

Accident and Crime Detection Using YOLO from Surveillance Video Simulation

Mr. Kunal Kanchankar¹, Dr. Ravindra Kale²

¹Student of M. Tech Final Year Department of computer Science Engineering

G H Rasoni Skill Tech University, Nagpur, Maharashtra.

²Assistant Professor & Guide, Department of computer Science Engineering,
G H Rasoni College of Engineering and Management, Nagpur, Maharashtra

Abstract

The rapid increase in urbanization and population has led to a significant rise in road accidents, criminal activities, and public safety threats. Traditional surveillance systems mainly rely on human monitoring, which is time-consuming, less efficient, and prone to errors due to continuous observation requirements. To address these challenges, this project proposes an intelligent system titled “Accident and Crime Detection from Surveillance Video Simulation Using YOLO” that automates real-time monitoring and incident detection using advanced Artificial Intelligence and Computer Vision techniques. The proposed system utilizes the powerful YOLO (You Only Look Once) deep learning algorithm for object detection and activity recognition from surveillance video feeds. The system is designed to detect various incidents such as road accidents, suspicious activities, fights, robbery attempts, crowd violence, fire incidents, and abnormal human behavior in real time. By processing video frames continuously, the model identifies objects like vehicles, persons, and emergency situations with high speed and accuracy. The detected incidents are then highlighted with bounding boxes and alert notifications, enabling faster response from authorities and improving public safety.

Keywords: Deep Learning, Smart City Surveillance, Incident Detection, Computer-Vision.

I. INTRODUCTION

With the rapid growth of urban populations and the development of smart cities, ensuring public safety has become a major challenge for governments and law enforcement agencies throughout the world. Urbanization has significantly increased the number of vehicles, public gatherings, transportation systems, and social activities occurring daily in cities. As cities continue to expand, maintaining security

and monitoring public environments have become increasingly complex tasks. To improve safety and reduce criminal activities, modern cities deploy thousands of surveillance cameras in public places such as roads, railway stations, airports, shopping malls, educational institutions, banks, parking areas, hospitals, and residential locations. These surveillance systems play a critical role in monitoring activities and maintaining law and order.

The deployment of surveillance infrastructure has grown rapidly over the past decade due to technological advancements and increasing security concerns. Modern surveillance systems continuously capture video streams and record activities for monitoring and analysis purposes. However, the widespread use of surveillance cameras creates another challenge associated with handling enormous amounts of video data. Large-scale surveillance networks generate massive volumes of video footage every day, making continuous analysis a difficult process. Human operators are traditionally responsible for monitoring these video feeds and identifying suspicious or dangerous events. However, manual monitoring of a large number of video streams becomes extremely difficult and inefficient over long periods of time.

Human operators often experience fatigue, reduced concentration, and slower reaction times when continuously observing multiple surveillance screens. In situations where several cameras are active simultaneously, important events can easily be overlooked. Delays in identifying incidents such as road accidents, thefts, violence, suspicious activities, or public disturbances can result in serious consequences, including delayed emergency response and threats to public safety. Therefore, dependence solely on manual surveillance methods is becoming increasingly impractical in modern security systems [8][14].

The limitations associated with traditional surveillance methods have encouraged researchers and industries to develop automated and intelligent monitoring systems capable of

analyzing surveillance data without continuous human intervention. Recent developments in Artificial Intelligence (AI), Machine Learning (ML), and Computer Vision technologies have significantly transformed the surveillance domain. Artificial Intelligence enables machines to mimic human intelligence by learning from data and making decisions based on observed patterns. Machine Learning algorithms allow systems to improve performance through experience, while Computer Vision enables machines to understand and interpret visual information obtained from images and video streams.

Deep learning, which is a subset of machine learning, has emerged as one of the most effective techniques for visual analysis applications. Deep learning models utilize artificial neural networks inspired by the structure and functioning of the human brain to learn complex patterns from large datasets. Among these models, Convolutional Neural Networks (CNNs) have achieved remarkable performance in image and video processing tasks. CNN-based models can automatically learn meaningful features from visual data without requiring extensive manual feature extraction processes.

Deep learning techniques have demonstrated significant success in various image and video analysis applications such as object detection, image classification, facial recognition, activity recognition, scene understanding, and anomaly detection. These technologies enable automated systems to process surveillance footage and identify events such as accidents, suspicious

activities, theft, crowd disturbances, and unusual behaviors in real time [1][4][6].

