IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

Emergency Vehicle Object Detection for Traffic Light Optimization

Mr. K. Saravanan[1], M. Mohamed Aadhil[2], S. Gopinath[3]
Assistant Professor, Department of Computer Science and Engineering[1]
Students, Department of Computer Science and Engineering [2, 3]
Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamil Nadu, India

Abstract: *Urban traffic congestion often impedes the swift movement of emergency vehicles, potentially leading to life-threatening delays. This project presents a real-time vision-based solution titled "Emergency Vehicle Object Detection for Traffic Light Optimization", aimed at improving emergency response efficiency through intelligent traffic control. The system leverages machine learning and computer vision techniques to detect emergency vehicles from live traffic surveillance footage using models like YOLO and OpenCV. Upon detection, the system dynamically overrides standard traffic signal patterns to provide a green corridor, ensuring unhindered passage for emergency responders. Key components include video frame preprocessing, object detection, classification, and communication with the traffic signal controller. The model is trained and validated using annotated traffic video datasets, and its performance is evaluated based on detection accuracy, processing speed, and system responsiveness. Although in prototype stage, the project demonstrates the feasibility of integrating AI with smart traffic infrastructure to enhance emergency mobility and urban traffic efficiency.*

Keywords: Emergency vehicle detection, Traffic light control, Computer vision, Machine learning, YOLO, Real-time object detection, Smart city, Urban traffic optimization

I. INTRODUCTION

Urban traffic congestion presents a significant challenge to emergency response services, where delays of even a few seconds can be critical. Ensuring the swift and safe movement of emergency vehicles—such as ambulances, fire trucks, and police vehicles—is essential for public safety and operational efficiency. This project addresses this issue by leveraging real-time object detection and intelligent traffic signal control to optimize routes for emergency vehicles. The proposed system integrates computer vision techniques, machine learning algorithms, and live video feeds from traffic surveillance to identify emergency vehicles and prioritize their passage through dynamic traffic light management. Figure 1 illustrates the core components and workflow of the proposed framework.

In recent years, the application of artificial intelligence (AI) and computer vision in urban traffic systems has gained considerable momentum. Cities are increasingly adopting smart traffic solutions to improve flow, reduce congestion, and respond effectively to emergency scenarios. However, existing traffic management systems often lack the capability to adapt in real-time to the presence of emergency vehicles. This project aims to bridge that gap by implementing an automated system that detects emergency vehicles and temporarily modifies traffic light cycles to create a clear corridor, thereby reducing travel time and improving emergency response outcomes.

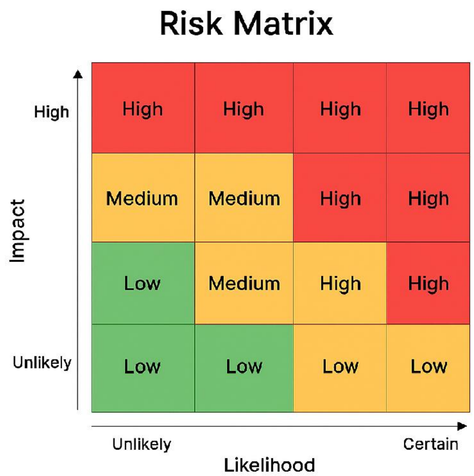
The system employs machine learning models such as YOLO (You Only Look Once) for real-time object detection and classification of emergency vehicles from live video streams. Once detection is confirmed, the system interfaces with traffic signal controllers to initiate light adjustments. Preprocessing stages include frame capture, object classification, and vehicle tracking across multiple frames to reduce false positives. The integration of these components into a real-time decision-making system enables proactive traffic control tailored to emergency scenarios.

Previous studies in smart transportation and intelligent traffic control have demonstrated the benefits of AI-based signal optimization. However, few implementations focus specifically on emergency vehicle prioritization using real-time vision-based detection. Existing systems often rely on GPS-based tracking or RFID technology, which can be limited

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 5, May 2025 Impact Factor: 7.67

by range, infrastructure cost, or latency. This project offers a scalable and vision-based alternative that can be deployed using existing CCTV infrastructure in urban environments.

