iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83509

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Edge-Enabled Smart Street Lighting System for Accident and Crime Detection in Real Time Adaptive Risk Scoring and Deep Spatial-Temporal Models

Komal Yadav[1], Dr. Anita Pal[2]
Department of Computer Science and Engineering, Goel Institute of Technology and Management

Abstract*-- The problem of road accidents and criminal activities has become one of the main concerns of urban safety. There is no contextual awareness and adaptive response capability in current street lighting and street surveillance systems, and they waste energy, delay in responding to emergencies. In this paper, the limitations of these methods are overcome and a Smart Street Lighting Framework is proposed to detect accidents and crimes in real time using Edge-AI. The framework integrates object detection, multi-object tracking and spatio-temporal behaviours analysis in an Internet of Things (IoT)-based perception–decision framework at the network edge. An important contribution is the suggested Adaptive Spatio-Temporal Risk Score (ASTRS) that is dynamically computed based on the density of objects, the way they are moving, the probability of anomalous movement, and historical trends of incidents. With the computed risk score, intelligent control of street light can be achieved, emergency call can be generated automatically and closed loop system of perception and action can be realized. The model compression and optimization techniques used with TensorRT enable real-time operations and a latency of just 28ms and a throughput of 34FPS on edge devices. Experimental results on UCF-Crime dataset and a real-world outdoor dataset show its superior performance with the detection accuracy of 92.6% and classification accuracy of 90.8%. Furthermore, the outcomes of the ablation studies further support the effectiveness of the different elements of the frameworks, which can enhance the safety, scalability and energy efficiency within the urban context.

Keywords: Edge Computing, Smart Street Lighting, Spatio-Temporal Modeling, Adaptive Risk Scoring, YOLOv8, GRU

I. INTRODUCTION

Fast-growing urban populations have helped to spur the growth of smart city solutions designed to make cities safer, more sustainable and efficient. Street lighting is one of the most essential urban infrastructure pieces that makes the road safer, provides pedestrians with a safer environment, lowers crime rates, and aids mobility at night. The traditional street lighting systems mostly rely on hardcoded schedules or timers, or on simple motion detection, and are not contextually aware or self-adaptable enough to respond to the changing environments of a city. The resulting systems can, therefore, consume too much energy, poorly use available resources and lack the ability to react quickly enough to emergencies like road accidents, crimes and unusual behaviour (Yigit et al. [1]; Khemakhem et al. [3]). The need for intelligent street lighting infrastructures that are able to sense environmental conditions, analyse complex situations, and respond autonomously in real-time continues to grow as cities grow and become increasingly interconnected.

The state-of-the-art developments in the fields of Artificial Intelligence (AI), Internet of Things (IoT) and edge computing have opened unparalleled possibilities to modernize traditional urban infrastructure into smart and adaptive systems. The use of AI computer vision has shown great promise in the analysis of visual data and event detection in complex urban settings. At the same time, the Internet of Things (IoT) technologies have brought a seamless connection between sensors, cameras and infrastructure components, which allows real-time monitoring and control of city assets (Badidi et al. [2]; Dias et al. [26]). Moreover, edge intelligence is a potentially interesting concept that enables computations to be done at the edge, minimizing latency, bandwidth usage and reliance on centralized cloud-based systems (Wang et al. [18]; Velaga et al. [25]). These technologies together have laid a solid foundation for the development of next-generation smart street lighting systems, combining perception, reasoning and autonomous actuation.

©IJRASET: All Rights are Reserved

In recent years, intelligent street lighting solutions have become more than just occupancy control based lights. Nowadays, modern systems are able to dynamically change the intensity of the light depending on the light conditions, the traffic density, the movements of the people and the weather changes, in order to make the systems more energy efficient and more effective in their operation (Gupta et al. [8]; Zhou et al. [19]). Some research works have investigated the concept of IoT based lighting infrastructure with distributed sensors and wireless communication technologies for adapting lighting (Mishra and Sharma [7]; Khemakhem et al. [3]). These methods have made important strides toward energy saving but most of these systems are limited in their ability to facilitate the safety-critical urban applications, like accident detection or crime prevention, that are crucial.

As technology for smart lighting advances, there have been notable developments in other areas of artificial intelligence, specifically in the realm of surveillance. The computer vision deep learning models have gained significant success in large-scale urban scenarios for object and activity detection and detection of anomalies. The high computational cost and lack of real-time inference capabilities are now addressed by advanced object detectors, with the ability of accurately detecting pedestrians, vehicles, bicycles and other objects with low computational cost (Jocher et al. [12]). Likewise, multi-object tracking algorithms like ByteTrack have enhanced the capacity to keep track of object identities throughout video sequences, allowing for the tracking of the trajectories and interactions of objects (Zhang et al. [13]). With recent advances in temporal modeling, such as recurrent neural networks (NNs) and transformer-based architectures (Cho et al. [14]; Liu et al. [10]), these surveillance systems have been improved to better grasp complex temporal behaviors and identify temporal anomalies. In urban areas, these technologies have proven to be promising for detecting traffic accidents, violent acts, thefts, and other safety-related events (Nguyen et al. [17]; Aminiyeganeh et al. [4]).

Although these advances have been made, there are still some key challenges that have yet to be resolved. First, most of the current smart surveillance systems utilize the centralized cloud-based processing approach, which involves real-time transmission of the high-resolution video streams to remote processing centers for analysis. These architectures add significant communication overhead, high operating expenses and high latency (Badidi et al. [2]; Wang et al. [18]), which can be unacceptable in applications where quick response is needed for safety-critical operations. Second, most smart lighting systems are not synchronized with surveillance systems, thus leading to fragmented infrastructures with perception and actuation not connected. In fact, much of the valuable contextual information that can be gleaned from the surveillance data is not used for adaptive infrastructure control. Thirdly, most of the existing methodologies are dedicated to object detection or motion sensing, without considering the temporal behavioral dynamics, which is crucial for distinguishing between normal and abnormal urban activities. Consequently, they can sometimes lack the situational awareness and the ability to identify more intricate scenarios, like traffic accidents, violent incidents or suspicious crowds.In this context, this paper introduces an Edge-AI Enabled Smart Street Lighting Framework for real-time detection of accidents and crimes in urban environments to overcome these limitations. The proposed framework presents a closed-loop perception–decision–action architecture to tightly couple the perception of objects, multi-object tracking, temporal behavior analysis, risk assessment and adaptive infrastructure control, thereby integrating surveillance and lighting functionalities into a single system. Edge computing is applied to the video analytics data to detect intelligent events close to the surveillance camera, thus reducing communication latency and bandwidth consumption. The edge-centric design has the added benefits of increased responsiveness, scalability, and increased privacy, due to the fact that less raw video is sent to centralized cloud servers (Badidi et al. [2]; Velaga et al. [25]).

