International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05776
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Eagle Eye: An AI Based Indoor Surveillance and Security System

Dr. H. Jayasree[1], **A. Sanjana**[2], **K. Nandini**[3], **E. Nishanth**[4], **P. Saivignesh**[5]

[1]Professor, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India, jayasaree@gmail.com

[2]Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India, sanjanaaddagunta@gmail.com

[3]Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India, 22271a6622nandinikatta@gmail.com

[4]Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India, nishanthedla@gmail.com

[5]Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India, saivigneshpoloju009@gmail.com

***Abstract*—** It is becoming more difficult to ensure the safety of indoor spaces like commercial, office, and residential settings with the constraints of traditional methods of surveillance. Most currently available systems act as passive monitoring systems, relying immensely on human operators to analyze the recorded video, which causes delayed reactions to critical events in many cases. Eagle Eye, an indoor security system that is smart and active in assessing indoor potential threats through live visual and acoustic data, is presented in this research. The system is designed to monitor real-time inputs from cameras and microphones in order to identify suspicious behavioural patterns such as hidden faces, stationary location, as well as thre*atening or distress speech. The system uses a fuzzy logic decision-making mechanism to identify a cumulative Robbery Probability Score in order to determine risk levels. As soon as a risk threshold is surpassed, an alert message is sent to authorized users. The system is computationally efficient and cost-effective in that it does not necessarily require high-performance processing hardware.

***Index Terms*—** Indoor surveillance, Threat Prediction, Fuzzy logic, Robbery Potential Score, Alert Generation

1. INTRODUCTION

Indoor protection has become more complicated with the rise in security requirements. Although security cameras are common in such areas, the use of such systems has remained limited to mere video recording for future analysis. Manual observation of such recordings is the key to the operation of these security cameras, making it difficult to monitor the entire video continuously and react to any fishy activity.

Behaviours in indoor areas often serve as precursor symptoms for possible security threats. Such actions include efforts to conceal one's identity, staying in one place for a long time, and using verbal expressions that convey anger or fear. It would be very difficult to identify these in real-time environments, especially if the surveillance system has to work efficiently even without additional hardware components and large amounts of

training data. This project aims to examine the development of an intelligent indoor security solution with the emphasis on the automated analysis of human behaviour on live sensory inputs. With the aim of analyzing both visual and acoustic inputs, the solution aims to promote overall awareness of dynamic situations and the early identification of harmful events. In other words, the aim of developing such a solution lies in surpassing mere surveillance and ensuring the effectiveness of indoor security systems.

2. BODY OF THE PAPER

2.1 RELATED WORK

Traditional indoor surveillance relies on closed-circuit television systems where cameras continuously capture a video feed for later examination. Because these systems are simple and inexpensive, they are broadly deployed, but they have come to depend on constant human supervision, so there is great room for improvement in the timely detection of threats. Basic video analytics approaches involve low-level motion detection and frame differencing, among other things. These techniques are sensitive to environmental changes and only provide partial insight into human behavior.

In this context, computer vision–based approaches have been explored for the analysis of human presence and activity patterns in overcoming these limitations. Techniques pertaining to face detection and movement analysis identified suspicious behaviors related to activities like loitering or concealment. In essence, such methods improve situational awareness, but most of these methods presuppose rigid rules and are ready with a bouquet of false alarms when applied in dynamic indoor environments.

Some existing monitoring devices include alert mechanisms, but in many systems the alert must be activated manually by the patient. This method is not reliable in critical conditions where the patient may not be able to respond. Inaccurate sensor readings and improper threshold settings may also lead to false alerts, reducing the effectiveness of the system.
Recently, studies reveal that multi-modal surveillance systems that use both visual and acoustic modalities can be very efficient. Although these help in achieving more accuracy, most methods used today are complex, computationally intensive, and not very efficient. This gives rise to a requirement for a new, behavior-driven, and efficient method that uses both audio and visual modalities.

