Survey Paper on Anomaly Detection in Surveillance Videos

Prof. Aparna Kalaskar[1], Siddharth Pershetty[2], Vaishnavi Mote[3], Sahilsingh Rajput[4], Tejal Patare[5]

[1]Assistant Professor, Department of Computer Engineering, Sinhgad College of Engineering, Pune, Maharashtra, India
[2,3,4,5]B.E(Computer Engineering), Sinhgad College of Engineering, Pune, Maharashtra, India

---***---

Abstract - *Videos represent the primary source of information for surveillance applications and are available in large amounts but in most cases contain little or no annotation for supervised learning. Analysis of the Surveillance videos captured using these cameras can play effective roles in event prediction, online monitoring and goal-driven analysis applications including anomalies and intrusion detection. The monitoring capability of law enforcement agencies has not kept pace. That results in an unworkable ratio of cameras to human monitors. Nowadays, various Artificial Intelligence techniques have been used to detect anomalies, amongst them convolutional neural networks using deep multiple instance ranking framework technique by leveraging weakly labeled training videos, i.e. the training labels (anomalous or normal) are at video level instead of clip-level improved the detection accuracy significantly. This intelligent algorithm for automatic video anomaly detection alleviates the waste of labor and time. A robust anomaly detection method should have low false alarm rates on normal videos. In our approach, we consider normal and anomalous videos as bags and video segments as instances in multiple instance learning (MIL), and automatically learn a deep anomaly ranking model that predicts high anomaly scores for anomalous video segments. This reduces the low false alarm rate.*

Key Words: **Multiple Instance Learning, Surveillance Videos, Anomaly Detection, Convolutional Neural Network, Artificial Intelligence**

1. INTRODUCTION

In the Last Few Years, the Security and Safety concerns in public places and restricted areas have increased the need for visual surveillance. As in Today's World everyone wants to live in a Secure Environment. Visual analysis of suspicious events is a topic of great importance in video surveillance. A critical issue in anomaly analysis is to effectively represent an event to al-low for a robust discrimination. Large distributed networks of many high quality cameras have been deployed and producing an enormous amount of data every second. Monitoring and processing such huge information manually are infeasible in practical applications. As a result, it is imperative to develop autonomous systems that can identify, highlight, predict anomalous objects or events, and then help to make early interventions to prevent hazardous actions (e.g., fighting or a stranger dropping a suspicious case) or unexpected accidents. Video anomaly detection can also be widely-used in variety of applications such as restricted-area surveillance, traffic analysis, group activity detection, home security to name a few. The recent studies show that video anomaly detection has received considerable attention in the research community and become one of the essential problems in computer vision. However, deploying surveillance systems in real-world applications poses three main challenges: a) the easy availability of unlabelled data but lack of labelled training data; b) no explicit definition of anomaly in real-life video surveillance and c) expensive hand-crafted feature extraction exacerbated by the increasing complexity in videos .One of the challenges in analysing video data is objects detection in video frames. Also, video anomaly detection has been one of the controversial research topics within the recent years. In the last few years, deep learning approaches have also been introduced for the implementation of anomaly detection methods. In all anomaly detection approaches, learning is achieved solely through normal data. Another important point regarding the anomalies is that abnormal events are usually rare events that occur comparatively less than other normal incidents. As a result, developing a good anomaly detector to detect unknown anomalous objects is a very challenging problem.

2. METHODOLOGY

The proposed approach begins with dividing surveillance videos into a fixed number of segments during training. These segments make instances in a bag. Using both positive (anomalous) and negative (normal) bags, we train the anomaly detection model using the proposed deep MIL ranking loss[3].

1. Multiple Instance Learning[1]

In standard supervised classification problems using support vector machine, the labels of all positive and negative examples are available and the classifier is learned. To learn a robust classifier, accurate annotations of positive and negative examples are needed. In the context of supervised anomaly detection, a classifier needs temporal annotations of each segment in videos. However, obtaining temporal annotations [2] for videos is time consuming and laborious.

