Original Article

SYSTEM FOR DETECTING POTENTIAL WEAPON THREAT ON SURVEILLANCE

Pranjal Tripathi, Priya Chaudhary, Kumari Ritika, Vijay

Department of Computer Science and Engineering Meerut Institute of Engineering and Technology, Meerut, U.P., India.
{pranjal.tripathi.cs.2019, priya.chaudhary.cs.2019, kumari.ritika.cs.2019}@miet.ac.in, vijay.movfrose@gmail.com

DOI: 10.47750/pnr.2022.13.S07.732

Abstract

According to numerous statistics, it can be inferred that the rate of violence using dangerous firearm use is increasing annually, preventing law enforcement agencies from quickly addressing this issue. There are several locations where there are high rates of crime with Handguns, particularly in areas with lax gun restrictions. For the safety of the public, the early detection of violent crime is crucial. Using surveillance footage to spot the presence of dangerous weapons like Handguns can help avoid these situations. Currently used surveillance and control systems still need manual oversight and management. For monitoring and control reasons, the proposed technology can automatically identify the weapons in the video. Real-time video weapon detection is done using the YOLOv3 method. Utilising prior techniques, such as R-CNN , which may be time-consuming and challenging to optimise, millions of network assessments must be carried out in order to forecast a single image. This method teaches every element individually while concentrating on a particular region of a picture. In contrast, The YOLO architecture only transmits the picture once over the neutral network. We used the YOLOv3 algorithm since real-time video requires speed. Authorities will be alerted when a weapon is discovered so they may respond appropriately and Before violent crimes occur, stop them.

Keywords: Weapon detection, Surveillance system, YOLOV3

1. Introduction

Since crime rates increase in congested or suspicious-looking remote regions, security is typically the top issue in all companies. Right now, protecting person's existence from mass shootings is a concern. The most fundamental human right, the right to life, is at danger in response to mass shootings. The misery of mass shootings affects people's lives all across the world on a regular basis. And over 500 people pass away as a consequence of mass shootings each day. One of main causes of the growing crime rate in America continues to be the easy access to guns. The United States has a very strong gun culture with a long historical background. Approximately 249 million weapons, roughly one-third of which are handguns, are in use in America.

Shootouts result in 50,000 deaths on average every year, including 12,000 homicides. The domestic handguns purchased for self-defense have also been proven in tests to be relatives are more likely to be hurt than saved. There's many variations since India has some of the tightest firearms prohibitions in the world. In this country, access to weapons is a privilege instead of a constitutional right. The 2016 Arms Rules stipulate that permits are necessary, even for small weapons. But getting a licence is a difficult process that might take weeks. Only after a thorough review that includes background checks are they issued. It is challenging to estimate the number of weapons that are unlawfully in possession, but examining the licencing status of earlier weapon seizures gives a decent idea of the scale of the issue. This raises serious concerns about the impact of these lethal weapons on public safety [1]

Due to the growing need for the protection of health, privacy, and personal land, the demand for the installation of video surveillance that can identify and analyze the picture and also abnormal events are necessary [2].

2. Literature Review

It is challenging worldwide to provide adequate security and reduce behaviour that endanger lives. In order to monitor various actions and behaviour, multiple researchers have contributed to using object detection. A intelligent video system's design frequently has 3 dimensions: first, to harvest limited data for object detection and tracking and feature extraction; second, to detect unusual human behaviour, behaviour, or detection of any weapons; and third, to make judgments such as unfortunate event detection [3].

Harsh Jain uses SSD and Faster RCNN convolution neural network (CNN) methods to develop automated gun (or) weapon identification. Two different dataset types are used in the suggested implementation. One dataset had photographs that were already labelled, and the other includes a collection of images that needed to be manually labelled. Both methods produce high accuracy in the results tabulated, but their practical use may depend on the trade-off between time and precision [4].

Shenghao created a weapon identification system based on the SSD’s, a well-known item recognition method , MobileNet and the High-level features are produced using CNN. An open-source machine learning framework is called TensorFlow.

