Armament Detection And Alert System Using YOLOV3

Shashank Shet K
K S Institute of technology
(of Affiliation)
Visveswaraya Technological University
Bengaluru, India
shashankshetk@gmail.com

Parth P Shah
K S Institute of technology
(of Affiliation)
Visveswaraya Technological University
Bengaluru, India
p.parthshah99@gmail.com

Rohith.R
K S Institute of technology
(of Affiliation)
Visveswaraya Technological University
Bengaluru, India
rohi1999@gmail.com

Assistant professor Mr. Roopesh Kumar B.N
Bengaluru, India
roopeshkumarbn@ksit.edu.in

***Abstract*— Based on current situation around the world, there is major need of automated visual surveillance for security to detect arms and alert the public. The objective of this paper is to visually detect the arms in real time videos and alert upon successful detection. The proposed method is using YOLO-V3 algorithm. To improve the result, we have created our own dataset of weapons with all possible angles and merged it with ImageNet dataset. The merged data was trained using YOLO-V3 algorithm. The detector performed very well to detect weapons in different scenes with different rotations, scales and shapes. The results showed that YOLO-V3 can be used as an alternative of Faster RCNN. It provides much faster speed, nearly identical accuracy and can be used in a real time environment.**

Keywords*—*YOLOV3; Armament detection;

I. INTRODUCTION

The number of crimes involving arms violence is increasing throughout the world, especially in that places where the possession of guns is legal. From the below statistics, it is evident that violence rate concerning guns are increasing and becoming a challenge for law enforcement agencies to overwhelm this issue timely.

Rank	Country	Deaths per 100,000	Total Deaths
1	Brazil	19.4	43,200
2	United States	10.6	37,200
3	India	2.1	26,500
4	Mexico	11.8	15,400
5	Colombia	25.9	13,300
6	Venezuela	38.7	12,800
7	Philippines	8.3	8,020
8	Guatemala	32.3	5,090
9	Russian Federation	2.6	4,380
10	Afghanistan	14.2	4,050

The solution of this problem is monitoring and early detection using control camera or surveillance systems and alerting which can help to prevent this kind of violence and assist policemen or security agents to act timely Recently, the area of machine learning mainly in detection and classification of objects and image segmentation has been revolutionized by deep learning. You look only once (YOLO) outperformed other detection algorithms at prediction. [1].

In this paper, we present an automatic armament detection and alerting system using YOLOv3. Detecting gun in a scene is very challenging issue because of various subtleties linked with it. The challenges in gun detection is particularly caused by occlusion. Gun to scene and gun to object are the two types of gun occlusion. Real time processing is another main problem in gun detection system that arises during detection. The rest of the paper is structured as follows: section 2 describes related work, section 3 describes the methodology, section 4 describes the results and discussions and section 5 describes the conclusion.

II. RELATED WORK

The research work in [2] proposed a visual gun detection system based on Harris interest point and SIFT. They used color based segmentation to select dissimilar object from an image by deploying K-mean cluster algorithm. In [3], researchers used 3D millimeter wave imaging technique to detect the weapon concealed in the body and other hidden location.

In another work [3], researchers deploy a real time gun classifier that detect and classify guns. They also used ImageNet dataset for training their model and acquired a mAP of 89% using overfeat-3 algorithm. The research work [4], is based on terahertz human dataset used in deep learning to detect the concealed weapon. The achieved a competitive accuracy compared to other concealed weapon detection system.

III. METHODOLOGY

Methodology has been organized as follows:

A. DATASET

We have created our own image dataset containing arms with different position with ImageNet dataset. YOLO v3 algorithm has been trained and validated to evaluate our armament detection system for better results. Figure 1 shows some of the examples from the dataset being used for training our model.

Figure-1 shows some of the images from ImageNet datasets including the image we merged.

B. YOLO-V3 ALGORITHM

In this work, we have used transfer learning for training YOLOv3 model for gun detection and used the weight trained on ImageNet by YOLOv3 instead of starting from zero.

YOLOv3 is an object detection algorithm widely used for real time processing. Input image is divided into $\boldsymbol{M} \times \boldsymbol{M}$ grids. A single object is then predicted by this grid cell.

Logistic regression is used to predict an object scores for each bounding box by YOLOv3 and changes the method to compute the cost function. If a ground truth object is early overlapped by a bounding box prior more than others, the resultant object score should be 1. No cost is experienced for other prior with overlap greater than predefined threshold 0.5.

Figure 1: images used for training YOLOv3 model.

