International Journal of Computer Science and Engineering (IJCSE)
ISSN (P): 2278–9960; ISSN (E): 2278–9979
Vol. 13, Issue 2, Jul–Dec 2024; 19–28

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

ENHANCING VIOLENCE DETECTION ON SURVEILLANCE CAMERAS USING YOLOV7

Teja Krishna Sai Uppugunduri[1], Batlanki Freesia Grace[2] & Andluru Sai Punith Reddy[3]

[1,2,3]School of Computer Science and Engineering, Lovely Professional University, Punjab

ABSTRACT

In this research, a unique real-time gun identification technique based on one more YOLOv7 deep learning algorithm is presented. The primary objective of the system is to promptly identify firearms in live streams in order to enable prompt responses to potential security threats. When the gun is identified, the system dynamically extracts the data from the streaming video using the SMTP library and sends out an email alert with important details and images to support the claim. By harnessing the potent capabilities of YOLOv7, this technology achieves exceptional precision and efficacy in firearm recognition, strengthening security standards in a range of contexts. With its use of state-of-the-art object detection technology, this approach represents a significant advancement in proactive threat mitigation strategies. This violence detection system needs several key parts in order to integrate and function properly. First off, guns can be swiftly and latency-freely identified thanks to the YOLOv7 model, which is the basis for real-time object detection. With constant refinement and enhancement, the model performs remarkably well in identifying weapons of various kinds, dimensions, and orientations. Additionally, the system has sophisticated post-processing techniques to reduce incorrect detections and enhance the accuracy of firearm location. Advanced screening methods and contextual analysis also improve identification results, ensuring consistent alerts and minimizing unnecessary disturbances. The integration with the SMTP libraries also makes it possible to promptly notify security personnel or authorized authorities of any suspected firearm incidents. Based on all the information provided in the email alerts—including the location, time, and cropped photo of the discovered firearm—respondents can act quickly. All things considered, real-time gun detection technology enhances social safety and security by taking precautionary steps. The system provides a robust defense against the threats posed by weapons in numerous scenarios by employing advanced detection methods and YOLOv7. This research eventually aids ongoing efforts to develop effective surveillance and intervention methods by safeguarding communities and lowering the risks connected with firearm-related occurrences.

KEYWORDS: *YOLOv7, Deep Learning, Object Recognition, Security System, SMTP Email Alerts, Surveillance Technology, Proactive Safety Measures*

Article History

Received: 26 Jul 2024 | Revised: 30 Jul 2024 | Accepted: 31 Jul 2024

INTRODUCTION

The ever-changing landscape of security threats demands creative solutions for swift detection and removal, particularly in handling the increasing risks associated with the pervasiveness of guns. The "Notifying the authorities: YOLOv7-based Weapons Identification on Video" method is one creative response to this pressing issue. This novel approach addresses the urgent need for rapid weapon detection in video streams, bolstering security measures and enabling timely responses. It accomplishes this by making use of the complex YOLOv7 deep learning algorithm [5]. Being able to recognize and respond to threats of this nature is crucial in a setting where public safety is paramount. Traditional threat identification methods often rely on manual, error-prone human surveillance. Additionally, because security threats are evolving, greater adaptability is needed. Artificial intelligence (AI) technology presents unmatched opportunities to improve existing security frameworks and optimize threat detection processes. Because of its exceptional speed, precision, and versatility, the YOLOv7 detection system is a formidable competitor for real-time detection tasks. The primary objective of the project is to develop a weapon detection system that is both dependable and effective, especially for use in video surveillance scenarios. The system precisely locates and detects guns in real-time using the advanced algorithms of YOLOv7. When this happens, an alert system is triggered, giving authorities pertinent data, including the weapon's location and time stamp. The system's scalability is one of its key benefits; it enables a variety of deployment scenarios and easy integration with the existing security infrastructure [8]. To assist law enforcement in swiftly neutralizing threats and preserving public safety, the identification process can be automated. However, the experiment also highlights how important privacy rights and moral considerations are in the design and application of AI-powered surveillance technologies. Finally, the project "Notifying Police: YOLOv7-based Violence Detection on Video" is an important advancement in security technology. By employing YOLOv7 and AI-driven object identification to proactively identify firearms in video feeds, the method enhances safety for everyone in the face of evolving threat situations [3]. This initiative demonstrates how innovation may be utilized responsibly and transparently to address challenging security challenges while upholding ethical norms.

