Anomaly Detection in Videos for Video Surveillance Applications usingNeural Networks

Ruben J Franklin, Mohana, Vidyashree Dabbagol
Department of Electronics & Telecommunication Engineering,
RV College of Engineering® Bengaluru- 560059, Karnataka, India.

Abstract— Security is always a main concern in every domain, due to a rise in crime rate in the crowded event or suspicious lonely areas. Abnormal detection and monitoring have major applications of computer vision to tackle various problems. Due to growing demand in the protection of safety, security and personal properties, the needs and deployment of video surveillance systems can recognize and interpret the scene and anomaly events play a vital role in intelligence monitoring. Anomaly detection is a technique used to distinguish various patterns and identify unusual patterns with a minimal period, this pattern is called outliers. Surveillance videos can capture a variety of realistic anomalies. Anomaly detection in video surveillance involves breaking down the whole process into three layers, which are video labelers, image processing, and activity detection. Hence, anomaly detection in videos for video surveillance application gives assured results in regards to real-time scenarios. In this paper, we anomaly was detected in images and videos with an accuracy of 98.5 %.

Keywords – Convolutional neural network, Common objects in context, Feature pyramids networks, Masked region convolutional neural network, Visual object classes.

I. INTRODUCTION

Anomaly detection is the identification of irregular, unexpected, unpredictable, unusual events or items which are not considered as a normally occurring event or a regular item in a pattern or items present in a dataset and thus different from existing patterns. An anomaly is a pattern that occurs differently from a set of standard patterns. Therefore, anomalies depend on the phenomenon of interest. Nowadays protection for personal and personal property becoming very important. Video surveillance gives a good role in real-time. Because of these needs deployment of cameras take place at every corner, video surveillance system understand the scene and it automatically detects abnormal activities [1]. The main aspect understands the action then reports the operator or users automatically when an unexpected event happens. Video surveillance performs efficiently improve in the application of safety and security for the management of personal life and public area [2]. It also develops an automatic surveillance system to replace human observed oriented services with a reduction in the workload of an observer. In the process of anomaly detection, cameras are used to collect the data of varies events representing the behavior of anomaly in an environment under surveillance. The system performance performs feature extraction on the data collected, process it and after that the resulting features become varies inputs to the specified algorithm. Mainly there are three anomaly detection techniques such as supervised, semi-supervised and unsupervised anomaly detection.

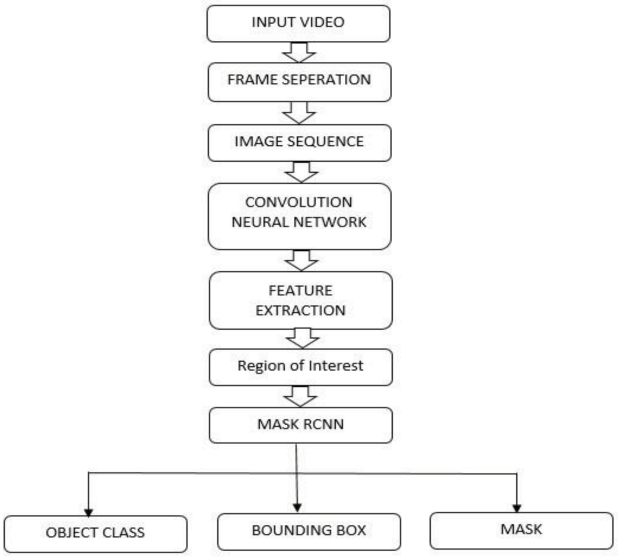


Fig.1. Block diagram of methodology.

Figure 1 depicts the block diagram of the methodology. Anamoly objects are detected for images of COCO dataset and as well as real-time video [3] [4]. Frames of a video extracted and objects are detected using CNN by extracting features in the required region of interest [5] [16]. Usually, the features are measured and compared based on considered patches such as motion and appearance. The obtained output in this step is feature representation, which is a very important aspect of anomaly detection. By using M-RCNN, three features are obtained such as a class of the object, bounded box and semantic segmentation [12] [13].

