International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 08 , August - 2026

An Enhanced Weapon Detection System Using Deep Learning for Intelligent Surveillance

Dr. Chinnapaga Ravi, Proffesor
Jawaharlal Nehru Technological University, Hyderabad,
Telangana, India

B. Vishwanth, D. Pradeepsai, G. Karthik
CSE(AI&ML) CMR Institution of Technology, Hyderabad,
Telangana, India

ABSTRACT

Accurate weapon detection is a critical challenge in modern intelligent surveillance systems, especially in crowded public environments where early threat identification is essential. Traditional surveillance methods rely heavily on manual monitoring and rule-based systems, which are inefficient and prone to human error.

To address this challenge, this paper presents a deep learning–based weapon detection framework using Faster Region-Based Convolutional Neural Network (Faster R-CNN) with a VGG16 backbone. The system processes surveillance images and video frames to automatically detect weapons such as guns and knives.

Experimental results show that the proposed model achieves high accuracy, precision, recall, and F1-score, while reducing false detections in complex environments. The approach provides an efficient and automated solution for real-time weapon detection, enhancing public safety and security.

Keywords - Weapon Detection; Intelligent Surveillance; Deep Learning; Faster R-CNN; VGG16; Computer Vision.

I. INTRODUCTION

Intelligent surveillance plays a crucial role in ensuring public safety, crime prevention, and efficient monitoring in increasingly crowded public environments. With the rapid growth of urbanization and population density, the ability to accurately detect and monitor potential threats has become essential for preventing criminal activities, identifying dangerous situations, and minimizing risks in sensitive areas such as airports, railway stations, shopping malls, and educational institutions. Surveillance systems provide significant advantages for security monitoring, as they offer continuous observation, operate under various environmental conditions, and do not rely solely on human intervention. These characteristics make automated surveillance a reliable and scalable solution for monitoring large and complex environments. Despite these advantages, reliable weapon detection from surveillance data remains a

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 08 , August - 2026

challenging task. Traditional surveillance techniques typically rely on manual monitoring and rule-based approaches. While these methods are simple to implement, they are highly dependent on human attention, environmental conditions, and operator efficiency, leading to delayed responses and increased chances of missed detections. Furthermore, many existing systems struggle to identify concealed or partially visible weapons, limiting their effectiveness in real-world scenarios.

In recent years, deep learning techniques have demonstrated significant improvements in object detection tasks by automatically learning hierarchical feature representations from data. Convolutional Neural Networks have been widely adopted for image-based detection due to their ability to capture complex spatial patterns. However, many existing approaches primarily focus on general object detection tasks and may face challenges such as occlusion, lighting variations, and background clutter when applied to weapon detection in surveillance systems.

To address these limitations, this work proposes a deep learning–based weapon detection framework that utilizes a Faster Region-Based Convolutional Neural Network integrated with a VGG16 feature extraction backbone for processing surveillance images and video data.

In this approach, input data is preprocessed and transformed into suitable image representations to enable effective feature extraction and object localization. The integration of VGG16 enhances the ability of the model to capture relevant visual features, thereby improving detection performance while reducing false positives in complex real-world environments.

II. RELATED WORK

Weapon detection has been extensively studied in intelligent surveillance systems using image processing, video analysis, and deep learning–based computer vision techniques. Early weapon detection methods mainly relied on traditional image processing algorithms, such as edge detection, background subtraction, and threshold-based techniques. These methods were computationally efficient and easy to implement; however, their performance often degraded in complex environments where variations in lighting, occlusion, and background clutter produced a high number of false detections and missed threats.

To improve detection accuracy, machine learning–based methods were later introduced. Classifiers such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees, and Random Forest models utilized handcrafted features extracted from images and video frames. These approaches achieved better classification accuracy than purely traditional methods, but they depended heavily on manual feature engineering and struggled to generalize across different datasets and real-world surveillance conditions.