In MIL, precise temporal locations of anomalous events in videos are unknown. Instead, only video-level labels indicating the presence of an anomaly in the whole video is needed. A video containing anomalies is labeled as positive and a video without any anomaly is labeled as negative. Then, we represent a positive video as a positive bag Ba, where different temporal segments make individual instances in the bag, (p1, p2, . . . , pm), where m is the number of instances in the bag. We assume that at least one of these instances contains the anomaly. Similarly, the negative video is denoted by a negative bag, Bn, where temporal segments in this bag form negative instances (n1, n2, . . . , nm). In the negative bag, none of the instances contain an anomaly. Since the exact information (i.e. instance-level label) of the positive instances is unknown, one can optimize the objective function with respect to the maximum scored instance in each bag:

$$\min_{\mathbf{w}} \left[\frac{1}{z} \sum_{j=1}^{z} max(0, 1 - Y_{\mathcal{B}_j} (\max_{i \in \mathcal{B}_j} (\mathbf{w}.\phi(x_i)) - b)) \right] + \|\mathbf{w}\|^2$$

Where initial part is hinge loss, yi represents the label of each example, φ(x) denotes feature representation of an image patch or a video segment, b is a bias, k is the total number of training examples and w is the classifier to be learned YBj denotes bag-level label, z is the total number of bags[1].

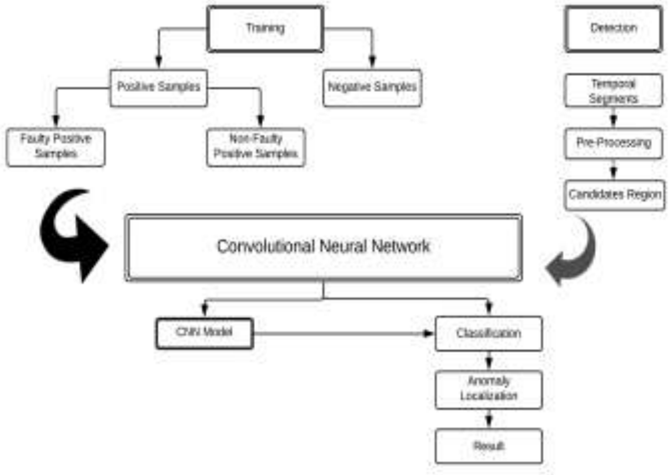

Fig- 1: Architecture Diagram of Anomaly Detection System

2. Deep MIL Ranking Model

In our proposed approach, we pose anomaly detection as a regression problem. We want the anomalous video segments to have higher anomaly scores than the normal segments. The straightforward approach would be to use a ranking loss which encourages high scores for anomalous video segments as compared to normal segments. We propose the following multiple instance ranking objective function:

$$\max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) > \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i),$$

where Va and Vn represent anomalous and normal video segments, f(Va) and f(Vn) represent the corresponding predicted scores, respectively. instance (anomalous segment). The segment corresponding to the highest anomaly score in the negative bag is the one looks most similar to an anomalous segment but actually is a normal instance. This negative instance is considered as a hard instance which may generate a false alarm in anomaly detection. We want to push the positive instances and negative instances far apart in terms of anomaly score[4].

There are two parameters that we have to keep in mind First, in real-world scenarios, anomaly often occurs only for a short time. In this case, the scores of the instances (segments) in the anomalous bag should be sparse, indicating only a few segments may contain the anomaly[8]. Second, since the video is a sequence of segments, the anomaly score should vary smoothly between video segments. Therefore, we enforce temporal smoothness between anomaly scores of temporally adjacent video segments by minimizing the difference of scores for adjacent video segments.[7] By incorporating the sparsity and smoothness constraints on the instance scores the loss function becomes:

$$l(\mathcal{B}_a, \mathcal{B}_n) = \max(0, 1 - \max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) + \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i))$$
$$+ \lambda_1 \overbrace{\sum_{i}^{(n-1)} (f(\mathcal{V}_a^i) - f(\mathcal{V}_a^{i+1}))^2}^{\textcircled{1}} + \lambda_2 \overbrace{\sum_{i}^{n} f(\mathcal{V}_a^i)}^{\textcircled{2}},$$

where 1 indicates the temporal smoothness term and 2 represents the sparsity term. In this MIL ranking loss, the error is back-propagated from the maximum scored video segments in both positive and negative bags [9].By training on a large number of positive and negative bags, we expect that the network will learn a generalized model to predict high scores for anomalous segments in positive bags.[Finally, our complete objective function is given by

$$\mathcal{L}(\mathcal{W}) = l(\mathcal{B}_a, \mathcal{B}_n) + \|\mathcal{W}\|_F ,$$

where W represents model weights.