Verma has also used the Faster RCNN model and deep learning to identify weapons. Imfdb, which I believe is a poor platform for training models for real-world settings, was employed for the assignment. They claimed to have a 93.1% accuracy on that dataset, but for the goal of identifying weapons, even higher accuracy is insufficient; precision and recall must also be taken into account. Siham Tabik et alwork .'s was directly related to the current state of affairs.

JEONG SEO AND HYE YOUNG PARK jointly train two deep neural networks in their framework for object detection in very low-resolution photos: To improve very low-resolution pictures into crisper, more informative images, The collaborative learning signals from the object identification network are utilized by the suggested picture enhancement network. Furthermore, it improves its training development by making use of the image enhancement channel's outputs to visual recognition with incredibly low definition more accurately.

Researchers have shown that the YoloV3 approach for real-time videos performs significantly better than other object recognition methods [5].

3. Methodology

The YOLOv3 algorithm is used in this suggested system to identify firearms in the manner that follows.

A dataset is first built that includes photos of weapons: Handguns. The YOLOv3 technique is used to train this dataset to categorise weapons. The system can identify the form of firearm that are used in simulated and actual video from the surveillance cameras together with the probability value of every firearm as once information has indeed been processed.An alarm will be issued to the authorities if the weapon is found [6].

Fig.1(Proposed Architecture)

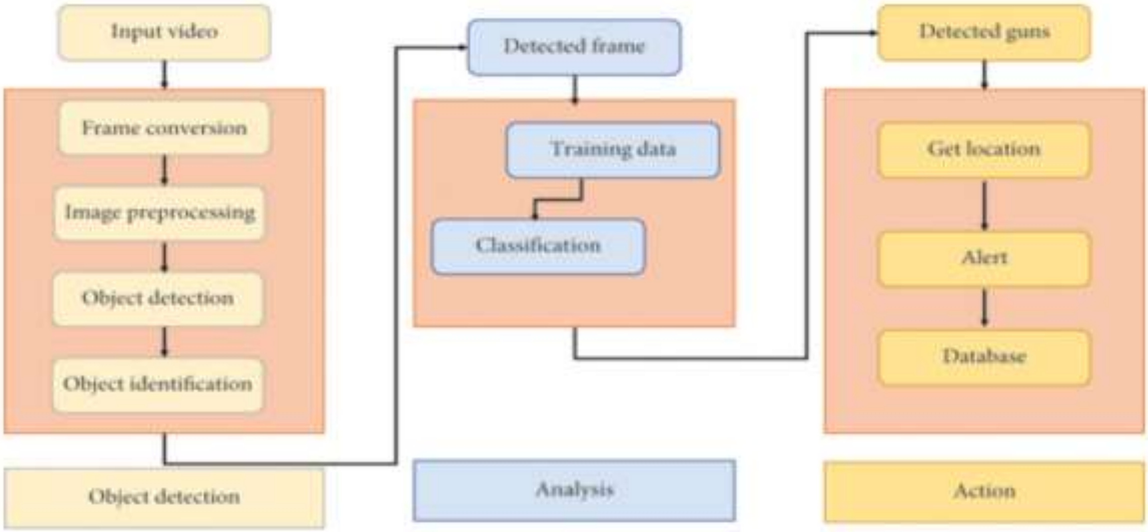
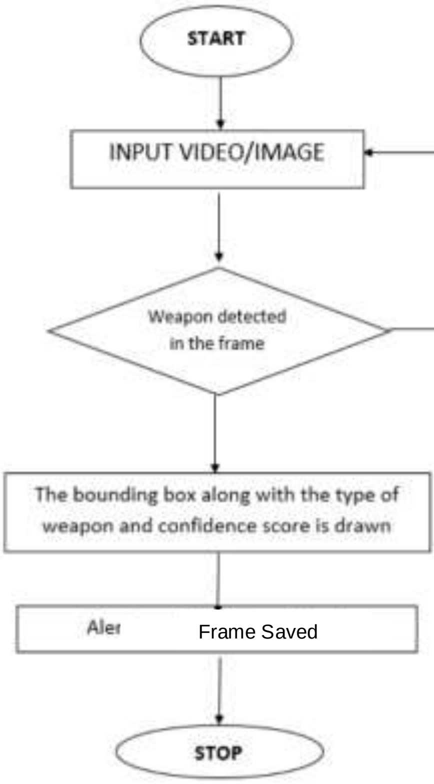


Fig.2(Flow Chart)



1. Dataset- Having a desired and appropriate dataset is essential to any application since it allows machine learning models to be trained. The dataset is developed by gathering high-quality weapon photos and preparing them for dataset production. The input datasets will also be taken from live CCTV surveillance. The placements of the weapons were documented on the images along with the different kinds of Handguns. The measurements for such indicators then generated for every image and stored in a text file [7][8].

