Anomaly Detection in Video Surveillance using Object Detection

KUSHAGRA GAUTAM
Student, Maharaja Agrasen Institute of Technology

Abstract- The automatic detection of anomalies captured by surveillance settings is essential for speeding the otherwise laborious approach. We try Dataset which can help in identification of weapon and give less false positives. In this paper, we introduce weapon Detection DASIC DATASET suitable and use a combination technique for human-related anomaly detection in real-time video surveillance system using specified object tracking.

I. INTRODUCTION

The detection of anomalous events in videos is a challenging task due to the broad definition of the term 'anomaly', as well as insufficient annotated data. Despite this, there has been much research in the field of video surveillance anomaly detection in the past years. Surveillance cameras are a widely used technology which aids law enforcement agencies in ensuring general public safety. Surveillance footage is also considered are liable piece of forensic evidence when the anomalies captured on the footage are identified as crimes. Related studies in this area show that early detection of security threats or risks is fundamental to mitigate the damage caused as much as possible. Anomaly involving firearms such as handgun attacks, mass shootings, gun fire incidents on school grounds or terrorist attacks are representative examples of this kind of threats, which unfortunately have become rather common nowadays. The development of intelligent systems capable of automatically detecting threats or risk situations as soon as possible can provide important advantages in terms of security. In these kind of unpredicted situations false negatives are also needed to be taken into consideration here, we use YOLOV5 object detection and UCF crime and DASIC weapon detection Dataset to detect sudden appearance of suspicious harmful objects like guns, knives which can be considered as anomaly in public places and remove false negatives. The main problems such as the high number of false negatives and false positives occur due to challenges such as the similar shape and handling of non-weapon objects which are commonly handheld. Another major challenge is ensuring that the model doesn't fail to detect the weapon and has a very low false negatives rate. The model should also be able to avoid false positives from the background of these images and videos. The already presented models have a very high rate of false negative when it comes to videos. Suppose if 10 people with weapons are interested to enter a building and out of these 10 only one person succeeded to enter in the building can lead to serious consequences. Therefore it is required to reduce the number of false negatives and false positives by improving techniques suggested by, while also expanding the range of weapons that can be detected to include rifles. To achieve the goal of better Anomaly detection and less false negative puts forth a model that can take advantage of the latest models such as YOLO which has very fast detection speeds. The main contributions of this work include reducing the number of false positives and negatives in the domain of Weapon Detection by using Gaussian blur to remove the background and only focusing on the area of Interest and its combined use with YOLO v5s with Stochastic gradient descent (SGD).

II. METHODOLOGY

The YOLO-v5s algorithm is designed to have deal with the rate of false negatives and has faster speed. A Situation can be considered as Anomaly in presence of gun or knife in the video frame therefore we use different types of Datasets to detect gun and knife in the video frame.The proposed framework consists of the following steps (A) Data Preprocessing (B)YOLOV5.

A. Data Preprocessing

Data pre-processing is an important phase of the data analysis activity which involves the construction of the final data set so it can be fed to deep learning algorithms. The pre-processing used in our model is tore size the images and change the images from pascal to coco128 format which is accepted by YOLO V5 the second pre-processing technique that we use on the frames/images is to blur or remove the background from the images using different algorithms. Which in this model is Gaussian Blur operation, we opted to use this pre-processing rather than any other because of its speed compared to other techniques such as median filter which require sorting and slow down the operation. The Gaussian filter is a low-pass filter that removes the high-frequency components; the pixels nearest the center of the kernel are given more weight than those far away from the center. This averaging is done on a channel-by-channel basis, and the average channel values become the new value for the filtered pixel. The Gaussian blur is a type of image-blurring filter that uses a Gaussian function, which also expresses the normal distribution in statistics for calculating the transformation to apply to each pixel in the image. The formula of a Gaussian function is:

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-x^2 \pm y^2}{2\sigma^2}}$$

In the above equationwhere x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and σ is the standard deviation of the Gaussian distribution.

