



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Study of Deep Learning for Vehicle Image Classification for Traffic Decongestion Control

Bhagyashree Patil¹, Rucha Sane², Vaishnavi Dedge³

DY Patil institute of MCA and Management Akurdi Pune

¹bhagyashree.patil@dypimca.ac.in, ²ruchasane28@gmail.com, ³dedgevaishnavi67@gmail.com

ABSTRACT

Now-a-days Traffic congestion is one of the most pressing problems in modern cities, causing delays, fuel wastage, and environmental pollution. Conventional traffic monitoring systems often depend on manual observation or sensor-driven methods, which generally suffer from limited accuracy and flexibility. With recent progress in deep learning, more effective solutions have emerged for real-time vehicle detection, classification, and traffic control. This study explores use of advanced deep learning architecture and Convolutional Neural Networks (CNNs), YOLO, and Faster R-CNN—for vehicle image classification and their potential integration into intelligent traffic management and congestion reduction systems.. The methodology includes dataset collection, preprocessing, model selection, training, and real-time deployment for adaptive traffic signal control. Experimental insights suggest that deep learning can significantly improve classification accuracy and support congestion reduction strategies, especially when combined with smart traffic signal systems.

Keywords: Deep Learning, CNN, YOLO, Traffic Decongestion, Vehicle Classification, Intelligent Traffic System

1. Introduction

Urban traffic congestion is a global concern, particularly in developing countries with rapidly increasing vehicle populations. Traditional systems such as inductive loop sensors, RFID, and manual monitoring are not effective for managing constantly changing traffic conditions. Deep learning-based computer vision provides new ways to detect, classify, and make intelligent decisions about vehicles in real time

This study focuses on applying deep learning techniques for vehicle image classification, with the ultimate goal of reducing congestion through adaptive traffic control.

Problem Statement:

Traffic congestion occurs due to inefficient traffic signal management and lack of vehicle-specific classification.

Need for Research:

- Real-time vehicle classification for smart traffic signals.
- Automation to replace error-prone manual monitoring.
- Application of advanced deep learning models (CNN, YOLO, Faster R-CNN).

Objectives:

1. Study deep learning methods for vehicle classification.
2. Develop a classification model for cars, buses, trucks, and motorcycles.
3. Integrate classification results with traffic signal decongestion strategies.

Hypothesis:

If vehicle classification using deep learning is implemented effectively, traffic signals can be optimized dynamically, reducing congestion and improving traffic flow.

2. Literature Review

- **Traditional Approaches:** Background subtraction, edge detection, and SVM classifiers. Limitation: poor accuracy in varied lighting/weather.
- **Deep Learning Approaches:** CNN models achieve high accuracy for classification tasks. Object detection frameworks like **YOLO (You Only Look Once)**, **SSD**, and **Faster R-CNN** balance speed and accuracy.
- **Transfer Learning:** Models like ResNet, VGG16, and MobileNet have been used to improve performance on small datasets.

Comparative Analysis:

Method	Pros	Cons
Traditional (SVM, Edge Detection)	Simple, low-cost	Low accuracy, sensitive to environment
CNN Classification	High accuracy	Requires large dataset
YOLO	Fast, real-time detection	Slightly less accurate
Faster R-CNN	Very accurate	Slower for real-time use

Research Gap:

- Most works focus only on detection, not on **traffic decongestion control**.
- Lack of datasets covering **Indian traffic conditions**.

3. Research Methodology

3.1 Research Design

experimental methodology was utilized to analyze classification performance and real-time deployment efficiency.

