International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

Weapon Detection in Surveillance System

Dr. N. Geetha, Akash Kumar. K. S, Akshita. B. P, Arjun. M
Coimbatore Institute of Technology

***Abstract*: From many statistics, it can be assumed that the violence rate concerning guns and harmful weapons is increasing every year, becoming a challenge for law enforcement agencies to deal with this issue on time. There are many places where the crime rate caused by guns or knives is very high, especially in places where there are no gun control laws. The early detection of violent crime is of paramount importance for citizens' security. One way to prevent these situations is by detecting the presence of dangerous weapons such as handguns and knives through surveillance videos. Present surveillance and control systems still require manual monitoring and intervention. Here, we present a system of automatic detection of weapons in the video, which is appropriate for surveillance and control purposes. We have used the YOLOv3(You look only once) algorithm for the detection of weapons in real-time video. The YOLO models are end-to-end deep learning models and are well-liked because of their detection speed and accuracy. Previous methods, like region-based convolutional neural networks (R-CNN), require thousands of network evaluations to make predictions for one image which can be time-consuming and painful to optimize. It focuses on a specific area of the image and trains each component separately. A YOLO model on the other hand only passes the image once through the neural network. Since, in a real-time video, speed is of paramount importance, we have used the YOLOv3 algorithm. The dataset is trained for classifying three classes of weapons – Handgun, Knife, and Heavy Guns. Once the weapon is detected, an alert will be sent to authorities who can act accordingly and reduce the violent crimes before they take place.**

Keywords- Weapon detection, Surveillance system, YOLOv3

I.INTRODUCTION

Security is always a main concern in every domain, due to a rise in crime rate in crowded or in suspicious lonely areas. Gun violence is a contemporary global human rights issue. Gun-related violence threatens our most fundamental human right, the right to life. Gun violence is a daily tragedy affecting the lives of individuals around the world. More than 500 people die every day because of violence committed with firearms. The easy availability of guns always remains a big contributory factor behind the spike in crime and lawlessness. This is typically illustrated by the crime scene in America. In the US, gun culture is very strong and has historical roots. There are around 249 million guns in America and about one-third of them are handguns, which are easy to conceal. On average, each year, shootouts account for 50,000 deaths, including 12,000 murders. Research studies have also shown that household hand-guns procured for self-defence are more likely to kill family members than save their lives.

In India, which has one of the strictest gun laws in the world, things are different. Obtaining weapons is a privilege rather than a constitutional right in this country (like in the US). Even for light weapons, licenses are required under the 2016 Arms Rules. However, procuring a license is a complex procedure that can take months. They are only granted after a rigorous examination, which includes background checks. It's difficult to put a figure on illegally possessed firearms, but a look at the license status of previous weapon seizures provides a good picture of how widespread the problem is. This possesses a major concern for these dangerous weapons on the security of the public.

Due to the growing demand for 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.

II.RELATED WORKS

JIANYU XIAO SHANCANG L et al.[1] developed advanced forensic video analysis techniques to assist the forensic investigation. An adaptive video enhancement algorithm based on contrast limited adaptive histogram equalization (CLAHE) is introduced to improve the closed-circuit television (CCTV) footage quality for the use of digital forensic investigation. To assist the video-based forensic analysis, deep learning-based object detection and tracking algorithm are proposed that can detect and identify potential suspects and tools from the footage.

JEONG SEO AND HYE YOUNG PARK et al.[2] proposes a framework for recognizing objects in very low-resolution images through the collaborative learning of two deep neural networks: The proposed image enhancement network attempts to enhance extremely low-resolution images into sharper and more informative images with the use of collaborative learning signals from the object recognition network.. It also utilizes the output from the image enhancement network as augmented learning data to boost its recognition performance on very low-resolution objects.

Harsh Jain et. al[3] implements automatic gun (or) weapon detection using a convolution neural network (CNN) based SSD and Faster RCNN algorithms. The proposed implementation uses two types of datasets. One dataset, which had pre-labelled images, and the other one is a set of images, which were labelled manually. Results are tabulated, both algorithms achieve good accuracy, but their application in real situations can be based on the trade-off between speed and accuracy.