B. YOLOV5

YOLO-v5 as a learner to detect weapon (as objects) with their centroids and sizes. The YOLO-v5 DL based object detection method isused in this paper to construct a model that can make the clustering algorithm free of initialization. This model finds automatically the appropriate clustering initialization parameters such as the number of weapon, possible centroids and weapon sizes. The input of the proposed model is a transformed data in 2Dfeature space. The proposed solution assumes that a high-dimensional data-set can be represented in a lower dimension space, pair wise distance is preserved among weapon. Given a transformed data-set of n points X =x1, . . . , xn, where xiR2 , the clustering algorithms, such ask-means,partitionthedata-setintokweaponC1,C2,...,Ck, where each weapon Cj is represented by a weapon center cj. Let k deno test he correct number of weapon in a data-set. To build accurate clustering models 6 that can achieve high clustering accuracy, k and other clustering parameters (e.g. no. of weapons and weapon sizes) are required as input. Getting a good estimate of these parameters and especially k is not a straight forward task. Several metrics were presented previously to determine the number of weapon. However, computing these metrics requires a high computation and resources overhead. The input of the proposed solution is an image representing the weaponed data. Thus, with large data-set, the computation overhead of the proposed solution will remain the same, contrarily to the existing techniques that require an exponential computation complexity in function of the data size The YOLO-v5 model was trained using a generated data-set.

III. MODELLING AND ANALYSIS

A. MODEL ARCHITECTURE

YOLO an acronym for 'You only look once', is an object detection algorithm that divides images into a grid system. Each cellin the grid is responsible for detecting objects within itself. YOLO is one of the most famous object detection algorithms due to its speed and accuracy.YOLOv5 is a family of compound-scaled object detection models trained on the COCO dataset, and includes simple functionality for Test Time Augmentation (TTA), model ensembling, hyper parameter evolution, and export to ONNX, CoreML and TFLite. YOLOv5 has four different models including YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. Generally, YOLOv5 respectively uses the following architecture:-

1. Backbone: Model Back bone is mostly used to extract key features from an input image. CSP (Cross Stage Partial Networks) are used as a backbone in YOLO v5 to extract rich in useful characteristics from an input image.
2. Neck: The Model Neck is mostly used to create feature pyramids. Feature pyramids aid models in generalizing successfully when it comes to object scaling. It aids in the identification of the same

object in various sizes and scales. Feature pyramids are quite beneficial in assisting models to perform effectively on previously unseen data. Other models, such as FPN, BiFPN, and PANet, use various sorts of featurepyramidapproaches.PANetisusedasaneckin YOLOv5togetfeaturepyramids.

3. Head: The model Head is mostly responsible for the final detection step. It uses anchor boxes to construct final output vector with class probabilities