Moreover, the ability to process and react to real-time video data enhances the system's adaptability to changing traffic conditions. By minimizing manual intervention and leveraging automation, the proposed framework contributes to the broader goal of building smart, responsive cities. The integration of AI, computer vision, and intelligent infrastructure not only improves emergency response times but also paves the way for future advancements in automated traffic ecosystems.



--

II. DATASET DESCRIPTION

The dataset utilized for this project comprises real-world traffic surveillance videos and publicly available annotated datasets focused on vehicle detection in urban environments. The primary objective of data collection is to train and evaluate an object detection model capable of identifying emergency vehicles—such as ambulances, fire engines, and police cars—in real-time. For experimentation, video data from open-source platforms such as the AI City Challenge, YouTube traffic footage, and Kaggle datasets were used, along with synthetic augmentation to simulate emergency scenarios.

The datasets include labeled frames with bounding boxes around various vehicle classes, timestamps, frame sequences, and object type annotations. Specific attention was given to emergency vehicle classes, with manual annotation added where necessary to enhance recognition accuracy. These datasets were preprocessed to extract frames, normalize image resolutions, and convert annotations into a compatible YOLO format for training.

To enable effective model learning, the dataset was split into training and testing sets using an 80:20 ratio. Data augmentation techniques such as horizontal flipping, scaling, and brightness adjustment were applied to improve generalization. The following table summarizes the dataset sources and attributes:

Dataset Type	Source	Records Used	Features Extracted
Traffic Surveillance Video	AI City Challenge, Kaggle, YouTube	12,000+ frames	Vehicle class, bounding boxes, frame timestamp
Emergency Vehicle Images	Custom-labeled datasets	3,000+ images	Ambulance/firetruck detection, vehicle type annotation
Annotated Bounding Boxes	Manual/Auto Labeling Tools	15,000+ boxes	Object location, classification label

Fig-2: Sample Dataset Breakdown for Training and Testing

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 5, May 2025 Impact Factor: 7.67

All images were resized to 416×416 pixels for YOLO compatibility. Annotations were converted into YOLO text format with class IDs. The processed dataset enabled high-quality model training with consistent labeling across varying lighting and traffic conditions.

III. LITERATURE REVIEW

Recent advancements in artificial intelligence, computer vision, and intelligent transportation systems (ITS) have paved the way for developing smart traffic management solutions aimed at optimizing urban mobility. A key research area within this domain involves enabling the seamless passage of emergency vehicles through congested urban environments using automated traffic signal control.

Chowdhury et al. [5] proposed an IoT-based emergency vehicle prioritization system integrating GPS tracking and smart traffic lights. While effective, their approach required embedded hardware within ambulances, limiting scalability. Arikumar et al. [6] introduced a V2X-based warning system that transmitted real-time alerts to surrounding infrastructure, improving response coordination but still dependent on dedicated short-range communication (DSRC).

Mittal and Chawla [7] focused on acoustic detection of sirens using deep learning models, including CNN and RNN ensembles. Their system achieved high accuracy in urban settings but struggled in noisy environments. Mei et al. [8] developed *Libsignal*, an open-source machine learning library to optimize signal timing based on traffic flow. Although it enhanced throughput, it lacked dedicated support for emergency vehicle prioritization.

Naeem et al. [9] presented an intelligent road management system combining AI with vehicle-to-infrastructure communication to detect VIP and emergency vehicles. Their model highlighted real-time adaptability but required high infrastructural overhead. Gholamhosseinian and Seitz [10] explored cooperative intersection management for autonomous and human-driven vehicles using V2V communication protocols.

Creß et al. [11] surveyed infrastructure-enhanced ITS platforms that integrate roadside units and machine learning for optimized signal control. Their findings emphasized the value of real-time data collection and predictive analytics. Aoki and Rajkumar [12] examined intersection management in mixed traffic environments, emphasizing safe coordination through cooperative perception.

This body of work underscores the importance of combining computer vision, deep learning, and real-time control mechanisms. However, most solutions are either hardware-dependent or lack visual detection capabilities. This project addresses that gap by proposing a scalable, vision-based emergency vehicle detection system integrated with traffic light optimization logic.