Object detection is considered one of the most important tasks in intelligent surveillance systems. The purpose of object detection is to identify objects present within an image or video frame and determine their locations accurately. In surveillance applications, object detection can be used to identify people, vehicles, animals, bags, weapons, or any other objects relevant to security monitoring. Traditional object detection methods relied on manually designed features and image processing techniques, which often resulted in limited accuracy and poor performance under varying environmental conditions.

II. LITERATURE REVIEW

AI and advanced deep learning methods have greatly enhanced the functionality of contemporary monitoring systems. Traditional CCTV monitoring systems depend heavily on human operators, which can lead to delays and errors in detecting suspicious events. Recent research focuses on automated surveillance using computer vision, machine learning and technologies related to the IoT to facilitate immediate observation and recognition of incidents.

Proposed a real-time surveillance system using the YOLO deep learning framework to detect objects in video streams efficiently. The model was trained on surveillance video datasets and optimized to handle problems such as low lighting conditions, occlusion, and crowded

environments, achieving high detection speed suitable for real-time security applications [1].

Agarwal. developed an AI-based smart surveillance system that detects anomalies and suspicious activities in monitored environments. Their approach analyzes behavioral patterns using artificial intelligence models to automatically identify abnormal events and improve public safety monitoring [2].

Manu. presented an AI-powered smart surveillance camera capable of detecting unusual activities through computer vision algorithms embedded within the camera system. The system performs local data processing, reducing network load and enabling faster response times in security monitoring environments [3].

Suggested a surveillance system powered by artificial intelligence that combines image analysis and automated technology. learning techniques for automatic monitoring of video feeds. Their system focuses on identifying suspicious behaviors and reducing the dependency on human supervision in security systems [4].

Villegas. explored the role of artificial intelligence in intelligent IoT monitoring systems. Their study highlighted how AI models integrated with IoT devices enable real-time analysis, prediction, and monitoring of events in smart city environments [5].

Developed an AI-based surveillance system aimed at crime control and security enhancement. The system analyzes video feeds

using machine learning techniques to detect criminal activities and send alerts to authorities for quick response [6].

Introduced an AI-based surveillance model specifically designed for robbery and theft detection in banking environments. The system uses smart CCTV cameras and AI algorithms to analyze suspicious movements and trigger alerts when abnormal activity is detected [7].

Saranya carried out a study on intelligent CCTV monitoring systems that utilize artificial intelligence, emphasizing the application of technologies including face identification, movement detection, and behavior assessment. Their work emphasized the growing importance of automated surveillance for improving monitoring efficiency in large-scale environments [8].

Proposed an AI-enabled Industrial Internet of Things (IIoT) framework for smart city monitoring. The system integrates surveillance cameras and sensors to monitor events such as traffic congestion, accidents, and public gatherings in real time [9].

Qaraqe. developed a secure surveillance framework called PublicVision for recognizing crowd behavior in public spaces. The system securely transmits CCTV video to a central hub where deep learning models analyze crowd size and violence levels using advanced architectures such as Swin Transformer [10][11].

Ahmed and Nyarko introduced Smart-Watcher, a surveillance system utilizing artificial intelligence and Internet of Things technology,

specifically crafted for small and medium-sized properties. The system integrates sensors and intelligent video analysis to monitor activities and detect potential security threats automatically [12].

Overall, the reviewed literature demonstrates that artificial intelligence and computer vision technologies have significantly improved surveillance systems by enabling automated object detection, anomaly detection, and behavior analysis. However, most existing systems focus on specific tasks such as crowd monitoring, anomaly detection, or theft detection. There is still a need for an integrated surveillance framework capable of detecting multiple types of incidents—including accidents, violence, fires, and abnormal events—in real time within a unified intelligent monitoring system. This gap motivates the development of advanced multi-incident detection systems for smart city surveillance.

III. METHODOLOGY

The proposed project “Accident and Crime Detection from Surveillance Video Simulation Using YOLO” follows a systematic methodology for detecting incidents from surveillance video streams using Artificial Intelligence (AI) and Computer Vision techniques. The system is designed to automatically monitor video footage, identify abnormal situations, and generate alerts whenever suspicious activities or accidents occur. The complete process consists of multiple stages including data collection, preprocessing, object detection, activity

analysis, incident detection, alert generation, database storage, and dashboard visualization.

