AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 1(1)(2026)

Improving Security Surveillance with Weapon Detection via Faster R-CNN

Mohammed Atif Jamal[1], *Abdul Aziz*[2], Syed Sirajul Altamashi[3], *Jamir Shaik*[4], Venkatesh Bhandari[5]
[1]Assistant Professor, Department of CSE (Data Science), Lords Institute of Engineering and Technology, Hyderabad, Telangana, India.
[2,3,4,5]UG Students, Department of CSE (Data Science), Lords Institute of Engineering and Technology, Hyderabad, Telangana, India.

***Abstract*— Security remains a critical concern across multiple domains due to the increasing incidence of crime in densely populated areas, public events, sensitive zones, and remote locations. The detection and monitoring of abnormal activities have emerged as key applications of computer vision in addressing these challenges. Weapon or anomaly detection involves identifying rare, irregular, or unexpected patterns that deviate from established norms within visual data. With the growing demand for public safety and the protection of assets, surveillance systems have become integral to modern security infrastructures, enabling the analysis of suspicious behaviours and unusual events for effective decision-making. Object detection techniques rely on feature extraction and deep learning models to accurately recognize various objects, where minimizing false positives is essential to prevent inappropriate responses. Consequently, selecting a robust and efficient methodology is crucial for ensuring high performance. In this study, an automated firearm detection system is proposed using convolutional neural network-based models, specifically Single Shot Detector (SSD) and Faster R-CNN. The proposed approach is evaluated using two types of datasets, one comprising pre-annotated images and the other consisting of manually labelled data. Experimental results demonstrate that both models achieve promising accuracy, while their real-world applicability depends on an optimal trade-off between detection speed and precision.**

***Keywords*— Object detection, Weapon Detection, Faster RCNN (FRCNN), video surveillance systems, convolution neural network (CNN), Single Shot Detector (SSD), Anomaly detection.**

I. INTRODUCTION

The worldwide crime rate has increased mostly owing to the widespread use of portable weapons; during violent incidents, people have used deadly weapons capable of causing harm and perhaps leading to deaths [1]. For a country to progress, the maintenance of law and order is essential. Whether a government seeks to attract investors or generate revenue via tourism, a peaceful and safe environment is crucial to guarantee that tourists and investors feel at ease in their interactions. The prevalence of firearm-related crime has markedly heightened in several areas worldwide, particularly in regions where private firearm ownership is permitted [6]. The world now operates as a global village, and our spoken or written communication impacts others worldwide. Although the material is inaccurate and lacking in truth, its rapid dissemination via media, especially social media, causes irreparable damage, as users often trust any content presented on news platforms and social networks. Modern individuals often exhibit reduced patience and limited emotional control, making them more susceptible to provocation and violent reactions. Some individuals are vulnerable to manipulation, and psychological studies suggest that access to weapons can significantly increase the likelihood of violent behavior.

Modern society faces growing threats related to weapons in public spaces, necessitating the rapid deployment of advanced surveillance systems for real-time crime detection [5]. Traditional surveillance systems often struggle with detecting small objects and suffer from delayed response times, reducing their effectiveness in dynamic and crowded environments. CCTV cameras play a crucial role in addressing these challenges and are widely recognized as an essential component of modern security systems [3]. Initially developed for military surveillance using basic camera networks connected via coaxial cables, CCTV systems have evolved significantly. Today, they are integrated with artificial intelligence and advanced analytics [9]. AI techniques such as facial recognition, object detection, person re-identification, and behavior analysis enable automated monitoring, real-time alerts, and efficient data retrieval, thereby enhancing situational awareness for security personnel. The deployment of surveillance systems across public areas such as shopping malls, highways, transportation systems, educational institutions, and airports has significantly strengthened security infrastructure. These systems generate large volumes of video data, which are valuable for tracking activities and supporting forensic investigations. For instance, China operates one of the largest surveillance networks globally, with millions of cameras capable of rapid identification using facial recognition technologies; a BBC reporter was reportedly identified within minutes using this system [8].

