World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Research Article)

Enhanced object detection in videos using bio vision

Rizwana Kallooravi Thandil * and Umira Parvi KP

Sullamussalam Science, College Areekode, University of Calicut, India.

World Journal of Advanced Research and Reviews, 2024, 24(01), 1720–1727

Publication history: Received on 02 September 2024; revised on 06 October 2024; accepted on 19 October 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.1.3121

Abstract

Video surveillance systems enhance security and monitor various environments. They act as a deterrent to potential threats and provide vital surveillance to maintain safety. A key component of these systems is object detection, which identifies objects of interest and generates critical data for tracking and analysis. However, detecting objects in low-light conditions is particularly challenging due to reduced visibility, low contrast, and shadows, which can obscure objects or cause them to blend into the background. To address these challenges, this research introduces an enhanced object detection model that combines bio-inspired vision techniques with traditional object detection methods. The YOLO model is augmented with bio-vision principles, and specialized CNN models are developed to improve detection accuracy, particularly in low-light scenarios.

Keywords: Object Detection; CNN; Bio-Vision; Video surveillance systems; Neural Networks

1. Introduction

Video surveillance systems are essential today, enhancing security, protecting public safety, and improving operations in various settings. One key advancement in these systems is object detection, which allows for the automated identification and tracking of objects in video feeds. However, detecting objects in low-light environments is particularly challenging due to poor visibility and low contrast.

This research aims to develop an improved method for detecting objects in low-light conditions. The study focuses on five main types of objects important for security: humans, animals, vehicles, fire, and weapons. We propose a deep learning approach to detect these objects using different models tailored to each category. Humans, animals, and vehicles are detected using the YOLO model, which is trained on the widely used COCO dataset for better accuracy. To boost detection in low-light situations, we incorporate bio-inspired vision techniques into the YOLO model. These techniques involve image processing methods that enhance the model's ability to detect objects when lighting is poor. The model is tested using the ExDark dataset, which includes various low-light images.

A dedicated Convolutional Neural Network (CNN) model is used specifically for fire detection, designed to identify fire instances accurately. For weapon detection, we use both YOLO and CNN models; YOLO first identifies potential weapons, and the CNN model helps refine these predictions for better accuracy. This combination of models uses their strengths to improve the detection and tracking of objects across different environments.

2. Literature Review

The rapid advancement of video surveillance systems has significantly improved security, public safety, and operational efficiency in various sectors. A key component of these systems is object detection, which enables automated identification and monitoring of objects within video feeds. However, the challenges of detecting objects in diverse and

* Corresponding author: Rizwana Kallooravi Thandil

complex environments, such as low-light conditions, have driven ongoing research into enhancing object detection models. Recent studies have focused on integrating deep learning techniques, such as CNNs, YOLO models, and bio-inspired vision approaches, to improve detection accuracy and performance. This literature review explores the current state of object detection and tracking technologies in video surveillance systems, highlighting the latest advancements and methodologies utilized to address challenges in real-time and low-light detection scenarios.

Chandan G., Ayush Jain, Harsh Jain, and Mohana applied Deep Learning and OpenCV for real-time object detection and tracking using SSD and MobileNets, enhancing object detection with frame differencing, optical flow, and background subtraction. Their model, trained on 21 object classes, showed reliable real-time performance using SSD [1].

M. Monika et al. proposed a two-step object detection approach using CNN for feature extraction and YOLOv8 for detection, significantly outperforming previous models like YOLOv5 and DenseNet on a diverse set of datasets, demonstrating notable advancements in real-time detection [2]. Muhammad Uzair et al. introduced a bio-inspired vision technique for small target detection at long distances, improving detection through statistical pixel-based background modeling and a multilayer bio-inspired mechanism. Their method demonstrated enhanced detection performance across various metrics [3].

V. Kakulapati et al. explored using OpenCV and night vision techniques for surveillance cameras, utilizing Haar Cascade and enhancing detection accuracy with the LBP and HOG descriptors. Their approach offered a cost-effective solution for intelligent video surveillance [4]. Kuo-Feng Hung and Kang-Ping Lin used a biologically inspired approach for low-light object detection, employing YOLOv8 with enhanced image processing techniques like gamma correction. Their model, ExDark with Dark Adaptation, significantly improved recognition rates in challenging conditions [5].

