Urban Traffic Management: Detect Vehicles Categories using Modified YOLO Model

Vandana Machhi [1], Yuvraj Chauhan [2], Jenil Gohil [3], Khushali Pandya [4], Dhaval Nimavat [5]

[1,2,3,4] Students, [5] Professor,
Department of Computer Science & Engineering, Parul Institute of Engineering & Technology, Vadodara, India.

Abstract
The significant rise in congestion on traffic lanes is one of the most significant challenges that slows the growth of a metropolitan metropolis. The reason for this is that there are a growing number of cars on the roadways, which is causing significant delays at junctions where traffic is concentrated. Throughout the years, a variety of strategies and approaches have been developed to find a solution to this issue and to make the systems that manage traffic more dynamic. Static traffic control systems relied on predetermined timings that were assigned to each traffic lane and could not be modified in any way. These timings were used to govern the flow of traffic. In addition, there was not a system for the counting and detection of vehicle with different categories, nor was there a facility for the identification of emergency vehicles while they were moving through traffic. In this research article, we will review different machine learning and deep learning models for the detection of vehicles. We will assess the viability of these models in terms of cost, dependability, accuracy, and efficiency, and we will also add some new features to improve the overall performance of the current detection system.

Keywords: Multi-categories Vehicle, Traffic Management, Transfer Learning, You Only Look Once (YOLO), Convolutional Neural Network

1. Introduction

Having a reliable public transportation system is crucial to the growth of every major city. Fig. 1 is a simplified schematic depicting the layout of a typical traffic light junction. Congestion, however, is a serious issue as well. Multiple factors have contributed to the recent spike in traffic. The growing number of cars on the road is a major contributor to the problem, as is the expanding human population. Ineffective capacity management, inadequate infrastructure, bad road conditions, etc. are other contributing factors to traffic jams.

It takes an increasing amount of time for people to wait at traffic lights. Traffic signals are often monitored by humans or mechanical timers. Having a traffic officer manually direct traffic is a waste of manpower. However, manual timers are distributed according to typical wait times. For example, a set timeframe of 30 seconds will be allocated regardless of whether the traffic density is low or high. This causes commuters to lose a lot of time and money on wasted travel. An rise in collisions is a direct result of drivers' growing impatience at the lights.

What a typical traffic management system in a big city looks like. After looking at the current methods available, researchers have concluded that wireless traffic control systems, which include object detection algorithms and adaptive traffic signals, are on the rise because of their ability to alleviate congestion, shorten average wait times, and give priority to emergency vehicles.

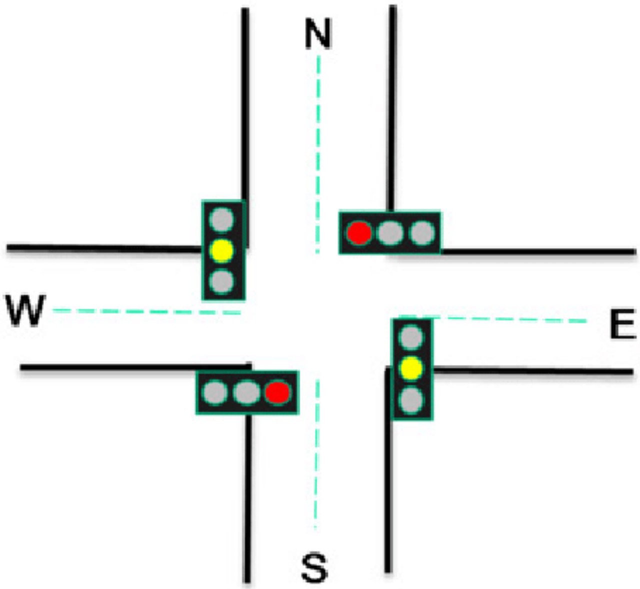

Fig. 1. 4-way Traffic Lights

While improvements have been made in road travel, several issues remain unaddressed. In the case of zebra crossings, no mechanism exists for detecting and counting pedestrians. In addition, there is no solution proposed for identifying and tallying the number of people and cars who cross paths. Vehicle identification and counting in varying environmental circumstances is likewise a significant undertaking.