Performance Comparison of Object Detectors				
Table 1 presents a comparison of four object detection algorithms—Haar Cascade, SSD, YOLOv3, and Mask R-CNN—based on their processing time and vehicle detection accuracy across three videos from the GRAM dataset: <i>M-30</i> , <i>M-30-HD</i> , and <i>Urban1</i> .				
Detector	Metric	M-30	M-30-HD	Urban1
Haar Cascade	Time (s)	0.08–0.13	0.30–0.44	0.02–0.06
	Accuracy	43%	75%	40%
SSD	Time (s)	4–7	11–14	2.6–5.6
	Accuracy	22%	70%	69%
YOLOv3	Time (s)	1.0–1.8	1.0–1.8	1.0–1.8
	Accuracy	82%	86%	91%
Mask R-CNN	Time (s)	2.4–3.0	2.4–3.0	2.4–3.0
	Accuracy	89%	91%	46%

Data Collection

- **Public Datasets:**
 - **UA-DETRAC:** Provides 140,000 frames with vehicle bounding box annotations under varying conditions.

- **MIO-TCD:** Includes over 600,000 vehicle images across 11 categories (cars, trucks, buses, motorcycles, etc.).
- **KITTI Vision Benchmark:** Offers real-world urban driving data with images, LiDAR, and GPS, useful for traffic-related AI research.
- **Local Traffic Data:**
 - Surveillance video collected from urban intersections using closed-circuit cameras.
 - Focused on heterogeneous traffic environments (e.g., two-wheelers, auto-rickshaws, buses, pedestrians intermingling with vehicles), which better reflects Indian traffic conditions.

3.2 Data Preprocessing

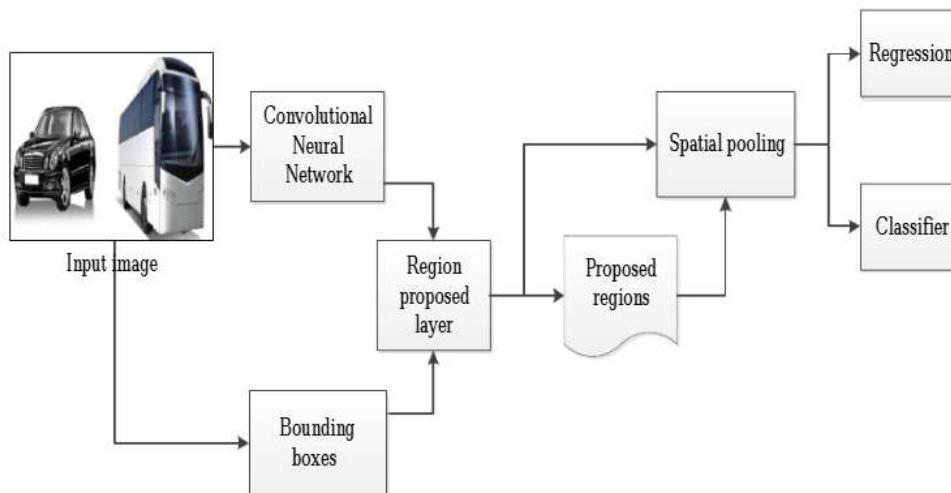
- **Resizing & Normalization:** Ensures input consistency and model stability.
- **Data Augmentation:** Random noise, Gaussian blur, and motion blur (to simulate moving vehicles) in addition to rotation, flipping, and brightness adjustments.
- **Annotation:** Bounding boxes labeled for cars, trucks, buses, two-wheelers, and non-motorized vehicles using annotation tools like LabelImg.
- **Balancing Datasets:** Oversampling underrepresented classes (e.g., buses) to avoid class imbalance.

3.3 Preprocessing

- Resizing, normalization, and augmentation.
- Annotation for bounding boxes (YOLO/Faster R-CNN).

3.4 Model Development

- **CNN (Baseline Model):**
 - The model is built with three convolutional layers using ReLU activation, along with max-pooling and fully connected layers that perform the classification.
 - Used as a benchmark for performance comparison.
- **YOLOv8 (Real-Time Detection):**
 - Grid-based detection approach.
 - Capable of detecting multiple vehicles in a single frame with bounding box regression and classification.
 - Optimized for inference speed to meet real-time constraints.
- **Faster R-CNN (High Precision):**
 - An RPN is utilized to produce candidate regions likely to contain objects.
 - Stronger accuracy for small or occluded vehicles.
 - Used in controlled scenarios where accuracy outweighs speed.