3.1 Data Acquisition

The first stage of the methodology involves collecting surveillance video data for training and testing the system. The performance of any Artificial Intelligence model depends heavily on the quality and diversity of the data used. Therefore, video datasets are collected from multiple sources such as publicly available datasets, CCTV recordings, traffic videos, and simulated surveillance videos.

The collected videos contain different types of scenarios such as vehicle movements, road accidents, human activities, crowd behavior, suspicious activities, and crime-related events. Including different scenarios helps the system learn various activity patterns and improves its capability to identify abnormal situations accurately.

For example, a surveillance video recorded at a traffic signal may contain cars moving normally, pedestrians crossing roads, and crowd activities. Another video may contain accident situations or suspicious behavior. Such diverse data helps the system understand both normal and abnormal events.

3.2 Data Preprocessing

After collecting the videos, preprocessing is performed before sending the data to the detection model. Raw surveillance videos often contain unnecessary information and inconsistencies that can reduce model

performance. Therefore, preprocessing improves the quality and structure of input data.

The preprocessing process includes extracting video frames because the detection model processes individual images rather than complete videos. After frame extraction, all images are resized into a standard size so that the model can process them efficiently. Image normalization is performed to maintain consistent image quality and reduce variations in lighting conditions.

The data is then labeled and annotated by placing bounding boxes around important objects such as persons, vehicles, and suspicious objects. Finally, the dataset is organized properly into different categories for training and testing purposes.

For example, if a frame contains a car and a person, labels such as **Car** and **Person** are assigned to the corresponding objects.

3.3 Object Detection using YOLO

Once preprocessing is completed, the frames are passed to the **YOLOv8 (You Only Look Once Version 8)** object detection model. YOLO is used as the primary detection algorithm because it provides high detection speed and better real-time performance.

The YOLO model analyzes each frame and identifies different objects present in the surveillance footage. The system detects vehicles, people, emergency vehicles, suspicious objects, and crowd-related information.

After detecting objects, the model generates rectangular bounding boxes around the identified objects along with labels and confidence scores.

For example, if a frame contains two cars and three people, the YOLO model identifies each object separately and displays labels such as:

- Car → 97% confidence
- Person → 95% confidence
- Person → 92% confidence

The ability of YOLO to detect multiple objects at the same time makes it suitable for real-time surveillance systems.

3.4 Incident Detection and Activity Analysis

After object detection, the system performs activity analysis to determine whether the detected situation is normal or abnormal. Simply identifying objects is not sufficient because surveillance systems also need to understand the behavior and interactions between objects.

The system analyzes movement patterns, object interactions, and behavioral characteristics to detect incidents.

For accident detection, the system observes:

- Sudden vehicle collisions
- Vehicle overturning
- Abrupt stopping patterns

For crime detection, the system identifies:

- Fighting activities
- Robbery attempts
- Crowd violence
- Suspicious movements

The system also performs behavior analysis by monitoring unusual human movements and abnormal crowd gatherings.

For example, under normal conditions, people walk normally in a shopping mall. However, if sudden aggressive movement or fighting occurs, the system recognizes it as suspicious activity and marks it as a possible incident.

3.5 Alert Generation

Once an incident is detected, the system immediately generates alerts and notifications. Quick alert generation is important because delayed responses can increase risks and damages.

The generated alert contains important information such as incident type, detection time, location, and priority status. The alert also appears on the dashboard interface and the corresponding video evidence is saved.

For example:

Incident Type: Accident Detected

Time: 10:34:28 AM

Location: Camera 2

Status: High Priority

This allows authorities to take immediate action during emergency situations.

3.6 Database Storage

After generating alerts, all detected incident information is stored in a database for future use and analysis. Maintaining records allows users to review previous incidents and generate reports.

The database stores important information including:

- Incident ID
- Incident type
- Date and time
- Detection confidence score
- Camera source
- Video evidence path

The project can use **SQLite** or **MongoDB** for storing incident information efficiently.

3.7 Dashboard Visualization

The final stage of the methodology is dashboard visualization. The dashboard provides a user-friendly interface that displays all monitoring information and system results in real time.

The dashboard includes features such as incident history, statistical reports, alerts, vehicle counts, and analytics. Users can view detected incidents, analyze historical records, and monitor activities efficiently.

For example, if an accident occurs, the dashboard immediately displays the alert, incident details, and related information, allowing users to quickly understand the situation.

