Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

R&D Project Proposal

Comparative evaluation of unsupervised methods for video anomaly detection

Amine Bahlouli

Supervised by

Paul Gerhard Plöger
Santosh Thoduka

May 2020

1 Introduction

Anomaly detection is an important topic that deals with identification of events that deviate from the usual behavior of data.The classical approach of anomaly detection is for every new data instance an anomaly score is computed then this score is compared to the average data instance and if the difference is bigger than a predefined threshold the data instance is then considered an outlier[9]. And because the development of technology and machine learning algorithms it has now become possible to apply it in video to detect events that do not correspond with the regular behavior. Detecting abnormal events is considered a challenging task because of many reasons: (I) The number of possibilities this task might include. (II) The variability associated with complex environment. (III) It is quite though to distinguish between normal events and abnormal events because there are infinite number of normal normal events while number of abnormal events is limited compared to normal events secondly sometimes normal events can have the same behaviors and patterns of abnormal events like social gathering, sports events. Video anomaly detection has attracted lots of research recently and several approach have been proposed. The first approach are models that are based on supervised learning, they often require labeling both normal and abnormal events, however this task is time consuming as it requires a massive labeling effort. Moreover, it is impractical to try to label all the normal and abnormal events that exists as it can be an endless process. Second approach is semi-supervised approach, it only requires labeling the normal event then any event that diverge from the normal event is considered an abnormal event, but it is hard to label all the normal event that are used for training. The third approach is unsupervised learning which does not require any labeling, this type of learning can detect anomalies including the ones that has never seen before.The importance of studying anomaly detection in video is:

- Identifying anomalies gives important, insightful information that can be used for variety of tasks[10].
- One of the tasks where anomaly detection in videos can be used is action recognition

- Action recognition is used in CCTV surveillance cameras to analyze videos frames and detect humans abnormal behaviors that can be considered a threat to environment or other people. An other factor, CCTV cameras require a great staff and such jobs are often repetitive and abnormal are events are a bit rare. Therefore there is a need to automate this process
- Even though visual anomaly detection are often used for surveillance however it is possible to use them in autonomous robots, for example self-driving car should predict the behavior of humans so that it can adjust its behavior. Same applies for robot when it is operating its environment if something goes wrong it can recognizes and stops before thing get worse
- Using Unsupervised learning approach reduces the labeling effort

1.1 Problem Statement

- The focus is to use unsupervised algorithms to detect anomalies
- The main objective of this research is to conduct a comparative study of unsupervised algorithms applied for video anomaly detection
- The categorization of algorithms will be based on the data set used and application and their learning approach
- The evaluation of the unsupervised algorithm will be on benchmark data based on categorization mentioned above
- The final output should be given a video frame the algorithm should be able to detect anomalies in frames if they exist

2 Related Work

2.1 Supervised Methods

- These methods classify event into normal ad abnormal events based on trained model where a labeled data is fed to the model and the model learn to classify events based on that.there were several supervised learning algorithms proposed.[1] Used Gaussian mixture model trained using expectation maximization to describe probability function of normal behavior pattern.While [2] has focused on constructing a video representation that enables anomaly detection using MRFs, LDA. However these methods might mask the visual representation, that's why representations based on dynamic textures (DTs) was proposed to model the complex scenes.To detect abnormal behavior [3] has used tracked spatio-temporal interest points instead of detection and segmentation, then SVM algorithm was used to find abnormal activities. All well trained supervised algorithms generally produce accurate results however they tend to fail to generalize to new unseen scenarios, as anything that is not included in training phase as normal behavior is considered abnormal

2.2 Unsupervised Methods

- Unsupervised learning extract features from training data without need to label data which makes it an interesting idea for anomaly detection as we need a model that is able to capture unseen situations.[7] Tried to model crowd motion using flow motion, under the assumption that people usually escape from a place where abnormal event happens, he proposes an algorithm that detects a divergent centers.[8] Partitioned each frame into several grids and in each grid features are extracted and modelled by an Online Weighted Clustering (OWC) algorithm.

2.3 Deep Learning Methods

- Deep learning have shown a great performance in computer vision tasks. [4] used 3D convolutional to capture regularities in dataset.[5] Has proposed Temporally-coherent Sparse Coding (TSC) which is mapped to sRNN to facilitate parameters optimization and accelerate anomaly detection. [6] Used Generative Adversarial Networks (GANs) to learn normality of crowd behavior then at test time they generate motion information and appearance to be compared with the real test frame. However such methods are often use a huge training data and a great computational power to train the model.[11] have used CONV-LSTM for video prediction, however what destinguished their work is using three three motion prediction modules, namely:dynamic neural advection (DNA), convolutional dynamic neural advection (CDNA), and spatial transformer predictors (STP. These three modules were the responsible for the prediction of physical interactions. Although their word has shown a great performance and managed to predict up to 10 time steps sequence however it has suffered from blurry predictions

2.4 Subsection 2

3 Project Plan

3.1 Work Packages

The bare minimum will include the following packages:

1 Literature Search: in this phase the main focus will be about reading all the literature related to video anomaly detection