The objective is to create a weapon detection system capable of real-time categorization of different types of weapons, hence reducing the aforementioned incidents. These occurrences may be mitigated by an early warning system that notifies security personnel for immediate response. The proposed system employs a live video feed from CCTV cameras to classify five kinds of weapons (Axe, Knife, Pistol, Rifle, and Sword) using deep learning techniques. Historically, individuals have been responsible for security, either via security personnel or law enforcement agencies. Notwithstanding their assignment to oversee a checkpoint, security violations continue to occur. Moreover, after the security breaches, the assailant was not swiftly caught at the scene.

Forensic investigations require time to analyze incidents and identify suspects, which can delay justice. Since prevention is more effective than response, detecting threats before an incident occurs can help avoid loss of life. However, maintaining continuous human surveillance is resource-intensive and costly, and vulnerabilities may arise during personnel shifts [7]. To address these limitations, the proposed approach introduces an automated real-time weapon detection and classification system based on advanced deep learning models [4]. This work focuses on detecting weapons potentially carried by criminals or terrorists in real-time scenarios. The system classifies firearms such as pistols and revolvers under a common category, while distinguishing them from non-weapon objects such as mobile phones, wallets, and other similar items. Detecting weapons in real-time remains a challenging task due to their small size and visual similarity to other objects in complex environments. Convolutional Neural Networks (CNNs), a powerful class of deep learning models, have demonstrated exceptional performance in image recognition tasks, making them highly suitable for weapon detection applications [12].

II. RELATED WORK

Muhammad Tahir Bhatti et al. (2021) [1] presented a study that emphasizes the establishment of a safe environment via the use of CCTV video to identify dangerous weapons by using advanced open-source deep learning algorithms. We have executed binary classification with the pistol class designated as the reference class, and we have incorporated the idea of relevant confusion objects to mitigate false positives and false negatives. No standard dataset was accessible for real-time scenarios; therefore, we created our own dataset by capturing weapon photographs with our camera, manually sourcing images from the internet, extracting data from YouTube CCTV footage, utilizing GitHub repositories, and acquiring information from the University of Granada and the Internet Movies Firearms Database (IMFDB) at imfdb.org. Two methodologies are employed: sliding window/classification and region proposal/object detection.

CCTV surveillance systems are being used in businesses, metropolitan environments, and almost all public spaces. The quantity of CCTV video feeds exceeds a human operator's capacity to monitor and analyze conditions meticulously, particularly in relation to potentially perilous incidents, such as "Active Shooter Events." Incidents such as the tragedies in the Colorado movie theater and Oslo need a quick reaction. M. Grega et al. (2013) [2] developed and evaluated an algorithm designed to identify an individual carrying an exposed handgun and notify the CCTV operator of a potentially hazardous scenario. We outline the constraints and challenges associated with the image analysis application, elaborate on the development of the suggested method, and offer the numerical outcomes for sensitivity and specificity.

Confronted with the necessity to devise more effective security protocols and identify more economical crime prevention methods, law enforcement agencies worldwide are increasingly adopting technological systems to improve operational efficiency, broaden their scope, and decrease expenditures. W. Deisman et al. (2003) [3] assessed that the impact of CCTV on crime is both very varied and rather unexpected. The impacts of CCTV are not uniform over time and vary across various kinds of crime. CCTV systems seem to have the minimal influence on public disturbance offenses.

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 1(1)(2026)

The extent of deterrent effects seems to be contingent upon location, with the most significant impact seen in parking facilities. Moreover, the cameras need not be functional for deterrent effects to manifest. The deterrent benefits of CCTV are maximized when used with other crime reduction strategies and customized to the specific local context. Ultimately, while deterrent effects have been shown prior to the cameras being active, ongoing publicity is necessary to sustain these benefits.

Closed circuit video systems (CCTV) are becoming prevalent and are being implemented in many workplaces, residential complexes, and other public areas. Monitoring systems have been established in several locations throughout Europe and America. This imposes a significant burden on CCTV operators, since the quantity of camera feeds that one operator can oversee is constrained by human limitations. This study addresses the automatic identification and recognition of hazardous situations in CCTV systems. M. Grega et al. (2016) [6] We propose algorithms capable of notifying the human operator when a firearm or knife is detected in the picture. We have concentrated on reducing the incidence of false alerts to facilitate practical use of the system. The specificity and sensitivity of the knife detection surpass those of previously announced alternatives. We have successfully proposed a weapon detection algorithm that has a near-zero false alert rate. We have shown the feasibility of developing a system that may provide early warnings in hazardous situations, perhaps resulting in expedited and more efficient reaction times, hence decreasing the number of prospective casualties.

