Grenze International Journal of Engineering and Technology, January Issue

Traffic Density Estimation using Real-Time Video Processing

Aditya Ozalwar[1], Om Jangade[2], Tanmay Dhakate[3], Suhas Asabe[4] and Dr. Santosh Warpe[5]
[1-5]MIT Academy of Engineering, Alandi, Pune, India.
Email: tanmay.dhakate@mitaoe.ac.in, {aditya.ozalwar, om.jangade, suhas.asabe, tanmay.dhakate, stwarpe}@mitaoe.ac.in

***Abstract*—To tackle the growing problems of traffic congestion in cities, the "Real-Time Traffic Density Estimation System" offers a complete solution. Traffic management systems are unable to keep up with the fast urbanization, industry, and population growth of cities, which leads to inefficiencies in transportation and increased air pollution. For efficient traffic management, infrastructure planning, and policy-making, our project offers a cutting-edge method for estimating traffic density utilizing real-time video processing. The system analyzes live video feeds from carefully positioned cameras and applies cutting-edge methods in vehicle detection, tracking, image processing, and machine learning for traffic prediction. The system's real-time aggregation and processing of this data yields precise traffic density estimates, facilitating improved decision-making for congestion reduction and urban traffic flow optimization. By facilitating better traffic control and lowering pollution, this project supports healthier living environments and smarter, more efficient urban planning.**

***Index Terms*—Urban Traffic Management, Congestion Control, Traffic Monitoring, Traffic Prediction, Image Processing, Machine Learning, Real-Time Video Processing, and Vehicle Detection.**

I. Introduction

In the rapidly evolving landscape of urbanization and technological advancements, traffic management has become a critical aspect of city planning and infrastructure development.
The growing population, especially in cities, has resulted in a corresponding rise in the number of vehicles on the roads.
The surge has resulted in congested traffic conditions, necessitating innovative solutions for effective traffic management, pollution reduction and informed decision-making in transportation planning and policy. The contemporary era witnesses the convergence of computer vision and video processing technologies, presenting a promising opportunity for real-time data utilization in traffic monitoring.
By utilizing these technological advancements, our project seeks to create a reliable system for estimating traffic density through real-time video processing.
The project recognizes the significance of accurate traffic density estimation as a fundamental element in addressing the challenges posed by dense traffic conditions. By employing cameras strategically positioned at key locations, we intend to capture real-time video data, which can be processed and analyzed to derive meaningful insights. The insights derived from this system will contribute to long-term infrastructure planning, pollution reduction strategies and traffic flow forecasting.

Grenze ID: 01.GIJET.11.1.347
© Grenze Scientific Society, 2025

II. Related Work or Literature Studies

N. Navghare et al. [1] studied the use of Python and OpenCV for automating live surveillance through object detection, which involves identifying, classifying, and tracking objects or events in images and videos. It describes the process of using algorithms that match patterns in images against a training database, including steps such as image importation, analysis, manipulation (e.g., compression, enhancement, pattern recognition), and the generation of modified images or analysis reports. Training involves the use of positive images (containing objects to be detected) and negative images (backgrounds), with a larger dataset enhancing the classifier's accuracy. The approach can be used in traffic density estimation using real-time video processing, which will improve video quality and enable accurate detection and analysis of traffic patterns, thereby enhancing surveillance system performance. Another study shows by D. Y. Sakhare et al. [2] proposed a forest fire detection system based on remote sensing and deep learning, aimed at monitoring
forests and identifying fires in their early stages. Remote sensing is used for comprehensive data collection, which is then employed to train the detection model. A deep learning model ensures accurate fire detection. Additionally, the video processing techniques applied in this system can be adapted for traffic density estimation, allowing for effective real-time traffic pattern analysis and monitoring.
P. R. Rajarapollu et al. [3] analyzed a defined system that employs three primary methods for object tracking:

- Point Tracking: This method identifies points of interest in each video frame, allowing the system to follow specific features throughout the sequence.
- Kernel Tracking: This approach focuses on tracking object regions across video frames, often using algorithms that compare appearance models or templates to locate the object of interest.
- Silhouette Tracking: Utilizing shape information, this method tracks objects based on their outlines or contours, making it effective for monitoring complex shapes and varying object appearances.