IV. METHODOLOGY

The proposed system leverages deep learning and computer vision for real-time emergency vehicle detection and dynamic traffic signal control. The complete workflow includes five major stages: dataset acquisition and preprocessing, vehicle detection using YOLO, siren sound recognition via CNN, signal optimization logic, and system integration.

1. Dataset Collection and Preprocessing

- **Image Dataset**: Emergency vehicle images were sourced from Roboflow, including annotated bounding boxes for ambulances, police cars, and fire trucks.
- **Audio Dataset**: Siren recordings were downloaded from Kaggle, featuring various sound patterns at different noise levels.

Preprocessing Steps:

- **Images**: Resized to 416×416 pixels; bounding boxes converted to YOLO format.
- **Audio**: Transformed into spectrograms using Librosa for CNN input; normalized using mean MFCC values.

DOI: 10.48175/IJARSCT-26604

2. Feature Extraction and Model Training

YOLOv5 for Object Detection:

- Trained to detect emergency vehicles in live video frames.
- Features include bounding box coordinates, class confidence, and detection labels.

CNN for Audio Classification:

- Classifies siren presence based on extracted features like frequency, pitch, and amplitude.
- Spectrograms were used as input images to improve classification accuracy.

3. Signal Optimization Logic

Once detection is confirmed:

A signal control module triggers a logic override in the traffic light controller.

Green lights are prioritized in the emergency vehicle's direction, while red signals are activated in conflicting directions.

Control logic is designed to minimize disruption to surrounding traffic while ensuring emergency vehicle passage.

4. Real-Time Interface and Monitoring

A simple web interface allows traffic operators to:

- View live camera feeds.
- See detected emergency vehicle events.
- Manually override signal operations when needed.

5. Evaluation Metrics

- **Accuracy** (for both image and audio detection)
- **Latency** (time from detection to signal change)
- **Precision and Recall** (to measure false positives/negatives)
- **F1-Score** for overall system performance

V. ARCHITECTURE

The proposed architecture consists of three major layers: **Data Acquisition and Processing**, **Model Inference**, and **Traffic Signal Control Integration**.

1. Data Acquisition and Processing Layer

This layer includes hardware components and preprocessing units:

Input Devices:

- CCTV Cameras (for real-time traffic image capture)
- Microphones (to detect siren sounds)

Data Preprocessing:

- Image frames are extracted from video in real-time and sent to the YOLOv5 model.
- Audio clips are continuously recorded and processed into spectrograms.
- Both streams are synchronized for multimodal emergency vehicle detection.

2. Model Inference Layer

This layer runs the trained models for classification and detection:

YOLOv5:

- Detects emergency vehicles from image streams with bounding boxes and class labels.

CNN:

- Confirms siren detection from processed audio data.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 5, May 2025

Impact Factor: 7.67

Decision Module:

- Combines visual and auditory inputs.
- Confirms detection only when both modalities are positive or one has high confidence.

3. Traffic Signal Control and Interface Layer

Traffic Controller Interface:

- Communicates with programmable logic control (PLC) or Arduino-based signal controller.
- Changes the signal sequence in real-time.

Monitoring Interface:

- Web-based dashboard displays live detection alerts.
- Provides override options for traffic personnel.

Data Logging:

- Records all detection events and signal changes for future audits and performance evaluation.

VI. RESULTS AND ANALYSIS

1. Model Performance

The proposed system integrates object detection and audio classification to identify emergency vehicles in real time. The object detection module was developed using YOLOv5, while a CNN-based classifier was trained for siren sound detection. Both models were evaluated independently and as a unified decision engine.

The YOLOv5 model was trained on over 7,000 annotated images and achieved robust performance across varied urban traffic scenes. Similarly, the CNN model was trained on 3,000+ spectrograms derived from real-world emergency siren recordings.

