International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

# AI-Driven Real-Time Surveillance: Anomaly Detection and Notification System

**Vandana K H [*1], Nishitha S [*2], Sudeepthi Rao P S [*3], Varshitha D [*4]**
**Prof. Rakshatha S [*5]**
[*1*2*3*4] Student, Department Of ISE, Jyothy Institute Of Technology, Bangalore, India.
[*5] Assistant Professor , Department Of ISE, Jyothy Institute Of Technology, Bangalore, India.

## ABSTRACT

Surveillance technologies are rapidly transitioning from traditional passive monitoring systems to intelligent platforms capable of detecting anomalies in real-time. This study presents an AI-powered surveillance framework designed to identify violent behavior across diverse input formats, such as static images, recorded videos, and live webcam feeds. Leveraging advanced deep learning architectures, including YOLO and TensorFlow-based models, the system delivers high-precision violence detection with minimal latency. On identifying a threat, it initiates instant alerts through visual prompts and sound notifications, ensuring rapid situational awareness. The interface is intuitively built to support real-time configuration of detection parameters and the monitoring of performance indicators such as detection frequency, operational uptime, and frame processing speed. Engineered for scalability and resilience, the system demonstrates strong applicability in domains like public safety, institutional monitoring, and content regulation. Its core advantages include high detection accuracy, efficient real-time processing, and adaptability to various data sources. By combining cutting edge AI strategies with responsive caution instruments, the framework offers a reliable, computerized arrangement for improving danger location in observation situations.

Keywords: Real-Time Surveillance, Anomaly Detection, Violence Detection, YOLO, TensorFlow, Deep Learning, Smart Monitoring, Alert System

## INTRODUCTION

In recent years, the demand for intelligent surveillance systems has increased significantly, driven by growing concerns over public safety and the rising complexity of modern security threats. Conventional observation approaches, which depend fundamentally on manual observing, are regularly restricted by human weariness, moderate reaction times, and wasteful aspects in recognizing basic occurrences. To address these impediments, this inquire about proposes an AI-based real-time reconnaissance framework outlined to independently distinguish rough behavior and issue prompt alarms, in this manner upgrading the adequacy and responsiveness of present day security foundation.The framework coordinating progressed profound learning methods, particularly leveraging question discovery and classification models such as YOLO and TensorFlow-based convolutional neural systems (CNNs). It is able of preparing a wide extend of input groups, counting inactive pictures, video records, and live webcam streams, empowering wide appropriateness over different observation scenarios. A real-time caution instrument conveys moment visual and sound-related notices when savage movement is identified, guaranteeing provoke mindfulness. Besides, the framework gives a user-friendly interface created utilizing Streamlit, permitting clients to oversee location limits, screen framework execution measurements, and associated with live discovery yields.

By combining computerization, tall precision, and adaptable engineering, this work points to contribute seriously to the progression of keen observation innovations, advertising a commonsense arrangement for proactive risk location and reaction in real-world situations.

## LITERATURE SURVEY

Spatial and worldly designs. Different considers have investigated designs of wrongdoing happenrence. These designs can be separated into two categories:spatial and transient designs. Spatial designs relate to the physical areas where wrongdoings regularly occur—such as urban centers, private zones, or amusement zones. In differentiate, worldly designs center on the timing and recurrence of wrongdoings, counting particular times of day, days of the week, or regular patterns [1].One of the essential challenges in foreseeing criminal action lies in proficiently analyzing large-scale wrongdoing datasets. Information mining strategies play a basic part in revealing covered up patterns and designs inside these tremendous datasets, permitting for speedier handling and diminished blunders. As a result, these strategies offer assistance move forward the precision and unwavering quality of wrongdoing expectation models [2].Numerous studies have examined strategies for reducing crime and have proposed a variety of prediction algorithms. The precision of these models to a great extent depends on the sort of highlights and information they utilize [3][5]. The application of AI-driven methods in wrongdoing determining has gotten to be progressively common, with models custom fitted to particular wrongdoing categories and datasets.For occasion, Shah et al. (2021) presented an approach combining computer vision and machine learning to screen high-crime zones through CCTV film and recognize people utilizing voice acknowledgment. Such innovations have the potential to help law authorization in avoiding violations some time recently they happen.

