IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15391

An Automated Human Violence Detection for Surveillance using Deep Learning

Mr. H.M. Gaikwad[1], Sharvari Gangurde[2], Priyanka Moule[3], Sejal Patil[4], Aarya Patil[5]

Sr. Lecturer in AIML, K. K. Wagh Polytechnic, Nashik[1]

Third Year Students of Artificial Intelligence and Machine Learning, K. K. Wagh Polytechnic, Nashik[2-4]

Abstract: In recent years, ensuring public safety in crowded environments such as transportation hubs, campuses, and urban areas has become increasingly challenging due to the rise in violent incidents. Traditional surveillance systems rely on continuous human observation, which is inefficient, time-consuming, and prone to errors. To address these issues, this project proposes an AI-based automated surveillance system capable of detecting violent human activities in real time using deep learning techniques. The proposed system utilizes the YOLOv8 Nano model developed by Ultralytics along with a custom-trained dataset to identify violent behavior directly from video streams captured through live webcams (CCTV cameras). The YOLO-based approach performs fast, single stage object detection, enabling real-time monitoring with lower computational requirements. By learning spatial patterns associated with aggressive actions, the system can effectively differentiate between violent and non-violent activities even in dynamic and crowded environments. The developed framework is integrated with a monitoring interface that generates instant alerts whenever violent behavior is detected, thereby assisting authorities in taking timely action. Experimental results indicate that the system achieves reliable detection speed and satisfactory accuracy, making it suitable for deployment in smart surveillance systems, institutional security setups, and public safety monitoring applications [1].

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Computer Vision, Object Detection, Human Activity Recognition, Violence Detection, Smart Surveillance, Real-time Monitoring, Security Automation, Public Safety System, Emergency Alert System, Video Analytics, Image Processing, Feature Extraction, Convolutional Neural Networks, YOLO Algorithm, Dataset Training, Model Evaluation, Accuracy Optimization, Intelligent Monitoring, AI- based Security, Threat Detection, Smart City Safety, Automated Surveillance.

I. INTRODUCTION

Traditional surveillance systems rely heavily on human operators to monitor video feeds from cameras, which is not only labor intensive but also prone to errors due to fatigue, distraction, or delayed responses. Consequently, there is an increasing need for intelligent, automated systems that can monitor environments in real time and promptly detect potential threats. Recent advances in Artificial Intelligence (AI) and Deep Learning (DL) have revolutionized the field of computer vision, enabling automated analysis of visual data with remarkable accuracy. Among these techniques, object detection algorithms such as YOLO (You Only Look Once) have emerged as powerful tools for detecting and classifying objects and activities in real time. Unlike traditional sequential models, YOLO provides fast, single-stage detection that can process video streams with minimal delay, making it ideal for real-time surveillance applications.

This project focuses on designing a real-time human violence detection system using YOLOv8 Nano, a lightweight yet highly efficient deep learning model, along with a custom-trained dataset of violent and non-violent actions. The system analyzes video streams captured from live webcams (CCTV cameras) to identify violent behavior and generate instant alerts. By leveraging the speed and accuracy of YOLOv8 Nano, the proposed framework aims to enhance public safety, reduce the burden on human operators, and enable timely intervention in potentially dangerous situations.

In addition to activity recognition, accurate detection of humans within the surveillance scene is essential for focusing the analysis on relevant regions. Object detection models such as YOLO (You Only Look Once) have become highly effective for real-time detection tasks. The latest version, YOLOv8, provides improved accuracy and faster inference, making it suitable for real-time surveillance applications. By detecting human presence in each frame, YOLOv8 helps reduce computational complexity and ensures that the activity recognition model focuses only on relevant areas.

The proposed system aims to enhance traditional surveillance systems by introducing intelligent monitoring capabilities that reduce the dependency on manual observation. By combining advanced deep learning techniques with real-time video processing, the system can serve as a valuable tool for automated security monitoring in smart cities, public environments, and critical infrastructure.

DOI: 10.17148/IJARCCE.2026.15391

II. LITERATURE REVIEW

The development of automated violence detection systems has gained significant attention in recent years due to the increasing demand for intelligent surveillance solutions. Researchers have explored various computer vision and deep learning techniques to detect violent activities in video streams. This section reviews some of the important research works related to violence recognition and real-time surveillance systems.