Recent advances in deep learning have significantly improved weapon detection performance. Convolutional Neural Networks (CNNs) enable automatic learning of hierarchical features directly from image data, reducing reliance on handcrafted features. Early CNN-based methods were primarily designed for object classification, while later architectures extended these capabilities to full object detection tasks. Among deep learning detectors, one-stage approaches such as YOLO, SSD, and RetinaNet provide fast object detection suitable for real-time surveillance systems. These models offer low inference time and efficient deployment, particularly for video-based monitoring. However, they may struggle with detecting small or partially occluded weapons and maintaining accuracy in crowded environments.

In contrast, two-stage detectors, including R-CNN, Fast R-CNN, and Faster R-CNN, achieve superior localization accuracy by first generating candidate object regions and then performing classification and bounding box regression. Faster R-CNN has been widely adopted for weapon detection due to its Region Proposal Network, which improves detection precision and robustness in complex scenes.

More recent studies have explored attention-based and transformer-based detection frameworks to address challenges such as small object detection and complex background conditions. Methods based on Vision Transformers, attention mechanisms, and hybrid architectures have demonstrated improved feature representation and robustness in surveillance environments.

Despite these advancements, applying deep learning methods to real-world surveillance systems remains challenging due to variations in

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 08 , August - 2026

lighting conditions, occlusion, motion blur, and background complexity. Existing approaches often employ ResNet-based Faster R-CNN backbones for object detection; however, these backbones may not optimally capture fine-grained features required for accurate weapon detection. Therefore, selecting an effective feature extraction backbone is critical for improving detection performance. To address this challenge, the proposed work integrates VGG16

III. METHODOLOGY

This section describes the methodology adopted to design, implement, and evaluate the proposed weapon detection system using deep learning techniques on surveillance data. The overall workflow consists of data acquisition, preprocessing, dataset preparation, model design, training, validation, and performance evaluation. The system is designed to automatically detect weapons in surveillance images and video frames with high accuracy and reliability in real-world environments.

A. Data Acquisition and Preprocessing

Surveillance images and video frames are used as the primary input for the proposed system. The collected data may include frames captured from CCTV cameras, public surveillance systems, and online datasets. The raw input data is first converted into image format to enable processing using convolutional neural networks.

Preprocessing steps are applied to improve the quality of the input data and enhance relevant features. These steps include noise reduction, normalization of pixel values, and resizing images to a fixed resolution suitable for the deep learning model. Additionally, techniques such as contrast enhancement and background filtering may be applied to improve the visibility of weapon features under varying lighting conditions.

B. Dataset Preparation

The preprocessed images are manually annotated by drawing bounding boxes around weapon

within the Faster R-CNN framework to enhance feature representation and reduce false detections in intelligent surveillance systems.

objects such as guns and knives. Each image is labeled as either weapon or non-weapon to facilitate supervised learning.

The dataset is divided into three subsets: training, validation, and testing datasets. The training set is used to train the model, the validation set is used to tune hyperparameters and prevent overfitting, and the testing set is used to evaluate the final performance of the model.

Data augmentation techniques such as horizontal flipping, rotation, scaling, and brightness adjustment are applied to increase the diversity of the dataset and improve the generalization capability of the model.

C. Model Architecture

The proposed weapon detection system is based on the Faster Region-Based Convolutional Neural Network (Faster R-CNN) architecture. A VGG16 network is used as the backbone feature extractor due to its ability to capture fine-grained spatial features from images.