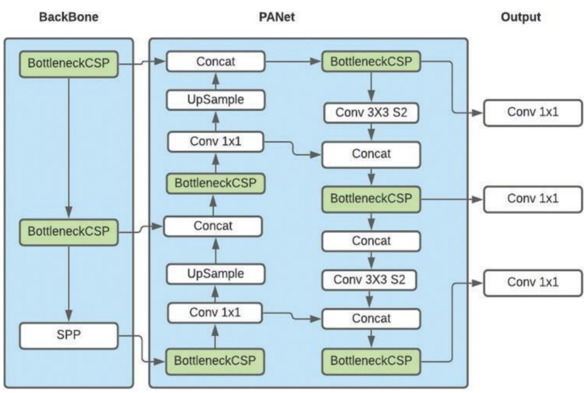


Figure 1:YOLOV5 Model

B. Dataset

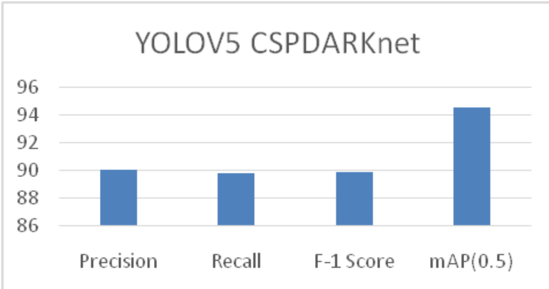
The datasets included in this section have been designed for the object detection task based on Deep Learning architectures with a CNN backbone. The selected images contain weapons and objects but also consider an enriched context of different background objects as well as the way objects are handled. After the training stage on these datasets, the detection models must locate and distinguish between weapons and different common objects present in the background or handled similarly. The data set contains 3000 images of pistols along with their boundary box files in Pascal-VOC XML format which was used for the positive class of Pistols. The Dataset was increased by the addition of 12, 887 Negative examples which are images that do not contain pistols. The resolution of images was 100x100 pixels and was sourced from open-sourced from the University of Grandala. The data set for the 3000 guns was mostly created from Internet Movie Firearms Database (IMFDB) dataset or Common Objects in Context (COCO) datasets. The dataset is divided into training, validation, and test set where 70% of the data isused for training and 20% used for the validation, and 10% is used for the test set. The total image count reaches 15,873, making it the largest existing dataset for pistol detection available as of date. We used Roboflow, which is an online website with collections of tools that help to organize, prepares, and improves your image and annotation training data. We use Roboflow to upload our dataset in VOCXML format and then Roboflow convert that data in to text format which was required by our YOLO-V5smodel. We also used Roboflow to pre-process our images into 416×416 which is recommended for the YOLO model.

IV. RESULT

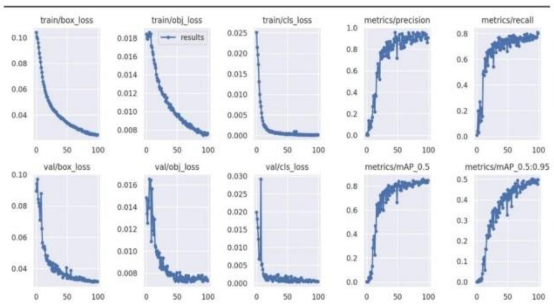
The YOLOV5 model was were trained at 100 epochs and with 3000 images. As result of Training model, the time spent was 2 hrs. When we test our model on 600 images where 416 images classified as Anomaly (Guns present in the frame) rest normal our model easily distinguish whether anomaly is present or not. By reducing the classes and categories over our model to simply two categories of Anomaly present or not we increased the speed and accuracy of our model to simply detect Anomaly. As a result of training and Valiation process, we found YOLOV5 model results to be good. After Evaluation our model had a Precision Score of 90.1%, Recall of 89.8% and mAP score is 94.6%.

Table 1

MODEL	Backbone	Precision	F-1 Score	mAP(0.5)
YOLO V5	CSPDARK net	90.1	89.9	94.6



(YOLOV5 Performance GRAPH)



The first three columns are the YOLOV5 model loss components, box loss, objectness loss and classification loss, train the leftmost row and validation second row, these are the indicators of how well algorithm predicts the object. These show that classes used like knife, guns, billets are accurately recognized during training process. The model is suitable for detecting gun and knifes, its perform well in open environment and can be used for presence of weapon in frame to be classified as Anomaly.

CONCLUSION

In this paper, we have proposed a model to detect anomaly by detecting the presence of weapon specifically knifes and guns in the video frame We used the latest light-weight model such as YOLOv5s that had very effective results and speed, we used it for the task of weapon detection and classify the detection of weapon as an Anomaly and rest as normal. As we further classified our data into only two categories as Anomaly and normal, this helped in lowering the false negatives. This model can be used in security video surveillance in public places like park etc where appearance of weapon can be considered as anomaly.

REFERENCES

[1] Dongkuan Xuand Yingjie Tian. Acomprehensive survey of clustering algorithms. Annals of Data Science,2 (2):165–193,2015.