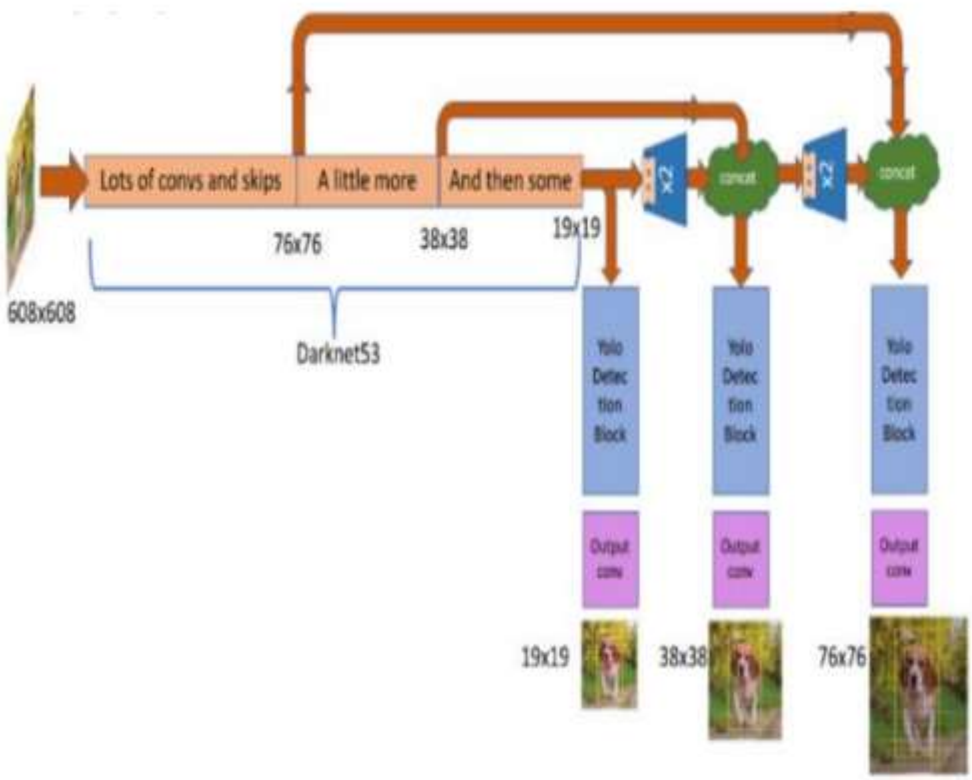
Fig.3(Input Datasets)

2. CNN-ModelConvolutional neural networks, often known as CNNs or ConvNets, excel in processing input like a cross grid pattern, kind of images A digital image is a digital depiction of image input. CNN is the most used deep learning technique for applications involving object recognition. This method performed admirably in the 2012 ImageNet Large-Scale Visual Recognition Competition. Since then, it has been applied to create new models for a variety of businesses [9].

3. YOLOV3 Algorithm - In movies, live streams, or still photographs, the YOLOv3 real-time object identification system can identify specific objects. Hundreds of network evaluations must be carried out in order to forecast a single image which may be tedious and hard to optimise. Object localization and feature extraction were combined into a single monolithic block in YOLOv3. They achieve an extremely quick inference time with their single-stage design, known as YOLO (You Only Look Once). In a single instance, it uses the entire image to forecast considering category chances & thresholded locations for such containers. YOLO's remarkable speed, which can analyse up to 45 frames per second, is by far its biggest advantage.

YOLO scans the full image through a convolutional neural network in one pass, as opposed to earlier methods that scanned images using a sliding window [10][11].