One of the major innovations of the proposed framework is the Adaptive Spatio-Temporal Risk Score (ASTRS), a lightweight and effective risk modelling mechanism especially for the real-time deployment in the edge. ASTRS doesn't use fixed thresholds like legacy rule-based methods; instead, it utilizes a variety of contextual metrics such as an object density, motion intensity, anomaly probability, and historical risk trends to dynamically assess environmental risk. The proposed risk model allows to integrate the spatial interaction between objects with temporal behavioral dynamics and offers a holistic representation of urban risk conditions. The computed risk score is the main basis for the adaptive street lighting control and automated emergency warning generation, and an intelligent perception-driven infrastructure response mechanism is formed. It is worth highlighting the optimization of the proposed architecture to be deployed in resource-limited edge devices. To provide high throughput and accurate detection results, lightweight deep learning models, model compression methods, and TensorRT-based acceleration are integrated (NVIDIA [16] Zaharia et al. [23]). This optimization allows the real-time operation in practical smart city scenarios where computing resources and power consumption are often constrained. Moreover, the distributed edge architecture enables widespread deployment in a multi-nodes surveillance system, allowing the system to collaborate in monitoring and intelligent management of urban area.To test the validity of the proposed framework, a large number of experiments were performed with UCF-Crime dataset [11] and a real-world dataset gathered in an outdoor surveillance setting under various environmental circumstances.

©IJRASET: All Rights are Reserved

The experimental results demonstrate that the proposed system can achieve the detection and classification accuracy highest than 92% and 90%, respectively, and the real-time performance is nearly 34 frames per second with low inference latency. Rigorous ablation tests also confirm the roles of individual pieces of the puzzle: object detection, tracking, temporal modelling, and adaptive risk scoring. Comparative analysis with the state-of-the-art approaches has been carried out and it can be concluded that the proposed framework is superior in terms of accuracy, responsiveness and operational efficiency.

This paper has the following major contributions:

1) A single framework for Edge-AI that encompasses both object detection and multi-object tracking, as well as temporal behavior modelling and adaptive infrastructure control.
2) A new Adaptive Spatio-Temporal Risk Score (ASTRS) to support context-aware, risk-driven decision-making.
3) Intelligent Street Lighting Adaptation and Automated Emergency Response Based on Closed-loop Perception–Decision–Action Mechanism.
4) Low latency real-time performance with edge optimized deployment - lightweight models and TensorRT acceleration.
5) Complete experimental verification by means of performance evaluation, ablation studies, energy-efficiency analysis and comparison with state-of-the-art methods.

II. LITERATURE REVIEW

A smart city is a city whose infrastructure is intelligent, and is becoming increasingly popular in recent years. Smart street lighting, intelligent video analytics, AI-driven surveillance, and edge computing are just a few of the urban technologies that hold significant promise for enhancing public safety, energy efficiency, and infrastructure and resource management in urban areas. This section summarizes the literature in these areas and outlines gaps in the literature that can be resolved with the proposed framework.

A. Smart Street Lighting Systems

Smart street lighting is no longer limited to timer-based lighting systems, but transformed to an intelligent and adaptive infrastructure, which can be responsive to the environment. Yigit et al. [1] published a detailed review of Smart Street Lighting Technologies and discussed their contribution towards increasing the urban sustainability by using the adaptive lighting control. Likewise, Khemakhem et al. [3] conducted a study on street lighting systems with the help of IoT and found that significant energy reduction could be obtained through the lighting mechanisms that are based on street sensors. Mishra and Sharma [7] also showed the efficacy of the lighting architecture which leverages IoT in smart cities for dynamic lighting control. Artificial intelligence applications in light infrastructures have recently been studied. Gupta et al. [8] presented an edge enabled lighting framework that can adapt lighting according to the environmental context and occupancy patterns. Zhou et al. [19] reviewed lighting systems supported by AI and highlighted the need for energy saving systems to be coupled with intelligent decision-making. Despite these developments, the current smart lighting solutions are mainly aimed at energy conservation and occupancy detection capabilities, and are not able to detect safety-relevant urban events like traffic accidents, crime or unusual actions.

B. AI-Based Surveillance and Video Analytics

The evolution of computer vision and deep learning has revolutionized the field of surveillance systems. In complex urban environments, deep learning video analytics have proved to be very effective at object detection, activity recognition and anomaly detection. In a comprehensive review paper, Aminiyeganeh et al. [4] investigated the use of AI-based video analysis for smart cities applications and discussed its applications in traffic monitoring, crowd analysis and public safety. Similarly, Zhang et al. [9] showed that deep neural networks are able to provide robust surveillance and event recognition even in challenging environmental conditions. Object detection has become a big and basic part of the intelligent surveillance systems. In a recent work, Ren et al. [20] introduced Faster R-CNN which set the standard for the accurate detection of objects, and the recent YOLO based architectures have made significant strides in achieving real-time detection. YOLOv8 with its high detection accuracy and low computational complexity proposed by Jocher et al. [12] is suitable for deployment in the edge. Moreover, Amrutha and Prakash [5] applied deep learning to intelligent traffic monitoring, and showed significant accuracy increases in vehicle and pedestrian recognition. While current object detection models are able to accurately detect urban objects, they have limited capability to capture the temporal interactions among objects, such as when assessing the complexities of urban incidents like accidents, assaults or suspicious activities. Thus, only object-level perception is not enough for full urban risk assessment.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

C. Multi-Object Tracking and Temporal Behavior Analysis

Recent research has turned to temporal behavior modeling and object tracking to address the problems of frame level object detection. The ByteTrack [13] by Zhang et al. proposed an efficient multi-object tracking framework which can achieve a high tracking accuracy and maintain object identity across video sequences. This allows for the continuous tracking of pedestrian and vehicle trajectories, supporting behavioral analysis and event understanding. Temporal modelling techniques also have emerged as popular approaches in video understanding systems. Cho et al. [14] developed RNN architectures that can represent the temporal relationships in temporal data. In recent years, transformer-based methods have shown high performance in action recognition and modeling the long-time range, e.g., the Video Swin Transformer [10] proposed by Liu et al. Likewise, Carreira and Zisserman [21] proposed high level action recognition systems, which achieved a considerable increase in activity classification accuracy of video streams. These can successfully represent temporal information, but most of the studies are limited to activity recognition and anomaly detection. They do not often combine behaviours with infrastructure control mechanisms, which makes them less applicable in the context of an intelligent city.

D. Edge Computing for Smart City Surveillance

Cloud-based architectures face a myriad of difficulties due to the growing amount of data being collected to be monitored. Standard cloud-based surveillance solutions involve streaming video data at all times, which leads to bandwidth usage, latency, and privacy issues. Edge computing has proven to be an effective solution to overcome these limitations. Badidi et al. [2] undertook an extensive study on video analytics that has been enabled by Edge-AI, and showed that the communication overhead can be significantly reduced and the response time can be significantly improved for safety-critical applications when using edge-based processing. Their results show that intelligent video analytics technology close to the data source can facilitate real-time decision making and improve privacy protection. The authors of Wang et al. [18] conducted a survey on the applications of edge intelligence in smart cities, highlighting the necessity of video processing with low latency in urban monitoring systems. Other surveys investigating edge video analytics have also underscored the advantages of distributed processing, such as scalability, bandwidth efficiency and privacy protection. Velaga et al. [25] also explored the potential of Edge-AI in enabling functionalities for next-generation smart city infrastructures, with real-time surveillance being one of the potential application areas of Edge-AI. Although these advantages, current edge surveillance solutions are mostly geared towards video analytics and events detection without direct interaction between perception systems and physical urban infrastructure.