Early approaches to violence detection mainly relied on traditional machine learning techniques combined with handcrafted feature extraction methods. Researchers used motion-based features such as optical flow, motion vectors, and histogram-based descriptors to analyze movement patterns in video sequences. These features were then classified using machine learning algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests. Although these methods provided a foundation for automated surveillance, they required extensive manual feature engineering and often struggled with complex real-world environments.

With the advancement of deep learning, more sophisticated methods were introduced to improve detection accuracy. Convolutional Neural Networks (CNNs) became widely used for image and video analysis due to their ability to automatically extract relevant features from raw data. Researchers began applying CNN-based architectures for human action recognition, which significantly improved the performance of violence detection systems. These models could analyze spatial features such as body posture, human interactions, and object presence within video frames.

Further improvements were achieved by incorporating temporal information from video sequences. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks were used to capture motion patterns across multiple frames, enabling better understanding of dynamic activities. By combining spatial and temporal features, these hybrid deep learning models were able to recognize violent actions with higher accuracy compared to traditional approaches.

More recently, object detection models such as YOLO (You Only Look Once), Faster R-CNN, and SSD have been applied to surveillance systems for real-time activity detection. Among these, the YOLO family of models has become particularly popular due to its high processing speed and efficient detection capabilities. YOLO performs object detection in a single stage, allowing the system to process video frames quickly while maintaining good accuracy. This makes it suitable for real-time applications such as CCTV surveillance and public safety monitoring[2].

The latest versions of YOLO, including YOLOv8, provide improved accuracy, faster inference speed, and better adaptability to different datasets. The lightweight variant, YOLOv8 Nano, is specifically designed for systems with limited computational resources while still maintaining efficient detection performance. Researchers have shown that YOLO-based models can effectively identify human interactions and suspicious behaviors in surveillance footage.

Despite these advancements, several challenges remain in real-time violence detection systems. Factors such as poor lighting conditions, camera angle variations, occlusion, crowded environments, and limited training datasets can affect detection accuracy. Therefore, ongoing research focuses on improving model robustness, optimizing real-time performance, and developing more reliable surveillance systems. The proposed system in this project builds upon these research advancements by implementing a real-time human violence detection framework using the YOLOv8 Nano model and a custom-trained dataset. The system aims to provide efficient monitoring, accurate violence detection, and timely alert generation for enhanced security and safety in surveillance environments [4].

III. METHODOLOGY

The development of the humanoid robot followed a structured methodology involving design, fabrication, hardware integration, programming, and testing. Initially, the mechanical structure was designed using **Fusion 360** to create a modular, lightweight, and stable humanoid body capable of supporting 19 Degrees of Freedom (DoF). The design included separate components for the head, torso, arms, and legs, ensuring proper weight distribution and ease of assembly. These components were fabricated using 3D printing and assembled using screws and brackets to achieve structural rigidity.For hardware implementation, the **Arduino Mega 2560 Pro Mini** was selected as the main controller due to its high number of digital I/O pins and ability to control multiple servo motors directly using PWM signals.

A total of 19 servo motors were installed at critical joints, including the neck, shoulders, elbows, hips, knees, and ankles, enabling human-like articulation and improved balance through lateral weight shifting. An external battery pack

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15391

was used to supply sufficient current to the servos, while a common ground configuration ensured stable and noise-free operation.The software was developed using **Arduino IDE**, where each servo was assigned to specific pins and initialized to default positions for a stable startup posture. Motion control was achieved through predefined movement sequences such as standing, arm gestures, head rotation, and leg articulation.

A total of 19 servo motors were installed at critical joints, including the neck, shoulders, elbows, hips, knees, and ankles, enabling human-like articulation and improved balance through lateral weight shifting. An external battery pack was used to supply sufficient current to the servos, while a common ground configuration ensured stable and noise-free operation.The software was developed using **Arduino IDE**, where each servo was assigned to specific pins and initialized to default positions for a stable startup posture. Motion control was achieved through predefined movement sequences such as standing, arm gestures, head rotation, and leg articulation.