Fig.4(YOLO Design)



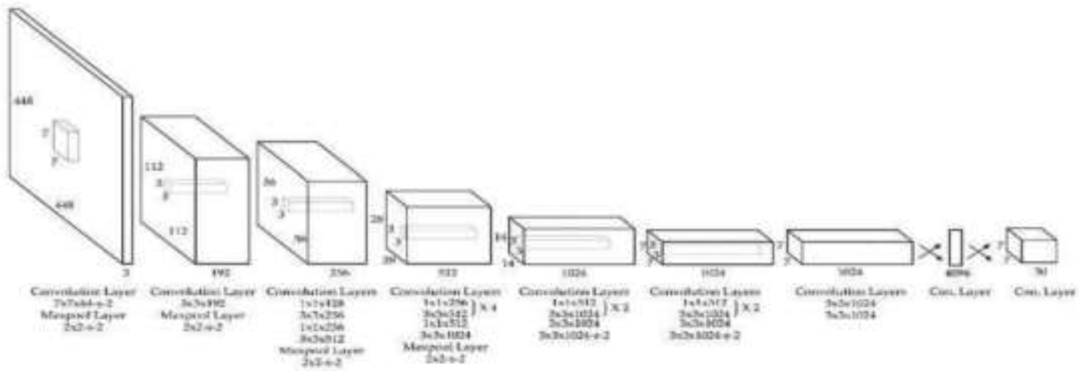
3.1. Architecture of YOLO

On a single semantic web, we combine the many components of object recognition. The webbing represents each bounded container by making use of all the attributes and traits present across the whole image. Additionally, this simultaneously displays every bounded container for a photo. This is a reference to the worldwide discussion on our website about the complete images and each and every thing in them. While maintaining the better overall accuracy, the YOLO sketch enables widespread education and real-time quick pace.

The image was divided into S x S frames by the system. If a framework cell contains the midpoint of an entity, then that framework cell will be the basis for recognising that entity.

The B bounding containers and specific levels for each container are displayed in each framework cell. These detailed levels demonstrate how precise the drawing is that the container portrays an entity and how specific it believed it was. The five depictions—x, y, w, h, and that specific level—are included in each bounding container. The midpoint of the container associated to the bound of the framework entity is shown by the lines X and Y. In relation to the whole picture, the width and depth are shown [12][13].

Fig.5 (YOLO Architecture)



3.2. Design

The Google Net design for photo classification serves as an inspiration for our model construction. Following the first two completely interlinked levels in our concept are twenty-four stated levels. Thanks to the distinctive modules utilised by Google Net, we may also employ 1 x 1 shrinkage levels and 3 x 3 completely linked levels. A quick version of YOLO developed to go beyond quick entity detection is another technique we train. Rapid YOLO employs a semantic web with fewer levels (nine as opposed to twenty-four), and these levels include less description overall. The only difference between YOLO and Rapid YOLO's teaching and checking is the web's dimensions. Our result is a tensor of predictions with dimensions 7 by 7 by 30. We keep gathering information at the levels we've set for the Image Net 1000 class. A meanpooling level, a connected level, and the first twenty particular levels from the image must be utilised for connection. After then, the model is altered to perform recognition. By combining the particular and related levels, the levels' accuracy may be increased.The precision of the levels can be improved by adding the specific and related levels together. We include 2 linked levels with randomly begun weights and 4 convolutional layers. We can increase the i/p dimensions of the web from 225 × 225 to 449 x 449 since recognition often requires fine-grained optic features. Class chances and bounding container locations are communicated by the output level. In order for it to lie between zero (0) and one (1), simplify the enclosing container's breadth, depth, and picture length (1).

3.3. Drawbacks of YOLO

Since every framework cell only informs about two containers and has the potential for a unit class, YOLO applies strict dimensional parameters to the bounding containers depicted. Smaller-dimensional things are difficult to anticipate. Because every framework is capable of predicting two bounding containers, it struggles to anticipate closer things. In comparison to the Faster R CNN model, it has a significantly lower recall and more localization errors [14][15].

3.4 Difficulties with Object Classification

Item recognition methods serve as a grouping of object positions and image categorization. It retrieves an image as input and outputs one or more bounding containers with a class flag associated with each one [16].