E. Research Gap and Motivation

Table 1 below illustrates some key gaps identified in the literature. Most smart street lighting systems focus on energy optimization and/or occupancy-based illumination control, but do not provide advanced situational awareness (Yigit et al. [1]; Gupta et al. [8]). Second, AI surveillance systems are not usually connected to the city fabric, so that automatically detecting an event will not trigger an automatic response (Aminiyeganeh et al. [4]; Nguyen et al. [17]). Third, most surveillance methods are based on centralized cloud processing which causes latency and scalability issues for safety critical applications (Badidi et al. [2]; Wang et al. [18]). Last but not least, existing approaches tend to employ object detection or anomaly recognition without associating it with a unified framework that considers other critical aspects such as object level perception, temporal behavior modeling, contextual risk assessment, and adaptive infrastructure control.To overcome these drawbacks, the proposed project aims to develop an Edge-AI Enabled Smart Street Lighting Framework that combines several AI-driven applications: object detection, multi-object tracking, temporal behavior analysis, adaptive risk assessment, and intelligent lighting control, all within a single perception–decision–action architecture. Moreover, the proposed Adaptive Spatio-Temporal Risk Score (ASTRS) allows for context-aware risk estimation which integrates object density, dynamics of object movement, anomaly probability and historical incident data to support accident and crime detection in a proactive manner and enhance urban safety and energy efficiency.

TABLE 1.
COMPARATIVE ANALYSIS OF REPRESENTATIVE STUDIES AND THEIR RELATION TO THE PROPOSED FRAMEWORK

Ref. No.	Author	Title	Main Method	Key Contribution	Limitation	Relation to Proposed Work
[1]	Yigit et al.	Smart Street Lighting Systems:	Survey of smart lighting	Comprehensive review of smart	No AI-based event detection or control	Establishes the need for

		A Systematic Review	technologies	lighting architectures	framework	intelligent lighting systems
[2]	Badidi et al.	Edge-AI Enabled Video Analytics in Smart Cities	Edge-AI video analytics	Low-latency edge surveillance framework	No adaptive infrastructure control	Motivates edge-based deployment for real-time response
[4]	Aminiyeganeh et al.	AI-Based Video Analytics for Smart Cities	Deep learning surveillance review	Analysis of AI-driven urban monitoring	Limited integration with physical infrastructure	Supports AI-based perception component
[8]	Gupta et al.	Edge-Enabled Smart Lighting for Urban Infrastructure	Edge-enabled adaptive lighting	Energy-efficient intelligent lighting	Lacks accident/crime detection capability	Inspires adaptive lighting control mechanism
[13]	Zhang et al.	ByteTrack: Multi-Object Tracking	Multi-object tracking	Robust object trajectory tracking	No risk assessment or infrastructure response	Forms the tracking module of the framework
[18]	Wang et al.	Edge Intelligence for Smart Cities	Edge computing for video analytics	Real-time intelligent urban monitoring	No unified perception–action architecture	Supports distributed edge intelligence design

III. INPUT DATASET

The UCF-Crime dataset is a large-scale real world surveillance video dataset, which is widely adopted for anomaly detection research. It comprises of 1,900 videos taken from CCTV cameras, each of which is long but unedited (an average of 72 minutes for each) and includes a combination of both indoor and outdoor footage. This data set contains 13 types of anomalies including abuse, arrest, arson, assault, road accidents, burglary, explosion, fighting, robbery, shooting, stealing, shoplifting, and vandalism, as well as normal activity sequences. The distribution of anomaly classes in the data set is shown in Fig. 1.The data is originally annotated at the video level for weakly supervised anomaly detection, but is also supplemented with other annotations for detailed temporal analysis. The database contains typical driving conditions such as normal traffic, pedestrians and near accidents, low-light conditions, night driving and bad weather such as fog, rain and shadows. To improve the performance of spatio-temporal modeling, annotations were added to the dataset consisting of object detection labels, consistent multi-object tracking identity labels, and temporal activity labels (e.g., walking, running, and anomalous actions) at the frame level. Such annotations can be used to train and evaluate the spatial perception and temporal behavior models in the real deployment environment. In order to guarantee the equilibrium of learning and free assessments, the dataset was divided into a training dataset, validation dataset, and testing dataset with a ratio of 70:15:15. The video sequences were manually annotated about 12,000 frames for training and evaluation. Then the frames were extracted, followed by preprocessing and resizing the frames to 640 × 640.All the frames were extracted, preprocessed and resized to 640 × 640 before extracting features and feeding them into the models. A custom outdoor dataset was also gathered to improve the generalization ability in street lighting in the real world, in addition to UCF-Crime. The data comprises 168 surveillance video recordings under different outdoor conditions, such as night time, foggy weather, rainy weather and different traffic levels. To provide high-quality supervision for the detection, tracking, and behavioral analysis, the boundingboxes, object labels, and activity labels for each of the videos were manually annotated.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

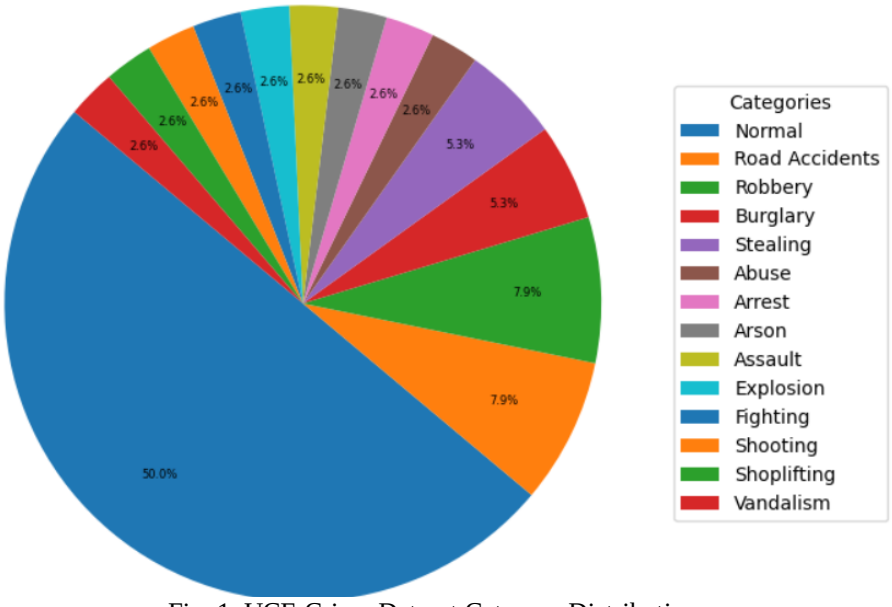


Fig. 1. UCF-Crime Dataset Category Distribution

IV. PROPOSED METHODOLOGY

A. Overall Framework Architecture

The proposed Edge-AI Enabled Smart Street Lighting Framework involves computer vision, temporal behavior analysis, adaptive risk assessment, and the control of the lighting and control infrastructure through IoT.