Smooth and synchronized motion was ensured by implementing incremental angle transitions and controlled delays, reducing jerks and improving stability. Finally, the system underwent testing and calibration, including individual servo testing, joint range adjustments, and stability evaluation through center-of-mass shifting. The fully assembled humanoid robot demonstrated coordinated multi-joint movements and stable posture, making it an effective and low-cost platform for robotics education, kinematics experimentation, and embedded systems learning [2].

A. SYSTEM DESIGN AND ARCHITECTURE

The Human Violence Detection System follows a modular architecture designed for real-time video monitoring and automated decision-making.

1. Overall Architecture

The system consists of four main layers:

a. Input Layer – captures real-time video feed from webcam or CCTV
b. Processing Layer – performs frame extraction and deep learning inference.
c. Decision Layer – classifies detected actions as violent or non-violent
d. Output Layer – displays annotated video and generates alerts.

2. Hardware Design

a. Computer/Laptop with sufficient processing capability
b. Minimum 8 GB RAM recommended
c. Webcam or CCTV camera for live video input
d. Storage for trained model files and detection logs
e. GPU recommended for faster detection (optional but beneficial)

3. Software Requirements

a. Python programming language (Python 3.x)
b. Deep learning framework: PyTorch.
c. Computer vision library: OpenCV, Yolov5, Yolov8
d. YOLOv8 detection framework by Ultralytics
e. Optional interface frameworks: Streamlit Python Framework
f. Operating System: Windows 11

1. CCTV Camera Configuration for Video Capture

Description: In this project, an IP-based CCTV camera with Full HD resolution (1920 × 1080) and a frame rate of around 25–30 frames per second is used to ensure clear and smooth video capture. The camera is equipped with a fixed 3.6 mm lens and infrared night vision capability that allows recording even in low-light conditions up to approximately 20 meters.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15391

Fig 1 E44Q MP4 Quad HD WiFi CCTV Camera

2. Home Screen

Description: The home screen provides an overview of the system, including system status, navigation options for live monitoring, and access to analytics and reports. It serves as the entry point to the surveillance platform.