Khan Muhammad et al. developed a CNN-based system for fire detection in surveillance videos, using a model inspired by GoogleNet. Their experiments demonstrated the superiority of their approach over traditional methods and other CNN architectures [6]. Pasquale Foggia et al. introduced a real-time fire detection method that combines color, shape, and motion analysis, showing superior true positive rates compared to individual methods [7].

Sanam Narejo et al. developed a smart surveillance system focused on weapon detection using YOLOv3, demonstrating improved performance over YOLOv2 with reduced computational requirements [8]. Harsh Jain et al. studied weapon detection using CNN-based SSD and Faster R-CNN, highlighting that SSD offers faster real-time detection while Faster R-CNN excels in accuracy [9].

The reviewed literature highlights the evolution of object detection and tracking techniques in video surveillance systems, emphasizing the integration of advanced deep learning models like SSD, YOLO, and CNNs. Various approaches have demonstrated the potential to enhance detection accuracy, especially in challenging conditions such as low-light environments and real-time scenarios. The adoption of bio-inspired vision techniques and the incorporation of image enhancement methods further contribute to improved performance, particularly for detecting small, faint targets and objects under adverse conditions. Collectively, these studies underscore the importance of continued innovation and optimization of object detection frameworks to meet the growing demands of modern surveillance applications, paving the way for more robust and efficient security solutions.

3. Dataset Description

This research focuses on detecting five key object classes critical for security: humans, animals, vehicles, fire, and weapons. The primary dataset used is the COCO dataset, sourced from Kaggle, which is widely recognized in object detection research. The COCO dataset contains over 200,000 images annotated with more than 80 object categories, providing a comprehensive resource for training models in detecting a variety of objects.

For model testing under low-light conditions, the ExDark dataset was employed. This dataset, also sourced from Kaggle, is specifically designed for extreme darkness scenarios, containing over 7,000 images. Each image is carefully annotated to support object detection, recognition, and segmentation, making it ideal for evaluating model performance in challenging visual environments. The weapon detection task utilized two distinct datasets from Kaggle. The first dataset includes images annotated with three classes of weapons for detection purposes, while the second dataset contains images with seven different classes of weapons for prediction tasks. These datasets enable a detailed assessment of the model's capability to identify various types of weapons accurately.

For fire detection, datasets from Kaggle were used, consisting of images depicting both real fire incidents and normal conditions. These images are organized into separate folders labeled ‘0’ for non-fire scenarios and ‘1’ for fire conditions, allowing the model to be trained effectively in distinguishing between these contexts. This structured approach supports the accurate training and evaluation of the model for fire detection in surveillance systems.

4. Methodology

This section outlines the methodologies employed in developing an advanced object detection system tailored for security-critical applications. The primary aim was to detect and classify objects such as humans, animals, vehicles, fire, and weapons in real-time, with a specific focus on enhancing detection accuracy in challenging conditions, such as low-light environments. The system utilizes state-of-the-art deep learning models, including YOLOv3 (You Only Look Once) and Convolutional Neural Networks (CNN), integrated with bio-inspired image processing techniques to improve performance under adverse conditions. YOLOv3 was chosen for its speed and high accuracy in object detection tasks, while CNNs were specifically utilized for specialized detection tasks like fire and weapon recognition.