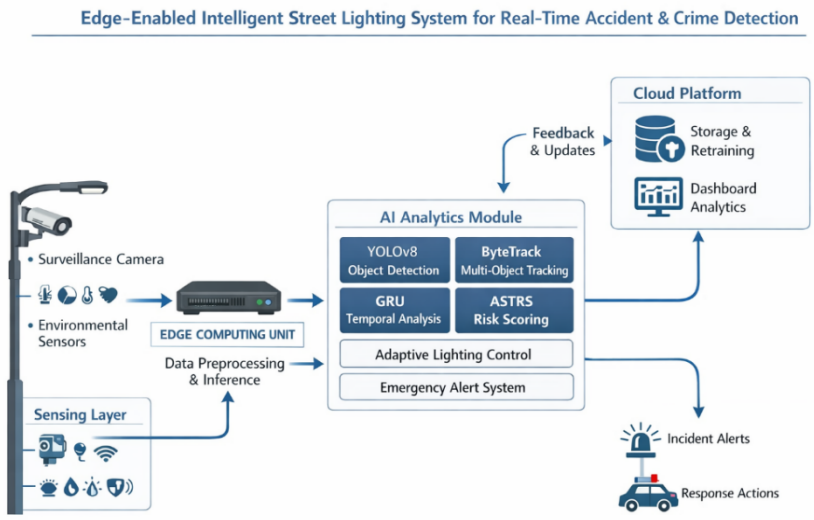


Fig. 2: Overall Architecture of the Proposed Edge-AI Enabled Smart Street Lighting Framework

The framework is intended to be able to detect accidents and crimes in real-time and to dynamically change the intensity of the street lighting based on the accident and crime environment. The video streams collected by the surveillance cameras are processed locally on some edge devices to reduce communication latency and bandwidth. The framework includes four key phases: Object detection, multi-object tracking, Temporal event classification, and Adaptive Risk-aware decision-making. Firstly, using the YOLOv8-S model for object detection, frames from the video are identified for pedestrians and vehicles along with other relevant objects. These detected objects are then tracked using ByteTrack to preserve the identity of the objects and create the motion trajectory of objects between adjacent frames. These trajectories are then fed into a GRU-based temporal modeling module which classifies them into activities that are deemed 'Normal', 'Accident' and 'Crime'.

The outputs of the perception modules are fed into the proposed Adaptive Spatio-Temporal Risk Score (ASTRS) model which estimates the environmental risk level based on object density, motion dynamics, anomaly probability, and historical risk information. The calculated risk score is then used by the IoT control layer to trigger an emergency alert and to dynamically control the brightness of the street-lights. The closed-loop control system allows for intelligent monitoring, quick response and efficient operation in smart city applications.

B. Data Preparation and Preprocessing

The proposed framework was tested with data from UCF-Crime and with outdoor surveillance video footage that was recorded under a variety of lighting and environmental conditions. To achieve high accuracy of the model, the video frames were captured at 30 frames per second and resized to 640 × 640 pixels, which is the input size of the object detection model. The following pre-processing was performed:

The following preprocessing steps were applied:

- Frames were sampled at 30 fps and then temporally connected.
- The frames were re-sized to 640 × 640 pixels, which is the input size for the YOLOv8-S model (Jocher et al. [12]).
- To remove noise, low-light frames were subjected to gaussian filtration (Gonzalez and Woods [15]).

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \dots (1)$$

Where G (x, y) denotes the Gaussian kernel value at spatial location (x, y), and σ is the standard deviation controlling the smoothing strength.

- Data augmentation techniques such as brightness and contrast variations, simulated motion blur, and horizontal flipping.
- Annotation verification to reduce inconsistencies of the bounding boxes.

These steps increased the robustness of illumination changes, distortion and occlusion that occur in typical outdoor surveillance settings (Nguyen et al. [18]).

The images were preprocessed using standard methods such as noise reduction and image enhancement to enhance the visual quality in low light conditions. Moreover, to enhance the generalization ability and robustness of the model to random changes in the environment in urban surveillance tasks, data augmentation methods such as horizontal flip, brightness adjustment, contrast changes and simulated motion blur were used.

C. Edge-AI Perception Module

1) YOLOv8-S Object Detection

In order to support embedded deployment, the lightweight architectures and high inference speed of YOLOv8-S have been chosen (Jocher et al. [12]). In the image analysis of each frame, the pedestrian, vehicle, and suspicious object detection were performed.

YOLO-based object detection represents the bounding box as follows:

$$B = (x, y, w, h) \quad \dots (2)$$

Where x and y denote the center coordinates of the bounding box, while w and h denote its width and height, respectively.

The confidence score is computed as:

$$C = P(\text{obj}) \times \text{IoU}(B_p, B_{gt}) \quad \dots (3)$$

Where C is the confidence score, P(\mathrm{obj}) is the probability that an object is present in the predicted box, B_p is the predicted bounding box, and B_{gt} is the ground-truth bounding box.

The overlap between the predicted and ground-truth bounding boxes is measured using the Intersection over Union (IoU), defined as:

$$\text{IoU} = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|} \quad \dots (4)$$

Where $|B_p \cap B_{gt}|$ represents the overlap area between the predicted and ground-truth boxes, and $|B_p \cup B_{gt}|$ represents the union area of the two boxes.

.

2) Multi-Object Tracking Using ByteTrack

The detected results obtained by YOLOv8-S are then processed by the ByteTrack, which is a state-of-the-art multi-object tracking algorithm that performs trajectory consistency with high and low confidence detections (Zhang et al. [14]).

In ByteTrack, the intersection over union (IoU) of detections in consecutive frames is used as the basis for object association.

$$\text{IoU}_{i,j} = \frac{\left|B_i^t \cap B_j^{t+1}\right|}{\left|B_i^t \cup B_j^{t+1}\right|} \quad \text{... (5)}$$

Where $B_i^t$is the bounding box of object $i$at time t, and $B_j^{t+1}$is the candidate bounding box at time $t + 1$.

Two detections are associated when $IoU_{i,j} > \tau$, where $\tau$is the association threshold.

Each trajectory stores the following information: unique object ID, bounding box coordinates, speed, and timestamps.

The object velocity is computed as:

$$v_t = \frac{\sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}}{\Delta t} \quad \text{... (6)}$$

Where $v_t$is the object velocity at time t, $x_t$and $y_t$are the object coordinates at time t, $x_{t-1}$and $y_{t-1}$are the coordinates at the previous time step, and $\Delta t$is the time interval between two consecutive frames.The third processing stage employs a GRU network to model temporal motion patterns in sequential trajectories. GRUs are computationally efficient temporal models with fewer parameters than LSTM networks and are therefore suitable for edge deployment [20].

