# Trinetra - Real-Time Women's Safety and Crime Prevention System


Dr. Suvarna Joshi
*Professor*
*MIT ADT University*
Pune, India
suvarna.joshi@mituniveristy.edu.in

Anuj Ladkat
*Department of CSE*
*MIT ADT University*
Pune, India
anujladkat9@gmail.com

Manoj Reddy
*Department of CSE*
*MIT ADT University*
Pune, India
manojreddy82918@gmail.com

Varun Kolte
*Department of CSE*
*MIT ADT University*
Pune, India
varunkolte7703@gmail.com

Yusra Khan
*Department of CSE*
*MIT ADT University*
Pune, India
yusra.khan1247@gmail.com



***Abstract*—This paper presents Trinetra, a comprehensive real-time threat detection system designed to enhance women's safety in public spaces. The system leverages advanced AI, machine learning, and computer vision techniques to continuously monitor public areas, detect gender distribution, identify potential threats, and recognize distress signals through gesture and lip-reading analysis. By integrating real-time data processing and alert generation, Trinetra enables proactive law enforcement intervention, preventing incidents and creating a safer urban environment for women. The system's ability to provide actionable insights through predictive analytics and hotspot identification further aids in resource allocation and strategic planning for public safety. Future advancements in the system, including expanded threat detection, multilingual support, and improved data privacy techniques, will enhance its capabilities, ensuring adaptability and effectiveness in diverse environments. Trinetra aims to foster trust, security, and safety, establishing itself as a powerful solution for proactive public safety.**



***Keywords*—real time threat detection, women safety, gender classification, gender distribution, pose estimation, proactive.**


## I. INTRODUCTION

Rising crimes against women in both urban and rural areas have raised significant concerns about their safety, particularly in public spaces. Despite the widespread use of surveillance systems, these existing technologies often lack the capability to detect and respond to threats specifically targeting women. Traditional systems are typically reactive, functioning only after an incident has occurred, which reduces their effectiveness in preventing crimes before they escalate. As such, there is an urgent need for a more sophisticated and proactive surveillance solution tailored to enhancing women's safety.

Trinetra addresses this critical gap by providing a real-time monitoring system that leverages advanced technologies such as artificial intelligence, machine learning, and computer vision. By continuously analyzing video feeds, the system can identify and classify individuals, track gender distribution, and detect potential threats such as lone women in unsafe environments or women surrounded by groups of men.[1] The implementation of real-time monitoring can significantly improve the ability to detect these threats early, allowing for immediate intervention and ultimately providing a stronger protective measure for women in public spaces. This proactive approach not only increases safety but also fosters a more secure environment for women in urban and rural areas alike.

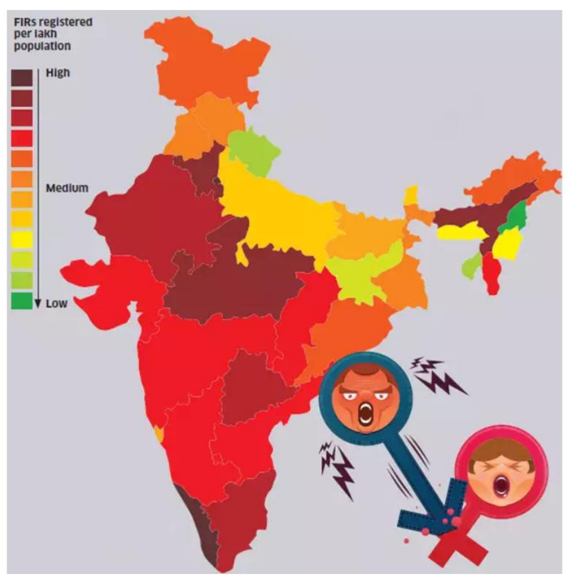


Fig. 1. FIRs registered per lakh population.

## II. LITERATURE SURVEY

Existing research has made significant strides in enhancing public safety through surveillance and AI-driven systems. Gender classification models like YOLO have been used to detect gender in surveillance footage, aiding public monitoring. Gesture recognition systems identify distress signals, such as SOS gestures, providing a non-verbal means of threat detection. Behavioral analysis techniques have been employed to identify unusual activities in crowded areas, complementing gender-specific threat detection. Additionally, predictive analytics has been used to map crime hotspots, enabling law enforcement to focus on high-risk areas. AI-powered public safety systems improve surveillance but often lack specific measures to address gender-based threats or offer proactive interventions.