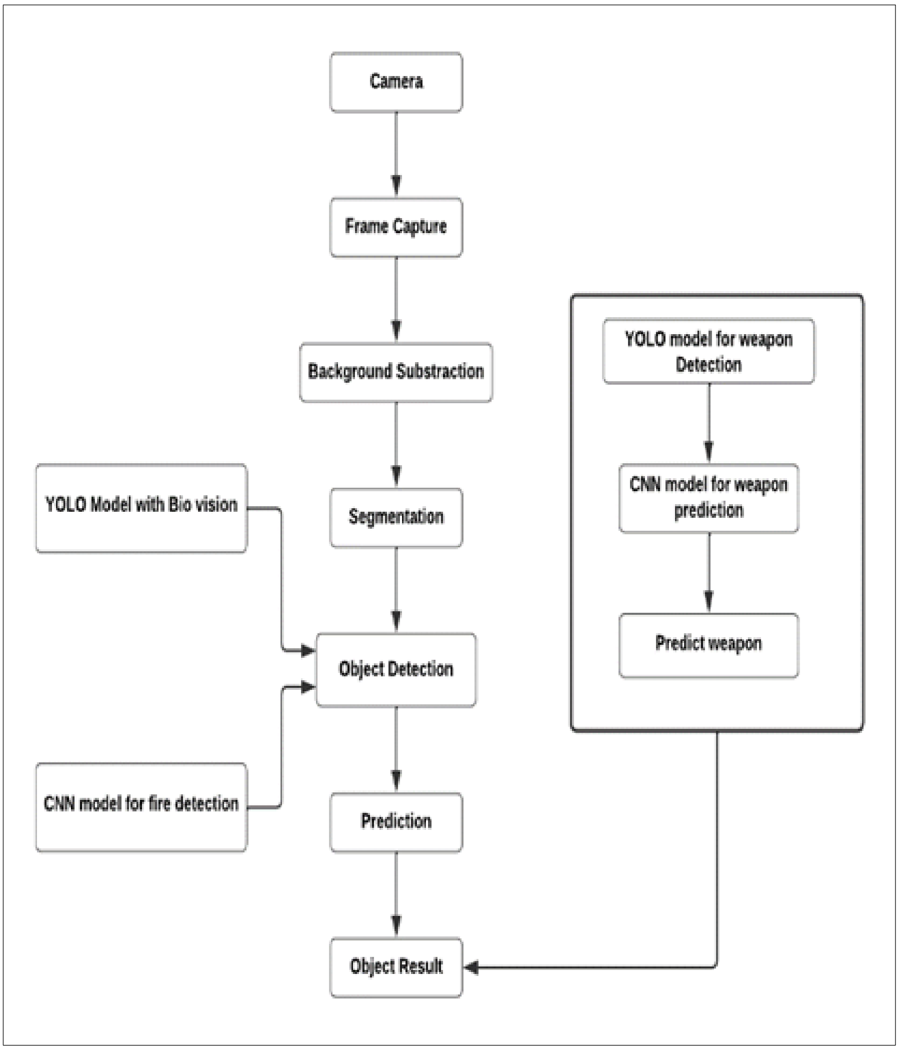

Figure 1 System Architecture

The methodology encompasses the training and testing phases of these models using diverse datasets, such as the COCO dataset for general object detection, ExDark for low-light scenarios, and specialized datasets for fire and weapon detection. This section will detail the integration of bio-vision principles into the YOLOv3 model, the architectural specifics of the CNN models, and the workflow of the overall system, highlighting the innovative approaches used to address the challenges posed by security-focused object detection in real-time surveillance. The combination of these advanced techniques aims to enhance detection reliability, accuracy, and efficiency in complex environments, ultimately contributing to the development of robust security systems.

The overall system architecture shown in Figure 1 begins with a camera capturing real-time frames, which are then processed for object detection tasks. The captured frames undergo background subtraction to isolate the objects of interest, followed by segmentation, which further refines the objects within the frame. This segmented output is then fed into distinct object detection models based on the task requirements, including human, animal, vehicle, fire, and weapon detection. The YOLO (You Only Look Once) model was selected for its high-speed and accurate object detection capabilities. YOLOv3, specifically, was trained using the COCO dataset, a large-scale dataset comprising over 200,000 images annotated with 80+ object categories, which provided a diverse set of objects for training purposes.