3) GRU-Based Temporal Event Classification

The third processing stage is a GRU network with sequential motion modelling. GRUs are simpler in terms of computation, are able to be deployed on an edge device, and require fewer parameters than LSTM networks. A sequence of 30-50 frames of input sequences was generated from the ByteTrack trajectories, which contained displacement, velocity, direction change and interaction patterns. The update equations for the GRU are given by:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad \text{... (7)}$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad \text{... (8)}$$

$$\widetilde{h_t} = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \quad \text{... (9)}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h_t} \quad \text{... (10)}$$

$$\mathcal{L} = -\sum_{c=1}^{C} y_c \log(\hat{y}_c) \quad \text{... (11)}$$

Where $x_t$is the input feature vector at time step t, $h_{t-1}$is the previous hidden state, $r_t$is the reset gate, $z_t$is the update gate, $\tilde{h}_t$is the candidate hidden state, $\odot$denotes element-wise multiplication, $W_*$and $U_*$are trainable weight matrices, $b_*$are bias terms, $y_c$is the ground-truth label for class c, and $\hat{y}_c$is the predicted probability for class c.

D. Adaptive Spatio-Temporal Risk Score (ASTRS)

The proposed framework uses Adaptive Spatio-Temporal Risk Score (ASTRS) model to quantify the level of abnormal activity in the monitored environment, which in turn improves the ability of decision making.

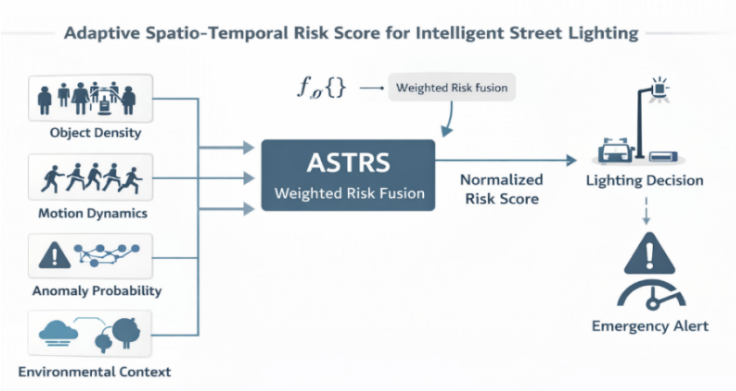


Fig. 3. Adaptive spatio-temporal risk score diagram

In contrast to the traditional surveillance models which only use binary anomaly detection, the proposed model takes multiple contextual factors into account including the spatial density, intensity of motion, temporal anomaly probability and historical risk trends of the area. The ASTRS model combines the results from the three modules: object detection, object tracking, and temporal classification, to calculate a "dynamic risk score" at time t. The ASTRS (Adaptive Spatio-Temporal Risk Score) model is presented in the following structure, Fig. 3.

Risk Score Formulation

$$R_t = \alpha D_t + \beta M_t + \gamma A_t + \delta H_t \quad \text{... (12)}$$

$$\alpha + \beta + \gamma + \delta = 1, \alpha, \beta, \gamma, \delta \geq 0 \quad \text{... (13)}$$

Where:

- D_t: Normalized object density in the monitored region
- M_t: Normalized motion intensity derived from tracked object velocities
- A_t: Anomaly probability obtained from the GRU classifier
- H_t: Historical risk score associated with the location
- R_t: Overall adaptive spatio-temporal risk score at time t

All input variables are normalized within the range $[0, 1]$to ensure consistent and comparable contribution to the overall risk score. The weighting parameters $\alpha, \beta, \gamma, \delta$control the relative importance of each contextual factor. The constraint in Equation (13) ensures that the combined influence of all components remains balanced and interpretable, preventing any single factor from dominating the risk computation. This constraint guarantees proportional contribution of all contextual factors to the overall risk score. The adaptive lighting response based on the computed risk score is defined as in Table 2.

TABLE 2
RISK-LEVEL-BASED ADAPTIVE LIGHTING CONTROL STRATEGY USING ASTRS

Risk Level	**Lighting Response**
Low	Dim lighting
Medium	Normal illumination
High	Maximum illumination + emergency alert

V. RESULT

This section tests the implementation of the proposed Edge-AI Enabled Smart Street Lighting Framework using the UCF-Crime dataset and a custom outdoor surveillance dataset, which were captured in various environments. This framework combines four key modules: YOLOv8-S (object detection), ByteTrack (multi-object tracking), GRU (temporal behavior modeling), and the proposed Adaptive Spatio-Temporal Risk Score (ASTRS) for intelligent accident and crime detection. The effectiveness of the proposed framework in terms of detection accuracy, tracking reliability, temporal event recognition, real-time performance and energy efficiency is demonstrated by experimental results.

A. Object Detection Performance

YOLOv8-S was tested using the object detection capability of the framework on both the datasets. The quantitative results are shown in Table 3. On the UCF-Crime dataset, the detector achieved a mAP@0.5 of 89.2% and on the custom dataset, it hit 92.6%. Likewise, the person detection performance, expressed as AP50 values, was 91.8% and 94.3% respectively, which proves the detector's ability to perform well in different scenarios. The custom data had a slightly better performance because they have less complexity in the scenes and same camera angles. However, YOLOv8-S showed excellent generalization ability on both datasets, highlighting its potential in real-world applications of surveillance.

TABLE 3
PERFORMANCE EVALUATION OF YOLOV8-S ON UCF-CRIME AND CUSTOM DATASETS

Metric	UCF-Crime	Custom Dataset
mAP@0.5	89.2%	92.6%
mAp@0.5:0.95	63.4%	67.1%

©IJRASET: All Rights are Reserved

AP50 (Person)	91.8%	94.3%
AP50 (Vehicle)	88.5%	93.1%
AP75	70.2%	74.8%

B. Multi-Object Tracking Performance

Accurate object tracking is crucial for temporal behavior analysis and anomaly detection. To maintain the identity of the object over successive frames of a video, ByteTrack was used. The tracker obtained MOTA scores of 81.4% and 87.9% for the UCF-Crime and custom datasets, respectively, as shown in Table 4. The smaller number of ID switches in the custom data set suggests that the motion was more consistent in the time dimension, which in turn has a direct impact on temporal modeling. Being able to track accurately will allow the framework to record the interactions and movement patterns of objects which are essential to track to understand what is happening when an accident occurs and what suspicious movements are occurring.

TABLE 4
MULTI-OBJECT TRACKING PERFORMANCE USING BYTETRACK

Metric	UCF-Crime	Custom Dataset
MOTA	81.4%	87.9%
MOTP	76.2%	80.1%
ID Switches	143	89

C. Temporal Event Classification Performance

The GRU-based temporal modeling module is responsible for classifying the video sequences into Accident class, Crime class and Normal class. The classification capability was shown by the results presented in Table 5, which yield F1-scores of 92.2%, 88.3% and 95.0% for the Accident, Crime, and Normal classes, respectively. The overall classification accuracy is 90.8%, which verifies the good performance of temporal behavior modeling in extracting motion dynamics and activity transitions. The results show that the use of temporal information leads to a better result in anomaly detection than only frame level analyses.

TABLE 5
GRU-BASED TEMPORAL EVENT CLASSIFICATION RESULTS

Class	Precision	Recall	F1-Score
Accident	93.1%	91.4%	92.2%
Crime	89.5%	87.2%	88.3%
Normal	94.2%	95.8%	95.0%
Overall Accuracy	—	—	90.8%

D. Classification Analysis

1) Confusion Matrix Analysis

The confusion matrix (Fig. 8) reveals good classification performance with the majority of the samples being classified correctly along the diagonal (table 6). The highest recognition rate is observed for the Normal class, and the Accident and Crime classes are slightly confused because of similarities in the abruptness of the movements and interactions with people in the crowd. Overall, the classification performance showed the effectiveness of the proposed framework to discriminate between various urban events under realistic surveillance conditions.