2. Related Works

In [1] created a video-based traffic volume count and junction turning pattern system. This framework uses a CNN model and object tracking algorithm to recognise and track cars in the camera's pixel view. Traffic counts and turn calculations are possible after projecting vehicle spatial-temporal data onto an orthogonal real-scale map. Traffic volume counting approaches have 96.91 percent accuracy. Finally, the lack of study into weather conditions like rain may affect automotive aesthetics. Zebra crossings cannot count or identify pedestrians.

In [2] developed an efficient, real-time, density-based traffic light management system. This research uses an artificial neural network (ANN) model and image processing to predict outcomes using real-time data. Principal component analysis trained a neural network model by picking the best features and minimising their dimensionality. Finally, this study's limitation is weather consideration. Rain and fog may degrade images. To assess its efficacy in a multi-intersection model, the Neural Network (NN) technique may be expanded to a multi-agent network.

In [3] propose a deep reinforcement learning model to regulate traffic signal cycles to reduce vehicle wait times and fuel waste. Sensor networks show junction traffic data as state-level grids of small squares. Markov decision mapping maps traffic signal duration changes afterwards. A convolutional neural network maps rewards to these states afterwards (CNN). To widen the technique's behaviour, the Double Q-Network (DQN), target network, and prioritised experience replay form a dualling network. SUMO, a traffic light cycle management simulator, tests the model. This strategy may be used to complex road junctions with heavy traffic.

In [4] tested YOLOv5 model. They used the YOLOv5 model on their dataset for Traffic Sign Identification (TSR), a deep learning visual object identification task. YOLOv5 outperforms Single Shot multibox Detector (SSD) in identifying speed. This project used the NZ-TSR YOLOv5. Input, backbone, neck, and prediction layers comprise the model. They investigated 2,182 traffic sign pictures from a custom dataset with 8 classes. Google Colab's tremendous computing capacity enabled the experiment. YOLOv5 has a 97.70% accuracy advantage over SSD. The technique might incorporate all traffic sign categories in the datasets. The authors will also develop Mask R-CNN, CapsNet, and Siamese neural network visual object detection models.

Using YOLO and SORT object detection,in [5] created a system to avoid congested roads. The proposed model would recognise and track automobiles on the video stream and count them as they cross a line to provide a real-time street view. After monitoring automobiles on a video stream, the recommended approach counts how many cross a line. They used YOLO and SORT for frame-by-frame object recognition and tracking. The model recognises current traffic patterns. The proposed system develops a self-adaptive traffic control algorithm. The adaptive traffic system will include pedestrian input to reduce wait times. Another option is to add sensors to the controller to improve system resilience and flexibility.

In [6] suggested the Intelligent Transportation System (ITS) for real-time road infrastructure study and better traffic management. They predicted traffic conditions using traffic forecasting. Vehicles were detected using a road traffic forecast public dataset. They employed 5 regression models: Random Forest, Stochastic gradient descendant, Multilayer perceptron, Gradient boosting, and Linear. Multilayer perceptron regressors train faster (18 sec). This strategy will expand machine learning and deep learning.

In [7] introduced a Reinforcement Learning-based cyclic phase model for traffic management. The model will address real-world restrictions. Biased Pressure Coordination was developed by the authors (BP). This strategy considers phase pressure and approaching or waiting cars. The suggested technique outperforms cycle-based methods on synthetic and real-world data. This method will test its validity using genuine traffic signals in the future. Examine the reality gap between actual and simulated limitations.

In [8] developed and constructed an ATM based on IOT and ML. Infrastructure, cars, and events are its key elements. The proposed approach detects anomalies using DBSCAN clustering. Traffic and neighbouring crossing movements update the ATM model. The recommended model was implemented using MATLAB. Several linked autonomous vehicle (LAV) traffic scenarios were created to test the ATM model. The proposed system will include security and energy-saving features. Real-time traffic flow data might verify the model instead of simulator data.

In [9] proposed a hierarchical decision-making system for traffic signal phases. Parameterized Deep Q-Networks were designed for this (P-DQN). Their hybrid Deep Reinforcement Learning system can make discrete and continuous decisions. The proposed design reduces wait time by 22.2% and vehicle travel time by 5.78%. DRL designs include Deep Q-Networks (DQN), Dueling-DQN, and Double-DQN. The approach might control phase selection and length. This system will cover several junctions centrally and decentrally. Their future models would also incorporate real-time crossing data.