1. In object recognition, the bounding containers are consistently rectangular in form. Therefore, if the item has the curvature portion, it does not require assistance in anticipating the form of the object.
2. Some metrics, such an item's size or perimeter, could not be determined from an image using object recognition.
3. **Image Partition :** Image Partition is a further development of object detection in which we identify an object's presence using pixelwise spots created for each object in the image. This method uses more grains than the construction of bounding containers since it allows us to identify the shapes of each object in the image. [6]

This granularity is useful in a variety of applications, including satellite projection and the analysis of medical images.

3.5 Picture Segmentation: Bounding containers for the class in the image are created by the object recognition construct. However, it doesn't mention the entity's dimensions. Only the array of bounding container locations is required. We need to gather more intelligence, and this is not what we need at all. Every entity in the image receives a pixel-wise shroud as a result of picture categorization. This strategy offers the entity(s) in the picture a more expanding build [17].

3.6 Working: Along with the parameters, settings, and conditions needed for training & validate the proposed, the prototype also includes several computer programs and execution aids.

1. **Input:** Live CCTV footages are feeded to system for providing input to the system . In this experiment, identical video samples and original videos are used. The data set is also comprised of weapon images uploaded to system including pistol and Handguns along with their boundary box files.
2. **Processing:** The stages of implementing CNN and YOLOv3 algorithms are different, So the sizes of images was reformed to 240x240 for CNN using PIL image Library and for YOLOv3 the images are reformed to 416x416 which is perfect for this.
3. **Test and Output:** 668 datasets pictures are used to test the model in which 329 are of weapons and rest are not. These datasets the are feeded to CNN model after resizing to 240x240 pixels and the CNN model is trained. 0.05s is used by CNN model to get trained. For YOLOv3 model the images are resized to 416x416 pixels and it took average time of 0.10 sec to process the result.
4. **Frameworks used :** This study took use of Google Collaboratory, a system with a Google platform that enables followers to contribute basic documents and notes via Google Drive, run Python programmes in the browser, and gain full access to GPUs. The machine learning community frequently uses Colab for tasks like working with TPUs, model training, or Tensorflow, among others.

Google drive is a free google platform for storing files and provide storage. It was used to store datasets and Google Colab is used to provide datsets to code [18][19].

4. Result

A table is created with the accuracy and kind of weapon detected for each class of weapons, including handguns, from the video frame results. This shows that our method of finding guns in real-time surveillance footage is quite assured.

Table.1 (Result using CNN and YOLOv3)

Reference	Dataset	Model	TP	FN	TN	FP	Precision (%)	Recall (%)	F1 (%)	Time per sec
Proposed Model	Image dataset	YOLOv3	635	115	747	3	99.5	84.6	91.4	0.011s
Proposed Model	Video of 625 frames per sec	YOLOv3	434	123	14	56	77.9	88.5	82.8	0.010s
Proposed Model	Image dataset	CNN	304	0	247	57	84.2	100	91.4	0.19s

Proposed Model	Video of 625 frames per sec	CNN	467	-	-	11	98.7	37.5	74.3	0.19s

TP: True Positives

FN: False Negatives

TN: True Negatives

FP: False Positives

Fig.6(Result of detection of Handguns)