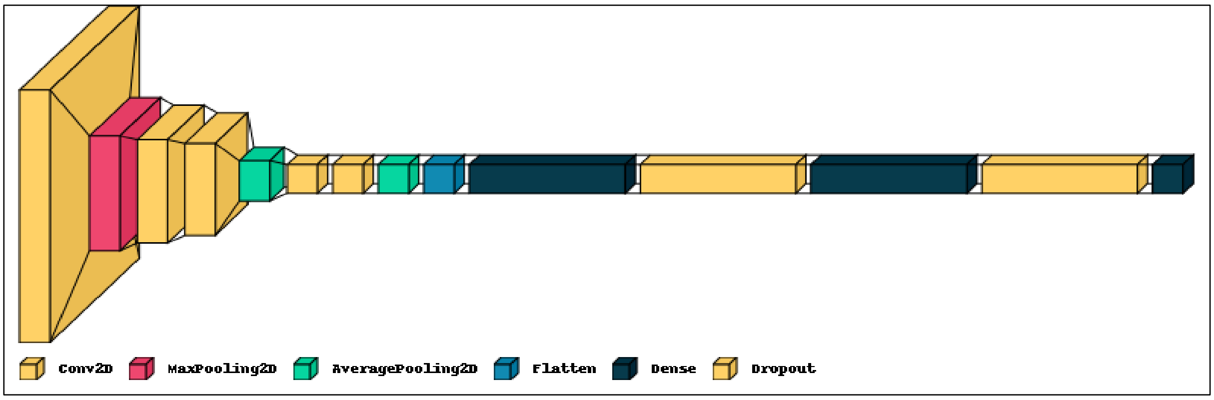

Figure 2 The CNN Architecture

To improve detection performance in low-light environments, bio-inspired techniques were integrated into YOLO. These techniques included image enhancement using contrast adjustment, smoothing with Gaussian blur to reduce noise, and image fusion to combine multiple processed images, enhancing the detection performance in challenging lighting conditions. The model's performance was then evaluated using the ExDark dataset, which focuses on extreme low-light scenarios to validate the bio-enhanced YOLO's efficacy.

Inspired by the biological visual system, bio-vision principles were applied to refine images in low-light conditions. Enhancement techniques, such as contrast enhancement and Gaussian blur, were used to elevate image quality, followed by image fusion to merge multiple processed images. This combined approach significantly improved detection performance, particularly in environments with limited lighting. The system utilized these enhanced images for detecting human, animal, and vehicle classes with high accuracy, as evidenced by the ExDark dataset evaluation.

A Convolutional Neural Network (CNN) model shown in Figure 2 was developed specifically for fire and weapon detection, utilizing multiple convolutional and pooling layers to extract and refine features from images. The model architecture begins with convolutional layers to detect low-level features, followed by pooling layers to reduce the spatial dimensions. A Flatten layer then converts the 2D feature maps into a 1D vector, which is processed through fully connected layers for classification. Dropout layers were included to prevent overfitting by randomly discarding neurons during training. The CNN model was trained using a specialized dataset containing images of fire incidents and achieved high detection accuracy. Table 1 shows the layered architecture of the CNN used in the study.

Table 1 The CNN Layers

Layer Type	Details
Input Layer	Input shape: (48, 48, 1)
1st Convolutional Block	2D Convolutional Layer: 64 filters, (5, 5), ReLU activation
	Max Pooling: pool size (5, 5)
2nd Convolutional Block	2D Convolutional Layer: 64 filters, (3, 3), ReLU activation
	Average Pooling: pool size (3, 3), strides (2, 2)
3rd Convolutional Block	2D Convolutional Layer: 128 filters, (3, 3), ReLU activation
	Average Pooling: pool size (3, 3)

Flatten Layer	Converts 2D feature maps into a 1D vector
Fully Connected Layer	2 Dense Layers: combines features learned by convolutional layers
	Dropout: 20% during training to prevent overfitting

Weapon detection within the system is handled by two models: YOLOv3 for initial detection and a CNN model for prediction and classification. The YOLOv3 model identifies potential weapons in the frame, while the CNN model further refines the detection by classifying the identified objects into specific weapon classes. The CNN model architecture, as shown in Figure 2, employs a combination of convolutional, pooling, and fully connected layers. The model was trained using datasets comprising multiple classes of weapons, ensuring robust weapon classification and detection.

5. Results and Discussion

5.1. Performance Evaluation of YOLOv3

The YOLOv3 model, enhanced with bio-vision techniques, was tested on the ExDark dataset to evaluate its performance in low-light conditions. The model achieved an accuracy of 96.51% over forty epochs, highlighting its ability to detect human, animal, and vehicle classes effectively even under extreme darkness. The integration of bio-inspired methods significantly improved detection accuracy, showcasing the potential of biologically inspired enhancements in computer vision tasks. Figure 3 shows the overall accuracy that was obtained for different models.