Shenghao Xu[4] developed a weapon detection system based on TensorFlow, which is an open-source platform for machine learning; the Single Shot MultiBox Detector (SSD), a popular object detection algorithm; and MobileNet, which is a convolution neural network (CNN) for producing high-level features.

From our other studies, we have inferred that the YoloV3 algorithm is significantly faster than other object detection algorithms for real-time videos.

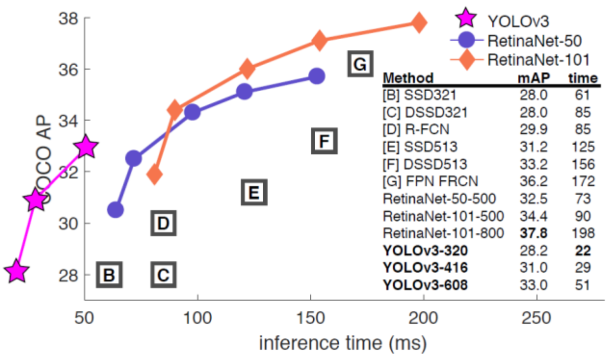


Fig.1. The comparison of various fast object detection models on speed and mAP performance on COCO 50 benchmark.

(Image source: focal loss paper with additional labels from the YOLOv3 paper.)

III.PROPOSED SYSTEM

In our proposed system, we have proposed the following method to detect weapons using the YOLOv3 algorithm. Initially, a dataset is created which consists of three classes of weapons – Handgun, Knife, and Heavy guns. This dataset is trained for the classification of weapons using the YOLOv3 (You Only Look Once) algorithm. Once the data is trained, the system can classify the type of weapon present in the real-time input video from the surveillance cameras along with the confidence score of each weapon. If the weapon is detected, an alert will be sent to the authorities.

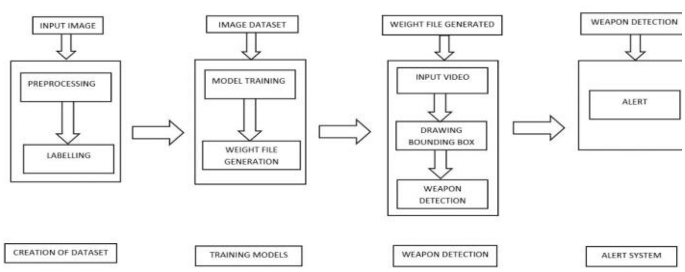


Fig.2. Architecture design

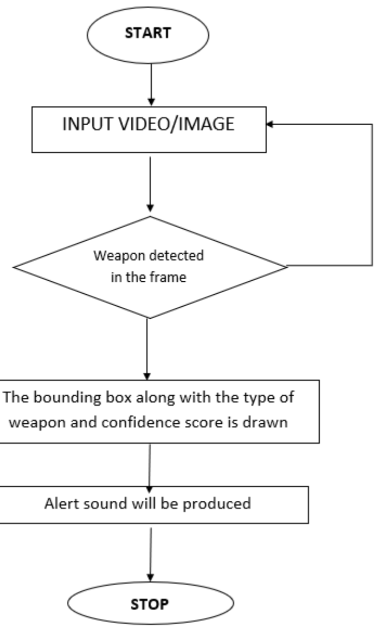


Fig.3. Project flow diagram

IV.METHODOLOGY

A. Dataset

Raw images are not appropriate for analysis purposes and need to be converted into the processed format, such as jpeg, jpg, and tiff for further analysis. The image size is reconstructed into a square image. The images were resized into 416px x 416px resolution to reduce the computational time and then the images were then retained in the RGB format. Dataset is created by collecting the good pixel weapon images and making them ready for the creation of the dataset.

Fig.4 Dataset of images containing the weapons – Handgun, Knife, and Heavy guns(416p x 416px)

Once the dataset of weapon images is collected from various sources, it is creating using the tool Labelimg toolbox. ***LabelImg*** is a graphical image annotation tool and labels object bounding boxes in images. It is a free, open-source tool for graphically labelling images. It's written in Python and

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

uses QT for its graphical interface. The position of the weapons was marked in the images based on the three classes of weapons – handgun, knife, and heavy gun. The coordinates for these markings were generated for each image and stored in a text file. The classes for which the images were marked were stored in a label file. This is used as the training dataset.