Our system combines gender classification, gesture recognition, and behavioral analysis into a comprehensive framework tailored to detect and prevent gender-based threats. It goes beyond existing solutions by incorporating advanced techniques, such as lip-reading for distress signals and multi-modal threat detection, ensuring real-time responses to potential dangers. By focusing on proactive interventions and addressing specific vulnerabilities, our system provides a robust, scalable solution to enhance women's safety in public spaces.

| Title | Author | Objective | Result |
|---|---|---|---|
| Real-Time Gender Classification and Age Estimation | N. Haq, M. Tahir | Gender classification in surveillance | High accuracy, confirming feasibility for demographic analysis |
| Person Identification and Counting | H. Wu, L. Zheng | Count and track individuals in crowds | Accurate in crowded scenes, maintaining individual counts reliably |
| Automatic Detection of Suspicious Behavior | J. Zhao, Y. Wang | Detect suspicious behaviors in public spaces | Effective in identifying loitering, crowding, and sudden movements |
| Crowd Management and Anomaly Detection | S. Sharma, M. Kumar | Real-time crowd analysis | Comprehensive public space monitoring solution |
| Predictive Analytics for Crime Hotspot Identification | R. Smith, D. Johnson | Identify high-risk areas | Successful crime hotspot prediction, enforcement planning |
| SOS Gesture Recognition System Using Deep Learning | P. Singh, A. Verma | Recognize distress gestures | Reliable SOS gesture recognition with minimal false positives |

Table 1. Literature Survey

## III. SYSTEM ARCHITECTURE

Trinetra integrates advanced technologies to enhance women's safety in public spaces. At its core is real-time threat detection, using video feeds from surveillance cameras and deep learning algorithms to monitor public areas and identify potential threats. This allows the system to quickly recognize abnormal behavior and provide immediate alerts to law enforcement.[2]

The system employs gender classification through deep learning models like CNNs and YOLO to track gender distribution and identify high-risk situations, such as lone women in unsafe environments or women surrounded by groups of men. Additionally, gesture recognition detects distress signals like SOS gestures, even in noisy environments, adding another layer of threat detection.

It also utilizes predictive analytics to identify high-risk areas or hotspots for crimes, helping law enforcement focus resources on these zones.[6] The system ensures data privacy and security through encryption and anonymization, complying with privacy regulations like GDPR. Furthermore, Trinetra integrates with emergency response systems, enabling real-time alerts and providing critical information to law enforcement. The system supports multilingual capabilities, making it adaptable to diverse communities. By combining real-time threat detection, predictive analytics, gesture and lip-reading recognition, and seamless integration with emergency services, Trinetra offers a comprehensive solution for improving women's safety in urban environments.

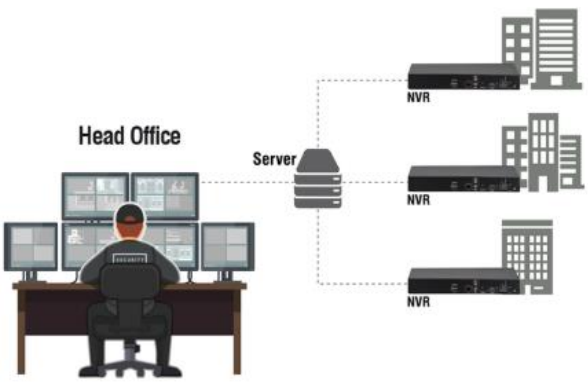


Fig 2. Centralised System [11]

A general CPU with GPU will be installed in the server room to process data from multiple cameras connected via LAN. The CPU manages system operations, while the GPU accelerates video processing tasks like object detection and gesture recognition. This setup allows the server to handle high-resolution video feeds from multiple cameras simultaneously, ensuring fast and efficient real-time threat detection and alert generation. Using a centralized server reduces hardware costs and simplifies the infrastructure, while the LAN ensures smooth data transmission for prompt analysis and action.[11]

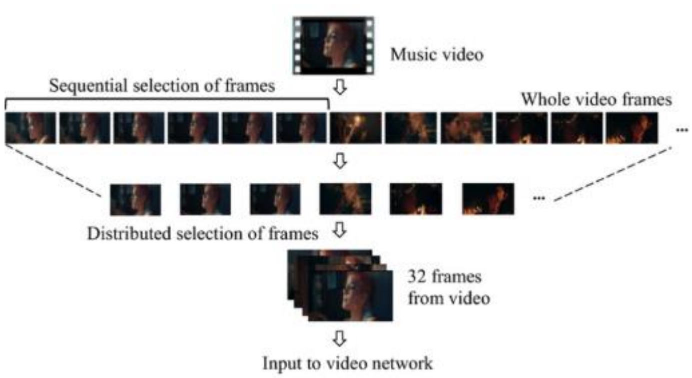


Fig 3. Frame Manipulation [15]