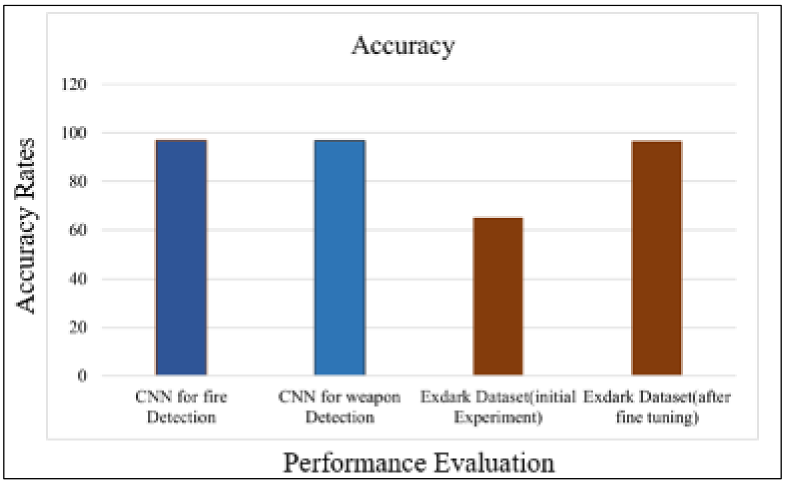

Figure 3 Performance Evaluation in Terms of Accuracy

5.2. Fire Detection Using CNN

The CNN model for fire detection underwent extensive testing with 20% of the specialized training dataset reserved for evaluation. The model demonstrated strong performance, achieving an accuracy of 96.75%. The architecture's use of multiple convolutional and pooling layers allowed it to effectively learn distinguishing features of fire incidents, confirming the model's reliability in detecting fire-related events.

5.3. Weapon Detection and Prediction

The combination of YOLOv3 for detection and CNN for prediction allowed the system to accurately identify and classify weapons. The testing phase involved 20% of the training data, and the model demonstrated high efficacy in distinguishing between various classes of weapons. This dual-model approach provided robust performance with an accuracy of about 96.62%, enhancing the system's overall security capabilities by accurately detecting and classifying potential threats. Figure 4 shows the overall loss that was obtained while training the models for detecting fire and weapons.

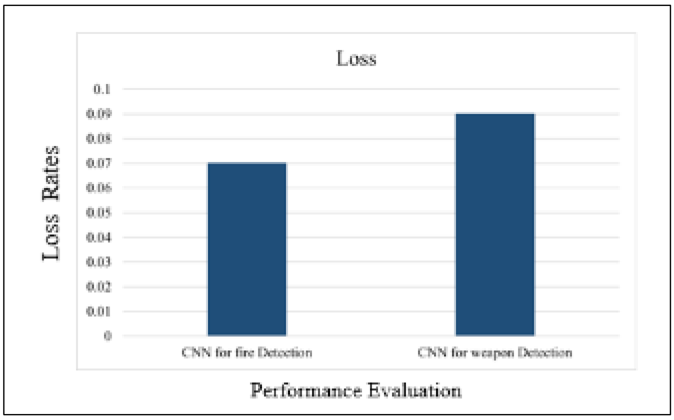

Figure 4 Performance Evaluation in Terms of Loss

Table 2 includes the comparison of this study with existing literature.

Table 2 Comparison with Existing Study

SI No.	Title	Year	Methodology	Accuracy
[2]	Real-time object detection in videos using deep learning models	2023	YOLOV8 & CNN	90% , 85%
[3]	Bio-Inspired Video Enhancement for Small Moving Target Detection	2021	Statistical pixel based techniques & Bio inspired multi layer mechanism	75.4%
[4]	A novelapproach object detection of video surveillance system using opencv	2024	LBPH, PCA & SVM	92%, 93%
[5]	Bio-Inspired Dark Adaptive Nighttime Object Detection	2024	YOLOV8	Above 55% (not explicitly mentioned)
[6]	Convolutional Neural Networks Based Fire Detection in Surveillance Videos	2018	CNN	94.43%
[7]	Real-time Fire Detection for Video Surveillance Applications using a Combination of Experts based on Color, Shape and Motion	2015	Expert based on color, shape, and movement	93.55 %
[8]	Weapon Detection Using YOLO V3 for Smart Surveillance System	2021	YOLOV3	98%
[9]	Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications	2015	SSD & Faster R-CNN	73.8%, 84.6%
[10]	Real-Time Vehicle and Pedestrian Detection Through SSD in Indian Traffic Conditions	2018	SSD	96%

[11]	A Video-Based Fire Detection Using Deep Learning Models	2019	Faster R-CNN & LSTM	88.3% , 95%
Proposed Study	Enhanced Object Detection in Videos Using Bio Vision	2024	Yolo v3 and CNN	Fire-96.75 Weapon-96.62 Yolov3 (initial)-65 Yolov3(after fine-tuning)-96.51

The integration of bio-vision techniques with traditional object detection models such as YOLOv3 demonstrated substantial improvements in detecting objects under challenging conditions, such as low light. The experimental results confirm the effectiveness of using enhanced image processing techniques to augment detection performance. Furthermore, the CNN models for fire and weapon detection proved highly effective, reinforcing the versatility and applicability of deep learning models in real-time surveillance systems. Future work could explore further integration of bio-inspired principles and hybrid models to tackle even more complex environmental challenges.