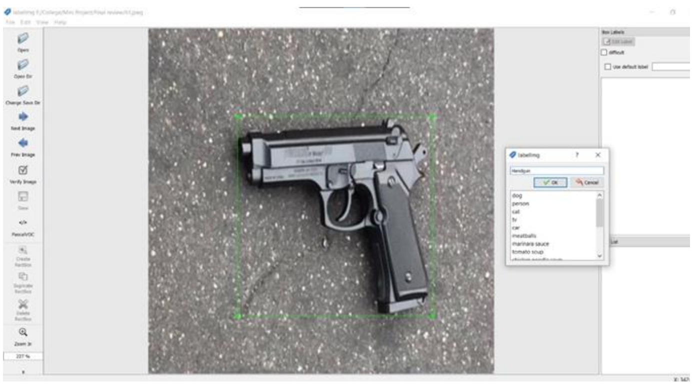


Fig.5 Labelling the images

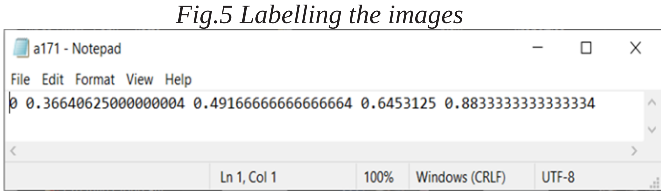

a171 - Notepad
File Edit Format View Help
0 0.36640625000000004 0.49166666666666664 0.6453125 0.8833333333333334
Ln 1, Col 1 | 100% | Windows (CRLF) | UTF-8

Fig.6 Coordinates of the bound Image

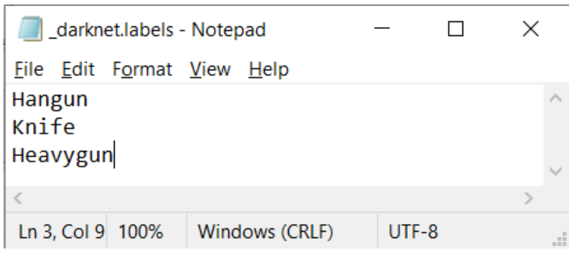

_darknet.labels - Notepad
File Edit Format View Help
Hangun
Knife
Heavygun
Ln 3, Col 9 | 100% | Windows (CRLF) | UTF-8

Fig.7 The classes used while labeling

B. Separation of the dataset into train and test data:

Once the image labeling process is completed, this complete data set is compressed into a zip file and upload into google drive. The dataset uploaded is then divided into 70% training data and 30% testing data and the images are separated into different folders which can be used for training the model.

C. Yolov3 Algorithm

YOLOv3 (You Only Look Once, Version 3) is a real-time object detection algorithm that identifies specific objects in videos, live feeds, or images. Previous methods, like region-based convolutional neural networks (R-CNN), require thousands of network evaluations to make predictions for one image which can be time-consuming and painful to optimize. In YOLOv3, The feature extraction and object localization were unified into a single monolithic block. Their single-stage architecture, named **YOLO** (You Only Look Once) results in a very fast inference time. It takes the entire image in a single instance and predicts the bounding box coordinates and class probabilities for these boxes. **The biggest advantage of using YOLO is its superb speed** – it's incredibly fast and can process 45 frames per second. Unlike other methods where images are scanned with a sliding window, in YOLO whole image is passed into a convolutional neural network and predicts the output in one pass.

For object detection using Yolov3, a pre-trained CNN network on image classification task from Alexey Darknet53 (https://pjreddie.com/darknet) is used in the background. We are applying the ***Transfer Learning*** technique by adding our layers to an already trained model. So, we download the pre-trained weights called darknet53.conv.74. Thus, our custom model will be trained using these pre-trained weights instead of randomly initialized weights which in turn will save a lot of time and computations while training our model.