3.5 Training & Evaluation

- Loss function: Cross-Entropy.
- Optimizer: Adam/SGD.
- Metrics: Accuracy, Precision, Recall, F1-score, mAP.

3.6 Integration with Traffic Control

- Classified results → dynamic adjustment of green/red light timing.
- Prioritization for buses, emergency vehicles. Evaluation Metrics

The boxes drawn by the vehicle detection model were compared with the actual labeled boxes (ground-truth data) to find out which model gave the most accurate results. The findings from this step are important because they affect all the later stages in the system.

The main performance measure, called **Correct Detections**, shows how well the model can correctly identify vehicles in each frame. This is calculated by checking how much the predicted box overlaps with the real box. The overlap is measured using something called the **Jaccard Index** or **Intersection over Union (IoU)**, which is given by the formula:

$$J(G, P) = \frac{|G \cap P|}{|G \cup P|}$$

Another important measure is **Computation Time**, which adds up how long it takes to process all the video frames. This helps decide which detector is best suited for different uses — for example, whether it's needed for real-time traffic monitoring, field analysis, or offline research.

Accuracy was also checked using this formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where:

- **TP (True Positive):** Cases correctly detected as part of a class.
- **TN (True Negative):** Cases correctly detected as not part of a class.
- **FP (False Positive):** Cases wrongly detected as part of a class.
- **FN (False Negative):** Cases that belong to a class but were missed.

All the tests were carried out on a computer running Ubuntu 18.04, with an AMD Ryzen Threadripper 1920x processor (12 cores), an NVIDIA Titan RTX graphics card (24 GB), and 64 GB of DDR4 RAM.

4. Expected Outcomes

- **Performance Goals:**
 - Vehicle classification accuracy of $\geq 90\%$.
 - Real-time detection at ≥ 30 FPS on surveillance video feeds.
 - **Traffic Impact:**
 - Reduced average waiting time by 15–25% at intersections.
 - Reduction in fuel wastage and CO₂ emissions through optimized signal timing.
 - Enhanced public transport prioritization by dynamically recognizing buses and allocating faster signal clearance.
 - **System Contribution:**
 - Prototype Intelligent Traffic Management System (ITMS) integrating AI-based classification with adaptive signals.
 - Scalable framework applicable to smart city infrastructure and future autonomous vehicle ecosystems.
-

5. Applications

- **Traffic Decongestion:** Smarter signal timing.
 - **Urban Planning:** Vehicle distribution insights.
 - **Law Enforcement:** Detecting violations.
 - **Environment:** Lower emissions due to reduced idling.
-

6. Challenges & Future Scope

Challenges:

- Occlusion of vehicles in dense traffic.
- Limited processing power in edge devices.
- Varying lighting and weather conditions.

Future Scope:

- Federated learning for privacy-preserving data sharing.
 - Reinforcement learning for adaptive traffic signal control.
 - Integration with 5G-enabled IoT systems for smart cities.
-

7. Conclusion

Deep learning provides advanced capabilities for **vehicle image classification** and **real-time traffic management**. By incorporating models such as **Convolutional Neural Networks (CNNs)** and **YOLO** into intelligent traffic systems, cities can effectively minimize congestion, enhance road safety, and promote sustainable urban transportation.

An initial comparison of various object detection algorithms, conducted on the **GRAM Road Traffic Monitoring** video dataset, demonstrated that **YOLOv3** is well-suited for real-time vehicle detection, achieving an accuracy rate exceeding **80%**.

For traffic state categorization, two methods were formulated and evaluated using the TrafficDB video dataset. The first method leveraged conventional computer vision methods plus machine learning classifiers, and the second used deep learning models that could automatically learn features while training on labeled data. In the traditional setup, the **Random Forest** classifier achieved an accuracy of **84%**, whereas the deep learning-based model attained an accuracy above **98%** under the same experimental conditions—representing a significant **14% improvement** in performance.