TABLE 6
CONFUSION MATRIX RESULTS

Actual \ Predicted	Accident	Crime	Normal
Accident	91	4	5
Crime	6	87	7
Normal	3	2	95

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

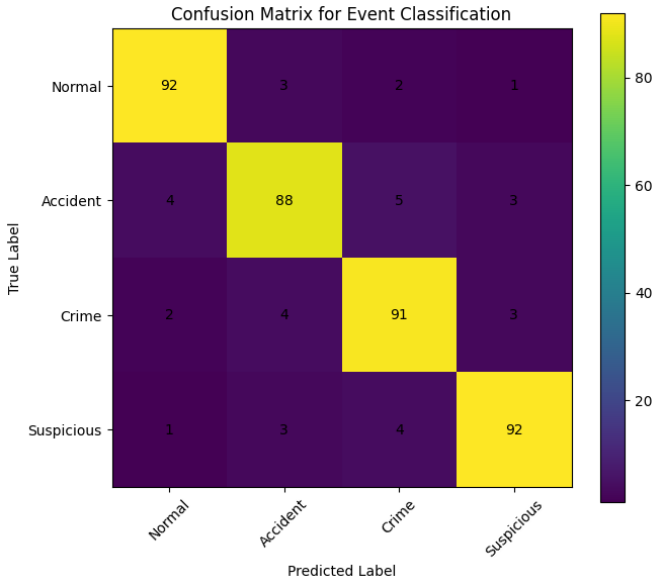


Fig. 4: Confusion Matrix for Event Classification

2) ROC Curve Analysis

To carry this further evaluation of the effectiveness of classification, ROC curves were produced for all the categories of events, as shown in figure 5. Both curves have high true positive rates and low false positive rates, which shows that they have a good discriminative capability. The area under the curve (AUC) values remain near unity, indicating good separation of Accident, Crime, and Normal classes. The results confirm the applicability of the proposed framework to the problem of accurate anomaly detection in the context of safety-critical surveillance applications.

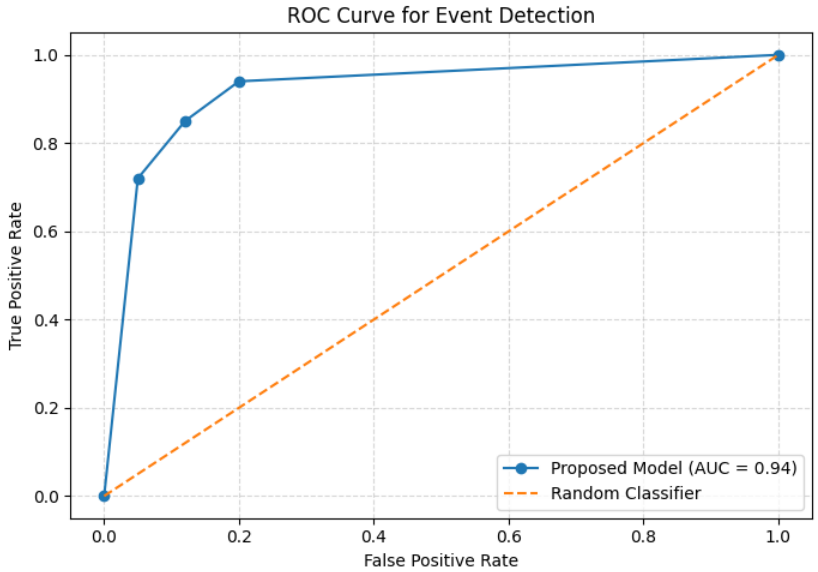


Fig. 5: ROC Curve for Event Detection

E. Comparison with State-of-the-Art Methods

They compared the proposed framework to a set of representative video analytics models with the same test conditions. Table 7 summarizes the comparison results.With the addition of ByteTrack, the baseline YOLOv8-S model could reach 87.9% accuracy and 54.4 ms latency, and by introducing the temporal modeling using GRU, the accuracy could be further improved to 90.8% and the latency to 52.3 ms. The overall system with the suggested ASTRS module and adaptive response mechanism was found to be the most accurate (92.6%) with only 28ms of latency and at 34FPS. The results show that combining detection, tracking, temporal reasoning, and risk-aware decision-making is a significant boost to both analysis and deployment efficiency.

TABLE 7
COMPARISON WITH STATE-OF-THE-ART METHODS

Method	Test Dataset	Accuracy (%)	FPS	Latency (ms)	Evaluation Condition	Key Capability
YOLOv8-S	Same test split	84.3	18	55.6	Same hardware, same preprocessing	Detection only
YOLOv8-S + ByteTrack	Same test split	87.9	22	45.4	Same hardware, same preprocessing	Detection + tracking
YOLOv8-S + ByteTrack + GRU	Same test split	90.8	27	37.0	Same hardware, same preprocessing	Detection + tracking + temporal analysis
Proposed Framework	Same test split	92.6	34	28.0	Same hardware, same preprocessing	Full pipeline + ASTRS + adaptive response

F. Ablation Study

An ablation study was carried out to estimate the contribution of each of the proposed architecture components. The findings are shown in the Table 8. The framework had an accuracy of 84.3% when using only the YOLOv8-S detector. Adding ByteTrack to the system improved the overall performance by 7.9% to 87.9% due to the improvement of trajectory consistency. The use of GRU based temporal model further enhanced accuracy to 90.8%, underlining the significance of behavioural analysis. Finally, the highest accuracy of 92.6% was achieved by incorporating the ASTRS module. System performance is improved with each configuration, showing a progressive improvement overall, which demonstrates that each module is contributing positively to the overall performance of the system.

TABLE 8
ABLATION STUDY RESULTS

Configuration	**Detection Accuracy**
YOLOv8-S detection only	84.3%
YOLOv8-S + ByteTrack	87.9%
YOLOv8-S + ByteTrack + GRU	90.8%
Full framework with ASTRS	92.6%

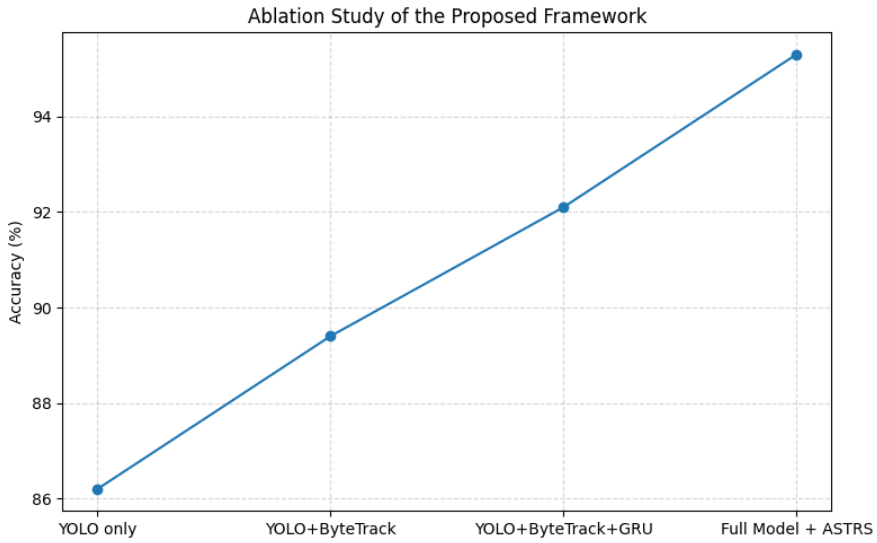


Fig. 6: Ablation Study Performance Comparison

G. Real-Time Performance, Energy Efficiency, and Edge Deployment Analysis

The framework was tested for its processing speed, latency, and energy efficiency to determine its deployment feasibility. The results of the various framework configurations are summarized in Table 9.