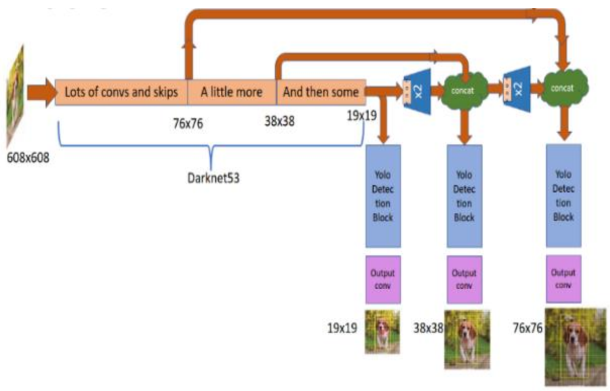


Fig.8. Architecture Of YoloV3

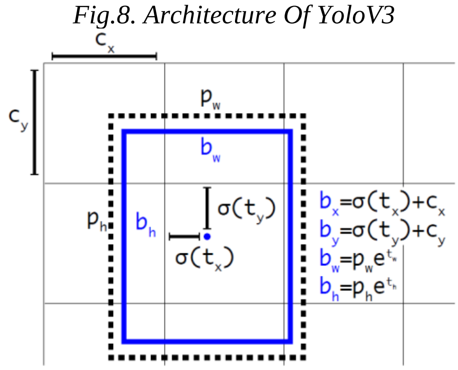


Fig.9. Bounding box location prediction

D.Working

Each image in the dataset is split into S×S cells. If an object's center lies into a cell, that cell is responsible for detecting the existence of that object. Each cell predicts the location of B bounding boxes, a confidence score, and a probability of object class conditioned on the existence of an object in the bounding box.

The coordinates of the bounding box are defined by a tuple of 4 values, (center x-coordinate, center y-coordinate, width, height) — (x,y,w,h), where x and y are set to be offset of a cell location.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

Moreover, x, y, w, and h are normalized by the image width and height, and thus all between (0, 1].

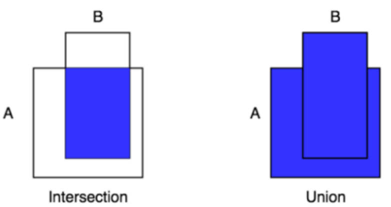
A confidence score indicates the probability that the cell contains an object: Pr(containing an object) x IoU(pred, truth); where Pr = probability and IoU = interaction under union.

If the cell contains an object, it predicts a probability of this object belonging to every class Ci,i=1,…, K: Pr(the object belongs to the class C_i | containing an object). At this stage, the model only predicts one set of class probabilities per cell, regardless of the number of bounding boxes, B.

In total, one image contains S×S×B bounding boxes, each box corresponding to 4 location predictions, 1 confidence score, and K conditional probabilities for object classification. The total prediction value for one image is S×S×(5B+K), which is the tensor shape of the final conv layer of the model. The final layer of the pre-trained CNN is modified to output a prediction tensor of size S×S×(5B+K). Here K=3 as there are 3 classes of weapons.

The objects in the image can be of different shapes and size, and to capture each of these perfectly, the object detection algorithms create multiple bounding boxes as shown in fig.11.Ideally, for each object in the image, it must have a single bounding box. To select the best bounding box, from the multiple predicted bounding boxes, these object detection algorithms use non-max suppression. This technique is used to "suppress" the less likely bounding boxes and keep only the best ones.