Description of System Architecture

The proposed framework presents a deep learning-based surveillance system developed to detect incidents from video footages in real-time. The system analyzes surveillance videos through several steps, starting with the initial video preparation, object detection, motion analysis, and incident evaluation. The entire

operation process of the system is depicted in the suggested architectural diagram in (fig1.1)

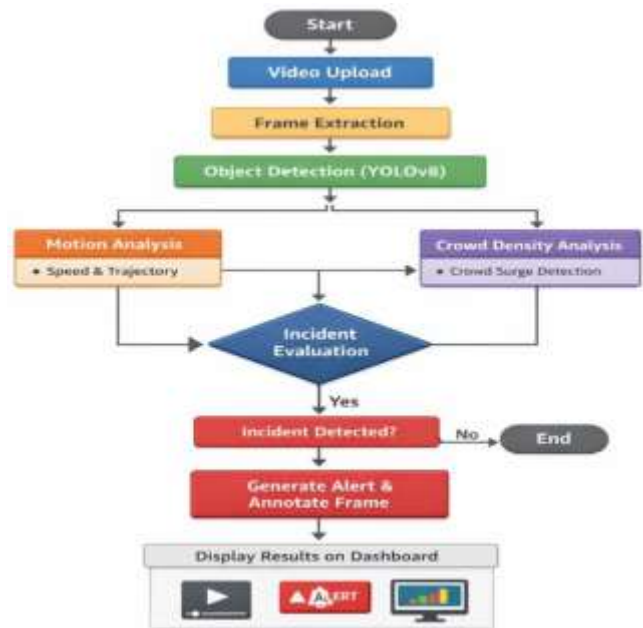


Figure 1(a) flowchart

Start : The process begins when the surveillance system is activated. At this stage, all necessary modules, libraries, and detection models are loaded into the system. The system becomes ready to receive surveillance video input for analysis.

Video Upload : After initialization, a surveillance video is provided as input to the system. The video may come from CCTV cameras, traffic monitoring cameras, public surveillance systems, or previously recorded video files.

For example, a CCTV camera installed at a road junction continuously records vehicle and pedestrian movements. This video is uploaded to the system for further analysis.

Frame Extraction : Since a video is a collection of many images shown continuously, the system first divides the uploaded video into

individual frames. This process is called frame extraction.

Each frame acts as a separate image that can be analyzed by the AI model.

For example, if a 30-second video is recorded at 30 frames per second (FPS), the system extracts approximately 900 frames for processing.

This step helps the system analyze video content more efficiently.

Object Detection using YOLOv8 : The extracted frames are then passed to the YOLOv8 object detection model. YOLOv8 identifies different objects present in each frame such as (Vehicles, Persons, Motorcycles, Buses, Crowd members, Suspicious objects). The model places bounding boxes around detected objects and assigns labels with confidence scores.

For example, if a frame contains two cars and three people, YOLOv8 identifies and labels each object separately.

This stage provides the basic information required for further incident analysis.

Motion Analysis : After object detection, the system performs motion analysis. In this stage, the movement of detected objects is analyzed. The system examines: Speed of vehicles, Movement direction, Object trajectory Sudden changes in motion. This analysis helps identify abnormal situations.

For example, if two vehicles are moving normally and suddenly collide, the system

observes the unusual change in movement pattern and marks it as a possible accident.

Crowd Density Analysis : At the same time, the system performs crowd density analysis. This stage is mainly used in public places such as railway stations, shopping malls, stadiums, and crowded streets. The system calculates: Number of people in an area, Crowd density level, Crowd movement behavior, Crowd surge situations. For example, if people are moving normally in a railway station, the system considers the situation normal. However, if a large number of people suddenly gather at one location or start running in different directions, the system detects abnormal crowd behavior. This helps in identifying emergencies, panic situations, or public disturbances.

Incident Evaluation : The information obtained from motion analysis and crowd density analysis is then combined in the incident evaluation stage. Here, the system checks whether the detected activities indicate any dangerous or abnormal event. Examples include: Road accidents, Fighting incidents, Robbery attempts, Suspicious movements, Crowd violence, Panic situations. The system evaluates all available information before making a final decision.

Generate Alert and Annotate Frame : When an incident is detected, the system immediately generates an alert. The alert contains information such as: Type of incident, Time of occurrence, Camera location, Priority level. At the same time, the system annotates the video

frame by highlighting the detected objects using bounding boxes and labels.

For example:

Alert: Accident Detected

Location: Camera 03

Time: 10:45 AM

Priority: High

This helps security personnel quickly understand the situation.