LITERATURE REVIEW

The primary objective of the study is to develop a real-time crime detection system that will swiftly spot any illegal activity by keeping an eye on CCTV video streams. By combining Conv LSTM's dynamic longitudinal modeling capabilities with YOLO v7's object identification ability, the system aims to significantly increase public safety. The first section employs Conv the LSTM with a light framework based on Mobile Net v2, which improves accuracy and efficiency in identifying aggressive postures. Since the system was trained on annotated surveillance footage that depicts an array of criminal acts, it can distinguish between aggressive and non-violent positions in real time. In addition, YOLO v7 detects weapons and classifies them into three groups: weapons, sticks, and sharp items. By using machine learning models to provide a comprehensive means of identifying and deterring illegal acts, this technology makes communities safer.

This proposed system aims to safeguard academic exam integrity. The system observes each student's behavior during the test and uses vision-based algorithms to spot any suspicious behavior. The recommended e-cheating detection procedure consists of these four crucial steps: 1) monitoring and recognizing individuals or groups of pupils; 2) spotting dubious activities; 3) sending out alerts; and 4) keeping track of attendance. The system classifies suspicious actions (such as swapping documents, sharing codes, etc.) using a CNN-RNN design with feature extraction based upon the inceptionV3 model. The technology records the student's attendance by recognizing and matching the faces of the children that are stored in the database. In the end, the proposed strategy might improve test integrity and create a fair learning environment for all students, which would elevate the bar for the educational system.

UAV detection is growing in popularity in domains like public safety because of the widespread usage of UAVs in factories and enterprises. As a result, algorithms for UAV object detection are also developing swiftly. Nevertheless, the small size of drones, complex airspace backdrops, and variable lighting still present significant challenges for this area of research. To enhance the device's capacity to identify small targets, an excellent-quality recognition head is initially added to the identification head component. Simultaneously, redundant networks and the large target recognition head are removed to minimize the number of network variables and speed up the detection of UAVs. Secondly, SPD-Conv is utilized instead of Conv to obtain multiple-scale characteristics during the feature extraction phase, minimizing the degradation of fine-grained data and enhancing the model's ability to extract features for small targets. Finally, the GAM attention module is added to the neck of the model to enhance its general efficacy in identifying UAVs and to improve its ability to merge target features. It reduces the number of parameters and the size of the model by 59.9% and 57.9%, respectively, in the interim.

This paper proposes YOLOV7-FM, an improved YOLOV7 deep learning recognition technique, to recognize pseudostems in bananas grown under different conditions. The loss optimization component of the YOLOV7 model is extended to include focal loss in order to improve the detection rate of challenging samples and optimize the challenging training for heavily shielded banana pseudostems. In this paper, the YOLOV7-FM method's AP (average precision) and inference time are contrasted with those associated with the YOLOX, YOLOV5, YOLOV3, and Faster R-CNN algorithms. Using pseudostems from bananas, the average inference time for each image is 8.0 ms, and the YOLOV7-FM technique surpasses the comparator model with respect to both AP and reasoning speed (81.45%). This improved YOLOV7-FM model can detect the banana pseudostem rapidly and precisely.

Limited but necessary movement can be provided by a wheelchair to an injured or disabled person. In order to create a smart wheelchair, a manual wheelchair can have its distinctive octascopic vision technology added to it. This paper outlines the first stage of this process. A 360-degree vision is provided by two monochromic lens arrays, each containing four cameras, mounted on each wheelchair's frame in this affordably priced autonomous wheelchair design. An autonomous wheelchair that could navigate an indoor environment both with and without human aid was the project's initial goal. It was to be controlled by an embedded CPU. Following testing of the wheelchair's functionality, (a) a large set of octascopic images was extracted from it, and (b) an object detection model based on YOLOv7 was developed to avoid obstacles and control its own mobility. Octascopic images are used in this study to illustrate the obstacle recognition model and camera placement. All design files may be reproduced publicly as long as they comply with the conditions of a license that is open-source.

RESEARCH METHODOLOGY

ALGORITHM

CNN

Convolutional neural networks (CNNs) are a popular class of deep learning techniques for image and video processing applications [9]. Through autonomous learning and feature extraction from incoming data, they are designed to effectively record spatial patterns. Among the layers that comprise a CNN are convolutional neural networks, pooling, and fully connected layers. To extract different types of information from input images at different spatial scales, CNN convolutional layers employ filters. As they go across the input, these filters perform element-wise multiplications to find patterns such as edges, textures, and shapes [10]. After that, the pooling layers sample the condensed feature maps, reducing the spatial dimension without losing any important information.