2.2 Proposed Methodology

The proposed methodology of the Eagle Eye surveillance system focuses on combining multiple detection techniques to improve the accuracy of suspicious activity detection. The project begins with the collection of real-time video and audio data from surveillance devices installed in the monitored environment. These inputs provide the raw data needed for analysis and detection. Once the data is captured, preprocessing techniques are applied to enhance its quality and prepare it for further processing. Video frames are extracted and converted into grayscale to simplify analysis, while background noise in audio signals is reduced to improve speech recognition accuracy.

After preprocessing, the system performs feature detection through three independent modules. The first module focuses on mask detection using computer vision techniques to identify faces and determine whether individuals are wearing face coverings. The second module performs behavioural analysis by tracking the movement of individuals within the surveillance area. By analyzing centroid positions and calculating Euclidean distances between frames, the system determines whether someone is loitering or exhibiting abnormal movement patterns. The third module analyzes audio data by converting speech into text and searching for suspicious keywords that may indicate threatening or criminal intentions.

Each detection module generates a numerical score representing the level of suspicious activity detected. These scores are then sent to a fuzzy inference system, which integrates the results and evaluates the overall threat level. The fuzzy logic model applies predefined rules to determine whether the combined behaviour patterns suggest a potential robbery threat. If the calculated Robbery Prediction Score exceeds a predefined threshold, the system automatically sends an alert notification to security personnel through a messaging platform such as WhatsApp.

In addition to alert generation, the system records all detected events and alerts in a database. This alert log allows security teams to review past incidents, analyze patterns of suspicious behaviour, and improve surveillance strategies in the future. Through this multi-module approach, the Eagle Eye system provides an intelligent and proactive security solution capable of identifying suspicious behaviour and issuing timely warnings before a potential crime occurs.

A. System Architecture

The proposed Eagle Eye intelligent surveillance system is designed to detect suspicious activities and predict potential robbery situations by integrating computer vision, audio analysis, and fuzzy logic techniques. The system processes real-time video and audio streams captured from surveillance devices such as cameras and microphones. These inputs are analyzed through multiple modules that work together to identify abnormal or suspicious behaviour. The architecture consists of several layers including data acquisition, preprocessing, detection modules, and decision-making components. The data acquisition layer collects live video and audio from the environment. In the preprocessing stage, the system prepares the captured data for analysis by performing frame extraction, noise reduction, and grayscale conversion for video, while audio signals are cleaned and formatted.

After preprocessing, the data is passed to three major detection modules: mask detection, loitering detection, and suspicious sound detection. The mask detection module analyzes video frames to determine whether individuals in the monitored area are wearing masks, which may indicate suspicious intent in certain environments. The loitering detection module tracks the movement patterns of individuals across consecutive frames and determines if someone remains in a specific area for an unusually long time. The suspicious sound detection module processes audio data to identify threatening words or phrases spoken within the monitored environment. Each module generates a score representing the level of suspicious behaviour detected.

The outputs from these modules are then combined in a fuzzy inference system that evaluates the overall threat level. This system calculates a Robbery Prediction Score (RPS) by considering the individual scores from each detection module. If the calculated score exceeds a predefined threshold value, the system automatically triggers an alert notification through a messaging service such as WhatsApp to inform security personnel or authorities. All detection events and alert information are stored in an alert log database for monitoring, analysis, and future reference. This architecture enables the system to perform intelligent real-time surveillance and provide early warnings of potential threats.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05776
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