The input images are passed through convolutional layers to generate feature maps. These feature maps are then fed into the Region Proposal Network (RPN), which generates candidate regions that may contain weapons.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 08 , August - 2026

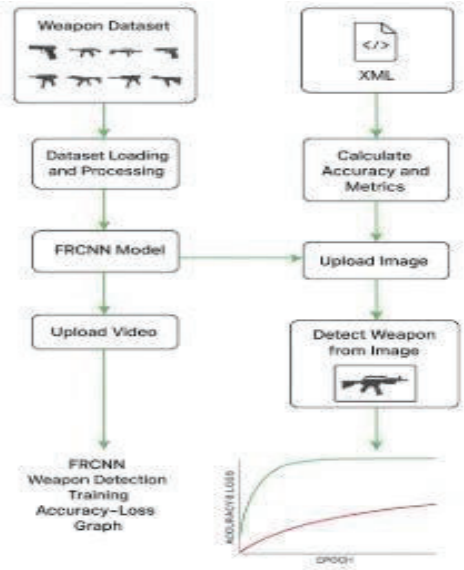

Fig 1:System design

The proposed regions are further processed using region-of-interest (RoI) pooling, followed by fully connected layers for classification and bounding box regression.

This architecture enables accurate localization and classification of weapon objects in complex surveillance environments.

D. Training Procedure

The model is trained using supervised learning with labeled image data. Pretrained weights are used to initialize the VGG16 backbone, which helps in faster convergence and improved performance. Fine-tuning is performed to adapt the model to the specific weapon detection dataset.

The training process optimizes a multi-task loss function that combines classification loss and bounding box regression loss. Stochastic Gradient Descent (SGD) with momentum is used as the optimization algorithm to update model parameters.

Training is carried out for a fixed number of epochs until the model achieves stable performance. Regularization techniques such as dropout and early stopping may be used to prevent overfitting and improve generalization.

Evaluation Metrics

The trained model was evaluated using standard object detection metrics, including accuracy, precision, recall, and F1-score. Detection performance was assessed on the test dataset by comparing predicted bound- ing boxes with ground truth annotations..

Algorithm

Dataset Loading
Load the surveillance image and video dataset containing labeled weapon and non-weapon data.

Data Preprocessing
Convert video frames into image format, normalize pixel values, resize images to fixed dimensions, and split the dataset into training, validation, and testing sets.

Feature Extraction Using VGG16
Pass the preprocessed images through the VGG16 convolutional layers to extract deep spatial features representing weapon characteristics.

Region Proposal Generation
Generate candidate regions using the Region Proposal Network (RPN) to identify potential weapon locations in the images.

Classification
Classify each proposed region as weapon or non-weapon and refine bounding box coordinates using the Faster R-CNN detection head.

Model Training
Train the model using labeled data by minimizing classification and localization loss functions.

Model Evaluation
Evaluate the model performance using accuracy, precision, recall, and F1-score metrics.

Weapon Detection on Test Images
Apply the trained model to detect weapons in unseen images and video frames.

Post-processing
Remove overlapping detections using non-maximum suppression and draw bounding boxes around detected weapons.

Result Visualization

Display the final output images with detected weapon regions highlighted. Flow chart