Only a single boundary box prior is associated with each ground truth object. If bounding box prior is not allocated, no classification and localization lost incurs, only confidence loss on objects incurs. We compute the loss by using the following equations.

$$l_x = \sigma(t_x) + c_x$$
$$l_y = \sigma(t_y) + c_y$$
$$l_w = p_w e^{t_w}$$
$$l_h = p_h e^{t_h}$$

$$Pr(\text{Object}) * \text{IOU}(\text{b}, \text{object}) = \sigma(t_o)$$

Where l_x and l_y are loss function and l_y and l_h are generated from sigma function with the range 0 and 1. The network calculates 5 coordinates for each bounding box t_x, t_y, t_w, t_h and t_o.

Note that loss function only corrects the classification error in the same grid cell.

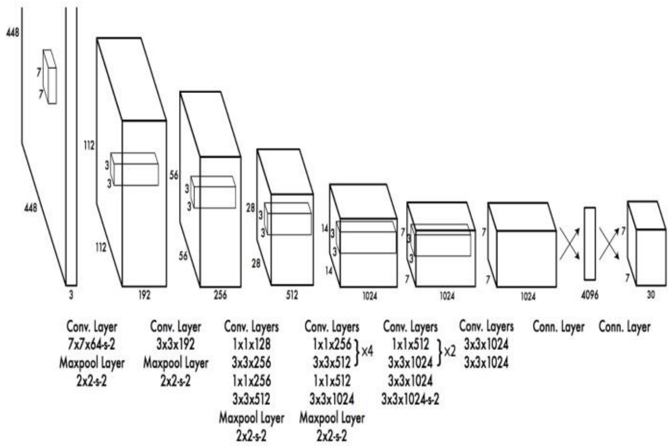

Figure 2: shows the block diagram of YOLO-V3 algorithm

The class prediction from the cell level is moved to the boundary box level. Now, each prediction includes 4 parameters for the boundary box, 1 box confidence score (objectness) and 20 class probabilities. i.e. 5 boundary boxes with 25 parameters: 125 parameters per grid cell. Same as YOLO, the objectness prediction still predicts the IOU of the ground truth and the proposed box.

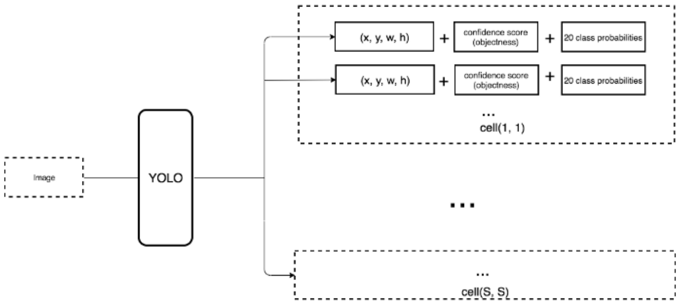

Figure 3: Class prediction from cell to boundary box level

IV. RESULTS AND DISCUSSION

Detecting guns in videos not only focus on gun detection but also emphasis on minimizing the number of false positives and providing real time detection.

We have estimated YOLO-V3 based classification model considering a single class to specify the presence of weapon. YOLO has been used by researchers for detecting different objects related to their interest. However, this paper is to check how the YOLOv3 neural network behave when used in object detection. The detection results are examined frame by frame in the videos during the experiments and measured a detection as true positive if the gun and bounding box overlapping is more than 50%. If the human eye recognizes the gun is considered as ground truth.

Figure 3: an example of accurate weapons detected in four different frames

False negative can be used to explain low recall in the videos frame which is due to the detection of weapons in obstruct environment or when the weapon is moved quickly.

The false negatives strongly rely on the frame quality and when the weapon is visible clearly.

V. CONCLUSION

The objective was to minimize the false positive using YOLOv3 algorithm. YOLOv3 based model has been trained with a dataset containing ImageNet and our own customized dataset. It is clear that YOLOv3 has a good detection performance even in low quality videos as than faster RCNN. The advantage of YOLOv3 over Faster RCNN is its speed..

VI. REFERENCES

[1] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788. 2016.

[2] Tiwari, Rohit Kumar, and Gyanendra K. Verma. "A computer vision-based framework for visual gun detection using harris interest point detector." *Procedia Computer Science* 54 (2015): 703-712.

[3]Arceda, V. Machaca, K. Fernández Fabián, and Juan Carlos Gutíerrez. "Real time violence detection in video." (2016): 6-7.

[4] Akcay, Samet, Mikolaj E. Kundegorski, Chris G. Willcocks, and Toby P. Breckon. "Using Deep Convolutional Neural Network Architectures for Object Classification and Detection Within X-Ray Baggage Security Imagery." *IEEE Transactions on Information Forensics and Security* 13, no. 9 (2018): 2203-2215.