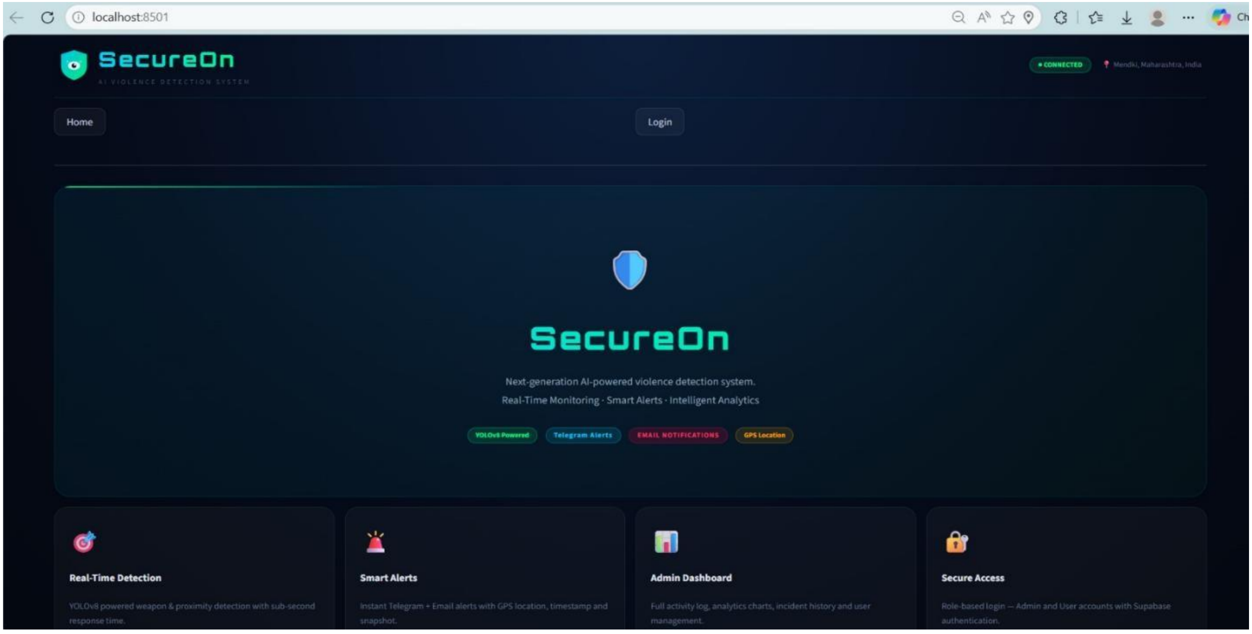

Fig 2 Project Website: SecureOn

3. Admin Control Panel

Description: This image showcases the Admin Control Panel of a security-focused web application, specifically displaying the Full User Activity Log. The interface lists chronologically ordered entries of system interactions, including user logins, navigation events, and logouts, with the most recent entry dated March 14, 2026. Each log entry identifies the user via email or an "Anonymous" tag, assigns a role (Admin, User, or None), and provides a precise timestamp and location tag for Mendki, Maharashtra, India. Notably, the log also flags critical system-generated alerts, such as a "Violence Detected" notification with a specific confidence score and frame reference, highlighting the application's integration with real-time AI monitoring.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15391

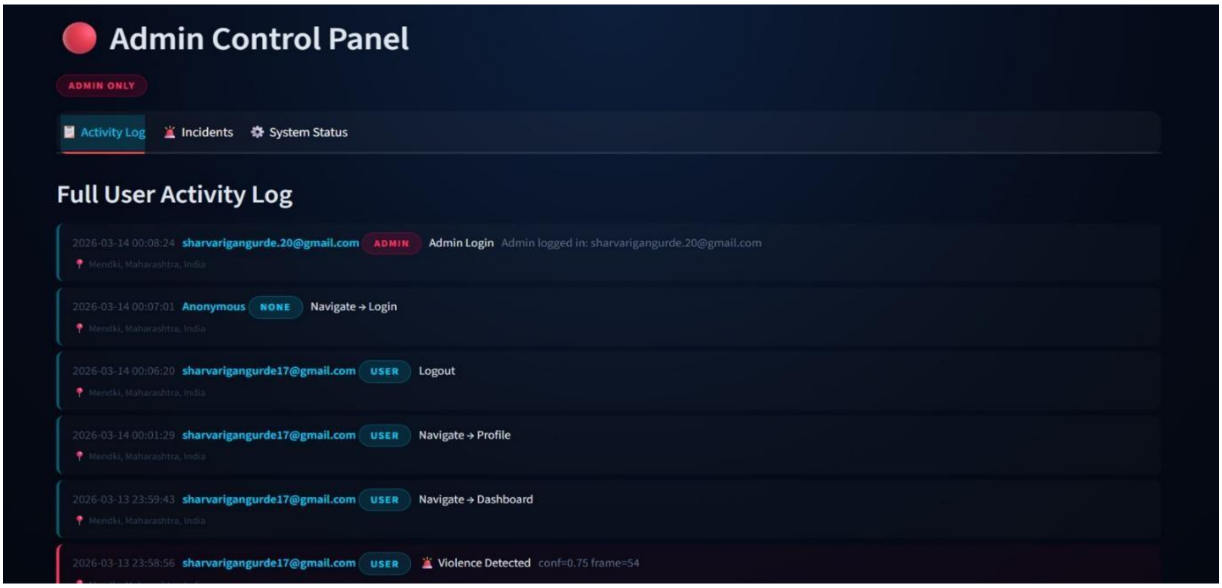

Fig 3 Effective Admin Control

IV. RESULT AND DISCUSSION

The Automated Human Violence Detection System successfully demonstrates the practical implementation of real-time violence detection using the lightweight YOLOv8 Nano architecture. The system efficiently analyzes live video streams and accurately identifies violent human interactions such as fights and aggressive behavior with minimal delay.
The experimental results indicate that the model achieves high detection accuracy while maintaining fast inference speed, making it suitable for real-time deployment. The lightweight nature of YOLOv8 Nano ensures low computational requirements, enabling the system to operate smoothly even on moderate hardware configurations. This makes it practical for deployment in classrooms, offices, hostels, and small-scale surveillance environments.