3. CONCLUSION

To detect Real World Anomalies in Surveillance is a complex thing. Due to the complexity of these realistic anomalies, using only normal data alone may not be optimal for anomaly detection. We attempt to exploit both normal and anomalous videos. To avoid labor-intensive temporal annotations of anomalous segments in training videos, we learn a general model of anomaly detection using deep MIL framework with weakly labeled data. MIL method for anomaly detection achieves significant improvement on anomaly detection performance as compared to the state-of-the-art approaches. This paper introduced an overview of Anomaly Detection in standard Surveillance Videos. Low Anomaly Recognition Rate of Unknown Anomalies as a consequence of a Very Challenging Dataset Opens more opportunity for Future

Work for improving the Recognition rate of Unknown Anomalies.

ACKNOWLEDGEMENT

We are profoundly grateful to Prof. A. S. Kalaskar for her expert guidance and continuous encouragement throughout to see that this project on "Anomaly Detection in Surveillance Videos" rights its target since its commencement to its completion .We are also extremely grateful to our respected H.O.D. Prof. M. P. Wankhade for providing us with all the facilities and every help for smooth progress of our project. We would also like to thank the committee staff of our group for timely help and inspiration for completion of this project .At last we would like to thank all the unseen authors of various articles on the Internet, helping us to become aware of the research currently ongoing in this field and all other staff of Department of Computer Engineering for providing help and support in our work.

REFERENCES

[1] Sultani, Waqas & Chen, Chen & Shah, Mubarak. (2018). Real-World Anomaly Detection in Surveillance Videos. 6479-6488. 10.1109/CVPR.2018.00678.

[2] Cheng, Kai-Wen & Chen, Yie-Tarng & Fang, Wen-Hsien. (2015). Video Anomaly Detection and Localization Using Hierarchical Feature Representation and Gaussian Process Regression. 10.1109/CVPR.2015.7298909. S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," IEEE Electron Device Lett., vol. 20, pp. 569–571, Nov. 1999.

[3] B Ravi Kiran, Dilip Mathew Thomas, Ranjith Parakkal, "An overview of deep learning based methods for unsupervised and semisupervised anomaly detection in videos", MDPI Journal of Imaging,arXiv:1801.03149v1, 2018

[4] Wang, Siqi & Zhu, En & Yin, Jianping & Porikli, Fatih. (2017). Video Anomaly Detection and Localization by Local Motion based Joint Video Representation and OCELM. Neurocomputing. 277. 10.1016/j.neucom.2016.08.156.

[5] Chong, Yong Shean & Tay, Yong Haur. (2015). Modeling video-based anomaly detection using deep architectures: Challenges and possibilities. 1-8. 10.1109/ASCC.2015.7244871.

[6] Xiao, Tan & Zhang, Chao & Zha, Hongbin. (2015). Learning to Detect Anomalies in Surveillance Video. IEEE Signal Processing Letters. 22. 1477-1481. 10.1109/LSP.2015.2410031.

[7] Kaur, Phulpreet & Gangadharappa, Mandlem & Gautam, Shalu. (2018). An Overview of Anomaly Detection in Video Surveillance. 607-614. 10.1109/ICACCCN.2018.8748454.

[8] Vu, Hung & Nguyen, Tu & Phung, Dinh. (2018). Detection of Unknown Anomalies in Streaming Videos with Generative Energy-based Boltzmann Model

[9] Revathi, A. & Kumar, Dhananjay. (2016). An efficient system for anomaly detection using deep learning classifier. Signal, Image and Video Processing. 11. 10.1007/s11760-016-0935-0.