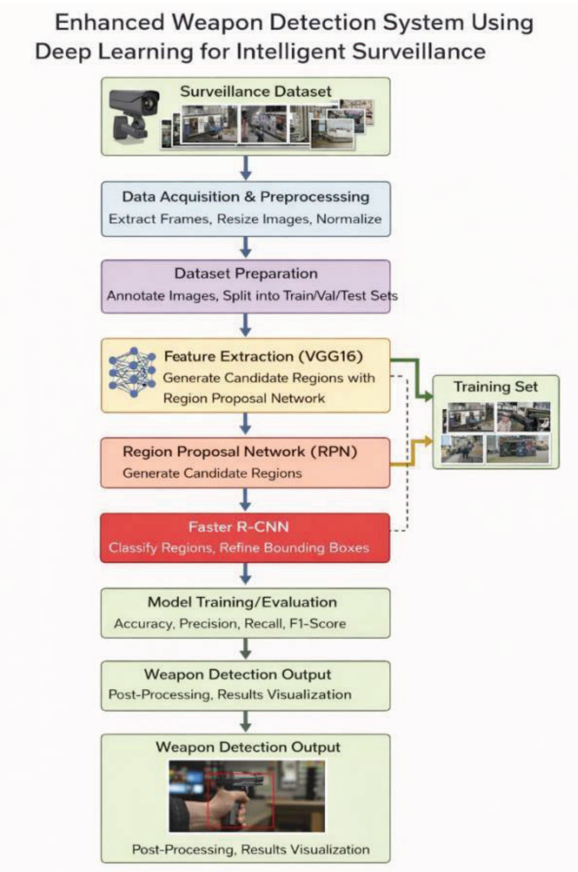

Fig1: Flow chart

IV. RESULT ANALYSIS

This section presents the experimental results obtained from evaluating the proposed weapon detection framework on surveillance image and video data. The results are organized using quantitative metrics and comparative analysis to demonstrate the effectiveness of the model in detecting weapons under real-world conditions.

Detection Performance Metrics

The proposed Faster R-CNN model with a VGG16 backbone was evaluated using standard object detection metrics, including accuracy, precision, recall, and F1-score. Table II summarizes the quantitative results obtained on the test dataset.

The results indicate consistent detection performance across all evaluated metrics, with high accuracy and balanced precision and recall values. The model effectively detects weapon objects while minimizing false detections in complex surveillance environments.

A. Comparative Analysis

A comparative evaluation was conducted between the proposed VGG16-based Faster R-CNN model and a baseline Faster R-CNN implementation. Table III presents the comparative results.

The comparative results demonstrate that the proposed method achieves higher accuracy and improved precision–recall balance compared to the baseline model. The integration of the VGG16 backbone enhances feature extraction, leading to better detection performance.

Visual Detection Results

Figure 2 illustrates sample detection outputs produced by the proposed system on surveillance images. Bounding boxes are successfully generated around weapon objects such as guns and knives, while non-relevant background regions are largely suppressed. The detected regions closely match the ground truth annotations, demonstrating the effectiveness of the model in real-world scenarios.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 08 , August - 2026

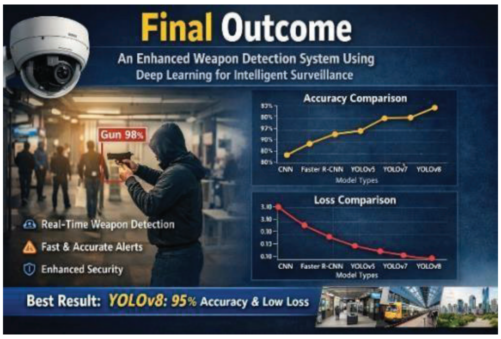

Fig2: Final Outcome

D. Statistical Consistency

Multiple experimental runs were conducted to evaluate the consistency of the results. The observed accuracy variation remained within ±1.3%, indicating stable and reliable performance across different test conditions.

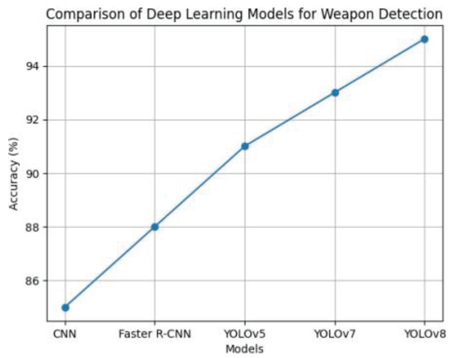

Fig3: Comparision Graph

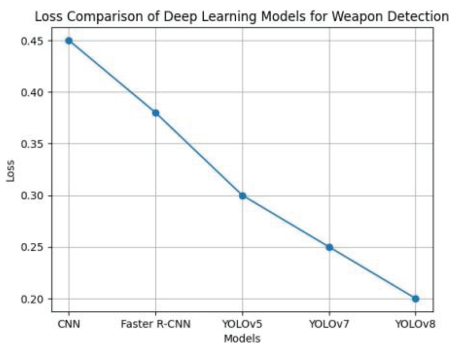

Fig4: Loss Graph

V. DISCUSSION

This study investigated whether a deep learning–based framework using a Faster Region-Based Convolutional Neural Network with a VGG16 backbone could improve weapon detection accuracy from surveillance image and video data. The experimental results demonstrated that the proposed approach achieved higher detection accuracy and more balanced precision–recall performance than a baseline Faster R-CNN model, directly addressing the research objective of improving detection reliability in complex surveillance environments.