G. Flitton et al. (2013) [10] conduct an experimental evaluation of three-dimensional feature descriptors used to threat identification in Computed Tomography (CT) airport luggage images. The detectors vary in complexity from a fundamental local density descriptor to local area histograms and three-dimensional (3D) extensions, including both the RIFT descriptor and the pioneering SIFT feature descriptor. We demonstrate that, within the complicated domain of CT images characterized by significant noise and imaging artifacts, a particular instance object identification method using simpler descriptors surpasses a more intricate RIFT/SIFT solution. Recognition rates of 95% are achieved with negligible false-positive rates for a collection of representative 3D objects.

Contemporary automated visual surveillance is essential for security, and this study is the first step toward automatic visual firearm identification. G. Kumar Verma et al. (2015) [12] authored an article aimed at developing a framework for visual gun identification in autonomous surveillance systems. Their methodology utilizes color-based segmentation to remove irrelevant objects from a picture using the k-means clustering technique. The Harris interest point detector and Fast Retina Keypoint (FREAK) are used to identify the item (gun) in the segmented pictures. Our approach is sufficiently robust regarding scale, rotation, affine transformations, and occlusion. We have executed and evaluated the system using example photos of firearms that we acquired. Our technology demonstrated good performance in gun detection. Moreover, our system exhibits exceptional performance across various picture appearances. Consequently, our system is invariant to rotation, scale, and form.

III. DATASET DETAILS

The selection of an appropriate dataset is crucial for model training; a suitable dataset is an essential component of our proposed study. Due to the absence of publicly accessible datasets for CCTV study on firearms including a enough number of instances, we opted to develop our own databases. Considering several methodologies. The weapon detection dataset was sourced from the internet. The pictures were extracted from the original frames via the sliding window technique. The dimensions of the window were deliberately configured to 100 × 100 pixels. These diminutive picture samples reflect the state and quality of actual CCTV recordings, which often exhibit low resolution and blurriness, with the objects seeming little. Furthermore, our prior study indicates that the dimensions of images within a dataset have a little influence on the ultimate categorization outcomes.

This dataset must undergo preprocessing, wherein the raw surveillance data will be converted into a structured format. This procedure ensures uniformity, enhances image quality, and addresses real-world issues prior to the processed data reaching the detection layer. A significant problem with CCTV video is the inconsistency in picture sizes and changing camera angles; to address this, it is necessary to scale the images to a uniform dimension for consistent processing. Low-light circumstances and overexposure provide additional challenges; these factors complicate the detection of firearms. To mitigate this, brightness enhancement and sharpening methods may clarify picture details, boost visibility, and improve feature clarity. Blurred pictures may introduce noise and disruptions that diminish detection accuracy.

Figure 1: Dataset for proposed work

IV. PROPOSED METHODOLOGY

At now, crime rates are increasing in both heavily populated and remote areas. Current monitoring and control systems need human intervention to resolve existing concerns; we have presented a research aimed at locating weaponry in heavily populated areas and remote locales. Weapon detection is the identification of atypical and suspicious behaviors that diverge from established norms within a certain place or dataset. Due to increased need for safety, security, and personal property protection, the use of video surveillance systems has become imperative in several industries.

Identifying the weaponry necessitates the use of algorithms that emphasize superior precision and rapidity. The proposed method for weapon identification used convolutional neural network (CNN)-based SSD and Faster R-CNN algorithms. Microsoft researchers have developed the Faster R-CNN. This is a deep convolutional network mostly used for object recognition; it functions as a unified, end-to-end model for the user. The network can accurately and rapidly predict the locations of diverse items. The suggested methodology employs two classifications of datasets. One dataset included pre-labeled photographs, whilst the other consisted of manually annotated pictures. SSD denotes single-shot detector. It lacks a delegated region proposal network and directly predicts border boxes and classes from feature maps in a single pass. SSD may be educated in an end-to-end manner to enhance accuracy. SSD generates a greater number of predictions and exhibits superior coverage for location, size, and aspect ratios. By eliminating the allocated region suggestion and use lower quality photos. The results are tabulated; both methods demonstrate commendable accuracy, however their practical applicability will rely on speed and precision.

