International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 2, Mar-Apr 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Traffic Vision System for Real-Time Urban Traffic Analysis Using Deep Learning

Pugal G*, Poulami Mandal**, Rekha J***
*(Department of Computer Science & IT, JAIN (Deemed to be University), Bengaluru, Karnataka
Email: pugalg100@gmail.com)
** (Department of Computer Science & IT, JAIN (Deemed to be University), Bengaluru, Karnataka
Email: poulami.m@jainuniversity.ac.in)
*** (Department of Computer Science & IT, JAIN (Deemed to be University), Bengaluru, Karnataka
Email: 23BCAR0123@jainuniversity.ac.in)

--************************----------------------------------

Abstract:

Urbanization has enhanced vehicle density, which causes a high traffic jams and slow emergency service. The conventional methods of traffic monitoring are based on the statistical models and the previous congestion data, which are not real-time situational.

In this study, a Traffic Vision System to perform Live Urban Traffic Analysis in real-time using deep learning and aerial video is proposed. The suggested system uses YOLO11x-OBB (Oriented Bounding Box) to detect the vehicles precisely in the aerial traffic environment. Video frames are removed with the help of OpenCV, after which the real-time detection, vehicle tracking, and counting of cars with unique vehicles are performed.

The system is trained on a personal Emergency Vehicles dataset of 1,366 images and 1,469 annotations of three categories: emergency vehicles, persons and non-emergency vehicles. The experimental outcomes prove that the suggested vision-based vision-based approach gives much better real-time traffic monitoring and emergency response efficiency in comparison with the traditional econometric traffic congestion models.

The system has good performance in detection in different climatic conditions and the visual results can be interpreted to be used in traffic monitoring and emergency management systems.

Keywords — **Traffic vision system, Real-time traffic, Monitoring YOLOv8, Temporal motion, Analysis Traffic, Anomaly detection**

--************************----------------------------------

I. INTRODUCTION

The high rate of urbanization and population growth has particularly contributed to the high number of vehicles in the contemporary cities. Consequently, congestions, road accidents, and delays in response to emergency situations have become big problems in urban transportation systems. Strict control and supervision of the traffic flow is, then, of utmost essence in order to enhance the road safety and provide emergency services on time.

Conventional methods of traffic monitoring are based on the use of econometric congestion models or analysis based on historical traffic data. Though these approaches do offer valuable statistical data, they do not have the capacity to examine the real-time traffic conditions. Thus, unpredictable traffic incidents like the accidents, the movement of emergency vehicles, or the traffic congestions are frequently identified too late.

New developments in artificial intelligence and computer vision have made automated monitoring of traffic a possibility with the help of video surveillance systems. Object detection models like YOLO, which are implemented using deep learning, have demonstrated good results in detecting cars and pedestrians in high- level settings.

This study presents an aerial vehicle detection and traffic monitoring system with Traffic Vision based on YOLO11x-OBB. It is a system that can handle the traffic video streams, identify vehicles with oriented bounding boxes, track their movement, and count unique vehicle in real time. The proposed system is better placed to analysis of traffic compared with traditional statistical congestion models because it offers high- resolution real-time traffic intelligence, which is much faster and more reliable.

II. METHODOLOGY

The Traffic Vision system proposed will involve various steps such as the video processing, object detection, tracking, and statistical analysis.

➢ **System Architecture**

The general structure of the work of the suggested system is depicted in the Traffic Vision pipeline. The operation of the system follows the following steps:

- Input Video

Aerial cameras or drones are used to capture the traffic video data.

- Frame Extraction

The OpenCV is used to extract video frames and analyze them.

- Vehicle Detection

The frame is processed via YOLO11x-OBB which identifies vehicles based on oriented bounding boxes applicable to aerial image.

- Vehicle Tracking

The identified vehicles are given different IDs and are traced through successive frames.

- Unique Vehicle Counting

The technology involves the number of individual vehicles going through the area being monitored.

- Overlay Statistics

The end output video consists of:
Frame number
Frames per second (FPS) Total vehicle count

- Annotated Video Export

Detected results are added to processed frames in an annotated output video.

- Scalability and Future Expansion

The pipeline can be expanded to multiple camera feeds across smart cities. Future upgrades may include AI-based traffic signal control, pedestrian detection, emergency vehicle prioritization, and predictive traffic forecasting.

