



Intelligent Traffic Management System using CNN-YOLOv7 with Integrated MQTT Protocol

Mann Manohar, Sai Chandra Mouli, Mithilesh Walde

Indian Institute Of Information Technology, Nagpur
Computer Vision & IoT Project

Abstract

Urban traffic congestion is a growing problem, leading to increased travel times, fuel consumption, and pollution. Traditional traffic management systems often rely on predefined schedules or simple loop detectors, lacking adaptability to real-time conditions. This poster presents an Intelligent Traffic Management System (ITMS) leveraging state-of-the-art computer vision techniques, specifically Convolutional Neural Networks (CNNs) and the YOLOv7 object detection model. The system analyzes real-time video feeds to detect and classify vehicles, pedestrians, and potentially other relevant objects. This information is used to estimate traffic density and flow. The Message Queuing Telemetry Transport (MQTT) protocol is integrated for efficient, lightweight communication between sensing nodes (cameras with processing units) and a central control system or other infrastructure components, enabling dynamic traffic signal adjustment and data dissemination.

Introduction

Effective traffic management is crucial for modern cities. Conventional systems struggle with dynamic traffic variations, leading to inefficiencies. Computer vision offers a powerful alternative by enabling direct observation and analysis of traffic scenes.

- Challenge:** Static signal timings and limited sensing capabilities of traditional systems.
- Solution Approach:** Utilize Deep Learning, specifically CNNs and YOLOv7, for accurate, real-time object detection from traffic cameras.
- Communication:** Employ MQTT for scalable and low-latency data exchange in a distributed system architecture.
- Goal:** Develop an adaptive ITMS capable of optimizing traffic flow, enhancing safety, and providing real-time traffic insights.

This system aims to replace or augment existing infrastructure with intelligent, vision-based monitoring and control.

Core Technologies

1. Convolutional Neural Networks (CNNs)

CNNs are a class of deep neural networks highly effective for image analysis tasks. They use convolutional layers to automatically learn hierarchical spatial features from input images, making them ideal for object recognition.

- Architecture: Composed of convolutional, pooling, and fully connected layers.
- Function: Feature extraction (edges, textures, shapes, objects).
- Application: Forms the basis for many object detection models, including YOLO.

2. YOLOv7 (You Only Look Once)

YOLOv7 is a state-of-the-art, real-time object detection algorithm. It processes entire images at once, predicting bounding boxes and class probabilities simultaneously.

- Type: Single-stage detector.
- Strengths: High speed (suitable for real-time video) and competitive accuracy. Architecture includes optimizations for efficiency and performance.
- Output: Bounding box coordinates (x, y, width, height), confidence score, and class label (e.g., car, truck, pedestrian, traffic light) for each detected object.
- Use Case: Detecting and classifying vehicles/pedestrians in traffic video streams.

3. MQTT (Message Queuing Telemetry Transport)

MQTT is a lightweight, publish/subscribe messaging protocol designed for constrained devices and low-bandwidth, high-latency networks. Ideal for IoT and distributed systems.

- Paradigm: Publish/Subscribe. Decouples message producers (publishers) from consumers (subscribers).
- Broker: A central server manages message distribution based on topics.
- Benefits: Low network overhead, scalability, asynchronous communication.
- Role in ITMS: Transmitting detected object counts, traffic density metrics, or control signals between cameras, processing units, and traffic controllers.

Important Point

Scan QR code for more details



System Architecture & Methodology

1. Video Acquisition: Cameras capture real-time video streams from traffic intersections or road segments.

2. Preprocessing (Optional): Frames may undergo preprocessing steps like resizing or normalization for optimal input to the detection model.

3. Object Detection (YOLOv7):

- Each frame is fed into a pre-trained or fine-tuned YOLOv7 model.
- The model outputs bounding boxes, class labels (car, bus, person, etc.), and confidence scores for all detected objects.

4. Data Extraction & Analysis:

- Information from YOLOv7's output is processed to:
 - Count vehicles/pedestrians per lane or region of interest.
 - Estimate traffic density or queue lengths.
 - Detect specific events (e.g., accidents, jaywalking - requires further logic).
- This data forms the basis for traffic management decisions.

5. MQTT Communication:

- The processing unit publishes extracted data (e.g., vehicle counts) to specific MQTT topics (e.g., 'traffic/junction1/lane2/count').
- Subscribers (e.g., traffic light controllers, central monitoring dashboard) receive relevant data from the MQTT broker.
- Example Topic Structure: 'sensor location/data type'

6. Control & Actuation / Monitoring:

- Based on subscribed MQTT data, traffic light controllers can dynamically adjust signal timings.
- Data can be visualized on dashboards for monitoring purposes.
- Alerts can be generated for incidents.

YOLOv7 Model Details:

- Typically uses models pre-trained on large datasets like COCO.
- Can be fine-tuned on specific traffic datasets for improved accuracy in target environments.
- Requires sufficient computational resources (GPU recommended for real-time performance).