YOLO Object Detection

You Only Look Once (YOLO) is a state-of-the-art object detection technology that revolutionized computer vision by allowing the real-time, high-precision identification of numerous objects in images or video frames [1]. YOLO generates a structure of cells out of the picture and predicts the probability of classes and box boundaries for items in every cell of the grid simultaneously, in contrast to standard approaches that require multiple passes across the image [14]. Thanks to its one-shot approach, YOLO can achieve incredible speed without compromising precision. Owing to its effectiveness and efficiency, YOLO is widely utilized in applications including medical imaging, surveillance, and driverless automobiles where accurate and fast object recognition is crucial.

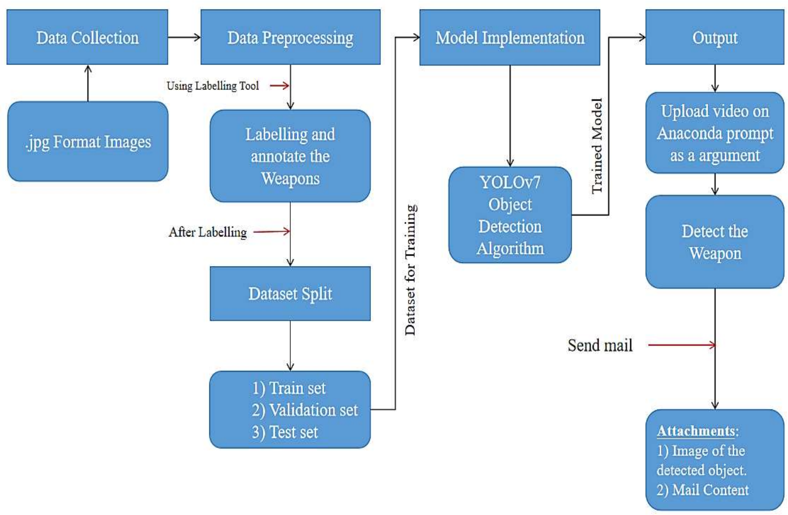

Figure 1: Methodology.

Data Collection and Annotation

By compiling a diverse library of weapon pictures from many sources, this module ensures a wide range of weapon sorts, orientations, and lighting conditions [2]. The bounding boxes are annotated using tools for annotation on the dataset, which then show the precise positions of weapons in each image. The annotated dataset serves as the foundation for training the YOLOv7 model.

Data Preprocessing

The dataset is processed beforehand before the model is trained in order to maximize training performance. To do this, the images must be consistently scaled, the pixel values must be standardized to a similar band (such as 0 to 1), and the dataset must be transformed with flips, shifts, and brightness adjustments [15]. These preprocessing steps improve the model's ability to recognize firearms in a variety of situations and aid in its improved generalization to fresh data.

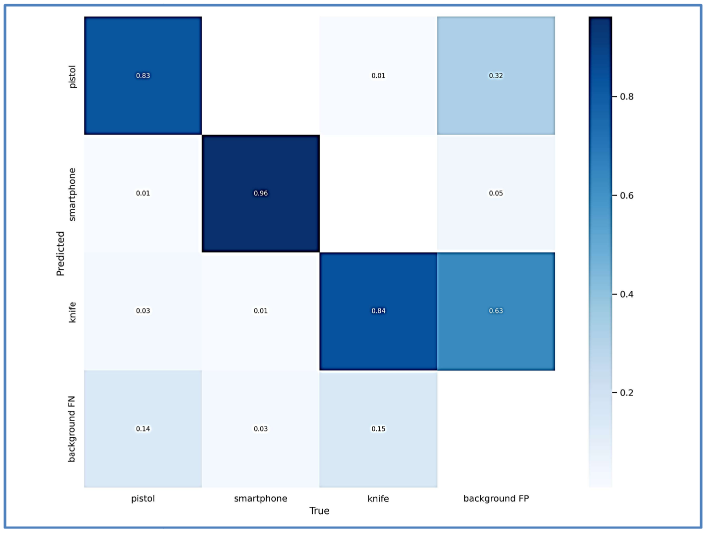

Figure 2: Confusion Matrix.

YOLOv7 Model Configuration

The module needs to set up the YOLOv7 algorithm's structure for weapon detection. Contemporary object recognition models, such as the YOLOv7 model, are well known for their speed and accuracy [19]. The model has the right number of categories (weapons, for example) and box anchors configured based on the properties of the dataset. Next, weights that were previously trained on a large dataset (such as COCO) are used to initialize the model, which speeds up training.