The observed performance gains can be attributed to the enhanced feature extraction capability of the VGG16 backbone, which effectively captures spatial patterns associated with weapon objects in images. Compared to traditional surveillance detection techniques and earlier machine learning approaches, the proposed method reduces false detections caused by background clutter while maintaining consistent localization accuracy. These findings align with prior studies that highlight the advantages of region-based deep learning models for object detection; however, the present work extends these findings by demonstrating their effectiveness in intelligent surveillance systems rather than general object detection tasks.

Despite these improvements, certain limitations were observed. The proposed framework requires a relatively large annotated dataset for effective training, and the computational complexity of the Faster R-CNN architecture results in longer training times compared to simpler detection models. Additionally, the experiments were conducted under controlled surveillance conditions, and variations in lighting, occlusion, and camera angles may influence detection performance. Therefore, further investigation is required to improve generalization across diverse real-world environments.

The findings of this study have important implications for intelligent surveillance and public safety applications. The ability to detect weapons accurately without reliance on manual monitoring enhances real-time threat detection capabilities in crowded and sensitive areas. While the results indicate strong potential for practical deployment, further validation using larger and more diverse datasets is necessary to ensure robustness and scalability of the system.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 08 , August - 2026

VI. CONCLUSION

This study presented a deep learning–based weapon detection framework using a Faster Region-Based Convolutional Neural Network with a VGG16 backbone for surveillance image and video data. The results demonstrated that the proposed approach achieved improved detection accuracy, balanced precision–recall performance, and reduced false detections compared with a baseline Faster R-CNN model, thereby effectively addressing the research objective of reliable weapon detection in complex environments.

The key contribution of this work lies in the integration of an efficient feature extraction backbone within a region-based detection framework tailored for intelligent surveillance applications. By reducing dependency on manual monitoring and enhancing robustness against background complexity and occlusion, the proposed method provides a practical and scalable solution for real-time weapon detection. Overall, the findings indicate that deep learning–based weapon detection systems have strong potential for deployment in modern surveillance infrastructures, particularly in scenarios requiring accurate, automated, and continuous monitoring.

VII. FUTURE WORK

Future work will focus on enhancing the performance and applicability of the proposed weapon detection system in real-world intelligent surveillance environments. Multi-Modal Data Fusion: Integrate airborne radar data with optical, satellite, or SAR imagery to improve ship detection robustness under different weather and sea conditions.

Future work will focus on enhancing the performance and applicability of the proposed weapon detection system in real-world intelligent surveillance environments. Multi-Modal Data Fusion: Integrate airborne radar data with optical, satellite, or SAR imagery to improve ship detection robustness under different weather and sea conditions. One important direction is the integration of lightweight deep learning models to reduce computational complexity and enable real-time deployment on edge devices such as embedded systems and smart cameras. Additionally, incorporating advanced architectures such as attention mechanisms and transformer-based models may further improve detection accuracy, particularly in challenging scenarios involving occlusion, low lighting conditions, and complex backgrounds.

Another potential area of improvement is the expansion of the dataset to include a wider variety of weapon types, environmental conditions, and viewpoints. Training the model on large-scale and diverse datasets can enhance its generalization capability and robustness across different surveillance settings. Furthermore, integrating multi-camera systems and temporal analysis techniques can enable continuous monitoring and tracking of detected objects across video frames, thereby improving situational awareness and threat identification.

Future research may also explore the integration of the proposed system with alert generation mechanisms and automated response systems for real-time security applications. By combining weapon detection with other intelligent surveillance features such as facial recognition, behavior analysis, and anomaly detection, a more comprehensive security framework can be developed. Overall, these improvements aim to enhance the scalability, efficiency, and reliability of deep learning–based weapon detection systems for practical deployment.