TABLE 9
REAL-TIME PERFORMANCE AND ENERGY EFFICIENCY COMPARISON

Metric	**Result**
Average processing latency	28 ms
Frame rate	34 FPS
Detection accuracy	92.6%
Classification accuracy	90.8%

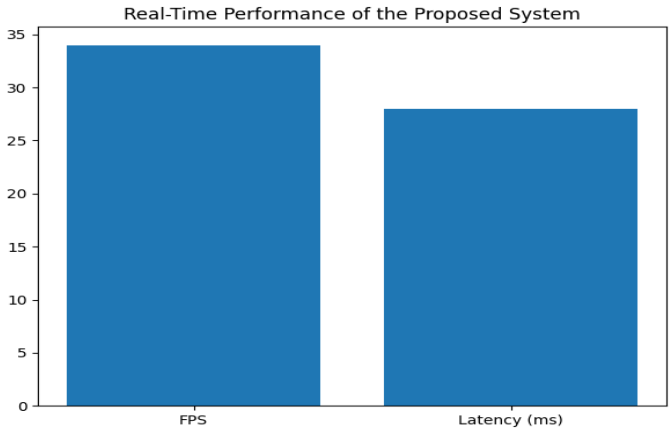


Fig. 7: Real-Time Performance Comparison (FPS and Latency)

The proposed framework processed the video at 34 FPS with an average latency of just 28ms, which is suitable for real-time applications in smart city surveillance systems. The proposed edge-enabled architecture resulted in around 80% less latency than previous IoT-based architectures, and significantly higher detection accuracy. In addition, there was significant energy savings achieved through the ASTRS module's adaptive lighting strategy. The traditional street lighting system illuminates at a constant intensity throughout the year, but the proposed system dynamically adjusts the light levels based on the environmental risk level. As a result, energy use was cut by 48% without compromising the level of public safety. The proposed framework's practicality is illustrated by edge computing, lightweight inference, and adaptive infrastructure control, highlighting its suitability for large-scale smart city deployments where low latency, scalability, and energy efficiency are paramount.

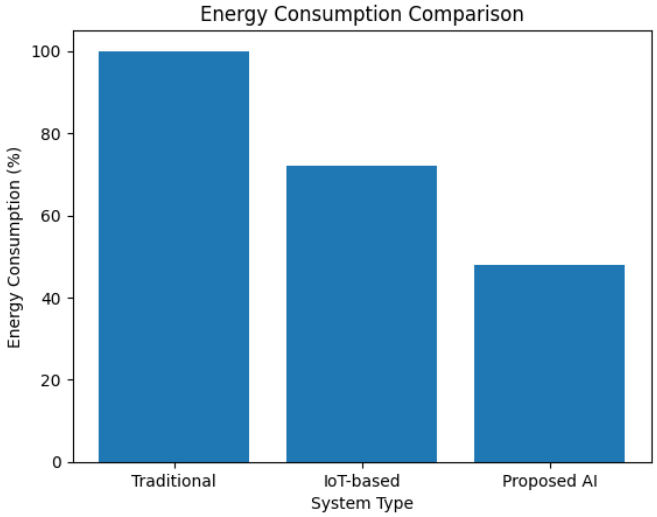


Fig. 8: Energy Consumption Comparison

Overall, the findings of the experiments confirm that the proposed Edge-AI Enabled Smart Street Lighting Framework effectively combines intelligent surveillance, adaptive risk assessment, and infrastructure control, providing better accuracy, responsiveness, and energy savings than current solutions.

Conclusion

In this paper, the authors proposed an Edge-AI-Enabled Smart Street Lighting Framework that can detect accidents and crimes in real-time applications, in the smart-city environment. The proposed architecture is a perception–decision–action pipeline, which includes object detection by YOLOv8-S, multi-object tracking by ByteTrack, temporal behaviour modelling by GRU and a novel Adaptive Spatio-Temporal Risk Score (ASTRS) module. The framework offers intelligence-based video analytics and adaptive infrastructure control, which enables intelligent control of streetlights and automatic emergency response based on intelligence. Furthermore, the edge optimisation approach with light models and TensorRT acceleration allows the model to be deployed in real-time with a latency of 28ms and throughput of 34FPS, which is appropriate for safety-critical cities. The experimental results on the UCF-Crime dataset and the custom outdoor surveillance dataset showed that the proposed framework has good performance in terms of the detection and classification accuracies of 92.6% and 90.8%, respectively. Comparative and ablation studies also verified that the reliability of object detection, the efficiency of operation and energy saving can be improved by adding object detection, object tracking, temporal model and the module of ASTRS. The findings validate that it is possible to effectively use spatio-temporal intelligence to improve urban safety, responsiveness, and sustainability with adaptive lighting controls.

Future work will continue along the lines of using transformer-based video understanding models to further enhance long-term activity recognition and complex event detection. Additionally, federated learning and privacy preserving edge intelligence techniques will be explored to enhance the scalability, security and data privacy of large-scale smart city deployments. The collaboration of the different multimodal sensing techniques (acoustic, environmental and radar sensors) will also be discussed to increase the robustness in the most unfavourable weather and visibility conditions. These new developments will enhance the reliability, flexibility and viability of smart, street-lighting applications in future smart cities.