Training the YOLOv7 Model

The annotated weapons dataset is utilized to train the YOLOv7 system via transfer learning. Throughout training, the model learns to recognize violence by adjusting its weights based on the annotated dataset's ground truth [17]. Stochastic gradient descent (SGD) and the Adam optimizer are two techniques used for maximizing the model's weights during training. Training stops when the predicted outcome performs well enough, and data from validation is used to monitor the model's performance.

Validation and Evaluation

After training on a separate validation dataset, the resulting YOLOv7 model's performance is evaluated [18]. Evaluation measures, including as accuracy, recall, and F1 score, are calculated to evaluate the model's efficacy in weapon detection. The evaluation's findings are used to enhance the model's functionality.

Real-time Violence Detection

After being trained and validated, the YOLOv7 version is utilized to identify violence in real-time within video streams. By examining each frame of the video feed, the model is able to identify firearms with accuracy [20]. The thing it has identified as a weapon is then cropped, and an email alert is sent out containing the cropped photo, the precise spot, and the time of the detection. Security measures are improved in a number of scenarios by the real-time detection capabilities, which enable prompt reactions to potential attacks.

RESULTS AND DISCUSSION

The setup of the "Notifying authorities about the YOLOv7-based Violence Detecting on Video" technology produced encouraging results, demonstrating the method's dependability and usefulness in real-world scenarios. Notable advancements were achieved in weapon identification accuracy, speed, and expandability after extensive testing and evaluation [13]. The technology demonstrated exceptional detection rates in identifying weapons and other armament with a high degree of precision. By leveraging YOLOv7's powerful object detection capabilities, we were able to achieve practically real-time detection performance, enabling prompt intervention and response to potential security risks. Furthermore, the system demonstrated robustness in various environmental conditions and video formats, demonstrating its versatility and adjustability in diverse surveillance scenarios. By consistently providing detection results in a range of outdoor landscapes with complex backdrops and inside locations with controlled illumination, our technique proved useful in real-world deployment scenarios. Furthermore, the integration of alert mechanisms facilitated the seamless relay of weapon incidents to relevant authorities, enabling timely and well-coordinated responses to potential threats.

Table 1

SI NO	weapon	F1- Score	Precision Curve	Recall Curve
1	pistol	0.87	0.91	0.89
2	Smart phone	0.91	0.88	0.80
3	knife	0.82	0.85	0.78
4	All classes	0.92	1.00	0.95

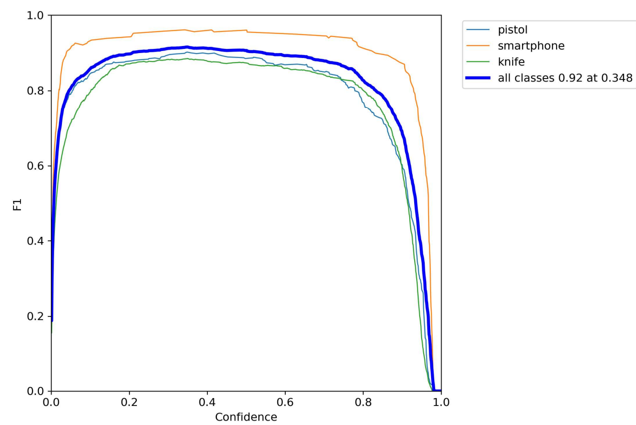

Figure 3: F1-Score.

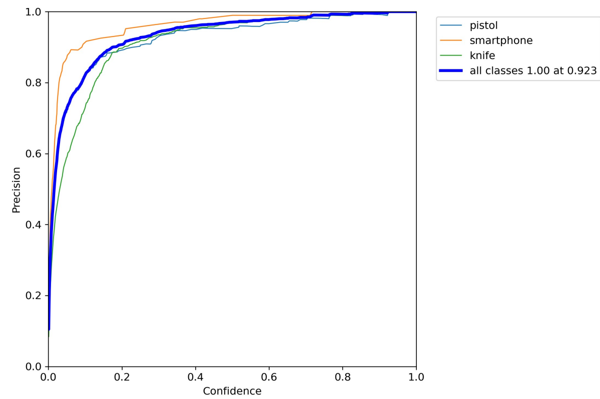

Figure 4: Precision Curve.

