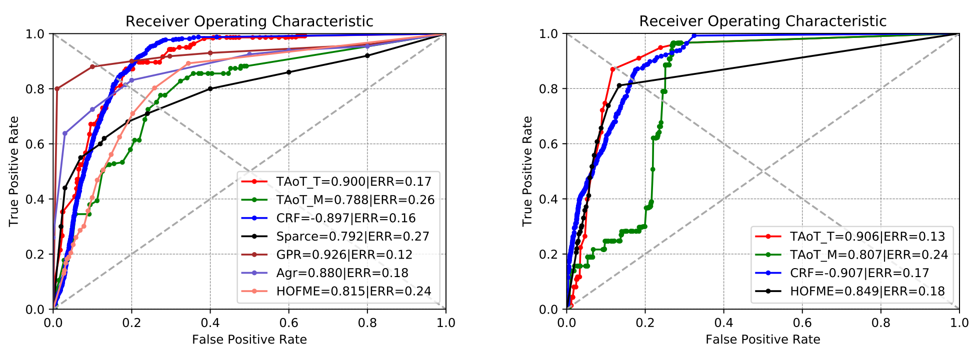
Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

Anomaly recognition models based on trajectories [25] are among first approaches in visual anomaly recognition. The main drawback of this model was the problem of people detection and trajectory building. However, with novel approaches and technologies, this issue has been progressively reduced. The model proposed by Cosar *et al.* [8] considers trajectories to build regions which are examined in a time lapse to find texture and movement information. The process is divided into two phases: description and filtering. Li *et al.*[13] proposed a technique that describes the scene using a sparse representation of overlapping trajectories, these trajectories are grouped and abnormal events are recognized when they differs much from any cluster. While Saini *et al.* [23] used trajectories to train a Hidden Markov Model (HMM) combined with genetic algorithm to detect anomalies by their low probability,the model proposed by Zhou *et al.* [29] developed a method based on HMM and feature clustering. An important difference between these approaches and ours is that our model does not segment the trajectories in parts or blocks, it focuses in complete trajectory. Furthermore, for surveillance purposes, region based models analyze motion characteristics, which are not meaningful without accurate localization of the targets. Thus, trajectories present the complete event that contains the anomaly.

In this work, we exploit high level information to create a robust representation for anomaly recognition. Our approach models people movements by leveraging from body skeletons obtained through a state-of-the-art pose estimator. The reference points are extracted from body skeleton and aggregated through time, building a trajectory. Each trajectory is then represented using deep neural networks to better encode its morphology. Our hypothesis is that trajectories are able to encode the necessary information from movement to recognize certain anomalous events. Since our proposed approach is based on trajectories, it is more robust to the aforementioned issues that affect movement and appearance approaches, because it, an advantage from using trajectories is that the localization of the particular individual performing an anomalous event is easily retrieved. In addition, trajectories allow other applications, such as people behavior analysis. We illustrate this application by using clustering models, such that it is possible to characterize the rarity of trajectories [28]. It is important to highlight that the proposed model is oriented to scenes where people detector and tracking algorithms may offer a good representation, thus, high crowds scenes are not considered in the scope of this research.

The novelty and contributions of this work are summarized as follows. (i) A spatial and temporal trajectory descriptor for anomaly event detection based on deep neural networks, aiming at describing trajectories by their morphology. (ii) A novel approach for anomaly recognition extracted from higher level information. (iii) A heuristic for multi-object tracking for data association based on Kalman filter. (iv) An experimental evaluation regarding trajectories and the relation between anomalies and rarity.

The pipeline of our proposed approach for anomaly recognition comprising four main steps: (i) reference point estimation, (ii) tracking building, (iii) feature extraction, and (iv) anomaly and rare trajectory recognition. Figures 1(a) and 1(b) present some of our experimental results in Subway dataset [1].



(a) (b)

Figure 1: Experimental results and comparison with the state-of-the-art on the *Entrance* and *Exit* sequences. (a) ROC results for Entrance clip; (b) results for the Exit clip.

References

[1] Amit Adam, Ehud Rivlin, Ilan Shimshoni, and Daviv Reinitz. Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 555–560, 2008.

[2] Margrit Betke and Zheng Wu. *Data Association for Multi-Object Visual Tracking*, volume 6. Synthesis Lectures on Computer Vision, 2016.

[3] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[4] Kai-wen Cheng, Yie-tarng Chen, and Wen-hsien Fang. Gaussian Process Regression-Based Video Anomaly Detection and Localization With Hierarchical Feature Representation. *IEEE Transactions on Image Processing*, pages 5288–5301, 2015.

[5] Kai Wen Cheng, Yie Tarng Chen, and Wen Hsien Fang. An efficient subsequence search for video anomaly detection and localization. *Multimedia Tools and Applications*, pages 15101–15122, 2016.