## METHODOLOGY

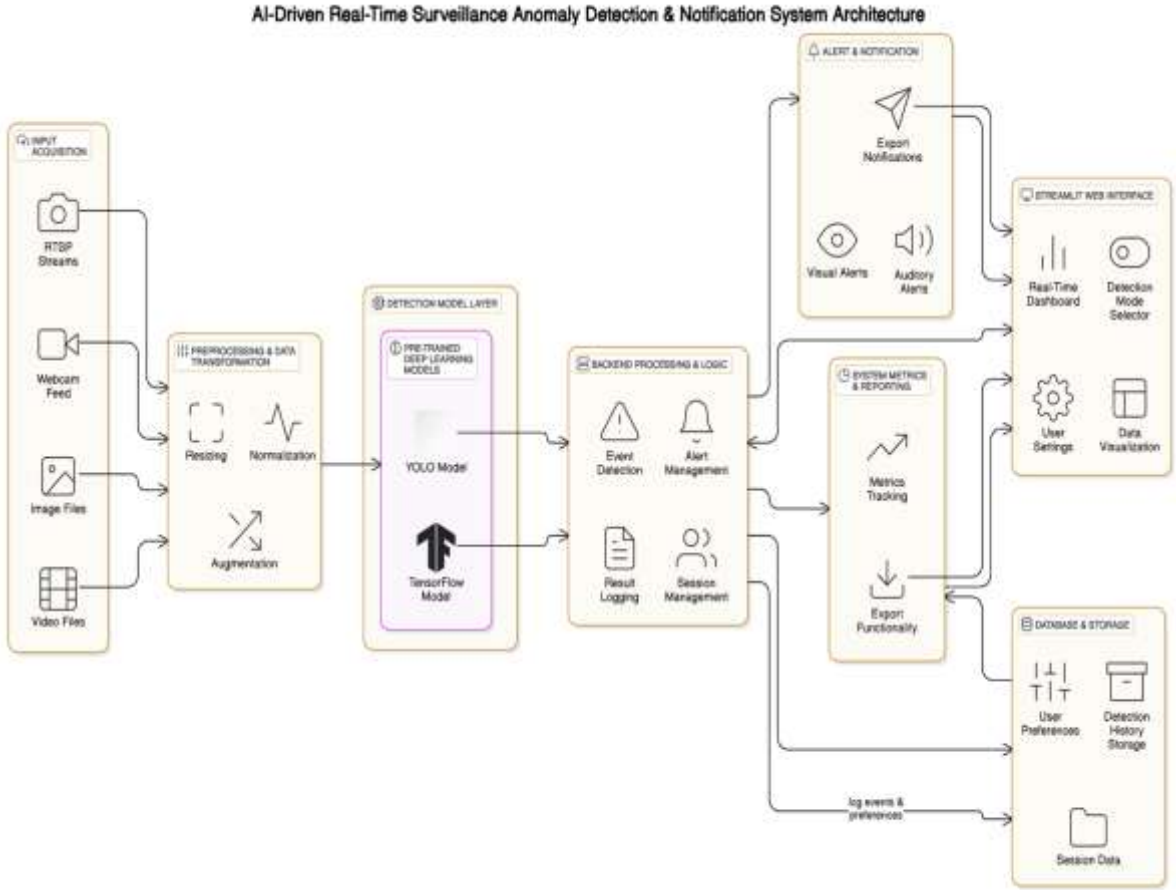


**Figure 1: System Workflow Diagram.**

The proposed system is structured using a modular design to ensure ease of development, real-time performance, and scalability. The entire workflow is divided into three core modules: (1) Data Acquisition and Preprocessing, (2) Model Training and Optimization, and (3) System Integration and Real-Time Deployment using a PyQt-based GUI.

### 4.1 Module 1: Data Acquisition and Preprocessing

To effectively train the system for detecting violent activities, datasets comprising labeled images and video clips were collected from open platforms like Kaggle and Roboflow. Both violent and non-violent scenes were present and clearly marked in the dataset.

Data Cleaning: Unwanted entries, including mislabeled data and corrupt files, were removed to ensure high-quality training input.

Exploratory Information Investigation (EDA):Visualizations such as histograms and dispersion plots were utilized to ponder the dataset's structure, distinguish course awkward nature, and analyze the in general information characteristics.