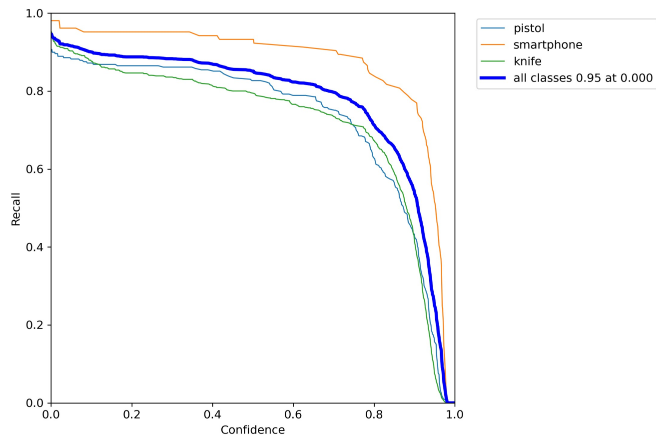

Figure 5: Recall Curve.

Important details, such as the location, date, and time, and visible evidence of the weapons that were discovered, were included in the email messages that the system generated. Responders were able to lower risks by taking the necessary action thanks to this information.

When presenting our findings, it is important to recognize the limitations and challenges we faced throughout the system's design and implementation.

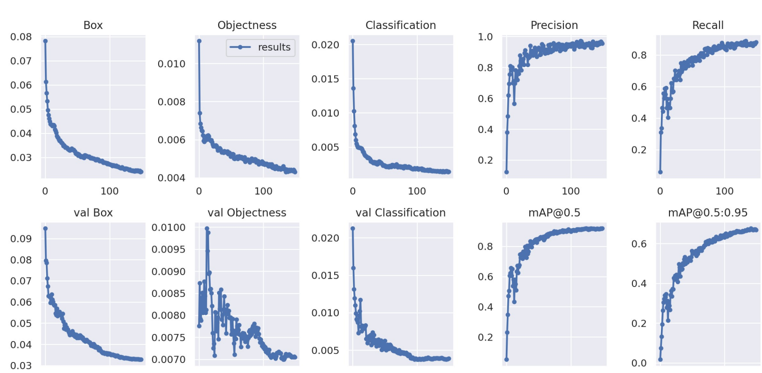

Figure 6: Final Results Graph.

Although the approach demonstrated exceptional accuracy and efficacy in identifying firearms, there may be room for further development, particularly with regard to lowering false-positive detection rates and enhancing performance in challenging circumstances. Furthermore, ethical concerns around privacy, security of data, and potential biases in AI-based surveillance devices need to be adequately addressed in order to ensure the responsible and suitable use of these technologies. Overall, the "Keeping Track of Authorities: YOLOv7-based Violence Detection on Video" approach is a significant development in the use of cutting-edge AI-driven technology to enhance public safety and security.

CONCLUSION

In conclusion, a notable advancement in the realm of security technology is the work being done on "Alerting Authorities: YOLOv7-based Violence Detection on Video." It offers a complete answer for enhancing security protocols in a range of contexts. Through the use of the YOLOv7 architecture, the system is able to identify firearms in real-time video feeds with remarkable speed and accuracy, enabling prompt and precise responses to potential threats. The inclusion of an automated email alert system guarantees that security personnel are notified immediately as weapons are discovered, thus enhancing the capabilities of the solution. Smooth integration streamlines the reaction process and makes it possible for immediate action and risk reduction in security.

FUTURE ENHANCEMENTS

Future performance and resilience of the gun detection system may be improved by incorporating advanced machine learning techniques like ensemble learning. In the context of ensemble learning, a multitude of machine learning techniques are fused together to yield predictions that surpass their individual accuracy. The integration of ensemble learning techniques into the present YOLOv7-based detection system should lead to even higher accuracy rates and a decrease in false-positive detections. It would also be beneficial to look at the integration of additional sensor data, like sound or infrared imaging, in order to increase the accuracy and reliability of firearm detection. To confirm the optical detection results from the video streams, auditory sensors might be used, for example, to detect gunfire.It may be possible to identify firearms more thoroughly and reliably through the combination of many sensor data modalities, especially in challenging circumstances when there is low light or obstruction of visual cues. Features for real-time monitoring could enhance the system's ability to trace the movement of weapons that have been found and provide crucial information to security experts for assessing potential threats and planning countermeasures. The basic objective of these impending developments is to strengthen the efficacy and dependability of real-time weapon detection technology, which will eventually enhance public safety and security in a range of situations.