[6] Yong Shean Chong and Yong Haur Tay. Abnormal event detection in videos using spatiotemporal autoencoder. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pages 189–196, 2017.

[7] Junyoung Chung, Caglar Gulcehre, Kyunghyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning, December 2014*, pages 1–9, 2014.

[8] Serhan Cosar, Giuseppe Donatiello, Vania Bogorny, Carolina Garate, Luis Otavio Alvares, and Francois Bremond. Toward Abnormal Trajectory and Event Detection in Video Surveillance. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 683–695, 2017.

[9] Igor R. De Almeida, Vinicius J. Cassol, Norman I. Badler, Soraia Raupp Musse, and Claudio Rosito Jung. Detection of Global and Local Motion Changes in Human Crowds. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 603–612, 2017.

[10] Dawei Du, Honggang Qi, Qingming Huang, Wei Zeng, and Changhua Zhang. Abnormal event detection in crowded scenes based on Structural Multi-scale Motion Interrelated Patterns. *Proceedings - IEEE International Conference on Multimedia and Expo*, pages 1–8, 2013.

[11] Mahmudul Hasan, Jonghyun Choi, Jan Neumann, Amit K. Roy-Chowdhury, and Larry S. Davis. Learning Temporal Regularity in Video Sequences. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–31, 2016.

[12] B. Ravi Kiran, Dilip Mathew Thomas, and Ranjith Parakkal. An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. *J. Imaging*, 4:36, 2018.

[13] Ce Li, Zhenjun Han, Qixiang Ye, and Jianbin Jiao. Visual abnormal behavior detection based on trajectory sparse reconstruction analysis. *Neurocomputing*, pages 94–100, 2013.

[14] Weixin Li, Vijay Mahadevan, and Nuno Vasconcelos. Anomaly detection and localization in crowded scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 18–32, 2014.

[15] C. Lu, J. Shi, and J. Jia. Abnormal event detection at 150 fps in matlab. In *2013 IEEE International Conference on Computer Vision*, pages 2720–2727, 2013.

[16] Rensso Victor Hugo Mora Colque, Carlos Caetano, Matheus Toledo Lustosa De Andrade, and William Robson Schwartz. Histograms of Optical Flow Orientation and Magnitude and Entropy to Detect Anomalous Events in Videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 27:673–682, 2017.

[17] 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*, pages 865–878, 2012.

[18] Mahdyar Ravanbakhsh, Moin Nabi, Enver Sangineto, Lucio Marcenaro, Carlo Regazzoni, and Nicu Sebe. Abnormal event detection in videos using generative adversarial nets. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 1577–1581, 2017.

[19] Joseph Redmon and Ali Farhadi. Yolo9000: Better, faster, stronger. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525, 2017.

[20] Manassés Ribeiro, André Eugênio Lazzaretti, and Heitor Silvério Lopes. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recognition Letters*, pages 1–10, 2017.

[21] M. J. Roshtkhari and M. D. Levine. Online dominant and anomalous behavior detection in videos. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2611–2618, 2013.

[22] Mohammad Sabokrou, Mohsen Fayyaz, Mahmood Fathy, Zahra. Moayed, and Reinhard Klette. Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. *Computer Vision and Image Understanding*, 2018.

[23] Rajkumar Saini, Partha Pratim Roy, and Debi Prosad Dogra. A segmental HMM based trajectory classification using genetic algorithm. *Expert Systems with Applications*, pages 169–181, 2018.

[24] Venkatesh Saligrama and Zhu Chen. Video anomaly detection based on local statistical aggregates. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2112–2119, 2012.

[25] Xiaogang Wang, Keng Teck Ma, Gee-Wah Ng, and W.E.L. Grimson. Trajectory analysis and semantic region modeling using a nonparametric Bayesian model. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[26] Andrei Zaharescu and Richard Wildes. Anomalous behaviour detection using spatiotemporal oriented energies, subset inclusion histogram comparison and event-driven processing. In *Proceedings of the 11th European Conference on Computer Vision: Part I*, pages 563–576, 2010.

[27] Dengsheng Zhang and Guojun Lu. Review of shape representation and description techniques. *Pattern Recognition*, pages 1–19, 2004.

[28] B. Zhou, X. Wang, and X. Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2871–2878, 2012.

[29] Shifu Zhou, Wei Shen, Dan Zeng, and Zhijiang Zhang. Unusual event detection in crowded scenes by trajectory analysis. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 1300–1304, 2015.

[30] Haibin Zhu, Dongning Liu, Siqin Zhang, Yu Zhu, Luyao Teng, and Shaohua Teng. Solving the Many to Many assignment problem by improving the Kuhn-Munkres algorithm with backtracking. *Theoretical Computer Science*, pages 30–41, 2016.