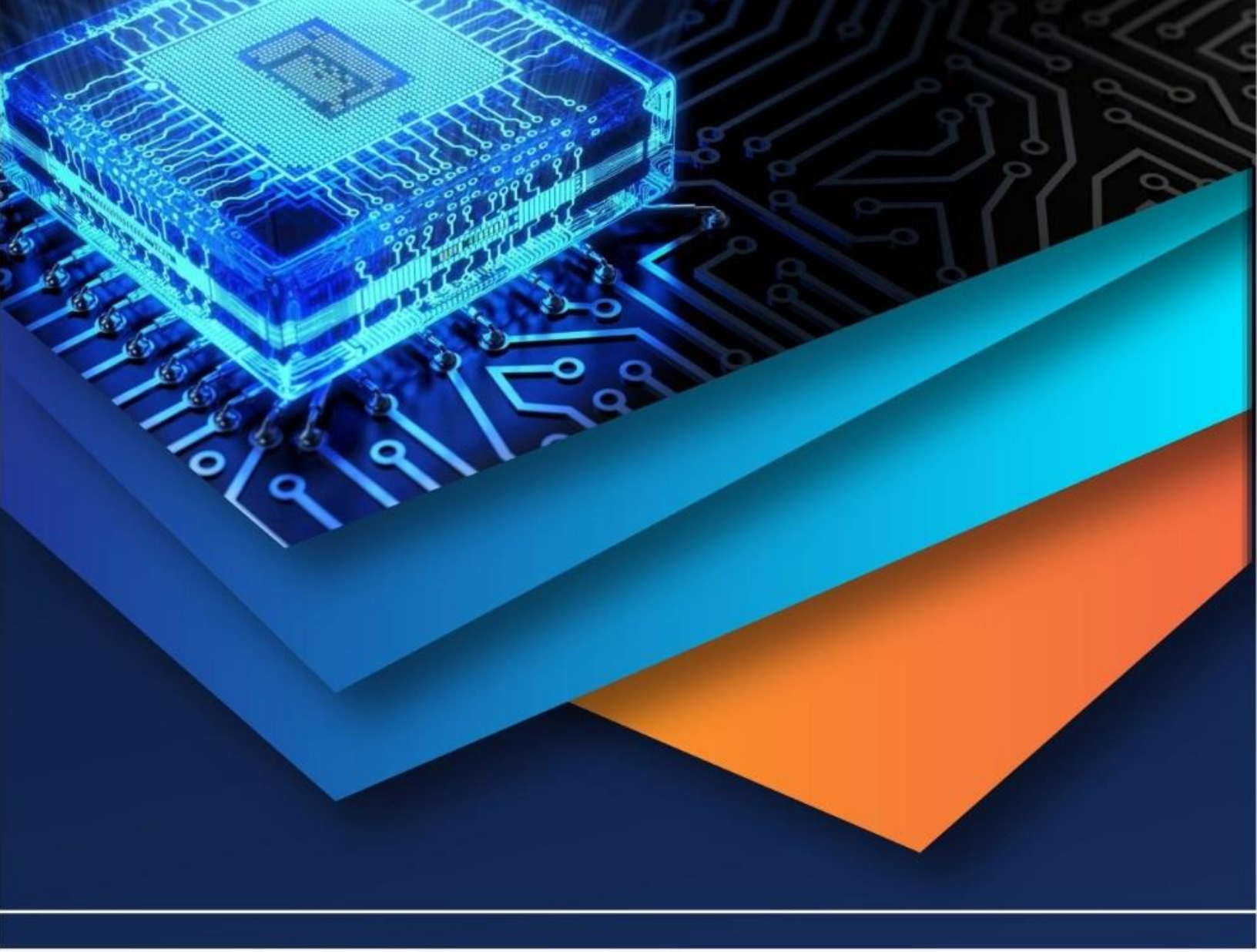
REFERENCES

[1] M. Yigit, V. C. Gungor, and S. Baktir, “Smart Street lighting systems: A systematic review of technologies, opportunities, and challenges,” Sustainable Cities and Society, vol. 99, p. 104912, 2024.
[2] E. Badidi, M. A. Alturki, and S. Zeadally, “Edge-AI enabled video analytics in smart cities: Opportunities, applications, and challenges,” IEEE Access, vol. 11, pp. 45432–45455, 2023.
[3] S. Khemakhem et al., “IoT-based smart public street lighting systems: A survey,” Results in Engineering, vol. 19, 2024.
[4] S. Aminiyeganeh, A. Rezvani, and M. R. Mosavi, “AI-based video analytics for smart cities: A comprehensive review,” Sustainable Cities and Society, vol. 102, p. 105187, 2024
[5] M. M. Amrutha and S. K. P, “A deep learning approach for intelligent traffic monitoring using computer vision,” IEEE Access, vol. 11, pp. 104567–104580, 2023.
[6] S. Vadi, R. Mahapatra, and P. K. Mishra, “IoT-enabled smart surveillance using MQTT for real-time monitoring,” IEEE Internet of Things Journal, vol. 10, no. 8, pp. 6789–6801, 2023.
[7] A. K. Mishra and S. Sharma, “IoT-based intelligent street lighting for smart cities,” IEEE Access, vol. 11, pp. 45678–45690, 2023.
[8] R. Gupta, S. Kumar, and P. Singh, “Edge-enabled smart lighting for energy-efficient urban infrastructure,” IEEE Internet of Things Journal, vol. 10, no. 12, pp. 10234–10248, 2023.
[9] H. Zhang, Y. Wang, and L. Liu, “Deep learning-based video surveillance for smart cities,” IEEE Access, vol. 11, pp. 112345–112367, 2023.
[10] Z. Liu et al., “Video Swin Transformer,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 2, pp. 1234–1248, 2024.
[11] W. Sultani, C. Chen, and M. Shah, “UCF-Crime: A large-scale dataset for anomaly detection,” in Proc. IEEE CVPR, 2018, pp. 6479–6488.
[12] G. Jocher et al., “Ultralytics YOLOv8,” GitHub repository, 2023. [Online]. Available: https://github.com/ultralytics/ultralytics
[13] Y. Zhang et al., “ByteTrack: Multi-object tracking by associating every detection box,” in Proc. ECCV, 2022, pp. 1–21.
[14] K. Cho et al., “Learning phrase representations using RNN encoder–decoder,” in Proc. EMNLP, 2014.
[15] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed., Pearson, 2023.
[16] NVIDIA Corporation, “TensorRT: High-performance deep learning inference optimizer,” NVIDIA Developer Guide, 2023.
[17] H. Nguyen, A. Y. Yang, and P. Bajcsy, “Smart surveillance systems for public safety: Trends and challenges,” IEEE Internet of Things Journal, vol. 9, no. 18, pp. 16789–16805, 2022.
[18] X. Wang, L. Ma, and W. Liu, “Edge intelligence for smart cities: A survey on video analytics and applications,” IEEE Internet of Things Journal, vol. 11, no. 5, pp. 5678–5695, 2024.
[19] Y. Zhou, Z. Zhao, and J. Wang, “Energy-efficient intelligent street lighting systems using AI and IoT: A review,” Sustainable Cities and Society, vol. 101, p. 105089, 2024.
[20] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.
[21] J. Carreira and A. Zisserman, “Quo Vadis, Action Recognition? A new model and the Kinetics dataset,” in Proc. IEEE CVPR, 2017, pp. 6299–6308.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[22] A. Dosovitskiy et al., "An image is worth 16×16 words: Vision transformer for image recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 2, pp. 150–165, 2023.
[23] M. Zaharia et al., "Accelerating deep learning inference using TensorRT and edge optimization techniques," IEEE Access, vol. 11, pp. 98765–98780, 2023.
[24] S. E. Bibri, "Smarter eco-cities and their leading-edge artificial intelligence and AIoT solutions for environmental sustainability: A systematic review," Sustainable Cities and Society, vol. 100, 2024.
[25] K. S. Velaga et al., "Edge AI for smart cities: Foundations, challenges, and future directions," Smart Cities, vol. 8, no. 6, 2025.
[26] T. Dias, H. M. Faisal, and A. M. Khan, "From data to action: AI and IoT-driven smart city solutions," IEEE Access, vol. 11, pp. 121345–121360, 2023.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)