The same process goes for the remaining boxes. This process runs iteratively until there is no more reduction of boxes. In the end, we will be left with only one bounding box. YOLOv3 uses Intersection over Union (IoU) for non-max suppression.



$$IoU = \frac{A \cap B}{A \cup B}$$

Fig.10 Non-max suppression using Intersection over Union(IoU)

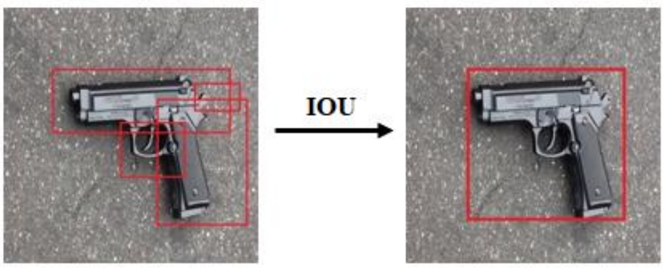


Fig.11 Reduction of bounding boxes using Non-Max Suppression

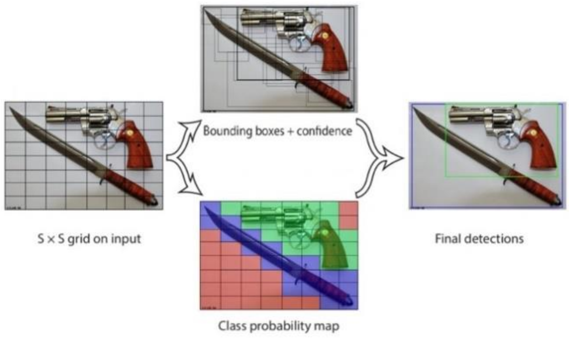


Fig.12. Working of YOLOV3

E.Training the model

The darknet repo is cloned from GitHub and using the darknet53.conv.74 file, the pre-trained files is used for transfer learning and the neural network is trained for our weapon data. The training takes place for 6000 iterations. Once the training is completed the training. weights file and yolov3.cfg file is generated which can be used for weapon detection.

```
# Change lines in yolov3.cfg file
!sed -i 's/batch=1/batch=64/' cfg/yolov3_training.cfg
!sed -i 's/subdivisions=1/subdivisions=16/' cfg/yolov3_training.cfg
!sed -i 's/max_batches = 500200/max_batches = 6000/' cfg/yolov3_training.cfg
!sed -i '610 s@classes=80@classes=3@' cfg/yolov3_training.cfg
!sed -i '696 s@classes=80@classes=3@' cfg/yolov3_training.cfg
!sed -i '783 s@classes=80@classes=3@' cfg/yolov3_training.cfg
!sed -i '603 s@filters=255@filters=24@' cfg/yolov3_training.cfg
!sed -i '689 s@filters=255@filters=24@' cfg/yolov3_training.cfg
!sed -i '776 s@filters=255@filters=24@' cfg/yolov3_training.cfg
```

Fig.13 Changing the parameters of yolov3.cfg to train 3 classes of images

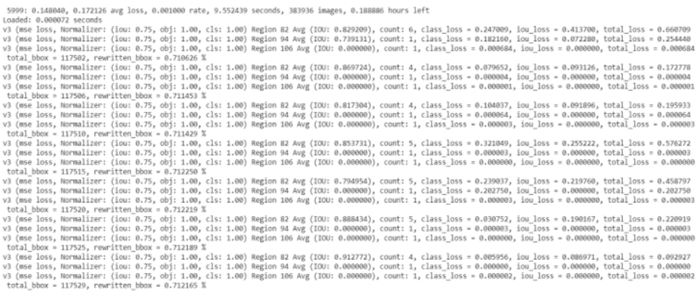


Fig.14 Training the dataset

F. *Implementation Using OpenCV*

After the training is completed, the training. weights file and yolov3.cfg file is generated which can be used for weapon detection. We used OpenCV for detecting the presence of a weapon in the live video. After the weight files are connected successfully in the code, the Input video is obtained through a web camera or internal file. The weapon detected in the video along with the confidence score is displayed. If a weapon is detected an alert sound is used to indicate the presence of a weapon in that detected frame.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