Preprocessing Techniques: Images and video frames were standardized by resizing and normalization (scaling pixel values to [0,1]). Data augmentation techniques—such as rotations, flips, and zoom—were applied to artificially expand the dataset and reduce overfitting.

Data Splitting: The cleaned and preprocessed dataset was divided into training, validation, and test sets to allow robust model evaluation and performance monitoring.

### 4.2 Module 2: Model Training and Evaluation

This module centers on building profound learning models competent of precisely recognizing rough exercises from visual inputs.

Model Selection: The system employed a pre-trained YOLOv11 model, fine-tuned for the specific task of violence detection. Additionally, a CNN-based model using TensorFlow was developed as an alternative to compare detection performance across architectures.

Training Process: The models were prepared utilizing the labeled dataset, where weights were upgraded through backpropagation. The preparing pointed to play down classification and localization mistakes.

Execution Assessment:The prepared models were evaluated utilizing measurements such as precision, exactness, review, F1-score, and IoU (Crossing point over Union) to guarantee both exact classification and exact localization.

Hyperparameter Tuning: Parameters including learning rate, batch size, and number of epochs were adjusted through systematic experimentation to identify the most efficient configurations.

Model Saving: Final trained models were saved in deployable formats (e.g., .h5, .pb) for efficient integration into the real-time detection system.

### 4.3 Module 3: Real-Time Arrangement and PyQt Interface Integration

To deliver a responsive and user-friendly experience, the final system was integrated into a PyQt-based GUI for real-time use.

Backend Integration: The backend was developed using Python to handle live video streams, webcam inputs, and RTSP feeds. Each outline was passed through the prepared show for discovery in real-time.

PyQt GUI: A custom graphical interface was created using PyQt, offering users the ability to upload video files, connect webcam feeds, and view detection outputs. The interface supports configuration of detection thresholds and displays real-time visual feedback.

Alert Mechanism: Upon detecting violence, the system triggers auditory alerts (e.g., playing a warning sound) along with visual cues (bounding boxes or highlighted frames) to immediately notify users of potential threats.

Checking and Announcing:System estimations such as number of disclosures, layout taking care of time, and in common system uptime are ceaselessly taken after.

**RESULT**

The effectiveness of the developed AI-based real-time surveillance system was assessed through multiple experiments, emphasizing classification accuracy, loss trends, prediction reliability, and overall system design.

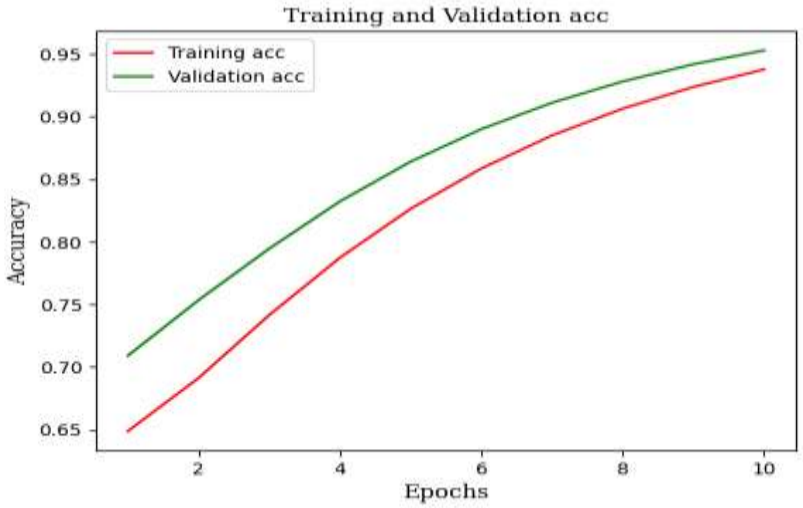


**Figure 2: Relationship between Training and Validation Accuracy.**

In this graph, the x-axis represents the number of epochs, while the y-axis shows the accuracy values, ranging from 0.60 to 1.00. The training accuracy gradually improves from 65% to about 94%, and the validation accuracy consistently performs better, reaching close to 96% by the final epoch. The proximity of both curves reflects the model's strong

generalization capabilities and indicates that it was not overfitting to the training data.