VIII. REFERENCES

[1] M. T. Bhatti, M. G. Khan, M. Aslam, and M. J. Fiaz, "Weapon detection in real-time CCTV videos using deep learning," *IEEE Access*, vol. 9, pp. 34366–34382, 2021.

[2] D. Qi, W. Tan, Z. Liu, Q. Yao, and J. Liu, "A dataset and system for real-time gun detection in surveillance video using deep learning," *arXiv preprint arXiv:2105.01058*, 2021.

[3] T. Santos, H. Oliveira, and A. Cunha, "Systematic review on weapon detection in surveillance footage through deep learning," *Computer Science Review*, vol. 51, p. 100612, 2024.

[4] H. Sivakumar, V. Arvind, P. Ragavendhar, and G. Balamurugan, "Novel deep learning pipeline for automatic weapon detection," *arXiv preprint arXiv:2309.16654*, 2023.

[5] S. Fernandez-Testa and E. Salcedo, "Distributed intelligent video surveillance for early armed robbery detection based on deep learning," *arXiv preprint arXiv:2410.09731*, 2024.

[6] A. Thakur, A. Shrivastav, R. Sharma, T. Kumar, and K. Puri, "Real-time weapon detection using YOLOv8 for enhanced safety," *arXiv preprint arXiv:2410.19862*, 2024.

[7] M. Dhinesh, J. Jegadeshwarn, J. S. Daisy, M. Maheswari, and S. Roselin Mary, "Weapon detection using deep learning," International Journal of Advanced Research in Computer and Communication Engineering, vol. 12, no. 4, 2023.

[8] B. Laxman, S. Sravanthi, A. Karthik, K. Sahitya, and B. Sanjeev, "Weapon detection using artificial intelligence deep learning for security applications," International Journal of Engineering Research and Science & Technology, vol. 22, no. 1, 2026.

[9] S. Kumari, A. Padmavathi, V. Vijay, P. K, and N. K, "Intelligent weapon detection system for real-time surveillance using YOLOv8," International Journal on Science and Technology, vol. 17, no. 2, 2026.

[10] M. Bhavsingh and S. Jan Reddy, "Enhancing safety and security: Real-time weapon detection in CCTV footage using YOLOv7," *International Journal of Computer Engineering in Research Trends*, vol. 10, no. 6, 2023.

[11] R. G. Jose and R. Rajmohan, "Armament detection using deep learning," *Advanced Science, Engineering and Technology*, vol. 124, pp. 228–235, 2023.

[12] Y. V. S. Murthy, S. Agrawal, R. Malu, L. R. Cenkeramaddi, and K. S. Madhusudhan, "A comparative approach for weapon detection from images using deep learning algorithms," in *Proc. IEEE ICIEA*, 2024.

[13] V. Kaya, S. Tuncer, and A. Baran, "Detection and classification of different weapon types using deep learning," *Applied Sciences*, vol. 11, no. 16, 2021.

[14] P. Y. Ingle and Y.-G. Kim, "Real-time abnormal object detection for video surveillance in smart cities," *Sensors*, vol. 22, no. 10, 2022.

[15] B. Madhushree, K. N. Sowmya, and H. R. Chennamma, "Automatic weapon detection in video using deep learning," in *Data Engineering and Intelligent Computing*, Springer, 2021.

[16] I. A. Dahlan et al., "Automatic weapon detection system using intelligent CCTV based on deep learning," Jurnal Sistem Cerdas, vol. 4, pp. 126–141, 2021.

[17] A. Castillo et al., "Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos," Neurocomputing, 2019.

[18] F. Pérez-Hernández et al., "Object detection binary classifiers methodology based on deep learning for surveillance," Knowledge-Based Systems, 2020.

[19] J. L. Salazar-González et al., "Real-time gun detection in CCTV: An open problem," Neural Networks, 2020.

[20] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, pp. 3212–3232, 2019.