II. PERFORMANCE METRICS USED IN ANOMALY DETECTION

This section describes some of the performance metrics used in anomaly detection.

Eyeball evaluation: This the easiest way to judge the performance of a detector is by considering visual aspects of results, or anomaly scores. For example, first, make use of histogram graphs to plot the different distribution of scores. Then compare two graphs by overlaying histograms in the same plot, one representing "normal" points and another one as the "anomaly". In this respect, eyeballing are fantastic because they are simple to create and interpret, and contain all of this information. However, if the data structure is complex, a lot of data points lie on the same point interpreting the graphs becomes much harder. Moreover, if the inspector misses the data and interpretation is different from the result will be completely different from the actual results, performance goes below an acceptable level. The efficiency of performance is very low when compared to the rest of the different methods [14] [15].

Percentiles: Percentiles describe the distribution of scores between data points labeled "normal" and "anomaly". This parameter gives us a well-defined value by painting the whole picture in a single value but insight aspect or in detail, results are not well defined. This is similar to “Yes” or “NO” with what percentage. Percentiles are easily comprehensible metrics, with the advantage of being able to indicate high false positive or false negative rates. There are indeed interesting conclusions to draw if percentile calculations are applied to the scores of data points from both “normal” and “abnormal” classes.

Usual Companion (AUC): AUC is a popular metric that is designed to represent the ranking capability of scorers. To what extent can we say that "unusual" data points get a higher score than "usual" ones?" It is the well-known Area under the ROC Curve, or ROC-AUC—simply AUC from now on. It practically means that if we randomly pick a "usual" and "unusual" user activity from our test set, the "unusual" one will have a higher score with a probability that equals to AUC. This is essential since we want to make sure that abnormal events are highlighted with an anomaly score larger than that of normal events. The AUC value of an anomaly scorer's performance ranges from 0 to 1. An AUC of 1 indicates a flawless anomaly scorer that perfectly separates the two classes ("usual" and “unusual” events). If the AUC is below 1, that means that some "usual" events have larger scores than "unusual" ones do. Although not being perfect, an AUC of 0.7–0.9 is considered acceptable, on the other hand, if the AUC is 0.5, the algorithm is as effective a classified as random guessing AUC fails to encapsulate this aspect of performance of an anomaly scorer.

An extension to AUC, AUC (probabilistic AUC) method aims to address the above issue. AUC stands for probabilistic AUC. Consider a score of 100 as the best possible anomaly score that an "abnormal" event can obtain and of score 0 as the optimal anomaly score for every "normal" event. AUC gives probabilistic scores. This is because AUC is much stricter in its assessment of those test cases that yielded too many highly false negatives or false positives.

RP distance: Percentiles in performance evaluation, the simplicity in interpreting and the amount of power it provides to express in terms of percentiles inspired our performance metric that we call reverse percentile distance. This concept is based on a reinterpretation of this quality as of how far the scores obtained by the “normal” and “abnormal” are from each other. With the new metric, we attempt to capture this distance numerically. This main assumption is that the greater this distance is, the accurate is the scores of each to its respective scores The RP distance at a given percentile p, or RP@p, is defined as:

RP@p = perc (d abnormal, 100—p)—perc (d normal, p)

Where: d abnormal refers to the score distribution of data points labeled “unusual” d normal (d, p) refers to the pth percentile of a score distribution d [6] [7] [8].