In [10] developed a WSN and visual analytics framework traffic control system. They examined average network performance, energy usage, and latency. Assessing the proposed system uses network lifetime, energy usage, access ratio, delivery ratio, and communication cost. The study's technique is inexpensive, unstable, and easy to implement. Future research should include wait length, occupancy, and traffic categorization. Double parking and roadside traffic are also difficulties.

In [11] propose image processing to measure traffic lane density. Image processing will give the lane with the most traffic a green light and the most time. The strategy uses four gathered and four reference images of four lanes. Reference pictures are unoccupied roads, while road photos are obtained. Edge detection and picture enhancement follow RGB-to-grayscale conversion. Traffic density and picture matching follow comparison. Timer assignment follows traffic density calculation. Traffic crossings with more than four lanes may use the proposed technology.

In [12] offer a computer vision-based traffic light controller concept. Live photos will be used to calculate traffic density and determine green signal timing. They employed YOLO for vehicle detection. The signal switching method refreshes the green and red signal timers using density. One CNN predicts multiple bounding boxes and class probabilities using YOLO. Vehicle detecting module testing yielded 75–80%

accuracy. The idea might be developed to identify traffic violators, detect accidents, adjust to emergency vehicles, and synchronise traffic signals at numerous crossings.

In [13] propose an image processing-based adaptive traffic signal control system that monitors highway density. Comparative edge detectors included Canny, Sobel, Log, and Roberts. Four-lane images are taken using high-quality 1080px cameras. MATLAB then processes these images. After converting RGB photos to grayscale, a dilation procedure will fill discontinuous edge parts. Finally, reference and test picture characteristics will be compared and the percentage of image matching computed. This effort will employ edge detectors other than Canny. This method might also be used on a four-lane traffic intersection.

In [14] propose a study to assess Computational Intelligence (CI)-based simulation approaches for enhancing Traffic Signal Timing (TST) and Traffic Signal Control (TSC) systems and give insights and suggestions for future research. This investigation will count and identify pedestrians. Driver behaviour impacts on the model should also be examined.

In [15] presented a wirelessly controlled traffic control system (TCS). Wireless traffic signals are easier to install. Wireless, this technology might be used at any junction. This system will incorporate lanes-flow sensors in the future. Another possibility is to create a green wave coordination system that adapts to traffic signal cycles.

3. Proposed Workflow

The technique is sketched out below as a visual flowchart for ease of understanding. It's safe to say that this setup will be more effective for studies in the future. The diagram's primary focus is on two detection processes: pedestrian detection and vehicle detection. There's an additional category inside vehicles, too.

Fig. 2. Workflow of the proposed framework

The interim framework's process is shown in Fig. 2.India is one of the countries where more advanced methods of traffic monitoring and regulation have been considered for some time. Congestion is an unavoidable byproduct of urban intersections when traffic management is inadequately efficient and flexible. By modifying a pre-trained item identification model using a transfer learning approach, we provide a framework for effectively dealing with this influx of data. As input, the framework takes a monitoring video and dissects it frame by frame. During the preliminary processing of these images, all identifying characteristics, such as faces and licence plates, are eliminated. The next stage is to feed each video frame into a convolutional neural network (CNN) model for vehicle recognition; one such model is YOLOv5 (you-only-look-once version 5). The YOLO convolutional neural network (CNN) model can extract bounding boxes that completely encompass the objects in a digital image.