These methods work together to enhance the accuracy and reliability of object tracking in dynamic video environments.
U. Verma et al. [4] presented video magnification methods, particularly Eulerian Video Magnification (EVM) and Phase-based Video Magnification (PVM), for detecting minor motions, colour shifts, and sound variations in videos. EVM offers speed and simplicity, making it ideal for smaller amplification factors, though it becomes noisy at higher levels, while PVM, although slower and more complex, handles larger amplification factors more effectively. The study proposes a hybrid approach to enhance performance, particularly in biomedical contexts. These techniques, assessed through metrics such as PSNR and execution time, can also be utilized in real-time traffic density estimation, improving the detection of subtle changes in traffic flow. Additionally, U. Verma et al. [5] addresses Eulerian Video Magnification (EVM), which enhances subtle, otherwise imperceptible signals in videos, particularly colour changes. However, EVM requires stable, jitter-free footage with the subject taking up most of the frame—conditions often not achievable in real-world or real-time scenarios, resulting in less-than-ideal performance. To address this, a new method called Region-based Stabilized Video Magnification (RSVM) was developed. RSVM stabilizes the video by preprocessing it, then detects, tracks, and crops the region of interest (ROI), producing a motion-free, modulated input that suits EVM's requirements. The study evaluates the performance of both EVM and RSVM across different inputs and amplification factors, using metrics such as PSNR, processing time, and the degree of colour and motion magnification. Findings show that RSVM outperforms EVM by handling motion more effectively and accelerating the process. Lim, D. et al. [9] assessed the Advanced Driver Assistance System (ADAS) method demonstrating that high precision in traffic density estimation can be achieved with a smaller number of probe vehicles compared to traditional methods requiring extensive datasets. The study uses ADAS data to estimate traffic density. ADAS is capable of capturing distance headway data and vehicle trajectories in real-time through sensors like LiDAR, radar, and GPS.

III. Motivation

Traffic Density Estimation Using Real-Time Video Processing project is driven by the increasing challenges of urbanization, including traffic congestion, the limitations of traditional traffic management, and the adverse effects on quality of life in densely populated cities. As urban populations grow and vehicle numbers rise, conventional systems struggle to provide accurate, real-time data on traffic conditions. Leveraging advancements in computer vision and video processing, this project aims to revolutionize traffic management by delivering adaptive, data-driven solutions that support smart city initiatives, enhance urban planning, and promote sustainable development.

IV. Problem Domain

The problem domain of the "Traffic Density Estimation Using Real-Time Video Processing" project lies in the intersection of urban traffic management, computer vision, and machine learning.
It focuses on addressing the growing challenges of traffic congestion in urban areas due to population increases and rising vehicle numbers. Current traffic systems lack real-time adaptability, and precision in dynamic conditions, and rely heavily on manual monitoring and historical data, which limits effective traffic control and long-term urban planning.
The project aims to develop an intelligent system that utilizes video processing and machine learning to provide real-time traffic density estimation, offering immediate insights for traffic management while contributing to urban infrastructure planning, pollution reduction, and sustainable transportation strategies.

V. Problem Definition

The "Traffic Density Estimation Using Real-Time Video Processing" project centers on the limitations of current traffic management systems in urban areas such as Real-Time Android Application for Traffic Density Estimation [6], which are increasingly burdened by traffic congestion due to population growth and rising vehicle numbers. These systems lack the capability for real-time precision, and adaptability to dynamic traffic conditions, and rely heavily on manual monitoring and historical data analysis such as DSRC-Enabled Probe Vehicles [7]. As a result, they struggle to provide effective traffic control or support long-term urban planning efforts. The project seeks to address these issues by developing an adaptive, intelligent system that leverages advanced video processing and machine learning to estimate traffic density in real time, thereby enhancing traffic management and contributing to sustainable urban development.