Display Results on Dashboard : The final step is displaying all results on the dashboard. The dashboard provides: Incident alerts, Annotated video frames, Live monitoring information, Vehicle statistics, Incident history, Analytical reports. Users can view incidents in real time and take immediate action when necessary. For example, if an accident is detected, the dashboard instantly displays the alert and related video frame so that authorities can respond quickly.



Figure 1(b) Block Diagram

i) Crowd Density Analysis

In addition to motion analysis, the system performs crowd density analysis. This module evaluates the number of people present in a given area and monitors crowd behavior.

The purpose of this stage is to detect events such as:

- Crowd congestion
- Crowd surge
- Abnormal crowd movements

Detecting crowd density patterns helps identify potential risks such as crowd panic situations or abnormal gatherings in public areas.

ii) Incident Evaluation

The results obtained from motion analysis and crowd density analysis are combined in the incident evaluation module. This module analyzes all extracted information and determines whether an incident has occurred. The evaluation process considers multiple factors such as:

- Sudden changes in object speed
- Collision patterns
- Abnormal crowd behavior
- Unusual movement trajectories

Based on these observations, the system classifies whether the situation represents a normal scenario or a potential incident.

iii) Alert Generation and Frame Annotation

When an incident is identified, the system automatically generates an alert. The alert mechanism may include:

- Highlighting the detected object
- Displaying warning messages
- Annotating the frame with bounding boxes or incident labels

This step helps authorities quickly identify the location and nature of the detected incident.

iv) Dashboard Visualization

Finally, the results are displayed on the monitoring dashboard. The dashboard provides:

- Real-time video visualization
- Incident alerts
- Analytical charts and monitoring statistics

This interface allows operators or authorities to easily monitor incidents and respond promptly.

IV. RESULTS



The examination of the chosen research articles indicates that advancements in artificial intelligence and deep learning methods have greatly enhanced the efficiency of contemporary surveillance systems. Many studies have successfully used computer vision models like CNN, YOLO-based object detection models, and sequential models like Recurrent Neural Networks and Long Short-Term Memory networks to analyze surveillance videos and detect abnormal events. These techniques allow automated systems to detect objects, monitor activities, and identify

suspicious behaviors in real time. The studies evaluated indicate detection rates that fall roughly between 85% and 93%, indicating the effectiveness of deep learning methods in surveillance applications.

However, several challenges still exist. Most current systems are designed to detect only a single type of incident, such as accidents, violence, or abnormal crowd behavior. In addition, issues such as high computational requirements, environmental variations, and false alarms can affect system performance. Therefore, developing integrated multi-incident detection systems is essential for improving real-time surveillance in smart city environments.



fig 5 ATM Crime

V. CONCLUSION

The project “Accident and Crime Detection from Surveillance Video Simulation Using YOLO” was developed with the purpose of improving traditional surveillance systems by introducing Artificial Intelligence and Computer Vision technologies. Nowadays, the increasing number of vehicles, population growth, and rise in criminal activities have made public safety an important concern. In many places such as roads, shopping malls, banks, railway stations, and public areas,

CCTV cameras are installed for security purposes. However, these cameras continuously generate a large amount of video

data that must be monitored by humans. Continuous observation of multiple camera feeds is difficult because a person can become tired, lose concentration, or miss important events. As a result, accidents and criminal activities may not be identified immediately, causing delays in taking action.

The proposed system attempts to solve this problem by creating an intelligent surveillance system that can automatically monitor video footage and detect incidents in real time. In this project, the YOLO (You Only Look Once) deep learning model was used because of its fast processing speed and efficient object detection capability. The system continuously analyzes surveillance videos, identifies objects such as vehicles and people, and determines whether any abnormal activity is taking place.

During the implementation and testing of the project, the system successfully detected different events such as vehicle accidents, suspicious human behavior, crowd-related situations, and crime-related activities. Whenever an unusual event was detected, the system generated alerts and stored information in the database. The dashboard also displayed live monitoring information, incident history, and related details, making it easier for users to understand and analyze events.

The project showed that Artificial Intelligence can significantly reduce human effort in surveillance monitoring. Instead of depending completely on manual observation, intelligent systems can automatically identify important incidents and assist authorities in making

quicker decisions. This can help reduce response time during emergencies and improve public safety.

Although the project achieved its intended objectives, some limitations were observed during testing. Factors such as poor video quality, low lighting conditions, crowded environments, and object overlap sometimes affected the detection accuracy. Despite these limitations, the overall results were satisfactory and proved that intelligent video analysis systems can be effectively used in surveillance applications.