6. Conclusion

The advancement of enhanced object detection in video surveillance systems, particularly for security and public safety, involved the implementation of three specialized models, each designed to detect specific object classes. A YOLO model, informed by principles of bio-vision, was developed to effectively identify objects in low-light environments, resulting in improved detection accuracy and heightened security measures. Additionally, a dedicated CNN model was created for fire detection, which excels in identifying and flagging instances of fire within surveillance footage. For weapon detection, two models—CNN and YOLO—were developed, each with distinct roles: YOLO executes the detection task within the frames, while CNN handles the prediction task. This collaborative approach demonstrates a significant reliability in predictions, underscoring the effectiveness of the training process. Overall, the integration of these models greatly enhances security measures, representing a proactive strategy to improve the effectiveness of surveillance systems.

Future efforts will focus on expanding the range of objects our models can detect, incorporating a wider variety of categories to create a more comprehensive surveillance approach that addresses an array of potential threats and scenarios. Additionally, we will explore techniques for processing captured images to produce high-quality visuals, particularly in low-light or adverse weather conditions. This enhancement aims to improve detection capabilities and overall situational awareness, thereby ensuring greater security.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] G. Chandan, A. Jain, H. Jain and Mohana, "Real Time Object Detection and Tracking Using Deep Learning and OpenCV," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2018, pp. 1305-1308, doi: 10.1109/ICIRCA.2018.8597266.

[2] M. Monika, Udutha Rajender, A. Tamizhselvi and Aniruddha S Rumale, “Real-Time Object Detection In Videos Using Deep Learning Models” ICTACT Journal on Image and Video Processing · December 2023.

[3] M. Uzair, R. S. Brinkworth and A. Finn, "Bio-Inspired Video Enhancement for Small Moving Target Detection," in *IEEE Transactions on Image Processing*, vol. 30, pp. 1232-1244, 2021, doi: 10.1109/TIP.2020.3043113.

[4] Kakulapati, Vijayalakshmi & Potla, Sridhar & Sista, Mayukha & Chelukali, Ashish & Dhabale, Rutuja & Katta, Sriya. (2024). AN NOVEL APPROACH OBJECT DETECTION OF VIDEO SURVEILLANCE SYSTEM USING OPENCV. 24. 376-406.

[5] Kuo-Feng Hung and Kang-Ping Lin, "Bio-Inspired Dark Adaptive Nighttime Object Detection", https://www.mdpi.com/journal/biomimetics, Biomimetics 2024.

[6] Khan Muhammad, Jamil Ahmad, Irfan Mehmood, Seungmin Rho, and Sung Wook Baik, "convolutional neural networks based fire detection in surveillance videos", special section on multimedia analysis for internet-of-things, 2018.

[7] Pasquale Foggia, Alessia Saggese and Mario Vento, "Real-time Fire Detection for Video Surveillance Applications using a Combination of Experts based on Color, Shape and Motion", DOI 10.1109/TCSVT.2015.2392531, IEEE Transactions on Circuits and Systems for Video Technology.

[8] Sanam Narejo, Bishwajeet Pandey, Doris Esenarro vargas, Ciro Rodriguez, and M. Rizwan Anjum, "Weapon Detection Using YOLO V3 for Smart Surveillance System", Hindawi Mathematical Problems in Engineering Volume 2021, Article ID 9975700.

[9] Harsh Jain, Aditya Vikram, Mohana, Ankit Kashyap, Ayush Jain, "Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications", in Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020) IEEE Xplore Part.

[10] Seyed Yahya Nikouei, Yu Chen, Sejun Song, Ronghua Xu, Baek-Young Choi, Timothy R. Faughnan, "Real-Time Human Detection as an Edge Service Enabled by a Lightweight CNN", in Proceedings of the International Conference on Edge Computing (2018) IEEE.

[11] Byoungjun Kim, Joonwhoan Lee, "A Video-Based Fire Detection Using Deep Learning Models", MDPI, 2019.