VI. Statement

Urban areas face increasing traffic congestion due to population growth and more vehicles. Existing traffic management systems lack real-time precision, adaptability, and rely on manual monitoring. This project aims to develop a system using video processing and machine learning to estimate traffic density in real time, improving traffic management and urban planning.

VII. Innovative Content

The project offers an innovative solution for real-time traffic management using video processing and deep learning models [CNN (Convolutional Neural Network), YOLO (You Only Look Once)] to accurately estimate traffic density. Key features include real-time monitoring, adaptable algorithms for dynamic traffic conditions, and user-friendly visualization for decision-making. The system supports smart city goals by providing data-driven insights for urban planning and sustainable traffic management, while also ensuring privacy and security of collected data.

VIII. Solution Methodologies or Problem-Solving

A. Object Detection Review

Object detection is identifying and locating objects in the image or video, the output consists of labeled boxes, classes and the confidence score(accuracy level) on them [10]. The primary challenge in this field is to create a model that efficiently detects objects with high accuracy and speed. Given the significance of real-time applications, such as traffic monitoring, autonomous vehicles, and surveillance systems, our focus is to optimize object detection algorithms to ensure precise detection.
YOLOv8, the latest in the YOLO (You Only Look Once) series, excels in precision and speed, making it ideal for real-time traffic density estimation and congestion management. The model integrates a novel anchor-free detection head and an advanced backbone network based on Darknet53.
The loss function is the summation of Classification Loss, Intersection over Union (IoU) and Confidence Loss. For the traffic density estimation task:

$$\text{Total Loss} = \text{Classification Loss} + \text{Localization Loss (IoU)} + \text{Confidence Loss}$$

This model achieves a mean Average Precision (mAP) of 53.7% on the COCO benchmark dataset, demonstrating its robustness for high-precision object detection tasks in traffic management.

In Convolutional Network Architecture, it consists of 24 Layers followed by 2 Fully connected layers. Its major part is feature extraction from the input images, and the task of the remaining layer is to predict the probability of the boxes which are created and the confidence score[8]. The network's architecture influences Prediction Accuracy

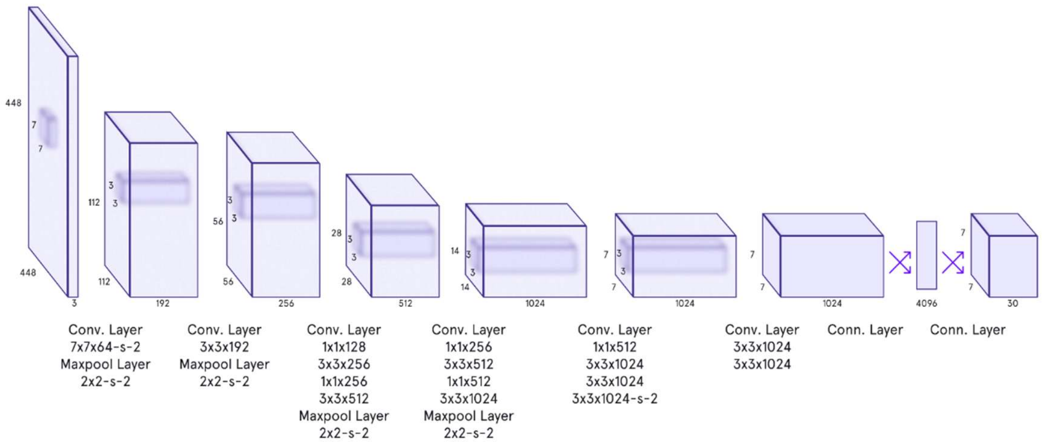

Fig. 1 Convolutional Neural Network of the YOLO Model [11]

B. System Architecture of Traffic Density Estimation

To implement the "Online Density Estimation and Vehicle Classification Management System," multiple databases are necessary. Video databases are used to store footage captured by mounted camera sensors at different locations over regular time intervals. These video databases function as distributed systems across designated areas of the city. The system's business logic follows a series of steps outlined in the system overview, with the vehicle classification engine and density estimation engine handling the processing, as illustrated in Fig. 2.