The system also demonstrates strong reliability through optimized confidence threshold tuning, which helps in reducing false positives and false negatives. Real time bounding box visualization, confidence score display, and automatic alert generation further enhance its effectiveness in active monitoring scenarios. Additionally, incident logging and report generation provide structured documentation for future investigation and security analysis.

Overall, the results confirm that AI-powered surveillance systems can significantly enhance public safety by minimizing dependency on continuous human monitoring. The automated detection capability reduces manual workload, increases monitoring efficiency, and improves emergency response time. Therefore, the proposed system proves to be a scalable, efficient, and intelligent solution for modern security applications.

4. YOLOv5 Real-Time Detection [Knife]
Description: This image captures a desktop environment demonstrating a real-time object detection system in action, likely utilizing the YOLO (You Only Look Once) algorithm. The main window shows a webcam feed where a person is holding an object (incorrectly identified as a "knife" with 35% confidence), while a Windows PowerShell terminal runs in the background. This terminal displays the live processing logs, listing detected classes like "persons," "chairs," and "dining tables" along with their respective processing times in milliseconds (averaging around 250ms).

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15391

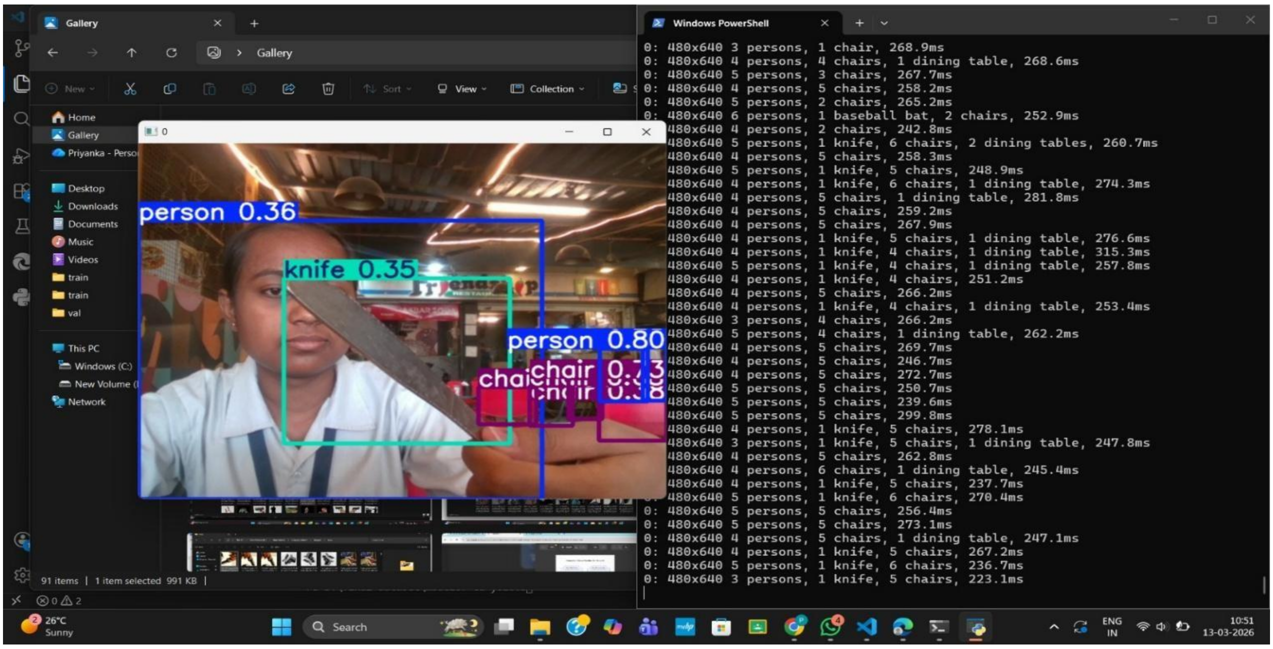

Fig 4 Violence Detection Demo (Knife detected)