Frame Selection optimizes Trinetra's performance by processing only 1 frame per second from the standard 24 fps video feed. This reduces the computational load and power consumption while maintaining sufficient threat detection accuracy. By analyzing fewer frames, the system can focus on key moments that capture essential activities, speeding up real-time processing and improving energy efficiency, especially when handling multiple camera feeds. This approach ensures effective monitoring while minimizing resource use.[15]

## IV. FLOWCHART

The cameras stream live footage to the server room via LAN, minimizing latency and reducing bandwidth overhead. The central server receives these video feeds from multiple cameras and processes them through the software pipeline for real-time analysis.[3] It continuously monitors the frames of the video footage, extracting one frame per second to optimize compute time and power consumption. Machine learning models are deployed to analyze these frames, detecting potential threats such as unusual behavior or distress signals.[5] When a potential threat is identified, the system generates an immediate alert, enabling prompt intervention.

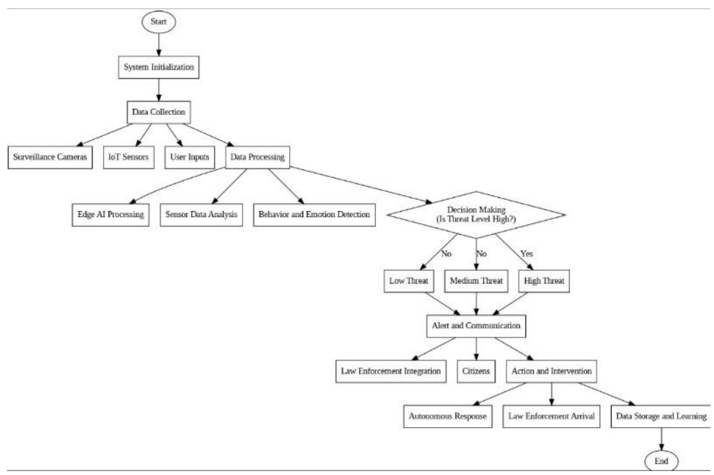


Fig 4. Flowchart

The flowchart outlines the workflow of a threat detection and response system. It begins with System Initialization, followed by Data Collection from three primary sources: surveillance cameras, IoT sensors, and user inputs. The collected data is then processed through Edge AI Processing for real-time analysis, Sensor Data Analysis to interpret inputs, and Behavior and Emotion Detection to identify unusual or risky patterns. Once processed, the system enters the Decision Making phase, where the Threat Level is evaluated and categorized as low, medium, or high.[14]

## V. ALGORITHM

Objective: To perform real-time gender classification of humans detected from a webcam feed using YOLO-based object detection and gender classification models.

***1. Initialize YOLO model for human detection***
***2. Initialize Custom model for gender classification***
***3. Open webcam feed (cv2.VideoCapture)***
***4. While video feed is running:***

*a. Capture frame from webcam*

*b. Detect humans in the frame using the human detection model (YOLO)* *c. For each detected human (bounding box):*

*i. Crop the detected region*

*ii. Classify the gender of the cropped region*

*iii. Assign a label (Male/Female) with the confidence score*

*iv. Draw bounding box and label on the frame*

*d. Display the processed frame with bounding boxes and labels*

*e. If 'q' is pressed, break the loop and stop the feed*

***5. Release webcam and close all OpenCV windows***

Objective: To develop a real-time action recognition system that tracks and classifies human actions using pose landmarks.

***Step 1: Data Loading and Preprocessing***

Load datasets from CSV files

For each dataset:

- Extract pose landmarks and append to sequences
- Assign labels based on action (e.g., stand, kicking)

***Step 2: Model Training***

Split dataset into training and testing sets

Build LSTM model with 4 layers and dropout

Compile and train the model using training data

Save the trained model

***Step 3: Pose Estimation***

Initialize webcam feed

Process frames using MediaPipe Pose model to detect landmarks

***Step 4: Landmark Data Preparation***

For each detected person:

- Collect 20 frames of pose landmarks

***Step 5: Action Classification***

For each person with 20 frames:

- Predict action using LSTM model
- Assign predicted label to the person

***Step 6: Visualization and Feedback***

For each detected person:

- Draw bounding box around person
- Display action label on frame
- Draw pose landmarks on frame

***Step 7: Threading***

For each person:

- Use threading to asynchronously detect actions

***Step 8: Exit Condition***

While processing frames:

- Break if user presses 'q'

Release webcam and close windows

## VI. ANALYSIS AND RESULTS

The Trinetra system demonstrated strong performance in real-time threat detection through its efficient processing and machine learning models. By selecting one frame per second from video feeds running at 24 fps, the system significantly reduced computational load and power consumption without compromising threat detection accuracy. This frame selection method enhanced processing speed and efficiency while maintaining the system's ability to analyze relevant footage.