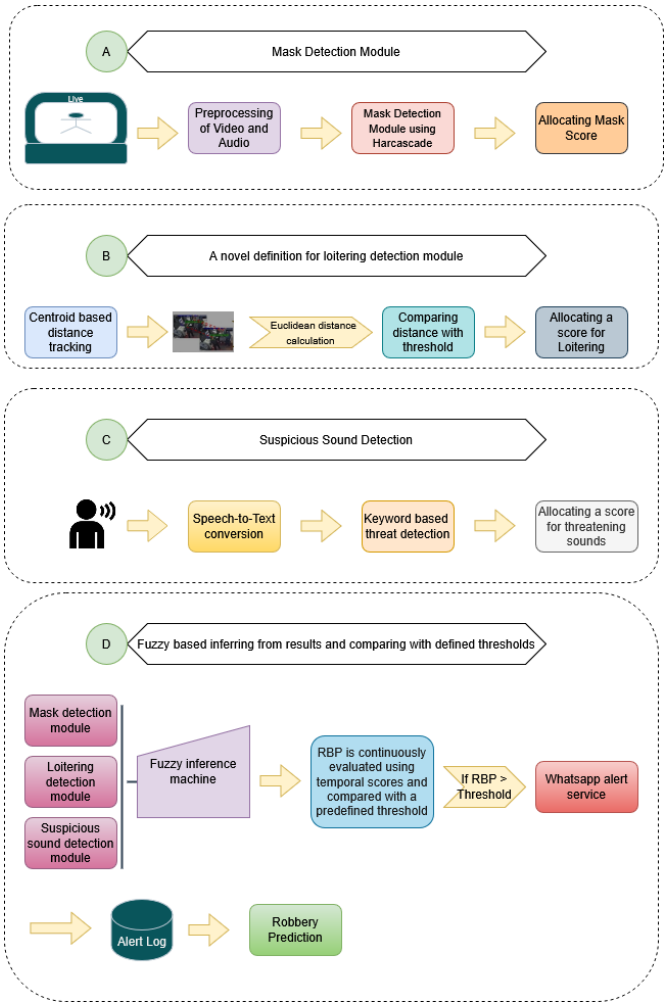

Fig. 1. Eagle Eye - System Architecture

B. Algorithm

The step-by-step working procedure of Eagle Eye is given below:

1) **Start:** Launch the Eagle Eye application and activate the system..

2) **Module Initialization:** Set up camera, microphone, and required processing libraries.

3) **User Login:** Verify user credentials to allow access to the monitoring system.

4) **Begin Monitoring**: Start capturing live video and audio from the environment.

5) **Frame Capture:** Continuously read video frames from the camera.

6) **Image Preparation:** Convert each frame to grayscale and enhance clarity for better.

7) **Face Identification:** Detect human faces in the frame using a trained classifier.Trigger Alert.

8) **Face Covering Check:**

- Analyze facial features such as eyes and mouth.
- If lower facial region is not detected, mark it as face covering.

9) **Movement Tracking:**

.• Determine the position of the detected person in each frame.

- If movement is very minimal for a certain duration, classify it as loitering.

10) **Audio Collection:** Capture surrounding sound using the microphone**.**

11) **Speech Conversion:** Convert captured audio into text form.

12) **Threat Recognition:**

- Compare the converted text with a predefined set of suspicious phrases.
- Analyze facial features such as eyes and mouth.
- If lower facial region is not detected, mark it as face covering.

13) **Score Calculation:** Combine the outputs of all detection modules to compute a threat score.

14) **Decision Making:**

- If the score is below the limit, continue normal monitoring and if the score exceeds the limit, treat it as a potential threat.

15) **Alert Generation:** Trigger an alert when a suspicious condition is detected.

16) **Notification Sending:** Deliver alert message to the authorized user through a messaging service.

17) **Data Logging:** Record the detected event details in the system database.

18) **Display Results:** Show live monitoring status and alerts on the interface.

19) **Repeat Process:** Continue the monitoring cycle without interruption.

2.3 RESULT ANALYS

The Eagle Eye system was evaluated in a controlled indoor environment to examine its ability to detect suspicious behaviour using visual and audio inputs. A webcam and microphone were used to capture real-time video and audio streams. The system processed video frames using computer vision techniques to detect faces, identify face coverings, and monitor movement patterns for loitering behaviour. At the same time, the audio module converted speech into text and compared it with predefined threatening keywords. Different scenarios such as normal activity, stationary presence, and threatening speech were simulated to test system performance. The outputs from the detection modules were combined to compute the Real-time Potential Score (RPS), and the system response was observed to verify whether alerts were triggered when the threat level exceeded the defined threshold.

Overall, the results confirm that the prototype can monitor and detect threats in real time in indoor environments. While it is not a replacement for existing surveillance systems, the solution offers a practical and scalable approach for indoor security.