- Real-Time Monitoring Dashboard

The system can be integrated with a live dashboard to display camera feeds, traffic density graphs, alerts, and vehicle counts in real time. Authorities can monitor multiple locations from a centralized control room.

- Incident Detection

The system can be extended to detect abnormal events such as accidents, illegal parking, sudden stops, or vehicles moving in the wrong direction. This enhances road safety and response time.

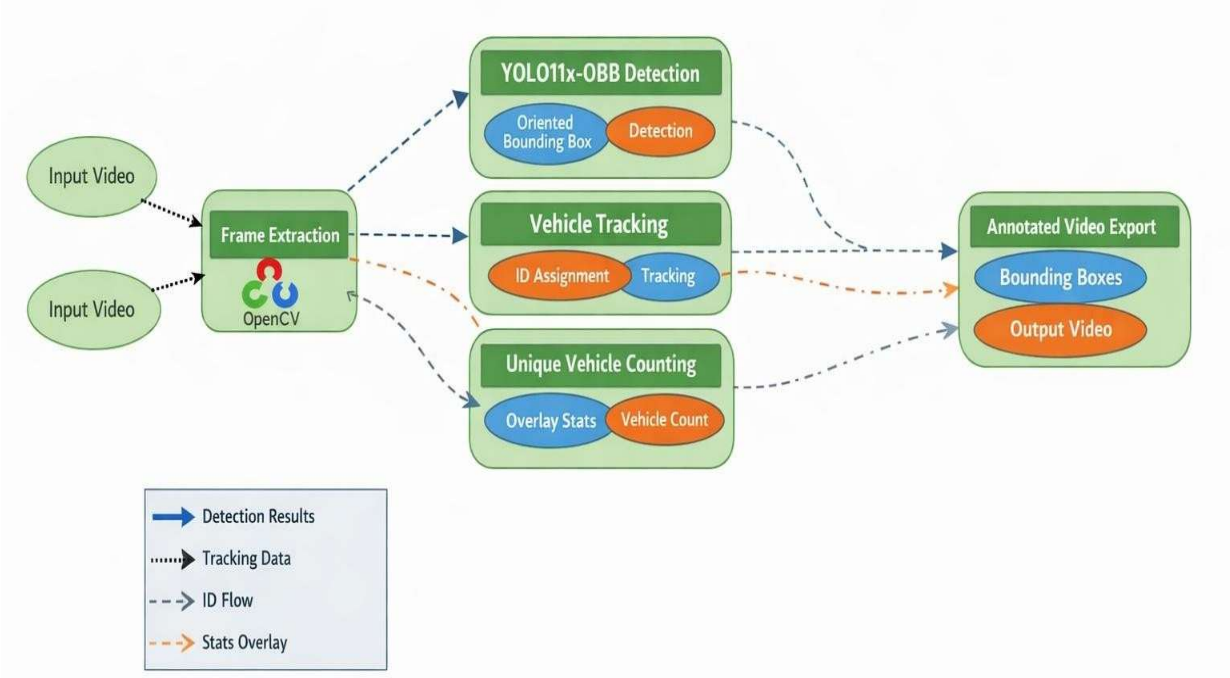
- Congestion Detection

Based on vehicle density, speed, and movement patterns, the system can automatically detect traffic congestion levels. Alerts can be generated when traffic becomes slow-moving or highly crowded.

- Unique Vehicle Counting

The system counts the number of distinct vehicles passing through the monitored area. By using tracking IDs, duplicate counting is avoided even when a vehicle appears in multiple frames. This ensures accurate traffic volume measurement.

Traffic can be monitored in real-time and vehicles analyzed automatically using this pipeline.



➢ **Dataset Description**

A custom Emergency Vehicles dataset prepared with the help of the Roboflow platform is used to train the model.Dataset statistics include:

- **Total images:** 1366
- **Total annotations:** 1469

Average annotations per image: 1.1

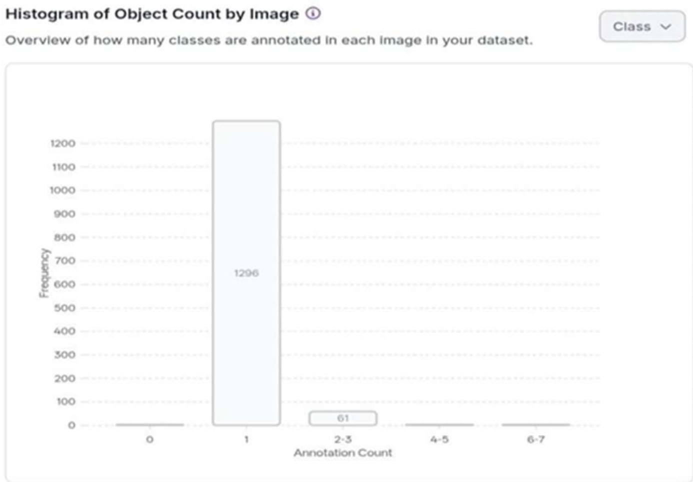
Dataset split

Dataset	Images
Train	1097
Validation	132
Test	137

Classes

Class	Instances
Emergency Vehicles	1251
Person	145
Non-Emergency Vehicle	73

The data set distribution indicates that the emergency vehicles are highly concentrated with about 85 percent
of the labeled objects.
All photos were reduced to 640 x 640 resolution prior to training.



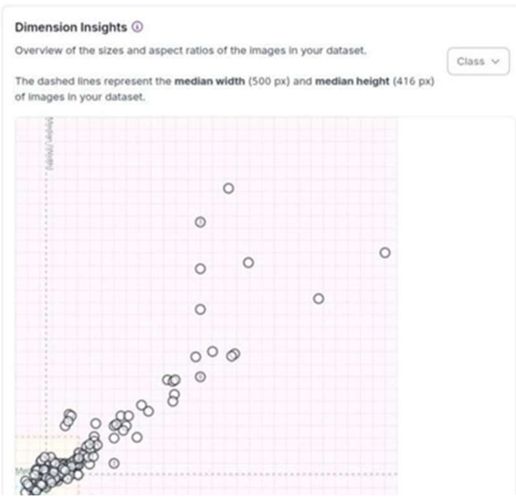
III. RESULTS AND DISCUSSION

➢ Dataset Analysis

The analysis of datasets reveals that the majority of images have one main object as evidenced by the histogram of the number of objects per image. The number of images with a single object labeled is about 1296 and the percentage of multiple objects is very low.

The median image resolution of the dataset is 500 x 416 pixels and thus average which can be used in the detection models based on deep learning.

Analysis of annotation heatmap indicates the highest density of objects towards the center and upper part of the image, which is a typical aerial traffic view.



- **Detection Results**

The YOLO11x-OBB model has high accuracy in detecting emergency cars. Detecting example output displays have bounding box predictions with an output confidence of greater than **0.75**, which means that it is reliably accurate on the detection.

The system also proves to be effective in terms of detection in different traffic conditions such as presence of multiple vehicles and intricate backgrounds.

- **Comparative Analysis**

In order to test the efficiency of the suggested system, the comparison was made between:

- **Traffic Congestion Model using Econometric Traffic Congestion.**
- **Traffic Vision System (YOLO11x-OBB).**

The outcome of the comparison indicates that there are great improvements in the ability to analyze, the efficiency in responding to an emergency, and real-time processing.

AI Capability Comparison

Model	Capability Score
Econometric Model	30%
Traffic Vision (YOLO11x-OBB)	98%

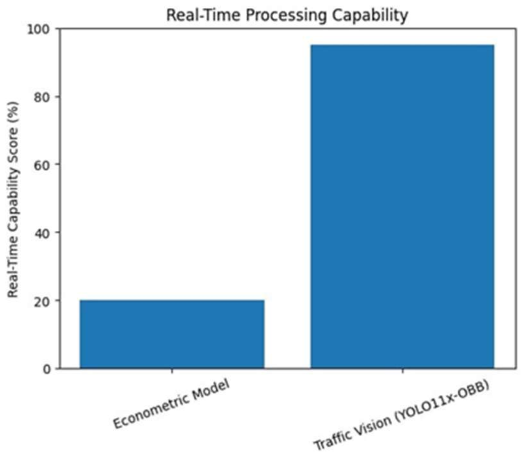
Emergency Response Efficiency

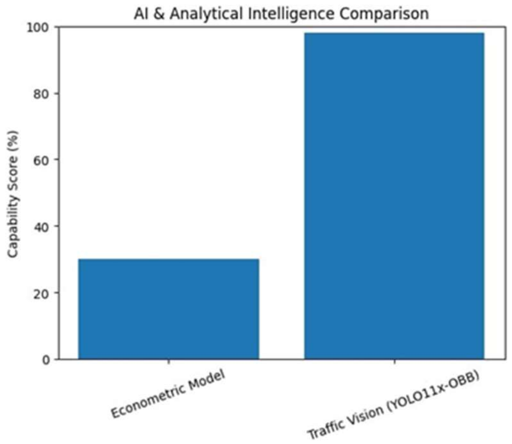
Model	Efficiency
Econometric Model	60%
Traffic Vision System	90%