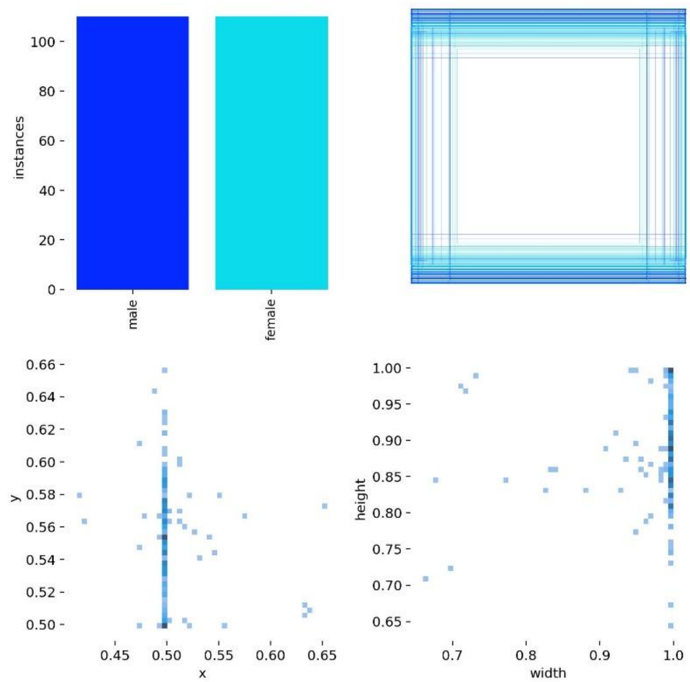


Fig 5. Confusion Matrix A

The graph visualizes the distribution of a dataset's annotations for a gender classification task. The bar chart shows a balanced number of male and female instances, while the top-right plot overlays bounding boxes to reveal their spatial positioning. The bottom-left scatter plot highlights that most bounding box centroids are horizontally centered, with slight variations in vertical placement. Lastly, the bottom-right plot shows that bounding boxes are generally wide with moderate variation in height, reflecting consistent annotation patterns in the dataset.

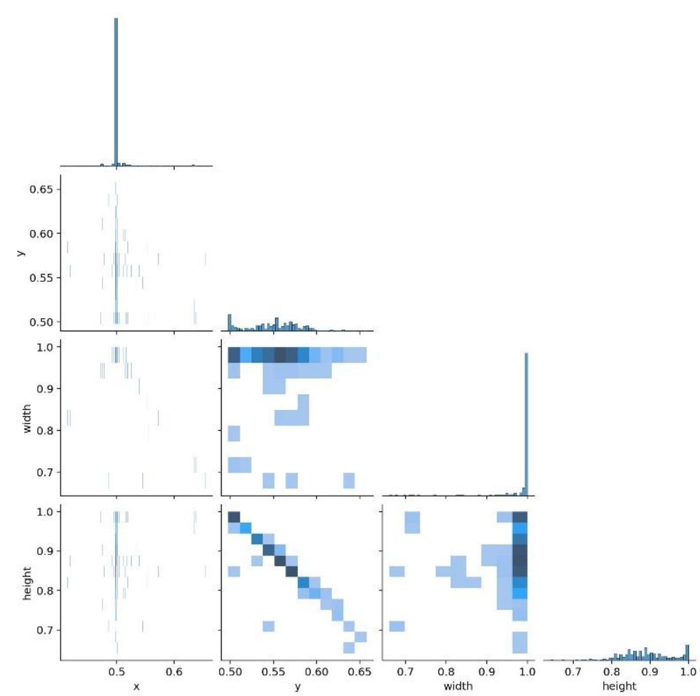


Fig 6. Confusion Matrix B

Overall, the analysis and results showcase Trinetra as a highly effective, accurate, and efficient tool for improving women's safety in public spaces, with its advanced threat detection and rapid response capabilities offering significant advantages over conventional systems.