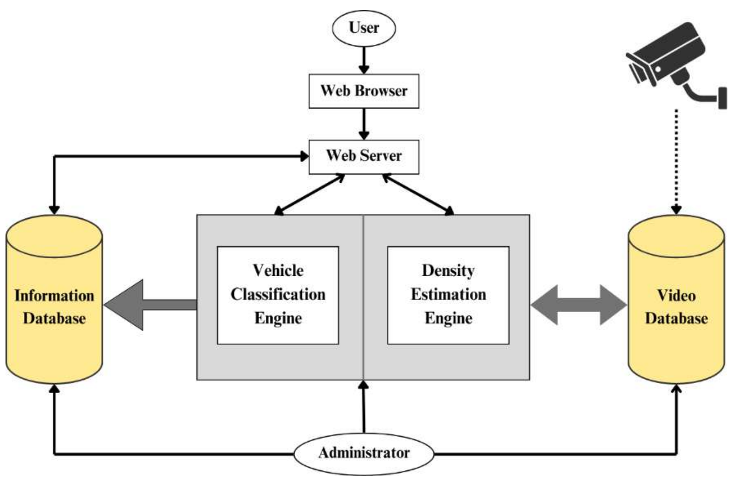

Fig. 2 System Architecture of Traffic Density Estimation

C. Dataset

To Fine-Tune our pre-trained model for vehicle detection, the top-view Vehicle Detection Image Dataset for YOLOv8 is prepared, Focusing on cars, Trucks and Buses. The Dataset includes 626 annotated top-view images, standardized to 640x640 pixels.

With 536 Augmented Training images and 90 unaugmented validation images, it enhances the model's generalization and evaluation. Sourced from pixels and annotated via Roboflow, this dataset is ideal for highway monitoring, with frames sampled at 1 Frame Per Second.

IX. Result And Analysis

A. Comparative analysis

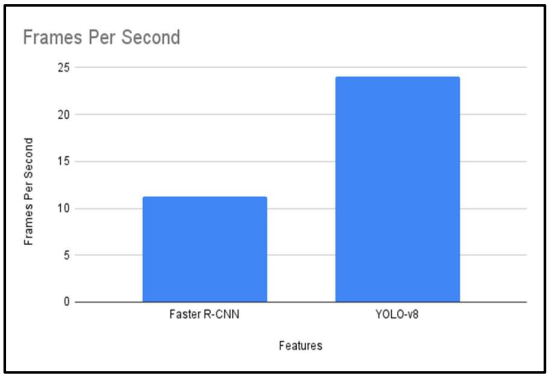

Fig. 3 Comparison Graph of FPS for Faster R-CNN & YOLO-v8

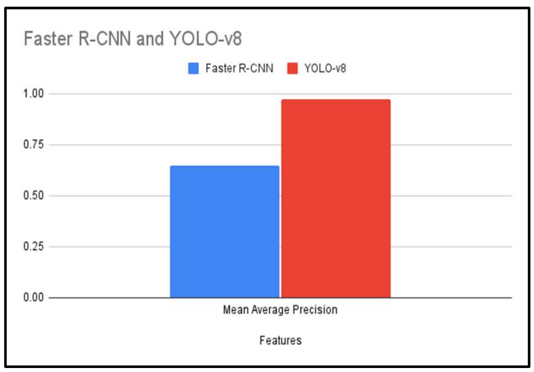

Fig. 4 Comparison Graph of mAP for Faster R-CNN & YOLO-v8

TABLE 1

Features	**Faster R-CNN**	**YOLO-v8**
Mean Average Precision	65%	97.2%
Frames Per Second	11.27	24.03

We carried out a comparison between the YOLOv8 and Faster R-CNN models on a dataset focused on traffic density estimation. In this analysis, YOLOv8 exhibited an impressive balance of speed and performance, achieving a mean Average Precision (mAP) of 97.20%. This demonstrates its ability to accurately detect vehicles across multiple frames. In contrast, Faster R-CNN outperformed in terms of accuracy, reaching an mAP of 65.00%, showcasing its strength in precise object detection. However, Faster R-CNN's significant drawback is its low processing speed, operating at only 11.27 frames per second (FPS), making it less suitable for real-time traffic monitoring applications.