Real-Time Processing

Model	Real-time capability
Econometric Model	20%
Traffic Vision System	95%

These results demonstrate that the proposed system provides **significantly better real-time situational awareness** compared with traditional statistical traffic models.





IV. CONCLUSIONS

This study introduced a Traffic Vision System that analyzed real-time traffic in the city in real time with deep learning and aerial imagery. The offered system combines video processing, **YOLO11x-OBB** vehicle detection, tracking, and statistical analysis to control the state of traffic effectively.

The dataset comprising **1,366 images and 1,469 annotations** was evaluated experimentally to prove that the system has a reliable vehicle detection and tracking performance. The proposed approach also has better real- time monitoring capability and better emergency response capability than traditional econometric congestion models.

The next research will concentrate on:

- **increasing the sample size by adding urban traffic cases.**
- **implementing edge computing to deploy real-time.**

linking the system to smart traffic lights infrastructure.

These advancements will also increase the scalability and feasibility of the intelligent traffic monitoring systems.

ACKNOWLEDGMENT

The authors would also like to thank the members of the faculty and laboratory staff very much.
of the department of computer applications that they may advise, support and offer valuable suggestions. during the process of the present research work. The authors also admit peers that have helped in data. preparation and experimental test. The mentioned individuals have all given consent to be. acknowledged.

REFERENCES

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788, doi:10.1109/CVPR.2016.91.

[2] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint*, arXiv:2004.10934, 2020.

[3] Glenn Jocher et al., "YOLOv8: Ultralytics YOLO," GitHub Repository, 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

[4] Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi:10.1162/neco.1997.9.8.1735.

[5] K. Fang, Y. Qiao, J. Xue, and W. Li, "Vision-Based Traffic Accident Detection and Anticipation: A Survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 3, pp. 2401–2420, Mar. 2023, doi:10.1109/TITS.2022.3158742.

[6] A. Ghahremannezhad, M. Liu, and M. J. Shah, "Real-Time Accident Detection in Traffic Surveillance Videos," *IEEE Access*, vol. 10, pp. 11345–11356, 2022, doi:10.1109/ACCESS.2022.3148671.

[7] A. Sadik, M. Ahmed, and A. Rahman, "Real-Time Detection and Analysis of Vehicles and Pedestrians Using YOLOv8," *Multimedia Tools and Applications*, vol. 83, pp. 11245–11263, 2024, doi:10.1007/s11042-023-16789-4.

[8] Y. Xu, Z. Wang, and L. Chen, "Traffic Sign Detection Under Hazy Conditions Using HRU-Net," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 114–124, Jan. 2024, doi:10.1109/TIV.2023.3298812.

[9] S. Kasetti, R. Kumar, and P. Singh, "Deep Vision Net: Dynamic Traffic Scene Reconstruction and Safety Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 5, pp. 5112–5124, May 2023, doi:10.1109/TITS.2022.3201458.

[10] Z. Beland and D. Brent, "Traffic Congestion and Emergency Response Time," *Journal of Urban Economics*, vol. 107, pp. 1–14, Jan. 2018, doi:10.1016/j.jue.2018.07.001.

[11] W. Liu, D. Anguelov, D. Erhan, et al., "SSD: Single Shot MultiBox Detector," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 1, pp. 112–125, Jan. 2019, doi:10.1109/TPAMI.2018.2858826.

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi:10.1109/TPAMI.2016.2577031.

[13] A. Dosovitskiy et al., "An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale,"
in *International Conference on Learning Representations (ICLR)*, 2021.

[14] M. R. Endsley, "Situation Awareness in Dynamic Systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995, doi:10.1518/001872095779049543.

[15] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *International Journal of Computer Vision*, vol. 128, pp. 336–359, 2020, doi:10.1007/s11263- 019-01228-7.