III. DESIGN AND IMPLEMENTATION OF ANOMALY DETECTION

Anomaly detection is used to distinguish anomaly patterns that do not conform to a set of expected behavior. The anomaly detection algorithm is designed by breaking into three separate yet correlated processes they are the convolutional neural network, mask recurrent convolutional neural network and spatial awareness using semantic segmentation. The design process is implemented in the anaconda environment

A.Convolution Neural Network (CNN)

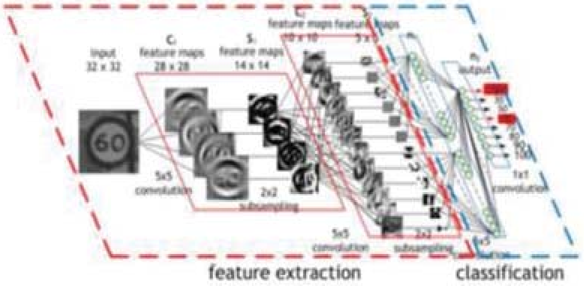


Fig.2.Layers of convo6lutional neural network

In anomaly detection, a Convolutional Neural Network (CNN or ConvNet) in figure 2, Layers of CNN is a class of Deep Neural Networks (DNN) to analyze images or videos. It is a class of multi-layer perceptron, refers to a Fully Connected Network (FCN) in which all neurons of one layer is connected to other neurons in the next layer [21] [26] [27][28].

B. Mask RCNN

Figure 3 shows mask RCNN structure, it is designed and built to solve instant segmentation problems in various commuter vision applications. It is capable to separate various objects in video or image.

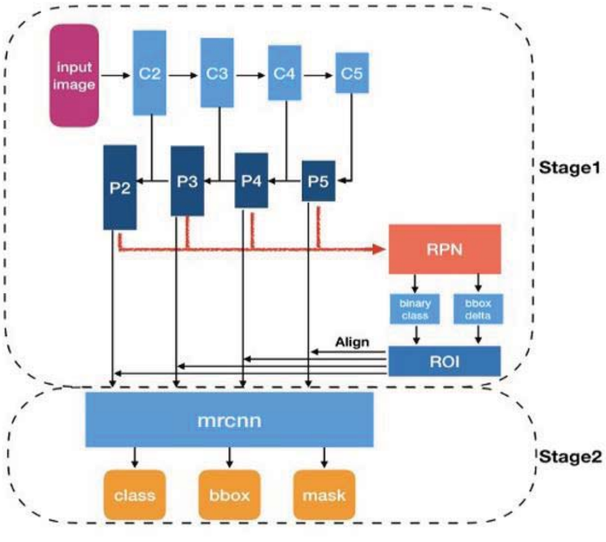


Fig.3. Mask RCNN Structure

It generates region proposals based on objects in an image and capable to predict the various class of objects by drawing a bounding box around an object to generate masks in pixel level. This stage acts as a backbone for Feature Pyramids Networks (FPN) to detect objects at different scales.

C. Semantic Segmentation

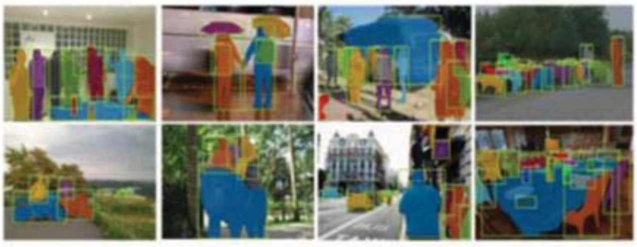


Fig.4. Mask R-CNN Semantic Segmentation

Figure 4. Shows semantic segmentation of Mask R-CNN. It is one of the high-level tasks that helps to paint a complete picture of understanding. The importance of understanding frame by frame is a core computer vision constrains which is highlighted by the fact that with a high increasing number of applications. Some of those applications are self-driving cars, human and computer interaction with each other, virtual-reality and so on. With many semantic segmentation changes are being solved using deep learning architectures, mostly convolutional neural networks, which yields much better results surpass other approaches by far better efficiency and accuracy [20].