On the other hand, YOLOv8, with its speed of 24.03 FPS, is much better suited for real-time video processing, where quick decisions are crucial. The high frame rate ensures smooth traffic monitoring and density estimation, even in dynamic environments. While YOLOv8 may have a slightly lower mAP than Faster R-CNN, its balance of Precision and Recall further highlights its suitability for traffic monitoring.

B. Experimental Results:

Fig. 5 shows the output of our YOLOv8 Model. The figures in this case depict various vehicles registered by the model, delivering each vehicle within an enclosed box. Each of the bounding boxes is written with the word vehicle comprising 'vehicle' class & confidence on detection. This kind of annotation showcases the wild capabilities of the model in terms of vehicle detection as well as for estimating the level of trust of the system in its output.

Fig. 5 Experimental Results of YOLO-v8 Model

X. CONCLUSION

Traffic Density Estimation Using Real-Time Video Processing offers an innovative solution to urban traffic challenges using machine learning and real-time video analysis. The system enhances traffic management, infrastructure planning, and environmental sustainability while ensuring user-friendly visualization and data security. It lays the foundation for future integration with smart city technologies, contributing to more efficient and intelligent urban environments.

REFERENCES

[1] Navghare, N., Kedar, P., Sangamkar, P., Mahajan, M. (2020). "Python and OpenCV in Automation of Live Surveillance". In: Swain, D., Pattnaik, P., Gupta, P. (eds) Machine Learning and Information Processing. Advances in Intelligent Systems and Computing, vol 1101. Springer, Singapore.

[2] S. Madkar, D. Y. Sakhare, K. A. Phutane, A. P. Haral, K. B. Nikam and S. Tharunyha, "Video Based Forest Fire and Smoke Detection Using YoLo and CNN," 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-5.

[3] A. G. Powale, V. D. Khade, A. L. Kamble and P. R. Rajarapollu, "Vision Based Object Tracking & Firing System," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-4.

[4] S. Yadav, P. Bhalkare, S. Shingde and U. Verma, "Performance Analysis of Video Magnification Methods," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1293-1301.

[5] Yadav, S. and Bhalkare, P. and Verma, U., "Region-Based Stabilized Video Magnification Approach," 2022 Lecture Notes in Electrical Engineering, pp. 381-394.

[6] F. Kerouh and D. Ziou, "Real-Time Android Application for Traffic Density Estimation," in IEEE Access, vol. 6, pp. 49896-49901, 2018.
[7] S. Chaturvedi, A. Ashok and B. R. Chilukuri, "Traffic State Estimation using DSRC-Enabled Probe Vehicles," 2021 International Conference on COMmunication Systems & NETworkS (COMSNETS), Bangalore, India, 2021, pp. 715-722.
[8] Jiang, Rui & Lin, Qian & Qu, Shuhui. (2016). Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio.
[9] Lim, D., Seo, Y., Ko, E., & Kim, H. (2021), "Spatiotemporal Traffic Density Estimation Based on ADAS Probe Data," Journal of Advanced Transportation, 2022(1), 5929725.
[10] Srivastava, S., Divekar, A.V., Anilkumar, C. et al. "Comparative analysis of deep learning image detection algorithms". J Big Data 8, 66 (2021).
[11] Pareto. (2024). Introduction to YOLO Object Detection: Understanding the Basics. Pareto https://pareto.ai/blog/yolo-object-detection