Table 1. YOLO Version Comparison

Layers	**YOLOv3**	**YOLOv4**	**YOLOv5**
Neural Network Type	FCNN	FCNN	FCNN
Backbone	Darknet53	CSPDarknet53 (CSPNet l/' Darknet)	CSPDarkent53 Focus structure
Neck	FPN (Feature Pyramid Network)	$p (Spatial Pyramid Pooling) and PANet (Path Aggregation Network)	PANet
Head	B x (5 + C) output layer B, No, of bounding boxes C: Class score	Same as YOIO v3	Same as Yolo v3

Loss Function	Binary Cross Entropy	Binary Cross Entropy	Binary Cross Entropy and Logit Loss Function

The newest version of YOLO, v7, introduces many architectural changes that boost its speed and precision. As with Scaled YOLOv4, YOLOv7 does not use ImageNet pre-trained backbones in its neural networks. Instead, just the COCO dataset is used to train the models. You may anticipate a resemblance between YOLOv7 and Scaled YOLOv4, an expansion of YOLOv4, as they are created by the same people. Some of the most notable changes in YOLOv7 are listed here. We'll examine each one separately.

- **Modernization of Buildings**

Model Scaling for Concatenation-based E-ELAN (Extended Efficient Layer Aggregation Network).

- **Educatable BoF (Bag of Freebies)**

Convolution with new parameters is in the works. Auxiliary Coarse, Lead Loss Fine.

In the YOLOv7 backbone, the computing block known as the E-ELAN may be found. It does this by drawing on earlier research that was done on the effectiveness of networks. The following parameters, which have an influence on speed and accuracy, were taken into consideration while designing it.

- Memory access cost
- I/O channel ratio
- Element wise operation
- Activations
- Gradient path

In simple terms, E-ELAN architecture enables the framework to learn better. It is based on the ELAN computational block.

After the images are processed, the subjects will be sorted into human and vehicle categories. Once the items have been sorted, the detection for the ones that overlap will be improved. The shapes of things that collide with one another have their contours overlap. Refining the objects' outlines and treating them as separate entities is important for accurate detection. The items will have their contours refined, and then they will be tallied so that a density map may be created from the data. When you look at a density map, you can see just how many and how loudly a lot of cars are driving along a certain stretch of road. The traffic management decisions will be based on the number of vehicles in each lane. This means that the busiest lane will get the longest allotted time.

The last step involves contrasting the provisional model with others using measures including accuracy, sensitivity, and specificity. Model training for the proposed framework will be place on Google Collab, which offers excellent mid- to high-range GPU performance. The project will be developed on the Spyder Anaconda distribution, since it provides quick access to hundreds of Python programmes and a thriving online community. TensorFlow, Keras, OpenCV, Scikit-learn, Matplotlib, and NumPy are some of the useful python libraries that may be used throughout the coding process. The traffic management system may benefit from the proposed framework's ability to adjust to real-world data.

Pseudo Code

```
Initialize CCTV.
Initialize CAM;
for
    Yolo Detector Is Initialized
    Tracker Window Annotation
    If
        Total Vehicle > V
    Then
```

Max V then on green light
Verify
Max V then on yellow light
Else
Min V then on red light
End
End

4. Results Analysis

The dataset is taken from Dhaka AI Traffic Challenge Weights – Yolov7. The considered vehicle classes are: ambulance, auto-rickshaw, bicycle, bus, car, garbage van, human hauler, minibus, minivan, motorbike, Pickup, army vehicle, police car, rickshaw, scooter, Suv, taxi, three-wheelers (CNG), truck, van, wheelbarrow.

Fig. 3. Dataset Frames

Fig. 4. Vehicle Annotation

Fig. 5. Multi-Categorical Detection

Conclusion

The development of a traffic signal management system that is flexible, real-time, and density-based is the primary objective of this line of study. This has given us a basic concept for designing a system that can manage the flow of traffic at crossings. A tracking algorithm and an object identification model, which we will refer to as YOLOv7, will be used in order to monitor things like autos and people that appear in video frames. The suggested framework will be more practical, adaptive, and cost-effective than the traditional methods of traffic monitoring, which rely on GPS data, in-situ human observation, and traffic sensors. In addition, the proposed framework will be more practicable. To be more specific, it might be hypothesised that the system under consideration would identify and count the movement of automobiles and other items on the traffic lanes in an effective manner. Detecting and counting items that intersect with one another would also be conceivable. In the long run, the provisional framework will be useful in lowering the amount of time spent waiting and increasing the amount of time spent travelling by cars.