D. Anomaly detection

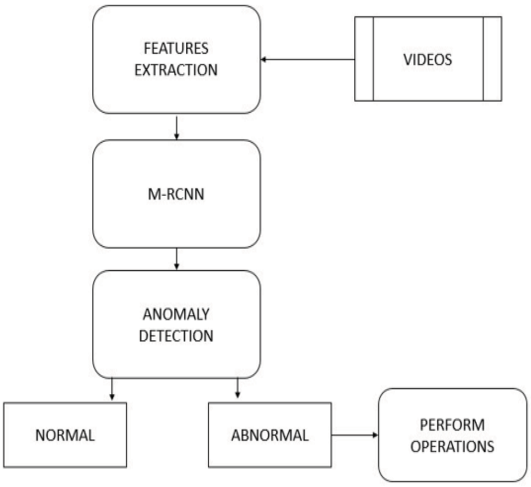


Fig.5. Block Diagram of Anomaly Detection

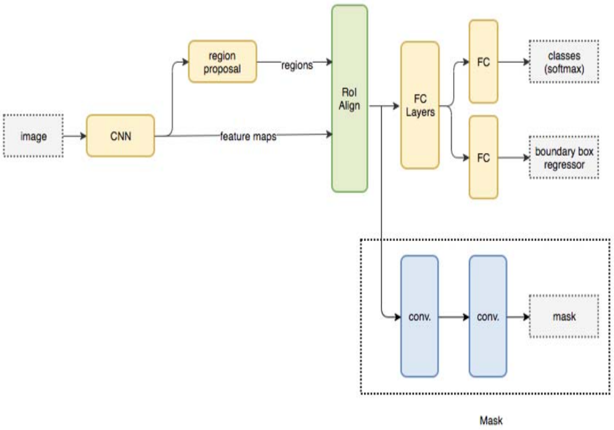


Fig.6. Flow diagram of M-RCNN

Figure 5 shows the anomaly detection block diagram. Features are extracted from a video. The frame per second is set to 20 frames for better results in a minimum amount of time. The features are extracted frame by frame, each video frame is divided into foreground and background region. Most of the time's background region information is unchanged and the foreground region contains moving objects [17] [22] [23] [25]. Features of both the region are extracted. Moving objects are detected and given as input to mask – RCNN. Mask R-CNN used there is nothing but faster R-CNN. Faster R-CNN has two outputs produced for each different object, one is a class label and another a bounding-box. Additionally, a third output is added which is an object mask. This additional mask is different from the other two outputs; extraction of information is much finer to construct a spatial layout of the object. Figure 6 shows the flow diagram of M-RCNN. The Pattern detection phase used in this model to identify patterns in the video sequence. Finally, the desired operations for the abnormal activities in the university campus and traffic flow monitoring is performed. The system can also be programmed to take some decisions upon pattern detection.

Some of the assumptions and constraints made during implementation are:

- The observed object brightness at any given time should be constant.
- The video is captured from a single stationary source
- The captured video should contain RGB frame structure
- Nearby objects in the image should move in a similar manner meaning smooth change in velocity.
- This camera unit should be connected to the computer and the video capture should be available.

IV. SIMULATION RESULTS AND ANALYSIS

Anomaly detection is the identification of various data points, events, and observations that deviate from the dataset's normal behavioral patterns. Anomaly detection also referred to as outlier detection, is used to find critical incidents, such as a technical glitch, fraud, or logistical obstacle, or potential opportunities.

A. Detection of Anomaly Object in the image

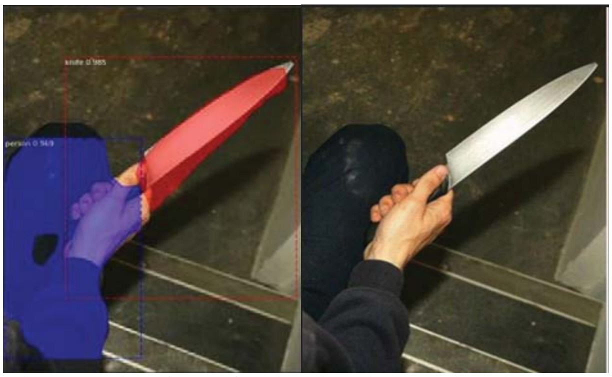
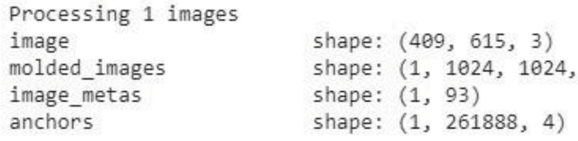