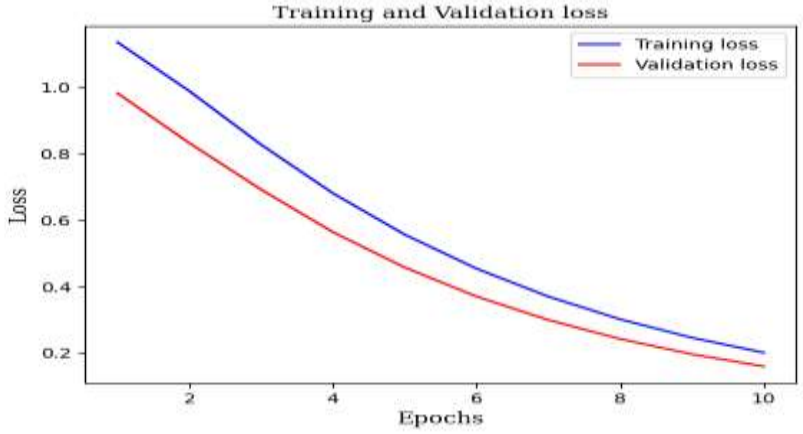


**Figure 3: Relationship between Training and Validation Loss**

The accuracy graph demonstrates a gradual increase in both training and validation accuracy across 10 epochs.The preparing precision starts at around 60% within the to begin with epoch and rises to over 90% by the tenth epoch. Similarly, the validation accuracy starts at around 55% and steadily improves, reaching just above 85%. The narrow gap between the training and validation accuracy curves suggests that overfitting is minimal, and the model's learning process remains stable. This reflects the model's strong capability to make accurate predictions on both the training dataset and previously unseen validation data.