2 comparative evaluation: This step includes artefact implementation and doing the necessary Comparative evaluation of selected algorithm performance on a one data set or two

3 Project Report: This phase includes the documentation of findings

3.2 Milestones

1 Literature search: This milestone will likely last until the end of semester due to the fact it is hard to start doing any coding now because of lack of time and because the topic is so broad and it consists of reading getting familiar with most famous unsupervised learning for anomaly detection in video

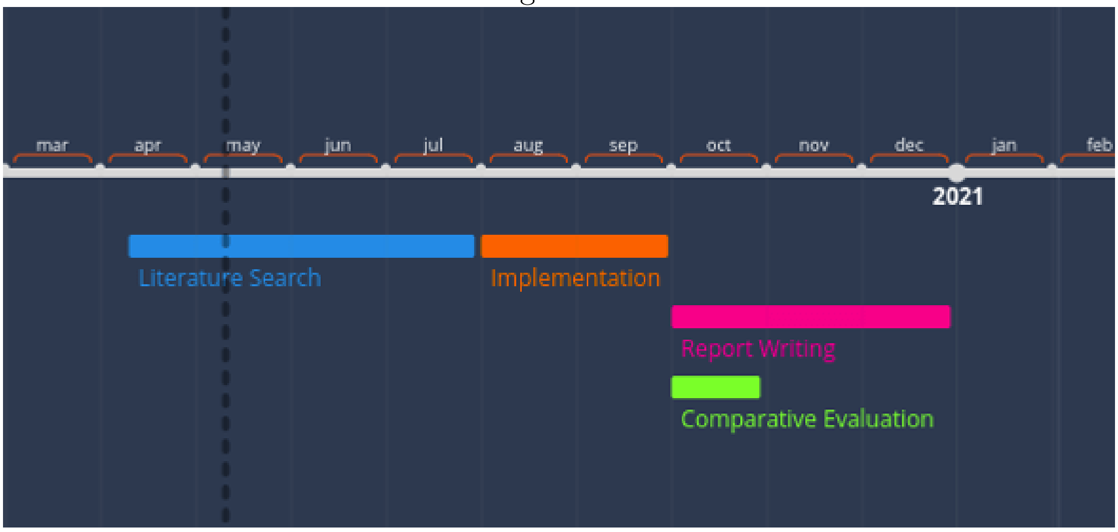
2 Implementation: This the phase is supposed to be the hardest as it requires first acquiring the basic requirement to start doing the actual coding then beginning the implementation process

3 Comparative Evaluation: This step starts after finishing the basic code implementation of all algorithms, in this phase comparing the algorithms will be the main focus by manipulating the hyperparameters

4 Report Writing: Report writing will be started first after finishing the literature search and choosing the algorithms by writing the theoretical part first

3.3 Project Schedule

Figure 1:



3.4 Deliverables

Minimum Viable

- Survey: Survey of the selected algorithms and their implementation
- Shallow comparison: It includes the implementation of algorithms with a superficial comparison between them without a deep explanation

Expected

- Comparison between state of the art algorithms: The expected final product is expected to be a full implementation and comparison of states of art in a different scenarios with various parameters

Desired

- Improving the state of the art: plus the expected output is some modification in algorithms architecture or parameters to achieve a better performance than the original algorithm performance in the

References

[1] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58, 2009.

[2] Chun-Yu Chen and Yu Shao. Crowd escape behavior detection and localization based on divergent centers. *IEEE Sensors Journal*, 15(4):2431–2439, 2014.

[3] Xinyi Cui, Qingshan Liu, Mingchen Gao, and Dimitris N Metaxas. Abnormal detection using interaction energy potentials. In *CVPR 2011*, pages 3161–3167. IEEE, 2011.

[4] Chelsea Finn, Ian Goodfellow, and Sergey Levine. Unsupervised learning for physical interaction through video prediction. In *Advances in neural information processing systems*, pages 64–72, 2016.

[5] Nico Görnitz, Marius Kloft, Konrad Rieck, and Ulf Brefeld. Toward supervised anomaly detection. *Journal of Artificial Intelligence Research*, 46:235–262, 2013.

[6] Mahmudul Hasan, Jonghyun Choi, Jan Neumann, Amit K Roy-Chowdhury, and Larry S Davis. Learning temporal regularity in video sequences. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 733–742, 2016.

[7] Weixin Luo, Wen Liu, and Shenghua Gao. A revisit of sparse coding based anomaly detection in stacked rnn framework. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 341–349, 2017.

[8] Ramin Mehran, Alexis Oyama, and Mubarak Shah. Abnormal crowd behavior detection using social force model. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 935–942. IEEE, 2009.

[9] 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. IEEE, 2017.

[10] Shandong Wu, Brian E Moore, and Mubarak Shah. Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2054–2060. IEEE, 2010.

[11] Yuan Yuan, Jianwu Fang, and Qi Wang. Online anomaly detection in crowd scenes via structure analysis. *IEEE transactions on cybernetics*, 45(3):548–561, 2014.