V. RESULT AND DISCUSSION

We have used SSD and FRCNN to develop and train a weapon detection model using the dataset we collected, this dataset will allow us to detect guns of various types. Our suggested approach yields the necessary outputs after the preprocessing stage, enabling real-time detection of firearms. The primary objective of our suggested approach was to provide an additional security and monitoring tool that may reduce the need for human intervention. Our model has a user interface, as seen in Figure [3].

Figure 3: User Interface of the proposed model

The graphic illustrates our model interface, which includes buttons for uploading a dataset, generating and loading a weapon detection model, uploading an image, and detecting from images and videos, as well as a graph showing training accuracy.

Figure 4: upload dataset UI

In figure [4], we can see that when we upload the dataset it is showing dataset loaded, and also giving the total number of images found in our dataset. After uploading we can see in the figure [5], that our proposed model is generated and gave the accuracy of 73.58%, along with the other metrics scores.

Figure 5: weapon detection model generates and load UI

Figure 6: Predicted results from image

After generating the model, here we are going to perform the prediction. In figure [6] we can see that one box with guns image and the model predicted them as weapon and gave a bounding box saying "weapon". Similarly, we can also predict from the video as shown in figure [7].

Figure 7: Predicted results from video

DISCUSSION

The suggested multi-model weapon detection approach, which combines Faster R-CNN with SSD, exhibits great precision in identifying and classifying weapons inside the CCTV dataset. Our proposed model exhibits enhanced accuracy relative to existing models, attaining an accuracy of 73.58% via the use of the SSD and FRCNN algorithms. This hybrid model achieves a recall of 50%, reflecting its challenges in real-time predictions. This highlights its capability to identify firearms in difficult surroundings, such as occlusions, fluctuating lighting conditions, and intricate backdrops, which hampers its use in real-time situations.

VI. CONCLUSION

This research introduces a hybrid model that integrates SSD and Faster R-CNN. This method enhances real-time weapon recognition in CCTV systems. This model achieves an accuracy of 73.58%, a precision of 36.79%, a recall of 50%, and an F1-score of 42.39%, differentiating between hidden and small-scale weapons. We have enabled both SSD and FRCNN for a pre-labeled image dataset for weapon detection. The amalgamation of both strategies produced enhanced outcomes; nevertheless, in real-time applications, this model may demonstrate diminished accuracy. The real-time accuracy depends on a trade-off between speed and precision. In terms of speed, the SSD method is efficient, attaining 0.736 seconds per frame; nevertheless, FRCNN exhibits more accuracy. By amalgamating both components, we constructed a hybrid model that produces encouraging outcomes; nevertheless, the accuracy in practical applications will depend on the velocity and precision of both algorithms.

REFERENCES

1. Muhammad Tahir Bhatti, Muhammad Gufran Khan "Weapon Detection in Real-Time CCTV Videos Using Deep Learning", March 4, 2021.
2. M. Grega, S. Lach, and R. Sieradzki, ''Automated recognition of firearms in surveillance video,'' in Proc. IEEE Int. Multi-Disciplinary Conf. Cognit. Methods Situation Awareness Decis. Support (CogSIMA), Feb. 2013.
3. W. Deisman, ''CCTV: Literature review and bibliography,'' in Research and Evaluation Branch, Community, Contract and Aboriginal Policing Services Directorate. Ottawa, ON, Canada: Royal Canadian Mounted, 2003.
4. Shanthi p and Manjula v "Weapon detection with FMR-CNN and YOLOv8 for enhanced crime prevention and security" 23 july 2025
5. J. Ratcliffe, ''Video surveillance of public places,'' US Dept. Justice, Office Community Oriented Policing Services, Washington, DC, USA, Tech. Rep. 4, 2006.
6. M. Grega, A. Matiolałski, P. Guzik, and M. Leszczuk, ''Automated detection of

firearms and knives in a CCTV image,'' Sensors, vol. 16, no. 1, p. 47, Jan. 2016.