REFERENCES

1. *C. Sakiba, S. M. Tarannum, F. Nur, F. F. Arpan, & A. A. Anzum (2023). Brac University doctoral dissertation on real-time criminal detection with convolutional LSTM and YOLOv7.*
2. *Kumar, P.; Manie, Y. C.; Yao, C. K.; Guo, B. L.; Nagi, S. K.; Peng, P. C. (2024). Shih, G. L. Improving the Safety of Smart Cities by Applying AI Expert Systems to Detect Violence.*
3. *Yadav, P., Gupta, N., & Sharma, P. K. (2023, November). Enhancing Video Surveillance with Deep Learning-Based Real-Time Handgun Detection and Tracking. In International Conference on Computer Vision and Image Processing (pp. 173-184). Cham: Springer Nature Switzerland.*

4. *Ghadekar, P., Jagtap, S., Sadmake, B., Mane, N., Singh, K., & Chavan, B. (2024). Suspicious Activity Detection in Adverse Weather Conditions using YOLOv7. Grenze International Journal of Engineering & Technology (GIJET), 10.*
5. *Chandrasekaran, K. P. (2023). YOLO ROBBERY: Video Surveillance based Robbery Prediction and Alert System Using YOLO V7.*
6. *Anandhi, R. J. (2023, August). Edge Computing-Based Crime Scene Object Detection from Surveillance Video Using Deep Learning Algorithms. In 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 1159-1163). IEEE.*
7. *Aung, Y. Y., & Oo, K. Z. (2024, May). Detection of Guns and Knives Images Based on YOLO v7. In 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT) (pp. 1-6). IEEE.*
8. *Song, J., & Nang, J. (2024). Pedestrian Abnormal Behavior Detection System Using Edge–Server Architecture for Large–Scale CCTV Environments. Applied Sciences, 14(11), 4615.*
9. *Low, B. K. S. (2023). Video surveillance: Front-yard monitoring (Doctoral dissertation, UTAR).*
10. *Pavithra, S., &Muruganantham, B. (2024). Utilizing enhanced deep learning methods for residential burglary detection via analysis of human behavior patterns. Multimedia Tools and Applications, 1-20.*
11. *Mounika, M. M., Prathyusha, V., Swetha, J., & Amitha, B. Deep Learning for Weapon Identification in Real-Time Camera Survillance.*
12. *Jadhav, V., Deshmukh, R., Gupta, P., Ghogale, S., &Bodireddy, M. (2023, August). Weapon Detection from Surveillance Footage in Real-Time Using Deep Learning. In 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA) (pp. 1-6). IEEE.*
13. *Hussain, A., Khan, S. U., Khan, N., Ullah, W., Alkhayyat, A., Alharbi, M., & Baik, S. W. (2024). Shots segmentation-based optimized dual-stream framework for robust human activity recognition in surveillance video. Alexandria Engineering Journal, 91, 632-647.*
14. *Poirier, F., Jaziri, R., Srour, C., & Bernard, G. (2023, December). Enhancing Anomaly Detection in Videos using a Combined YOLO and a VGG GRU Approach. In 2023 20th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA) (pp. 1-6). IEEE.*
15. *Dalal, S., Lilhore, U. K., Sharma, N., Arora, S., Simaiya, S., Ayadi, M., ... &Ksibi, A. (2024). Improving smart home surveillance through YOLO model with transfer learning and quantization for enhanced accuracy and efficiency. PeerJ Computer Science, 10, e1939.*
16. *Gomathi, G., & Niranjana, G. (2023). Recent Advances and Research Challenges of Deep Learning Framework for Automatic Object Detection in the Real-Time Videos. 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), 1-8.*
17. *VP, P. R., & Ramkumar, G. (2023, November). A Novel Detection of Hazardous Fires from CCTV Images using Novel YOLOv7 in Comparison with Faster R-CNN. In 2023 9th International Conference on Smart Structures and Systems (ICSSS) (pp. 1-5). IEEE.*

18. *Zahrawi, M. H. (2023). Developing a Framework for Weapon and Mask Detection in Surveillance Systems (Master's thesis, The British University in Dubai).*
19. *F. Choo and K. Tan, "Abrupt academic dishonesty: Pressure, opportunity, and deterrence," The International Journal of Management Education, vol. 21, no. 2, pp. 1-15, 2023.*
20. *S. Zhao et al., "VESPA: A General System for Vision-Based Extrasensory Perception Anti-Cheating in Online FPS Games," IEEE Transactions on Games, pp. 1-10, 2023.*