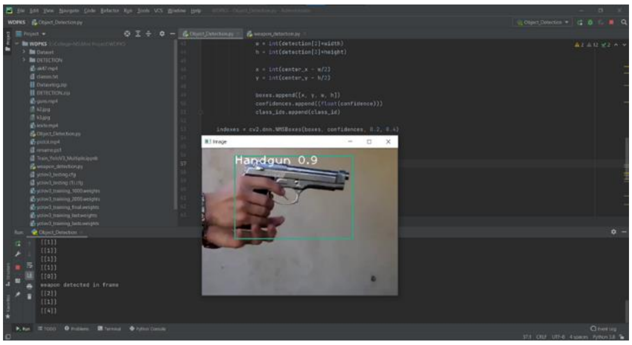


Fig.15 Implementation using OpenCV

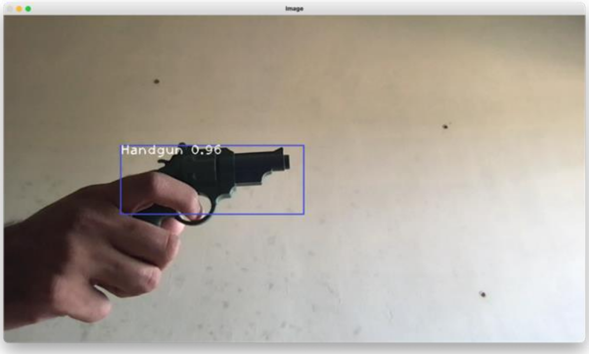


Fig.16 Handgun detected from webcam

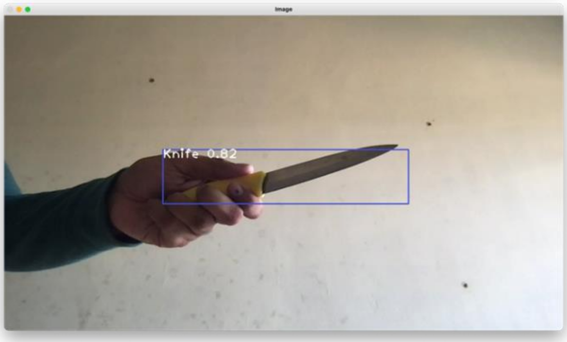


Fig.17 Knife detected from webcam

Fig.18 Multiple knives detected from an internal video

Fig.19 Heavy gun detected from an internal video

Once the weapon is detected in the frame, the program gives an alert sound and a message to indicate the presence of a weapon in that frame. This can be useful to police officers who are constantly on patrol to make them aware of the weapon in the video stream.

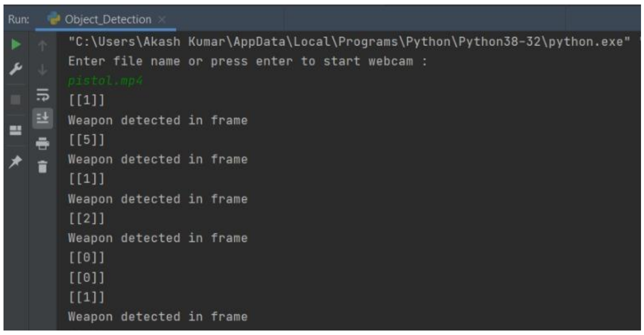


Fig.20 The weapon detected in the frame is shown along with an alert sound to indicate it

V.RESULTS

The results of the detected weapon from the video frame for each of the 3 classes of weapons – Handguns, Knives, and heavy guns with the accuracy and type of detected weapon is formulated in a table. Using this we can infer that our weapon detection system is fairly accurate to detect weapons in real-time surveillance videos.

TRAINING SET	TESTING SET	ACCURACY	RESULT
    		Handgun 95%	DETECTED
		Handgun 88%	DETECTED
		Handgun 87%	DETECTED

Fig. 21 Result of detection of Handguns

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

TRAINING SET	TESTING SET	ACCURACY	RESULT
    		Knife 75%	DETECTED
		Knife 87%	DETECTED
		Knife 85%	DETECTED

Fig.22 Result of detection of Knife

TRAINING SET	TESTING SET	ACCURACY	RESULT
   		Heavy Gun 84%	DETECTED
		Heavy Gun 80%	DETECTED
		Heavy Gun 82%	DETECTED

Fig.23 Result of detection of Heavy guns

VI.CONCLUSION AND FUTURE WORK

The weapon detection in surveillance system using yolov3 algorithm is faster than the previous CNN, R-CNN and faster CNN algorithms. In this era where things are automated, object detection becomes one of the most interesting field. When it comes to object detection in surveillance systems, speed plays an important role for locating an object quickly and alerting the authority. This work tried to achieve the same and its able to produce a faster result compared to the previously existing systems.