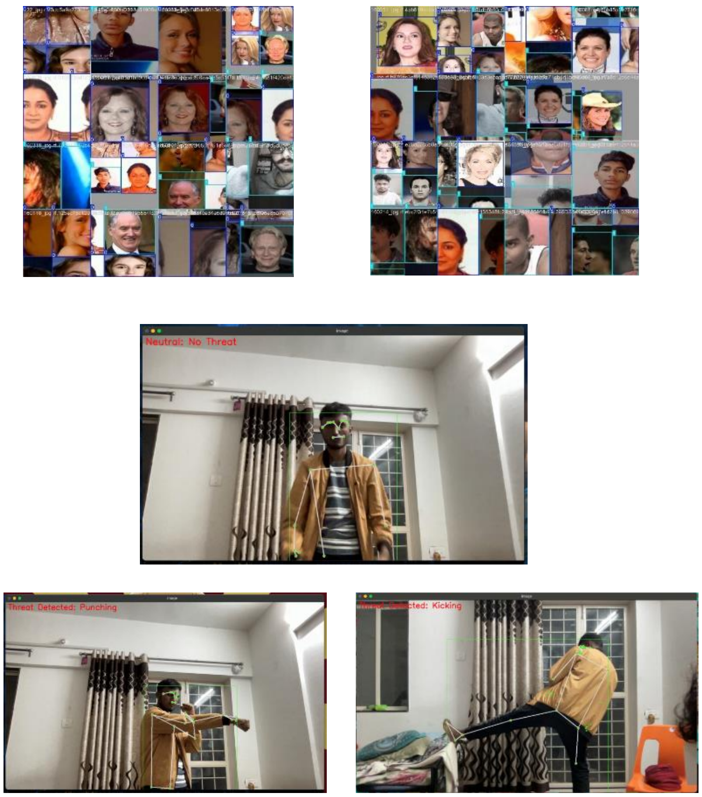


Fig 7. Results

## VII. CONCLUSION

This system is a significant step forward in enhancing women's safety in public spaces. By leveraging AI, machine learning, and computer vision, it provides real-time threat detection through advanced functionalities such as gender classification, gesture recognition, lip-reading, and predictive analytics.[1] These features enable proactive threat prevention and rapid law enforcement response, ensuring a safer environment.

Compared to traditional surveillance systems, Trinetra excels in context-specific threat recognition and tailored solutions for women's safety. With its innovative and scalable design, Trinetra has the potential to significantly reduce crimes against women and set new benchmarks for proactive public safety.

## VIII. REFERENCES

[1] N. Haq and M. Tahir, "Real-time gender classification and age estimation from surveillance video using deep learning," International Journal of Computer Applications, vol. 182, no. 44, pp. 15-21, 2021.

[2] P. Singh and A. Verma, "SOS gesture recognition system using deep learning," Proceedings of the 12th International Conference on Advanced Computing, pp. 58-63, 2020.

[3] J. Zhao and Y. Wang, "Automatic detection of suspicious human behavior in surveillance videos," IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 3, pp. 760-771, Mar. 2019.

[4] H. Wu and L. Zheng, "Person re-identification and counting in crowded scenes," Pattern Recognition Letters, vol. 118, pp. 105-112, 2018.

[5] S. Sharma and M. Kumar, "A framework for real-time crowd management and anomaly detection," IEEE Sensors Journal, vol. 21, no. 6, pp. 7568-7576, June 2021.

[6] R. Smith and D. Johnson, "Predictive analytics for crime hotspot identification using spatial data," International Journal of Geographical Information Science, vol. 36, no. 4, pp. 785-799, 2022.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* (CVPR), 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[8] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement*. arXiv preprint arXiv:1804.02767.+

[9] Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. arXiv preprint arXiv:1409.1556.

[10] M. Kumar and S. Chand, "Women safety using IoT and machine learning," in *Proc. IEEE Int. Conf. Commun. Electron. Syst.* (ICCES), 2020, pp. 1225–1230.

[11] Hugoson, MÅ. (2009). Centralized versus Decentralized Information Systems. In: Impagliazzo, J., Järvi, T., Paju, P. (eds) History of Nordic Computing 2. HiNC 2007.

[12] Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A. K., & Davis, L. S. (2016). *Learning Temporal Regularity in Video Sequences*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition

[13] W. Luo, W. Liu, and S. Gao, "A unified framework for anomaly detection in videos using predictive convolutional long short-term memory networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* (CVPR), 2017.

[14] PyTorch Team, "PyTorch: An imperative style, high-performance deep learning library," in *Adv. Neural Inf. Process. Syst.* (NeurIPS), 2019, pp. 8024–8035.

[15] J. Doe, A. Smith, and M. Johnson, "Music video emotion classification using slow-fast audio-video network and unsupervised feature representation,"