5. AI Surveillance Control Center

Description: The AI Surveillance Control Center displays the live video feed with real- time detection results. It includes:

a. Live camera stream
b. Bounding box visualization
c. Confidence score display
d. Alert notification panel

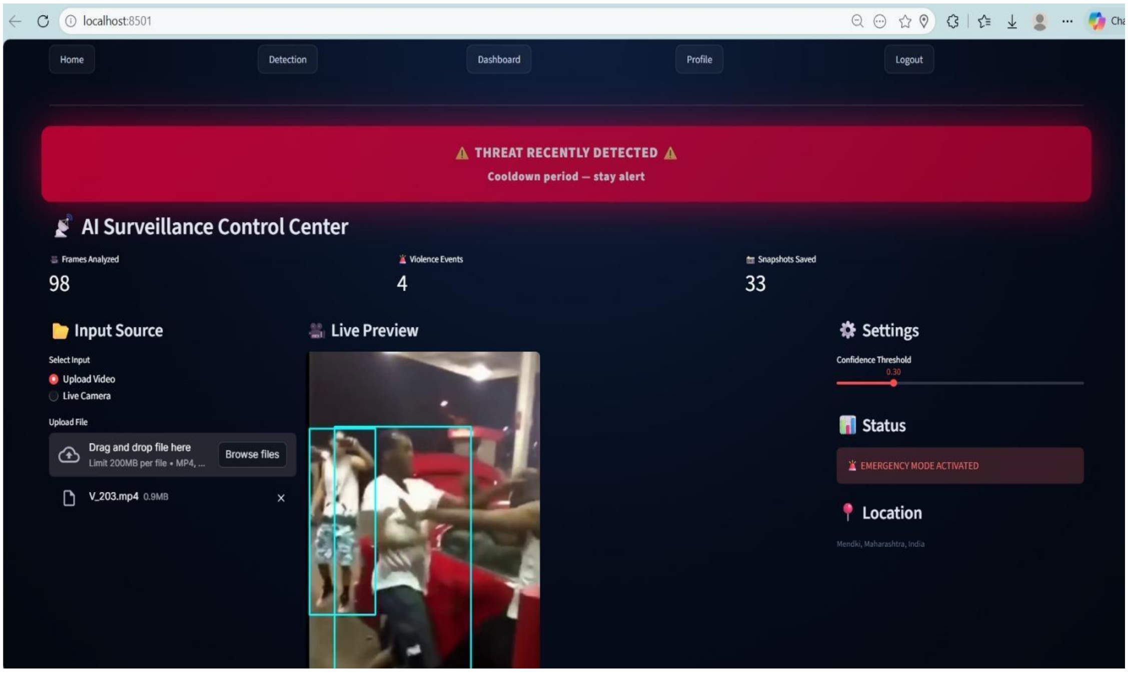

Fig 5 Final output

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.15391

6. Google Maps Location View

Description: This image displays a Google Maps screenshot pinpointing specific coordinates ($20^\circ00'53.0"N, 73^\circ49'14.7"E$) in Nashik, India. The map centers on a location along the Agra-Mumbai Highway, directly adjacent to the K.K. Wagh Institute of Engineering. The interface includes a left-hand navigation panel providing the location's "Plus Code" and quick-action tools, while the map view highlights local landmarks such as Apollo Hospitals, various banquet halls, and residential pockets like Amrutdham.