The longer-term work of the proposed system is to extend a greater number of types of weapons and classifying them. The accuracy of the weapon detected can be improved by using different types of algorithms. A possible way to improve this work is to detect a concealed weapon which cannot be detected using the normal camera. Also, analyzing the behavior of the people to find any suspicious activities like hiding the weapon can be done to improve this surveillance system. The alert system can also be improved to notify multiple users if the weapon is detected. A surveillance system with these features can be helpful to prevent violent crimes and provide security to the public.

VII.REFERENCE

[1] JIANYU XIAO SHANCANG L, (Member, IEEE), QINGLIANG XU11 School of Computer Science and Engineering at Central South University, China"Video-based Evidence Analysis and Extraction in Digital Forensic Investigation", IEEE Access (2020).
[2] JEONGIN SEO AND HYEYOUNG PARK School of Computer Science and Engineering, Kyungpook National University, "Object Recognition in Very Low Resolution Images Using Deep Collaborative Learning",IEEE Access (2020).
[3] Harsh Jain, Aditya Vikram, Mohana, Ankit Kashyap, Ayush Jain Telecommunication Engineering, "Weapon detection using AI and deep learning for security application", IEEE (2019).
[4] Shenghao Xu School of Science and Technology, The Open University of Hong Kong, Hong Kong, China, "Developement of AI based system for automatic detection and recognition of weapons in surveillance video", IEEE (2020).
[5] Dongdong Zeng, Xiang Chen, Ming Zhu, Michael Goesele and Arjan Kuijper, "Background Subtraction with Real-time Semantic Segmentation", IEEE (2019).
[6] Edge detection based boundary box construction algorithm for improving the precision of object detection in YOLOv3 Shaji Thorn Blue, M. Brindha Department of Computer Science and Engineering National Institute of Technology Tiruchirappalli, India, IEEE 2019
[7] Efficient Object Detection Method Based on Improved YOLOv3 Network for Remote Sensing Images Jintao Wang, Wen Xiao Key Laboratory of Unmanned Aerial Vehicle Development & Data Application of Anhui Higher Education Institutes 2Maanshan Wireless Sensor Network and Intelligent Perception Engineering Technology Research Center 3Wanjiang University of Technology Maanshan, China, 2020
[8] An Improved YOLOv3-based Neural Network for De-identification Technology Ji-hun, Dong-hyun Lee, Kyung-min Lee, Chi-ho Lin Won School of Computer Semyung University Chungcheongbuk-do, Korea,IEEE 2018
[9] Amrutha C.V, C. Jyotsna, Amudha J. Dept. of Computer Science & Engineering, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video", IEEE Xplore (2019).
[10] HYESEUNG PARK , SEUNGCHUL PARK, AND YOUNGBOK JOO, "Detection of Abandoned and Stolen Objects Based on Dual Background Model and Mask R-CNN", IEEE Access (2020).
[11] Maddula J N V Sai Krishna Asrith, K Prudhvi Reddy,Mrs. Sujihelen, " Face Recognition and Weapon Detection from Very Low Resolution Image", ICETITER (2018).
[12] Jose L. Salazar González a, Carlos Zaccaro a, Juan A. Álvarez-García a,∗,Luis M. Soria Morillo a, Fernando Sancho Caparrini, "Real-time gun detection in CCTV: An open problem", ELSEVIER (2020).
[13] Collins Ineneji, Mehmet Kusaf, "Hybrid weapon detection algorithm, using material test and fuzzy logic system",ELSEVIER (2019).
[14] Francisco Luque Sanchez, Isabella Hupont, Siham Tabik,Francisco HerreraF, "Revisiting crowd behavior analysis through deep learning: Taxonomy anomaly detection, crowd emotions, datasets, oppurtunities and prospects.", ELSEVIER (2019).
[15] Zhong-Qiu Zhao, "Object Detection With Deep Learning: A Review", IEEE (2018)