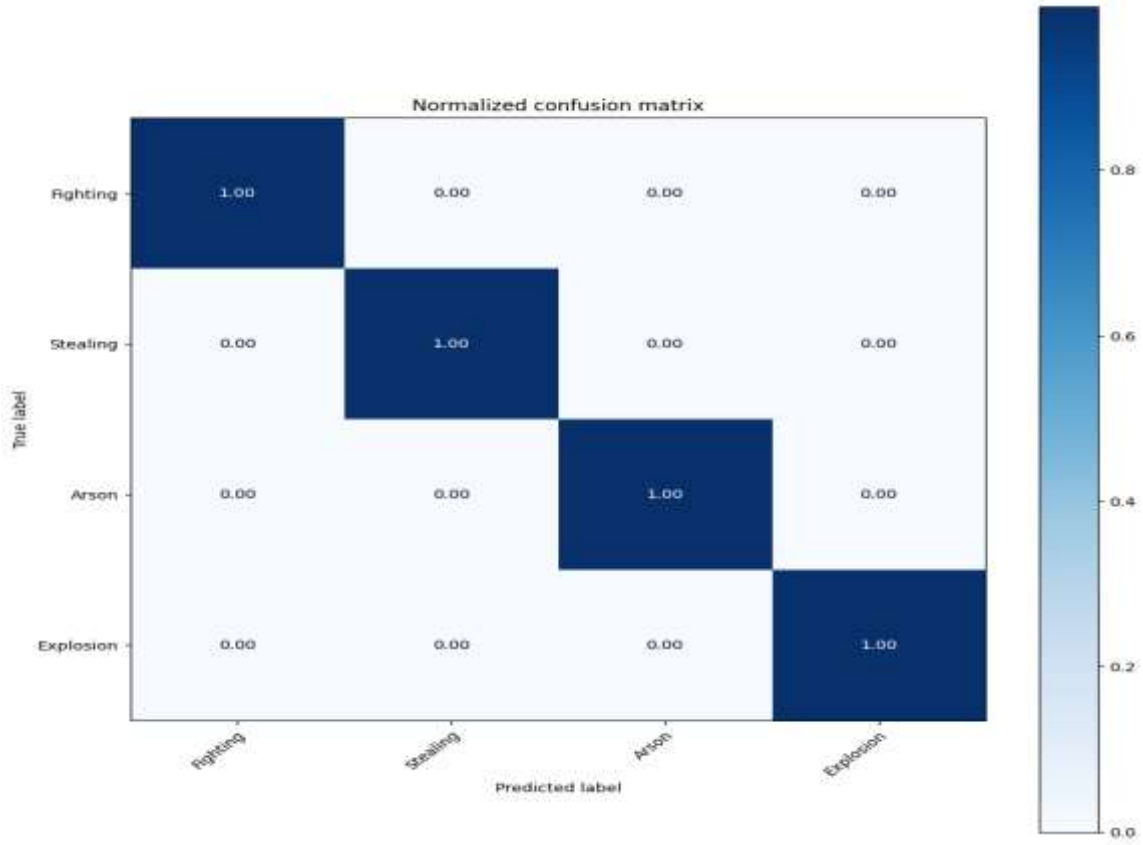


**Figure 3: Confusion Matrix between True label and Predicted label.**

A normalized confusion matrix generated from the model's predictions on the test data. The horizontal axis lists predicted classes, while the vertical axis lists the actual (true) classes. The matrix includes four classes: Fighting, Stealing, Arson, and Explosion. Each class shows perfect prediction accuracy with a score of 1.00 along the diagonal, and no off-diagonal entries, which confirms that the model classified each event type without error. This level of precision is critical for real-time detection systems, where accuracy directly affects user response and safety.

## CONCLUSION

For this extend, I outlined and executed an AI-powered real-time checking framework able of recognizing savage behaviors from distinctive sources, counting live webcam input, pre-recorded video film, and still pictures. The framework utilizes cutting-edge profound learning calculations such as YOLOv11 and TensorFlow-based convolutional neural systems to precisely distinguish occasions like battles, burglaries, fires, and blasts. Amid assessment, the show appeared dependable execution with negligible overfitting, as seen within the near coordinate between preparing and approval results. Moreover, I included an alarm framework that quickly informs clients through both visual and sound signals when savage movement is recognized. Future changes I arrange to work on incorporate extending the preparing information to cover more complex and unobtrusive activities, optimizing execution for gadgets with constrained assets, including multi-camera usefulness, and coordination prescient analytics to distinguish potential dangers some time recently they happen.

These upgrades will make strides the system's common sense, responsiveness, and by and large adequacy for real-time reconnaissance applications.

## REFERENCES

[1] **Kang, H.-W., & Kang, H.-B. (2017).** Prediction of crime occurrence from multi-modal data using deep learning. *PLOS ONE, 12*(4), e0176244. https://doi.org/10.1371/journal.pone.0176244

[2] **Kim, S., Joshi, P., Kalsi, P. S., & Taheri, P. (2018).** Crime analysis through machine learning. *Proceedings of the 9th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 8614828. https://doi.org/10.1109/IEMCON.2018.8614828

[3] **Wubineh, B. Z. (2024).** Crime analysis and prediction using machine-learning approach in the case of Hossana Police Commission. *Security Journal*. https://doi.org/10.1057/s41284-024-00416-6

[4] **Shah, N., Bhagat, N., & Shah, M. (2021).** Crime forecasting: A machine learning and computer vision approach to crime prediction and prevention. *Visual Computing for Industry, Biomedicine, and Art, 4*(9). https://doi.org/10.1186/s42492-021-00075-z

[5] **Jenga, K., Catal, C., & Kar, G. (2023).** Machine learning in crime prediction. *Journal of Ambient Intelligence and Humanized Computing, 14*, 2887–2913. https://doi.org/10.1007/s12652-023-04530-y

[6] **Kaur, M., & Saini, M. (2024).** Role of Artificial Intelligence in the crime prediction and pattern analysis studies published over the last decade: A scientometric analysis. *Artificial Intelligence Review, 57*, Article 202. https://doi.org/10.1007/s10462-024-10823-1

[7] **Nasaruddin, N., Muchtar, K., Afdhal, A., & Dwiyantoro, A. P. J. (2020).** Deep anomaly detection through visual attention in surveillance videos. *Journal of Big Data, 7*(87). https://doi.org/10.1186/s40537-020-00365-y

[8] Mariswari, R., & Narayani, V. (2024). Anomaly detection in surveillance videos using hybrid deep learning model DBNSSGAN. *International Journal of Intelligent Systems and Applications in Engineering, 12*(22s), 1859–1867. https://www.ijisae.org

[9] Singh, M. (2017). A survey on video anomaly detection. *International Journal of Engineering Research & Technology (IJERT), 5*(10), ICCCS-2017 Conference Proceedings. https://www.ijert.org/research/a-survey-on-video-anomaly-detection-IJERTCONV5IS10004.pdf