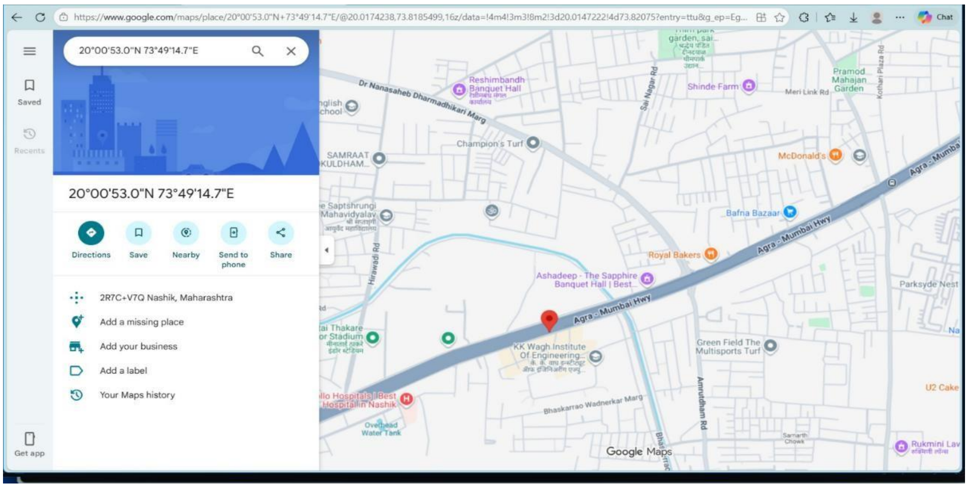

Fig 6 Displaying Live Location of the User

7. Telegram Alert Notification System

Description: The system integrates a Telegram chatbot to provide real-time alerts whenever a violent activity is detected in CCTV footage. When the detection model identifies a potential violence event, the system automatically captures the frame and sends a notification through the Telegram bot. The alert message includes important information such as the detection timestamp, user identification, confidence score of the model, frame number, and GPS location. A snapshot of the detected event is also attached to the message so that authorities or administrators can quickly review the situation. This real-time alert mechanism enables faster response to suspicious activities and improves the effectiveness of the surveillance system.

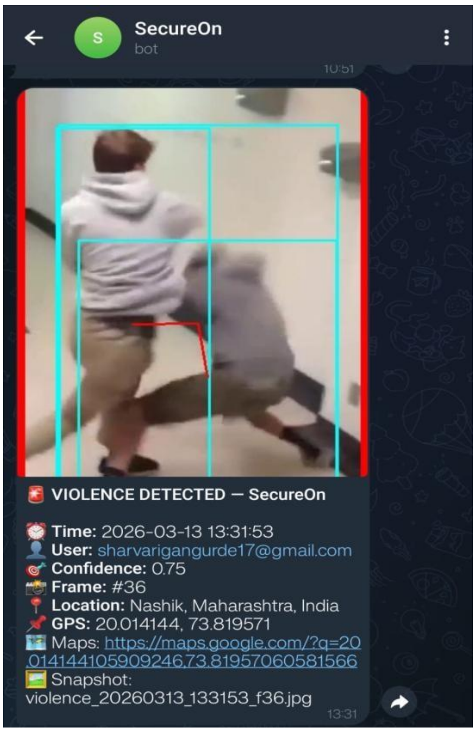

Fig 7 Telegram Alert when violence gets detected

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.15391

V. APPLICATIONS

1. **Educational Institutions –** Detects violent incidents in schools and colleges to ensure student safety.
2. **Public Surveillance Systems –** Monitors crowded public places like bus stations, railway stations, and markets to identify violent activities.
3. **Shopping Malls and Commercial Complexes –** Enhances security by automatically detecting fights or weapon threats in commercial areas.
4. **Smart City Security –** Integrates with smart city surveillance networks to improve real-time crime monitoring.
5. **Corporate Offices –** Helps maintain workplace safety by detecting aggressive or violent behavior.
6. **Hospitals and Healthcare Facilities –** Monitors emergency areas and waiting zones to prevent violent incidents.
7. **Residential Societies and Apartments –** Improves residential security by detecting suspicious or violent behavior.
8. **Parking Areas and Isolated Locations –** Provides safety monitoring in locations with limited human supervision.
9. **Law Enforcement Support –** Assists police and security agencies in early identification of violent activities.
10. **Event Security Management –** Monitors large gatherings such as concerts, festivals, and sports events to detect violence quickly

VI. CONCLUSION

The Automated Human Violence Detection using YOLOv8 Nano project demonstrates how modern artificial intelligence and real-time object detection technologies can significantly enhance security monitoring systems. By utilizing the lightweight and high-speed architecture of YOLOv8 Nano, the system efficiently processes live webcam video streams to detect violent activities such as fights, physical assaults, and aggressive interactions with improved speed and accuracy. The custom-trained model analyzes human movement and interaction patterns frame-by-frame and generates instant alerts when suspicious behavior is identified, enabling faster response and improved situational awareness in monitored environments.