References

[1] Pi, Yalong, et al., "Visual Recognition for Urban Traffic Data Retrieval and Analysis in Major Events Using Convolutional Neural Networks." Computational Urban Science, vol. 2, no. 1, 2022, https://doi.org/10.1007/s43762-021-00031-0

[2] Chandrasekhara, W.A.C.J.K., et al., "A Real-time Density-based Traffic Signal Control System". 5th International Conference on Information Technology Research (ICITR), pp. 1-6, 2020, https://doi.org/10.1109/ICITR51448.2020.9310906

[3] Sahu, Satya Prakash, et al., "Traffic Light Cycle Control Using Deep Reinforcement Technique." Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021, 2021, pp. 697–702, https://doi.org/10.1109/ICAIS50930.2021.9395880

[4] Zhu, Yanzhao, et al., "Traffic Sign Recognition Based on Deep Learning." Multimedia Tools and Applications, vol. 81, no. 13, 2022, pp. 17779–91, https://doi.org/10.1007/s11042-022-12163-0

[5] Sharma, Moolchand, et al., "Intelligent Traffic Light Control System Based on Traffic Environment Using Deep Learning." IOP Conference Series: Materials Science and Engineering, vol. 1022, no. 1, 2021, https://doi.org/10.1088/1757-899X/1022/1/012122

[6] Navarro-Espinoza, Alfonso, et al., "Traffic Flow Prediction for Smart Traffic Lights Using Machine Learning Algorithms." Technologies, vol. 10, no. 1, 2022, p. 5, https://doi.org/10.3390/technologies10010005

[7] Ibrokhimov, Bunyodbek, et al., "Biased Pressure: Cyclic Reinforcement Learning Model for Intelligent Traffic Signal Control." Sensors, vol. 22, no. 7, 2022, https://doi.org/10.3390/s22072818.

[8] Lilhore, Umesh Kumar, et al., "Design and Implementation of an ML and IoT Based Adaptive Traffic-Management System for Smart Cities." Sensors, vol. 22, no. 8, 2022, https://doi.org/10.3390/s22082908

[9] Bouktif, Salah, et al., "Traffic Signal Control Using Hybrid Action Space Deep Reinforcement Learning." Sensors, vol. 21, no. 7, 2021, pp. 1–15, https://doi.org/10.3390/s21072302

[10] Naveed, Quadri Noorulhasan, et al., "An intelligent traffic surveillance system using integrated wireless sensor network and improved phase timing optimization." Sensors, vol. 22, no. 9, 2022, https://doi.org/10.3390/s22093333

[11] Ijeri, Dakshayani, et al., "Traffic Control System Using Image Processing." Proceedings of B-HTC 2020 - 1st IEEE Bangalore Humanitarian Technology Conference, 2020, https://doi.org/10.1109/B-HTC50970.2020.9298014

[12] Gandhi, Mihir M., et al., "Smart Control of Traffic Light Using Artificial Intelligence." 2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering, ICRAIE 2020 - Proceeding, vol. 2020, 2020, https://doi.org/10.1109/ICRAIE51050.2020.9358334

[13] Meng, Belinda Chong Chiew, et al., "Smart Traffic Light Control System Using Image Processing." IOP Conference Series: Materials Science and Engineering, vol. 1088, no. 1, 2021, p. 012021, https://doi.org/10.1088/1757-899x/1088/1/012021

[14] Qadri, Syed Shah Sultan Mohiuddin, et al., "State-of-Art Review of Traffic Signal Control Methods: Challenges and Opportunities." European Transport Research Review, vol. 12, no. 1, 2020, pp. 1–23, https://doi.org/10.1186/s12544-020-00439-1

[15] De Oliveira, Luiz Fernando Pinto, et al., “Development of a Smart Traffic Light Control System with Real-Time Monitoring.” IEEE Internet of Things Journal, vol. 8, no. 5, 2021, pp. 3384–93, https://doi.org/10.1109/JIOT.2020.3022392

[16] D. Vyas, et al., "Visual Social Distance Alert System Using Computer Vision & Deep Learning," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1512-1516, https://doi.org/10.1109/ICECA49313.2020.9297510

[17] D. Vyas, et al., "Yolo-v4 Deep Learning Model for Medical Face Mask Detection," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 209-213, https://doi.org/10.1109/ICAIS50930.2021.9395857