Fig.7. Detection of anomaly objects in an image (a) input image (b) output image.



```
Processing 1 images
image                      shape: (409, 615, 3)
molded_images              shape: (1, 1024, 1024, 3)
image_metas                shape: (1, 93)
anchors                    shape: (1, 261888, 4)
```

Fig.8. Output parameters

Figure 7 shows the detection of an anamoly object (knife) in an image. Fig. 7(a) is an input image, Fig. (b) Shows detection of knife with an accuracy of 98.5 % and detection of a person in the masked area with an accuracy of 96.5%. Also, the corresponding output parameters as shown in figure 8. Feature Pyramids Networks (FPN) are a basic component in recognition systems for detecting objects at different scales, which is used in this part of the algorithm. It describes the frame format, size, type, and pixel dimension.

B. Detection of Anomaly Object in video

Case -1: For home security related video

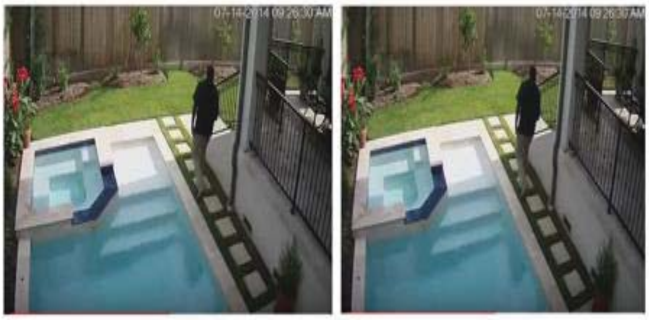


Fig.9. Detection of a person in a video. (a) Input image (b) Output image

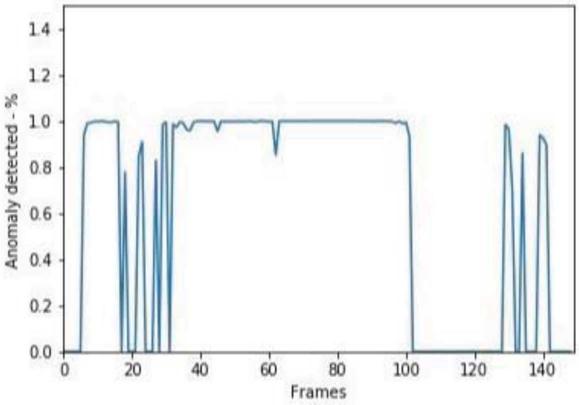


Fig.10. Detection of a person in a video plotted frame by frame

Figure 9 shows the detection of humans in a video, and the corresponding frame of video as shown in figure 10. Since this model is capable of detecting the presence of a knife in an image it is has been enhanced to detect the presence of a human in a video. The corresponding frame of the video is shown in fig.9a, which is the input frame of the video, the shown frame number in this video is 57 and the corresponding fig.9b. This is the output frame of the video, the shown frame number in this video is 56 where a person is detected and is masked with a color to locate the person in that particular frame. This is a real video surveillance camera footage model that can detect any presence of human activity around the home. In this scenario, if owners of the house were not present in the home. If a motion detector were installed to detect any presence, it would generate alarms even if wild birds and animals enter the area. This algorithm eliminates all the drawbacks mentioned and can detect selective anomalies. Figure 10, detection of a person in a video plotted frame by frame.