References

- Kanungo A., Sharma A., Singla C. Smart traffic lights switching and traffic density calculation using video processing; Proceedings of the 2014 Recent Advances in Engineering and Computational Sciences (RAECS); Chandigarh, India. 6–8 March 2014; pp. 1–6. [[Google Scholar](#)]

- Perkasa O., Widyantoro D.H. Video-based system development for automatic traffic monitoring; Proceedings of the International Conference on Electrical Engineering and Computer Science (ICEECS); Kuta, Indonesia. 24–25 November 2014; pp. 240–244. [[Google Scholar](#)]
- .Eamthanakul B., Ketcham M., Chumuang N. The traffic congestion investigating system by image processing from cctv camera; Proceedings of the International Conference on Digital Arts, Media and Technology (ICDAMT); Kuta, Indonesia. 24–25 November 2017; pp. 240–245. [[Google Scholar](#)]
- Xun F., Yang X., Xie Y., Wang L. Congestion detection of urban intersections based on surveillance video; Proceedings of the 18th International Symposium on Communications and Information Technologies (ISCIT); Bangkok, Thailand. 26–28 September 2018; pp. 495–498. [[Google Scholar](#)]
- Ke X., Shi L., Guo W., Chen D. Multi-dimensional traffic congestion detection based on fusion of visual features and convolutional neural network. *IEEE Trans. Intell. Transp. Syst.* 2019;20:2157–2170. doi: 10.1109/TITS.2018.2864612. [[DOI](#)] [[Google Scholar](#)]
- .Li W., Dai H.Y. Real-time road congestion detection based on image texture analysis. *Procedia Eng.* 2016;137:196–201. [[Google Scholar](#)]
- Haralick R.M., Shanmugnm K., Dinstein I. Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.* 1973;3:610–621. doi: 10.1109/TSMC.1973.4309314. [[DOI](#)] [[Google Scholar](#)]
- Xu Y. Crowd density estimation using texture analysis and learning; Proceedings of the IEEE International Conference on Robotics and Biomimetics; Kunming, China. 17–20 December 2006; pp. 214–219. [[Google Scholar](#)]
- Lozano A., Manfredi G., Nieddu L. An algorithm for the recognition of levels of congestion in road traffic problems. *Math. Comput. Simul.* 2009;79:1926–1934. doi: 10.1016/j.matcom.2007.06.008. [[DOI](#)] [[Google Scholar](#)]
- Wang P., Li L., Jin Y., Wang G. Detection of unwanted traffic congestion based on existing surveillance system using in freeway via a CNN-architecture trafficnet; Proceedings of the 13th IEEE Conference on Industrial Electronics and Applications (ICIEA); Wuhan, China. 31 May–2 June 2018; pp. 1134–1139. [[Google Scholar](#)]
- He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Las Vegas, NV, USA. 27–30 June 2016; pp. 770–778. [[Google Scholar](#)]
- Wang P., Hao W., Sun Z., Wang S., Tan E., Li L., Jin Y. Regional detection of traffic congestion using in a large-scale surveillance system via deep residual TrafficNet. *IEEE Access.* 2018;6:68910–68919. doi: 10.1109/ACCESS.2018.2879809. [[DOI](#)] [[Google Scholar](#)]
- Kurniawan J., Syahra S.G., Dewa C.K. Traffic Congestion Detection: Learning from CCTV Monitoring Images using Convolutional Neural Network. *Procedia Comput. Sci.* 2018;144:291–297. doi: 10.1016/j.procs.2018.10.530. [[DOI](#)] [[Google Scholar](#)]
- Guerrero-Gomez-Olmedo R., Lopez-Sastre R.J., Maldonado-Bascon S., Fernandez-Caballero A. Vehicle tracking by simultaneous detection and viewpoint estimation; Proceedings of the IWINAC 2013, Part II, LNCS 7931; Mallorca, Spain. 10–14 June 2013; pp. 306–316. [[Google Scholar](#)]