REFERENCE

- [1]. Sanjeev Kukreti; RVS Praveen; Saloni Bansal; Hemanth Raju; Navdeep Singh; Rowsonara Begum "Object Detection in Real-Time Surveillance Using Deep Learning-Based YOLO Framework" in 2025 International Conference on Computational, Communication and Information Technology (ICCCIT), February 2025
- [2]. Varad Agarwal, Yuvraj Singh Rathore, Shyam Singh, Indra Kishor "Smart Surveillance System Using AI for Anomaly Detection" in 2025 International Journal of Innovative Research in Technology (IJIRT), Volume 11 Issue 11 Apr 2025.
- [3]. Manu, Y. M., S. V. Shashikala, N. Darshan, A. P. Dheeraj, N. J. Hemanth, and B. Nishanth Gowda. "Smart Surveillance Camera Using AI." In 2024 Second International Conference

- on Advances in Information Technology (ICAIT), vol. 1, pp. 1-6. IEEE, 2024.
- [4]. Maheshwari, K.K., Hajano, S., Baloch, S., Zafar, Z. and Fatima, T., 2025, April. AI Based Surveillance System. In 2025 International Conference on Emerging Technologies in Electronics, Computing, and Communication (ICETECC) (pp. 1-6). IEEE.
- [5]. Villegas-Ch, W., García-Ortiz, J. and Sánchez-Viteri, S., 2024. Toward intelligent monitoring in IoT: AI applications for real-time analysis and prediction. *IEEE Access*, 12, pp.40368-40386.
- [6]. Siddiqui, Muhammad Zulqarnain, Kamran Khan, and Abdul Jameel Khan. "Smart Surveillance and Crime Control Using an AI-Based Model." *Journal of Computing & Biomedical Informatics* 6, no. 01 (2023): 216-221.
- [7]. Kakadiya, R., Lemos, R., Mangalan, S., Pillai, M. and Nikam, S., 2019, June. Ai based automatic robbery/theft detection using smart surveillance in banks. In 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 201-204). IEEE.
- [8]. Saranya E, Vanitha A, Selvanandhini A, Soundarya VN, Vinisha M. A Survey on AI Based Smart CCTV Surveillance System. In 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS) 2024 Mar 14 (Vol. 1, pp. 1-5). IEEE.
- [9]. Rahman, Md Abdur, M. Shamim Hossain, Ahmad J. Showail, Nabil A. Alrajeh, and Ahmed Ghoneim. "AI-enabled IIoT for live smart city event monitoring." *IEEE Internet of Things Journal* 10, no. 4 (2021): 2872-2880.
- [10]. Qaraqe M, Elzein A, Basaran E, Yang Y, Varghese EB, Costandi W, Rizk J, Alam N. PublicVision: A secure smart surveillance system for crowd behavior recognition. *IEEE Access*. 2024 Feb 16;12:26474-91.
- [11]. Qaraqe, M., Elzein, A., Basaran, E., Yang, Y., Varghese, E.B., Costandi, W., Rizk, J. and Alam, N., 2024. PublicVision: A secure smart surveillance system for crowd behavior recognition. *IEEE Access*, 12, pp.26474-26491.
- [12]. Ahmed, Ahmed Abdelmoamen, and Bernard Nyarko. "Smart-watcher: An ai-powered iot monitoring system for small-medium scale premises." (2024).
- [13]. Afreen H, Kashif M, Shaheen Q, Alfaifi YH, Ayaz M. IoT-based smart surveillance system for high-security areas. *Applied Sciences*. 2023 Aug 3;13(15):8936.

- [14]. Ardabili, B.R., Pazho, A.D., Noghre, G.A., Neff, C., Bhaskararayuni, S.D., Ravindran, A., Reid, S. and Tabkhi, H., 2023. Understanding policy and technical aspects of ai-enabled smart video surveillance to address public safety. Computational Urban Science, 3(1), p.21.
- [15]. Sikora P, Malina L, Kiac M, Martinasek Z, Riha K, Prinosil J, Jirik L, Srivastava G. Artificial intelligence-based surveillance system for railway crossing traffic. IEEE Sensors Journal. 2020 Oct 16;21(14):15515-26.
- [16]. Jun H. Smart Drone Police System: Development of Autonomous Patrol and Real-time Activation System Based on Big Data and AI. International Journal of Internet, Broadcasting and Communication. 2024;16(4):168-73.