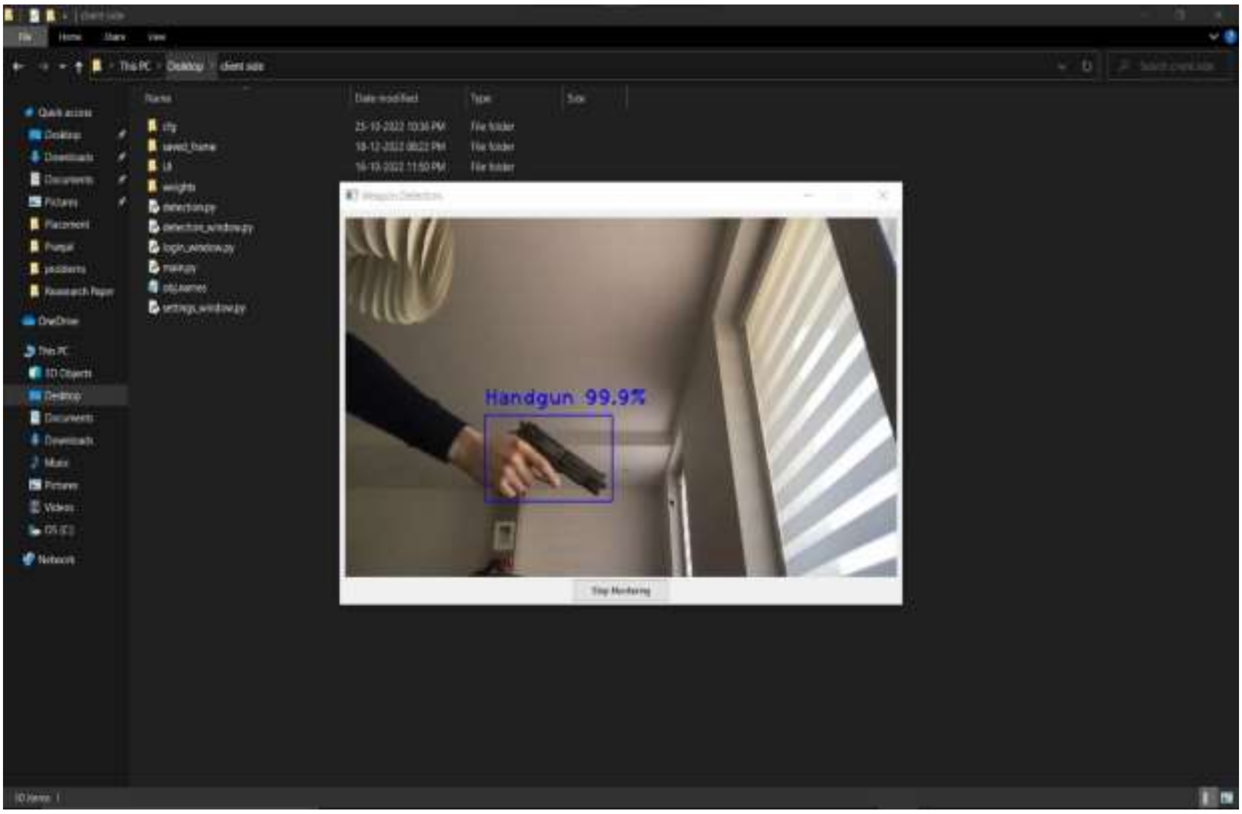
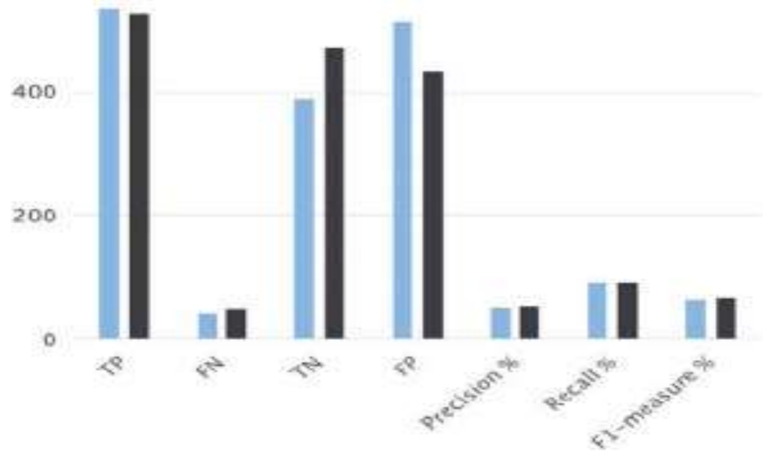


Fig.7(Frame Saved by the Detection System)

Fig.8 (Comparison of YOLO and CNN)



Comparison of execution time of CNN and YOLOv3 Algorithm

Fig.9(Speed Comparison of CNN and YOLOv3)

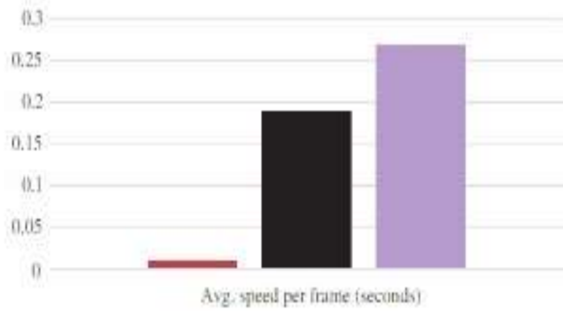


Fig.10(An Example of accurate weapons detected in a frame)