Case -2: Detection for surveillance video

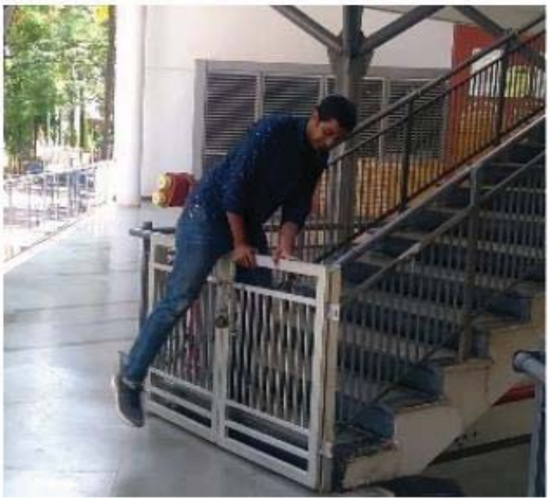


Fig.11. Detection of a person in an input video frame

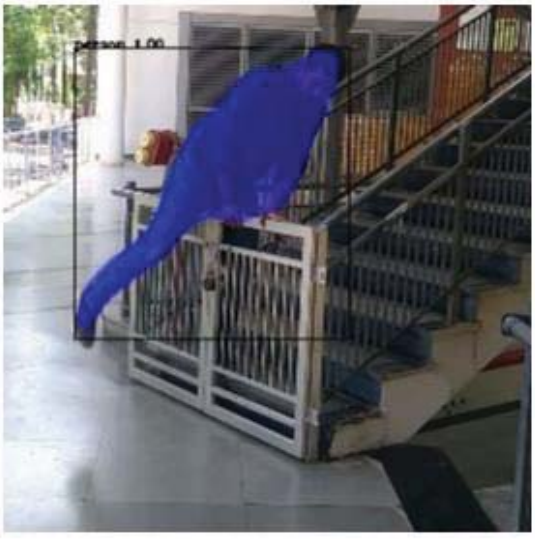


Fig.12. Detection of a person in an output video frame

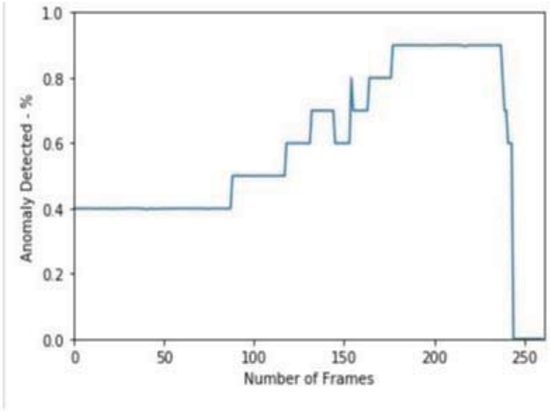


Fig.13. Detection of anomaly event in video frames

A particular anomaly is been set to detect if any anomaly is detected in and around a fence to monitor if any person is "jumping" the fence. This is achieved by extracting the masked binary value of the person detected in the video frame by frame and using the mathematical model. Which will generate a summed location values of the person in each frame will be stored to find the initial and current position of the person. This way a person trying to jump or jumping the fence location in that frame for the normal position is, higher this will be the direct indicator for anomaly detection. Figure 11 shows a person climbing across the gate and figure 12 shows the detection of the person with the probability values, finally figure 13 shows the anomaly value plotted for every frame from start until the end of the video. In fig.13, the percentage of anomaly detected for frames, we can conclude that an anomaly event occurs from frame 150 to 225, which is a person-climbing gate. Also from frame 0 to 88, we see the anomaly value is a constant 0.4% this is due to the person is on a flight of stairs which is relatively higher position the normal.

V. CHALLENGES IN ANOMALY DETECTION

Evolving "anomaly" – Different trade of, no universal model for an anomaly, timestamp difference, and change over time research in anomaly detection has been carried out for over several years in a wide area of disciplines. Even then still open research, the key reason is that the definition of anomaly is contextual, this is because what may be an anomaly in E-commerce may be different in the context of networking or computer vision. Furthermore, a different application needs different trade-offs to detect an anomaly. Moreover, what anomaly is today may not be the same tomorrow and the existing algorithm may not be useful later because the norm might change over time.