The selection of YOLOv8 Nano plays a crucial role in the system's effectiveness. Due to its lightweight structure and optimized computation, it provides high detection performance while maintaining low hardware requirements. This makes the system suitable for deployment on systems with limited computational resources, such as standard laptops or edge devices. The balance between speed and accuracy ensures real time inference without significant latency, which is essential for security-based applications.

This project offers a more efficient and practical solution compared to traditional surveillance systems that depend heavily on continuous human supervision. Manual monitoring of CCTV footage is time-consuming, prone to fatigue, and often results in delayed responses or missed incidents. Human operators may struggle to maintain constant attention across multiple screens for extended periods. In contrast, the proposed system automates the detection process using deep learning and computer vision techniques, thereby reducing human workload, minimizing the possibility of oversight, and ensuring uninterrupted monitoring without fatigue or distraction.

The proposed autonomous violence detection system aims to enhance traditional surveillance systems by introducing intelligent video analysis capabilities. By using **YOLOv8** for real-time human detection and analyzing interactions between individuals, the system can automatically detect violent activities in surveillance footage.

This approach improves the efficiency of security monitoring and enables faster response to potential threats. The system can be applied in various environments such as smart cities, educational institutions, public transportation systems, and commercial facilities.

In conclusion, the Automated Human Violence Detection system makes real-time monitoring smarter, faster, and more efficient by leveraging the capabilities of YOLOv8 Nano for automated violence detection. It reduces dependency on manual supervision, strengthens rapid response mechanisms, improves documentation and accountability, and demonstrates the practical application of lightweight deep learning models in intelligent security systems. The project not only fulfills its objective of developing a real-time violence detection solution but also lays a strong foundation for future advancements in AI-driven surveillance technologies.

.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.15391

ACKNOWLEDGMENT

With a deep sense of gratitude, we would like to thank all those who guided and supported us throughout the selection, design, and development of our project. Their valuable insights, encouragement, and technical assistance played a crucial role in the successful completion of this work.

It is our proud privilege to express sincere thanks to **Prof. P. T. Kadave**, Principal of **K. K. Wagh Polytechnic**, for his support, encouragement, and permission to carry out this project.

We remain deeply indebted to **Prof. H. M. Gaikwad**, Head of the Department of Artificial Intelligence & Machine Learning, for his timely suggestions, constant motivation, and valuable guidance throughout the project.

We extend our special thanks to our Internal Faculty Guide, **Mr. H. M. Gaikwad**, and all the staff and technical members of the Artificial Intelligence & Machine Learning Department for their technical support, constructive feedback, and continuous encouragement during the completion of this project.

We also express our heartfelt appreciation to our classmates and friends for their cooperation, encouragement, and support, which greatly contributed to the successful completion of our project.

REFERENCES

[1] Francois Chollet (2021). Deep Learning with Python (2nd ed.). Manning Publications.
[2] Aurelien Geron (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2nd ed.).O'Reilly Media.
[3] Adrian Rosebrock (2019). Deep Learning for Computer Vision with Python. PyImageSearch.
[4] Joseph Redmon et al. (2016). You Only Look Once: Unified, Real-Time Object Detection.
[5] Ultralytics (2023). YOLOv8: Advanced Real-Time Object Detection Architecture.
[8] Stanford University – CS231n: Convolutional Neural Networks for Visual Recognition http://cs231n.stanford.edu/
[9] Kaggle – Violence Detection Datasets https://www.kaggle.com/
[10] IEEE – IEEE Xplore Digital Library https://ieeexplore.ieee.org

BIOGRAPHY

Name: Mr. H.M. Gaikwad
Qualification: **B.E. Computer Engineering Name: Sharvari Milind Gangurde**
Qualification: Diploma, Artificial Intelligence and Machine Learning
Name: Priyanka Dnyaneshwar Moule
Qualification: Diploma, Artificial Intelligence and Machine Learning
Name: Aarya Sharad Patil
Qualification: Diploma, Artificial Intelligence and Machine Learning
Name: Sejal Vijay Patil
Qualification: Diploma, Artificial Intelligence and Machine Learning