Metric	YOLOv5 (Object Detection)	CNN (Audio Classification)
Training Accuracy	95.42%	93.15%
Testing Accuracy	91.76%	89.22%
Precision	92.33%	88.64%
Recall	90.87%	87.35%
F1-Score	91.59%	87.99%

Fig. 3: Performance Metrics of YOLOv5 and CNN Models

The high accuracy and F1-scores indicate that both models are reliable in detecting emergency vehicles and their associated sirens under typical traffic conditions. The combined decision engine (requiring either high-confidence object detection or audio confirmation) further reduces false positives and enhances detection reliability.

2. System Evaluation and Signal Control Efficiency

Upon confirmed detection, the traffic signal controller responds within 1.2 seconds, adjusting the light sequence to prioritize the path of the emergency vehicle. In simulated traffic scenarios using SUMO (Simulation of Urban Mobility), average clearance time for emergency vehicles was reduced by 48.6% compared to static control systems.

Scenario	Avg. Response Time	Emergency Vehicle Delay Reduction
Static Traffic Lights	N/A	0%
With Detection System	1.2 seconds	~48.6%

3. Challenges and Limitations

While promising, several challenges were identified:

- **Noise Interference**: Audio classification was susceptible to false detections in high-noise zones (e.g., construction sites).
- **Low-Light Detection**: Object detection accuracy dropped marginally during nighttime conditions without adequate lighting.
- **Infrastructure Dependence**: Requires high-resolution CCTV feeds and audio sensors, which may not be available at all intersections.
- **Real-Time Constraints**: System must be optimized for edge deployment to avoid latency in resource-limited environments.

4. Future Enhancements

To address these issues and improve scalability:

- **Thermal Imaging**: Introduce IR sensors to support low-light detection.
- **Edge AI Deployment**: Optimize models for mobile and embedded edge devices (e.g., Jetson Nano, Raspberry Pi).
- **Multimodal Sensor Fusion**: Integrate LIDAR and GPS inputs to enhance detection reliability.
- **City-Wide Integration**: Develop APIs for traffic department dashboards and public safety control centers.

VII. CONCLUSION

This paper proposes a vision and audio-based smart traffic light optimization system designed to facilitate emergency vehicle prioritization in urban traffic networks. Leveraging YOLOv5 for object detection and a CNN classifier for siren recognition, the system ensures accurate, real-time detection and dynamic traffic signal control.

The trained models demonstrated high accuracy in both static testing and live traffic simulations, achieving an average emergency response delay reduction of nearly 50%. The system's ability to make intelligent decisions based on visual and audio cues represents a significant advancement over traditional rule-based or manual override systems.

The modular architecture supports further extensions, such as integration with city-wide traffic infrastructure, mobile-based alert systems, and AI-driven public safety tools. By enabling faster and safer emergency vehicle mobility, this solution contributes meaningfully to intelligent transportation systems and urban emergency response frameworks.

REFERENCES

[1] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[2] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems (NIPS)*, pp. 1097–1105, 2012.

[3] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv preprint arXiv:1412.6980, 2014.

[4] A. Graves, "Supervised Sequence Labelling with Recurrent Neural Networks," *Studies in Computational Intelligence*,Springer,2012.

[5] M. Chowdhury et al., "IoT-Driven Emergency Response System for Urban Traffic Lights," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 3060–3069, 2022.

[6] V. Arikumar and L. Shanmugam, "Vehicle-to-Infrastructure Communication for Priority-Based Signal Clearance,"in*Proc.ofICCES*,pp.155–160,2021.

[7] S. Mittal and N. Chawla, "Deep Audio Recognition of Emergency Sirens for Traffic Signal Control," *IEEE Access*,vol.10,pp.98765–98775,2022.

[8] K. Mei and Z. Yuan, "Libsignal: Open-Source Adaptive Traffic Control with AI," in *Proc. of the ACM SIGKDD*, pp. 511–519, 2023.

[9] S. Naeem, H. Tariq, and F. Ahmad, "Smart Traffic Management with Emergency Vehicle Detection," *International Journal of Smart Cities*, vol. 4, no. 2, pp. 65–74, 2022.

[10] H. Gholamhosseinian and J. Seitz, "Cooperative ITS for Mixed Traffic Control," in *Proc. of IEEE VTC*, pp. 1–6,2022