Lack of Data: Due to the fact that data is necessary for the model to be trained or to be tested. Lack of labeled data in "class imbalance" adds more time factor to implement a well-defined working anomaly detector. Increasing in WSN and IOT without doughty generates large amounts of data which makes it even more difficult to label them and differentiates between "normal" and "abnormal" data set. Due to all the facts stated above makes the development of a new anomaly detection algorithm is challenging.

Continuous learning and Training: Dynamic nature of the data streams calls for anomaly detection based on continuous learning. This way the constantly changing anomaly definition can be kept track and system updates can be performed from time to time.

Veracity: To identify objects labels and their corresponding location in video frame and image requires a computational power in both hardware and software and a very large amount of computing time. The resolution of the interested detection has to improve localization accuracy on small objects under partial occlusions. Real-time analysis of data, centralized anomaly detection system is not realistic and feasible.

VI. ANOMALY DETECTION ISSUES

Number of Attributes that are to be used to define anomalies: Consider a scenario when there is one attribute for an object, for some data values are anomalous but other values of other attributes are normal. Whether we have to decide based on those values of attributes or not. For example, consider people who are 6 feet tall and people who weigh 90kgs. But it is uncommon to have people who are 6feet tall and 90kgs weight. So, there must be a definition of anomaly specifying how multiple attributes are used to identify whether an object is an anomaly or not.

Limits to consider when to detect anomaly: To decide whether an object is anomalous or not is based on a divisive decision in a binary format for some techniques. But it is important to know that, there are some extreme anomalies and very less in degree. This is possible by giving a score to each anomaly based on the degree of it being anomalous. This assessment is called anomaly or outlier score.

Masking and Swamping: In some anomaly detection techniques anomalies are identified one at a time. In such techniques, there is a chance of missing out anomalies because of masking. Masking means the presence of several anomalies masks all other objects. Whereas identifying multiple anomalies at once can the involvement of swamping. Swamping is where normal objects are considered as anomalies.

Precision, recall false positive rate: If class labels are available to identify anomalies and normal data, then it is usually normal that the class of anomalies is smaller than that of a normal class. In such cases measures like precision, recall and false-positive rate are more important than the accuracy [9][10][11].

Efficiency: There are some significant differences between anomaly detection techniques. It is very important to choose the right technique. Based on the time complexities choosing an anomaly detection technique is important as finding an anomaly, it is also important to reduce the cost too.

VII. CONCLUSIONS

Detections of an anomaly in video surveillance are a challenging aspect due to the fact of different factors affecting results like video noise, outliers, and resolution. In this paper segmentation and classification model using thresholds, classification model and graph-based segmentation algorithms are used to obtain results with an accuracy of 98.5%. By utilizing the features from normal and anomalous surveillance videos as well, to avoid long training time; a general model of anomaly detection using deep learning yields best results with very minimal time. To explore the importance of locality in anomaly detection, experimental results show that, locating anomaly in the frame helps to achieve good results over a long surveillance video and this method is very robust. Further, it can be implemented for larger datasets by training using GPUs and high-end FPGA kits [18] [19] [24].

REFERENCES

[1] Henry et al., "Anomaly Detection in Videos Recorded by Drones in a Surveillance Context", *International Carnahan Conference on Security Technology (ICCST), Madrid, 2017.*

[2] Mohammad Farhadi et al., *"AAD: Adaptive Anomaly Detection through traffic Surveillance videos", arXiv:1808.10044, 2018.*

[3] Abhiraj Biswas et.al.,"Classification of Objects in Video Records using Neural Network Framework," *International conference on Smart Systems and Inventive Technology (ICSSIT-2018).*

[4] Arka Prava Jana et.al.,"YOLO based Detection and Classification of Objects in video records," *International Conference On Recent Trends In Electronics Information Communication Technology,(RTEICT)* 2018.

[5] Nehashree M R et.al., "Simulation and Performance Analysis of Feature Extraction and Matching Algorithms for Image Processing Applications" *International Conference on Intelligent Sustainable Systems (ICISS-2019).*

[6] Y. Wang, et al., "Traffic Camera Anomaly Detection",*2014 22nd International Conference on Pattern Recognition*, *Stockholm, 2014, pp. 4642-4647.*

[7] Jefferson Ryan Medel et al.,"Anomaly Detection in Video Using Predictive Convolutional Long Short-Term Memory Networks", *arXiv:1612.00390,* 2016.