7. K.P. Vijayakumar, K. Pradeep, A. Balasundaram, A. Dhande "R-CNN and YOLOV4 based Deep Learning Model for intelligent detection of weaponries in real time video" Mathematical Biosciences and Engineering, 2023.
8. TechCrunch. (2019). China's CCTV Surveillance Network Took Just 7 Minutes to Capture BBC Reporter. Accessed: Jul. 15, 2019.
9. N. Cohen, J. Gattuso, and K. MacLennan-Brown. CCTV Operational Requirements Manual 2009. St Albans, U.K.: Home Office Scientific Development Branch, 2009.
10. G. Flitton, T. P. Breckon, and N. Megherbi, ''A comparison of 3D interest point descriptors with application to airport baggage
11. object detection in complex CT imagery,'' Pattern Recognit., vol. 46, no. 9, pp. 2420–2436, Sep. 2013.
12. R. Gesick, C. Saritac, and C.-C. Hung, ''Automatic image analysis process for the detection of concealed weapons,'' in Proc. 5th Annu. Workshop Cyber Secur. Inf. Intell. Res. Cyber Secur. Inf. Intell. Challenges Strategies (CSIIRW), 2009.
13. Madar Maske, Prasad Narsale, Jay Kingre, "Weapon Detection in Real-Time CCTV Videos Using CNN" 2023.
14. Pokala, H. K., & Gummadi, V. P. K. (2026). Autonomous AI-Powered Resource Management for Apache Flink on Amazon EKS. 2026 International Conference on Artificial Intelligence, Systems, and Emerging Technologies (ICAISET), 1–4. https://doi.org/10.1109/icaiset66439.2026.11541881
15. Babburi, S. Lightweight Distributed Provenance Framework for Edge and IoT Data Systems.
16. Gaddam, S. From Fixed Specifications to Self-Adapting Systems: A Machine Learning Perspective on Software Engineering.
17. Immadi, S. K. (2025). Optimizing ERP for Human Capital Management. Applied Research for Growth, Innovation and Sustainable Impact, 377–384. https://doi.org/10.1201/9781003684657-63
18. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
19. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
20. Mahimalur, R. K., Vasgam, M., & Manoharan, D. Devops Lifecycle Management And Cloud Migration Assessments: A Security-Driven CICD Perspective.
21. Viswanathan, V. Generative AI for Smarter Workforce Planning and Enterprise Resource Decisions.
22. Gajula, S. (2026, March). Two Pillars of Banking Intelligence: A Comparative Analysis of AI Techniques for Fraud Prevention and Churn Mitigation. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
23. Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 275–284.
24. Venkata Ramana, P. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Research in Information Technology and Computing, 8(4).
25. Srikanth Kavuri. (2025). AI-DRIVEN TEST AUTOMATION FRAMEWORKS: ENHANCING EFFICIENCY AND ACCURACY IN SOFTWARE QUALITY ASSURANCE. International Journal of Applied Mathematics, 38(10s), 699–710. https://doi.org/10.12732/ijam.v38i10s.990
26. Venkata Pavan Kumar Gummadi. (2026). Infrastructure Optimization Techniques for Enterprise Integration Platforms: A Comprehensive Analysis. Computer Fraud and Security, 37–44. https://doi.org/10.52710/cfs.875
27. Shashank, A. (2025). AI-Enhanced ETL Processes: Leveraging Artificial

Intelligence for Optimized Data Integration Systems. Journal Of Multidisciplinary, 5(8), 219-225.

28. Kandula, S. T. R., Susarla, R. S., & Boyapati, P. K. (2025, July). Enhanced Cyber Security Using Global Local Artificial Neural Network Based Intrusion Detection in Big Data Environment. In 2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC) (pp. 426-431). IEEE.
29. Boyapati, P. K. Building a centralized data operations hub for healthcare enterprise integration. IJSAT-Int. J. Sci. Technol. 16 (2). https://doi.org/10.71097/IJSAT.v16.i2.3219