5. Conclusion

For identifying weapons in surveillance systems, the yolov3 algorithm surpasses the old R-CNN and CNN methods . Object detection is one of the most fascinating fields in today's automated world.Speed is crucial for object recognition in monitoring systems for promptly identifying a thing and alerting the authorities. Our effort was quicker than previously employed systems in achieving the same objective. The order of assist monitoring the efficiency of human operators, it is urgently necessary to upgrade the present surveillance capabilities, As reduced memory, visual network, and cutting-edge camera computational power are much more readily available, intelligent monitoring systems will completely replace the present architecture. Whenever poor computing, multimedia facilities, high - tech, and speedier video processing become accessible, digitally detection systems inside this form of robots would completely replace the present surveillance systems. The recommended method's long-term objective is to extend the classification of different types of weapons. Various sorts of algorithms may be used to increase the accuracy of the weapon detection. Finding a hidden weapon that cannot be seen by the standard camera might help to enhance this job. To enhance this monitoring system, it is also possible to examine people's behaviour for any questionable activity, such as hiding a weapon. If a weapon is found, the alarm mechanism can be enhanced to alert numerous people. These qualities in a surveillance system can help deter violent crimes and guarantee public safety.

6. References

[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] Narayan, Vipul, and A. K. Daniel. "Design Consideration and Issues in Wireless Sensor Network Deployment."." *Invertis Journal of Science & Technology* (2020): 101.

[3] Pramanik, Sabyasachi, et al. "A Novel Approach Using Steganography and Cryptography in Business Intelligence." *Integration Challenges for Analytics, Business Intelligence, and Data Mining*. IGI Global, 2021. 192-217.

[4] Irfan, Daniyal, et al. "Prediction of Quality Food Sale in Mart Using the AI-Based TOR Method." *Journal of Food Quality* 2022 (2022).

[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] Narayan, Vipul, and A. K. Daniel. "Energy Efficient Protocol for Lifetime Prediction of Wireless Sensor Network using Multivariate Polynomial Regression Model." *Journal of Scientific & Industrial Research* 81.12 (2022): 1297-1309.

[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] Awasthi, Shashank, et al. "A Comparative Study of Various CAPTCHA Methods for Securing Web Pages." *2019 International Conference on Automation, Computational and Technology Management (ICACTM)*. IEEE, 2019.

[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] Narayan, Vipul, and A. K. Daniel. "CHOP: Maximum coverage optimization and resolve hole healing problem using sleep and wake-up technique for WSN." *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal* 11.2 (2022): 159-178.

[13] Paricherla, Mutyalaiah, et al. "Towards Development of Machine Learning Framework for Enhancing Security in Internet of Things." *Security and Communication Networks* 2022 (2022).

[14] Francisco Luque Sanchez, Isabella Hupont, Siham Tabik,Francisco HerreraF, "Revisiting crowd behavior analysis through deep learning: Taxonomy anomaly detection, crowd emotions, datasets, opportunities and prospects.", ELSEVIER (2019).

[15] Zhong-Qiu Zhao, "Object Detection With Deep Learning: A Review", IEEE (2018)

[16] NARAYAN, V., Daniel, A. K., & Chaturvedi, P. (2022). FGWOA: An Efficient Heuristic for Cluster Head Selection in WSN using Fuzzy based Grey Wolf Optimization Algorithm..

[17] Smiti, P., Srivastava, S., & Rakesh, N. (2018, January). Video and audio streaming issues in multimedia application. In *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 360-365). IEEE.

[18] Srivastava, S., & Singh, P. K. (2022). HCIP: Hybrid Short Long History Table-based Cache Instruction Prefetcher. *International Journal of Next-Generation Computing*, *13*(3).

[19] Srivastava, S., & Singh, P. K. (2022). Proof of Optimality based on Greedy Algorithm for Offline Cache Replacement Algorithm. *International Journal of Next-Generation Computing*, *13*(3).