[2] Christophe Guyeux, Ste´phane Chre´tien, Gaby Bou Tayeh, Jacques Demerjian, and Jacques Bahi. Introducing and comparing recent clustering methods for massive data management in the internet of things. Journal of Sensor and Actuator Networks, 8(4):56(25), dec2019.

[3] Kristina P Sinaga and Miin-Shen Yang. Unsupervisedk-means clustering algorithm. IEEE Access, 8:80716–80727, 2020.

[4] Adil Fahad, Najlaa Alshatri, Zahir Tari, Abdullah Alamri, Ibrahim Khalil, Albert Y Zomaya, Sebti Foufou, and Abdelaziz Bouras. A survey of clustering algorithms for big data: Taxonomy and empirical analysis. IEEE transactions on emerging topics in computing, 2(3):267–279, 2014.

[5] Maria Camila N Barioni, Humberto Razente, Alessandra MR Marcelino, Agma JM Traina, and Caetano Traina Jr. Open issues for partitioning clustering methods: an overview. Wiley Inter disciplinary Reviews: Data Mining and Knowledge Discovery, 4(3):161–177,2014.

[6] Mohammad Alhawarat and M Hegazi. Revisiting k-means and topic modeling, a comparison study to cluster arabic documents. IEEE Access, 6:42740–42749,2018.

[7] Yinfeng Meng, Jiye Liang, Fuyuan Cao, and Yijun He. A new distance with derivative information for functional k-means clustering algorithm. Information Sciences,463:166–185,2018.

[8] A.P. Dempster, N.M. Laird and D.B. Rubin. Maximum likelihood from incomplete data via theem algorithm. Journal of the Royal Statistical Society: Series B, 39(1):1–38, 1977.

[9] J. B. Mac Queen. Some methods for classification and analysis of multivariate observations. In Proceedings of 5-thBerkeley Symposiumon Mathematical Statistics and Probability, pages 281–297. Berkeley, University of California Press.

[10] Dan Pelleg and Andrew Moore. X-means: Extending k-means with e cient estimation of the number of clusters. In Proceedings of the 17th International Conference on Machine Learning, page 727.Citeseer.
[11] Joseph C Dunn. A fuzzy relative of the is o data process and its use in detecting compact well-separatedclusters.1973.
[12] James C. Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms. Springer US, 1981.
[13] Robert E Kass and Adrian E Raftery. Bayes factors. Journal of the american statistical association, 90(430):773–795, 1995.
[14] Hamparsum Bozdogan. Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. Psychometrika, 52(3):345–370, 1987.
[15] David L Davies and Donald W Bouldin. A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, (2):224–227,1979.15
[16] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20:53–65,1987.
[17] TadeuszCali´nskiandJerzyHarabasz.Adendritemethodforclusteranalysis.CommunicationsinStatistics-theoryandMethods,3(1):1–27,1974.
[18] Robert Tibshirani, Guenther Walther, and Trevor Hastie. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2):411–423,2001.
[19] Ere´ndira Rendo´n, Itzel Abundez, Alejandra Arizmendi, and Elvia M Quiroz. Internal versus external cluster validation indexes. International Journal of computers and communica-tions, 5 (1):27–34, 2011.
[20] Glenn Jocher. Yolo v5 in pytorch. https://github.com/ultralytics/yolov5,062020. (undefined28/4/202110:16).
[21] Aristidis Likas, Nikos Vlassis, and Jakob J Verbeek. The global k-means clustering algorithm. Pattern recognition, 36(2):451–461, 2003.
[22] Dan Pelleg, Andrew W Moore, et al. X-means: Extending k-means with efficient estimation of the number of clusters. InIcml, volume1, pages 727–734, 2000.
[23] Pradipta Majiand Sankar K Pal. R fcm: a hybrid clustering algorithm using rough and fuzzy sets. FundamentaI n format icae, 80(4):475–496, 2007.