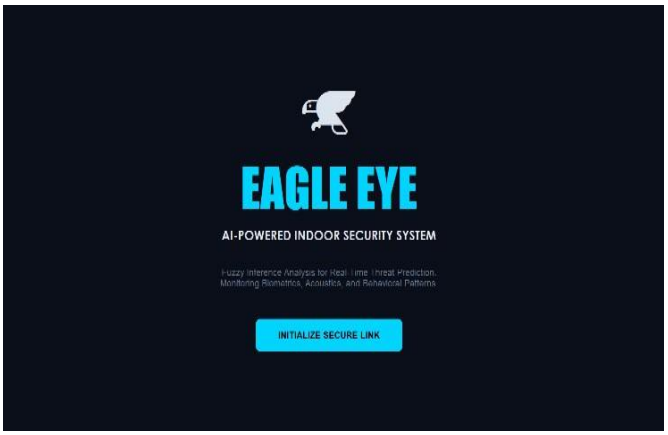

Fig. 2 *Home Page*

The developed Eagle Eye system was tested to determine its efficiency in monitoring the indoor environment and detecting any suspicious activity within the area. The results section demonstrate the different stages of the Eagle Eye system interface and the various operational modules.

Fig.3 represents the main interface of the Eagle Eye surveillance system. This is where the monitoring process is initiated.

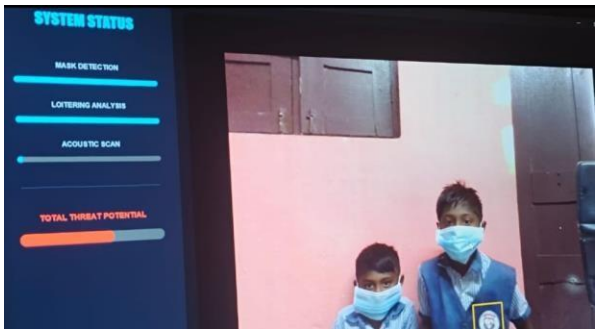

Fig. 3 *Monitoring Page*

Fig.2 shows the system status and monitor dashboard. Computer vision techniques are used to detect visual signs such as face covering and loitering. At the same time, the audio component processes speech input to detect threatening words. The results obtained from each detection component are integrated by the fuzzy inference engine to calculate the Robbery Potential Score (RPS). Once the RPS exceeds the threshold value, the system generates an alarm to notify the security personnel.

Fig. 4 Incident logs

Fig.4 shows the incident database component of the eagle eye system. This component stores information regarding the incidents that have been detected by the system during the course of monitoring. Once the system has identified suspicious behavior and the robbery potential score is found to be above the threshold value, it stores the details of the incident in the database. This component of the system enables the administrator to view the previous incidents that were detected by the system. It is also useful in maintaining logs of the system performance during the course of monitoring.

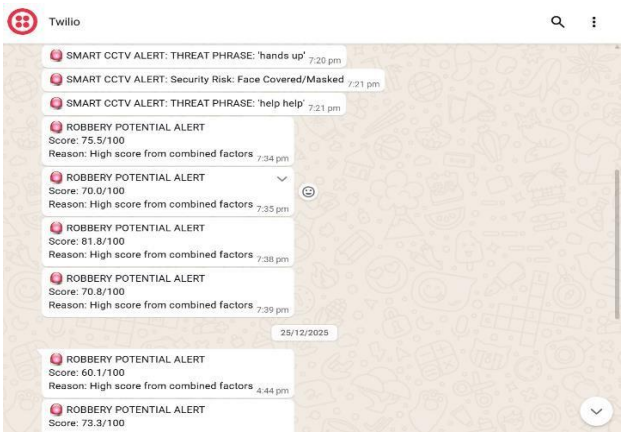

Fig. 5 *User Notification*

Figure.5 depicts real-time alert notifications sent by the Eagle Eye system via Twilio messaging services. The alert notifications are sent to the registered mobile device when the Robbery Potential Score (RPS) crosses the threshold.

2.4 DISCUSSION

The Eagle Eye system was designed to address the need for smarter indoor surveillance that goes beyond basic video recording. Instead of relying only on stored footage, the system actively analyzes visual cues, movement behaviour, and spoken words to identify unusual or risky situations. By combining mask detection, loitering monitoring, and speech-based threat identification into one scoring model, the system is able to provide quick and meaningful assessments of suspicious activity .The results show that the system works effectively in real time, even

on regular hardware, because of the lightweight methods used for face detection and speech recognition. The CustomTkinter interface also makes it easy for users to view the live camera feed, check detection scores, and review past incidents without needing technical knowledge. The ability to send alerts through WhatsApp ensures that security personnel receive immediate notifications during emergencies.