[8] Biao Yang et al., "Anomaly Detection in Moving Crowds through Spatiotemporal Autoencoding and Additional Attention" *Advances in Multimedia,2018.*

[9] G. Chandan et.al., "Real Time Object Detection and Tracking Using Deep Learning and OpenCV", *International Conference on Inventive Research in Computing Applications (ICIRCA), 2018.*

[10] Abhiraj Biswas et.al.,"Survey on Edge Computing - Key Technology in Retail Industry" *International Conference on Intelligent Computing and Control Systems (ICICCS 2019).*

[11] Krishna C V et.al., "A review of Artificial Intelligence methods for data science and data analytics: Applications and Research Challenges" *International Conference on I-SMAC (IoT in social, mobile, Analytics and cloud), 2018.*

[12] A. Raghunandan et. al., "Object Detection Algorithms for Video Surveillance Applications," *International Conference on Communication and Signal Processing (ICCSP)*, 2018.

[13] A. Mangawati et al., "Object Tracking Algorithms for Video Surveillance Applications," 2018 International Conference on Communication and Signal Processing (ICCSP), 2018.

[14] X. Ma, T. Lu, F. Xu and F. Su, "Anomaly detection with spatio-temporal context using depth images", *International Conference on Pattern Recognition (ICPR2012), Tsukuba, 2012, pp. 2590-2593.*

[15] Samir Bouindour et al., "An On-Line and Adaptive Method for Detecting Abnormal Events in Videos Using Spatio-Temporal ConvNet" *Applied Science, 2019.*

[16] Manjunath Jogin et.al., "Feature extraction using Convolution Neural Networks (CNN) and Deep Learning" *International Conference On Recent Trends In Electronics Information Communication Technology, 2018.*

[17] Meghana R K et.al., "Background Modelling techniques for foreground detection and Tracking using Gaussian Mixture model" *International Conference on Computing Methodologies and Communication (ICCMC 2019).*

[18] S. K. Mankani et al., "Real-time implementation of object detection and tracking on DSP for video surveillance applications," *International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 2016.

[19] S. Sajjanar et. al., "Implementation of real time moving object detection and tracking on FPGA for video surveillance applications," 2016 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER) 2016.

[20] W. Sultani et al., "Real-World Anomaly Detection in Surveillance Videos", *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 6479-6488.*

[21] Federico Landi et al., "Anomaly Locality in Video Surveillance", *arXiv:1901.10364, 2019.*

[22] Mohana et.al., "Elegant and efficient algorithms for real time object detection, counting and classification for video surveillance applications from single fixed camera" *International Conference on Circuits, Controls, Communications and Computing (I4C),2016,*

[23] Mohana et.al., "Simulation of Object Detection Algorithms for Video Survillance Applications", *2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud),2018.*

[24] V. P. Korakoppa et.al., "An area efficient FPGA implementation of moving object detection and face detection using adaptive threshold method," *International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 2017.

[25] H. V. Ravish Aradhya et.al., "Real time objects detection and positioning in multiple regions using single fixed camera view for video surveillance applications," *International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)*, 2015, pp. 1-6.

[26] Bashar, Abul. "SURVEY ON EVOLVING DEEP LEARNING NEURAL NETWORK ARCHITECTURES." *Journal of Artificial Intelligence* 1, no. 02 (2019): 73-82.

[27] Kong, Lingchao, Ademola Ikusan, Rui Dai, and Jingyi Zhu. "Blind Image Quality Prediction for Object Detection", In *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pp. 216-221. IEEE, 2019.

[28] Moosbauer, Sebastian, Daniel Konig, Jens Jakel, and Michael Teutsch. "A Benchmark for Deep Learning Based Object Detection in Maritime Environments", In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.