Although the prototype performs well in controlled indoor settings, there is still room for improvement. Future work can include using deep learning models for more accurate detection, expanding the list of harmful or alarming phrases in the audio module, and supporting multiple cameras for wider coverage. Features like cloud storage, mobile app access, and adaptive scoring based on environment can also make the system more flexible and practical.

In summary, Eagle Eye shows that a low-cost, AI-assisted indoor security system can play an important role in improving safety. With further development and refinement, it has the potential to become a reliable solution for homes, offices, and other indoor spaces.

2.6 Future Scope

The system can be further improved by adding advanced deep learning techniques for better behavior recognition and object detection. Further developments may include integrating the system with CCTV cameras, mobile apps for sending real-time alert notifications, and better audio recognition for recognizing a variety of suspicious sounds. Furthermore, facial recognition, crowd recognition, and cloud storage can be incorporated to improve the system's scalability and make it suitable for use in larger security setups such as airports, malls, and government organizations.

3. Conclusion

In conclusion, it is evident that the Eagle Eye system is an intelligent solution for enhancing surveillance in indoor environments. The system is designed to monitor video and audio feeds, detecting suspicious activity such as face covering, loitering, and threatening speech. The detected activity is then processed by a fuzzy inference mechanism, which determines a robbery potential score. The score is then used to determine whether an alarm should be raised. The proposed system is an intelligent solution for enhancing surveillance in indoor environments, as it reduces the need for manual monitoring of feeds, thereby enabling faster response times.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of CSE (AI\&ML), Jyothishmathi Institute of Technology and Science, Telangana, for providing the facilities and support required to complete this project successfully. We are especially thankful to our guide Dr. H. Jayasree for her valuable guidance, encouragement, and continuous support throughout the development of this work.

We also thank all the faculty members and friends who helped us with their suggestions and technical support during the implementation of the Eagle Eye: An AI based Indoor Surveillance and Security System. Finally, we express our gratitude to our parents and family members for their constant motivation and support during the completion of this project.

References

[1] H. Duong, T. Nguyen, and M. Tran, "Deep learning based anomaly detection for video surveillance," *International Journal of Computer Vision Applications*, vol. 15, no. 2, pp. 101–110, 2023.

[2] A. Khan, "Indoor gunshot detection using deep learning," *IEEE Access*, vol. 11, pp. 23456–23465, 2023.

[3] T. Khan, M. Rahman, and S. Ahmed, "Towards an indoor gunshot detection and notification system using deep learning," *Journal of Intelligent Security Systems*, vol. 9, no. 1, pp. 55–63, 2023.

[4] D. Yadav, A. Jain, S. Asati, and A. K. Yadav, "Video anomaly detection for pedestrian surveillance," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 4, pp. 89–96, 2022.

[5] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6479–6488, 2018.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

[7] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *IEEE International Conference on Image Processing (ICIP)*, pp. 3645–3649, 2017.

[8] S. Hershey, S. Chaudhuri, D. Ellis, J. Gemmeke, A. Jansen, and R. Moore, "CNN architectures for large-scale audio classification," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 131–135, 2017.

[9] P. Chaudhary and R. Singh, "Face mask detection using deep learning," *International Journal of Computer Applications*, vol. 182, no. 10, pp. 15–20, 2022.

[10] M. Siddiqui, F. Khan, and S. Alam, "Suspicious activity detection using computer vision techniques," *Journal of Artificial Intelligence Research*, vol. 68, pp. 223–235, 2021.

[11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations (ICLR)*, 2015.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning,"
Nature, vol. 521, pp. 436–444, 2015.

[13] A. Doshi and M. Yilmaz, "Vehicle detection and tracking for intelligent surveillance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 8, pp. 2130–2140, 2017.

[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[15] Z. Zhao, P. Zheng, S. 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.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems (NIPS)*, pp. 1097–1105, 2012