Visual Demonstration & Results

The system's core capability is demonstrated through processing sample traffic video frames.

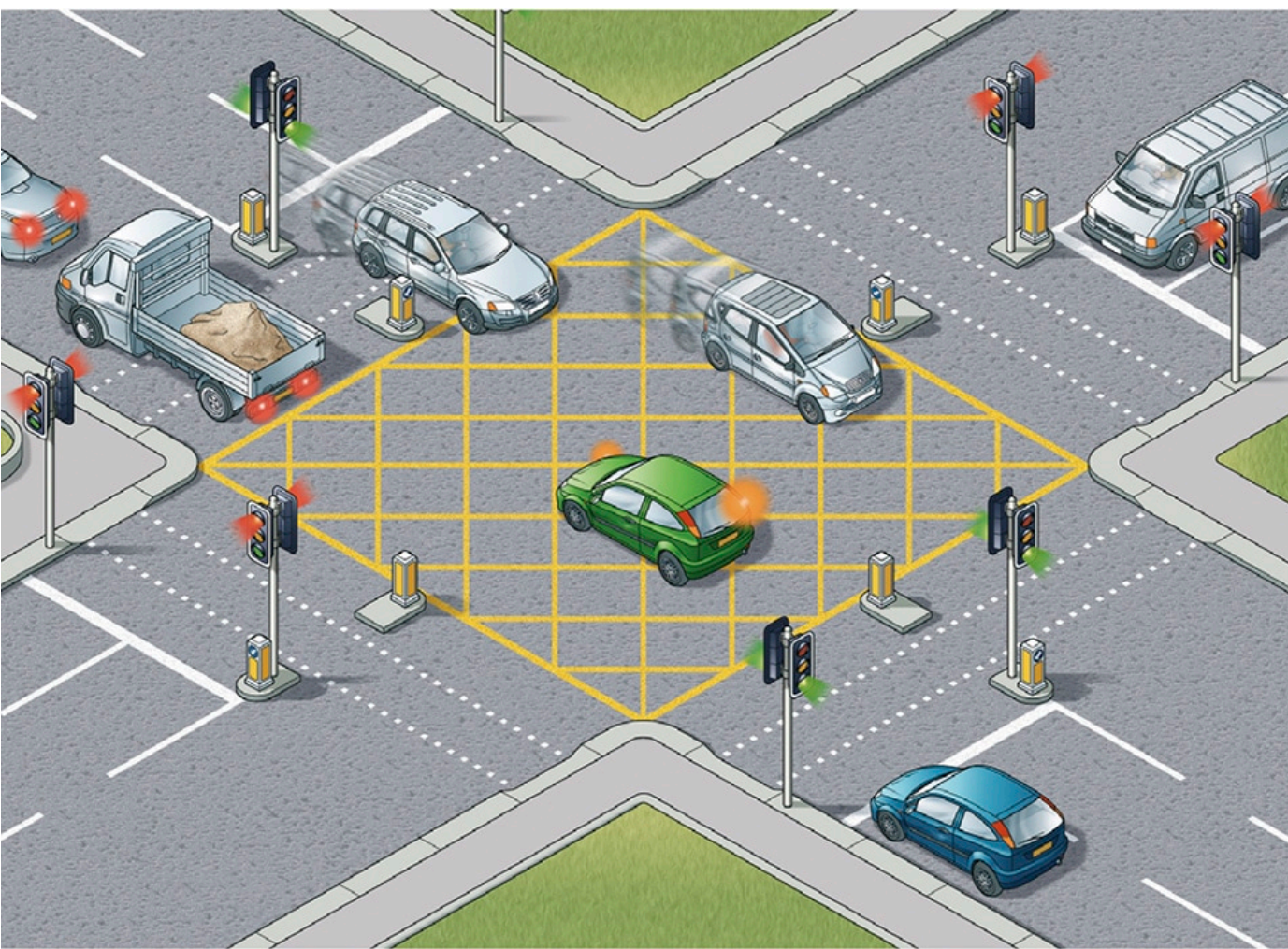


Figure 1. Original Traffic Scene Frame

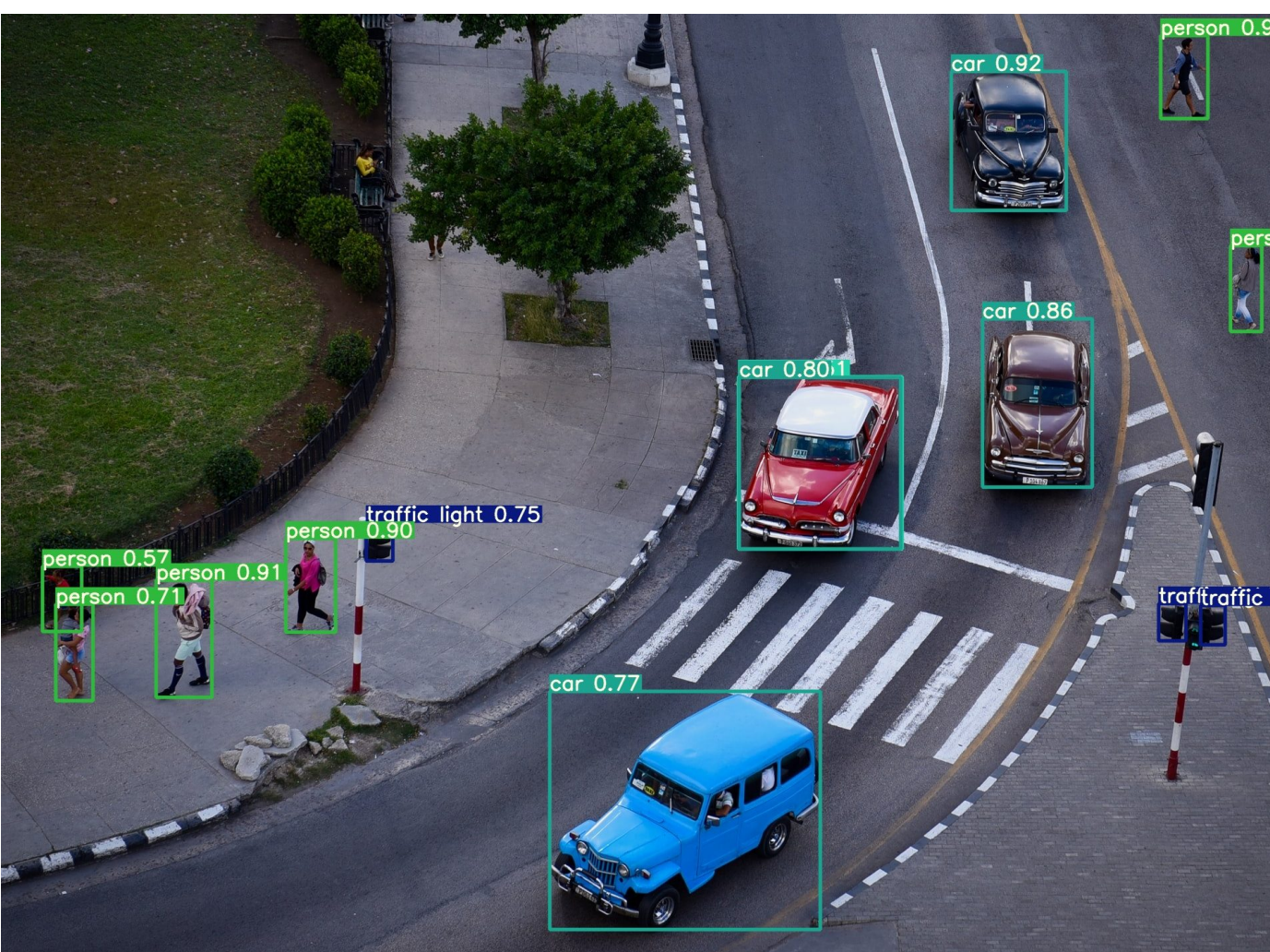


Figure 2. YOLOv7 Classification Results

Observations:

- YOLOv7 accurately detects various vehicles and pedestrians even in moderately complex scenes.
- Detection speed allows for real-time processing on suitable hardware.
- Extracted counts and locations provide valuable data for traffic analysis.
- MQTT enables flexible communication for integrating diverse system components.

Discussion

Advantages:

- Real-Time Adaptability:** Enables traffic signal control based on actual demand, reducing delays.
- Rich Data Acquisition:** Provides detailed information (vehicle types, counts, pedestrian presence) beyond simple vehicle detection.
- Accuracy & Speed:** YOLOv7 offers a strong balance for real-time traffic object detection.
- Scalability & Flexibility:** MQTT allows easy integration of multiple sensors and control units.
- Non-Intrusive Installation:** Camera-based system avoids road closures needed for installing in-ground sensors.
- Incident Detection Potential:** Can be extended to detect accidents, stalled vehicles, or traffic violations.

Limitations & Challenges:

- Computational Cost:** Real-time YOLOv7 processing often requires dedicated GPUs, increasing deployment cost.
- Environmental Factors:** Performance can be affected by adverse weather (rain, fog, snow), poor lighting (night, glare), and occlusions (large vehicles blocking smaller ones).
- Camera Requirements:** Needs well-positioned cameras with sufficient resolution and clear lines of sight.
- Training Data Dependency:** Model performance relies heavily on the quality and diversity of the training dataset.
- MQTT Broker Reliability:** System availability depends on the MQTT broker's uptime and performance.
- Privacy Concerns:** Use of cameras for monitoring requires addressing potential privacy issues.

Future Work & Enhancements:

- Integrate multi-camera views for wider coverage and handling occlusions.
- Incorporate tracking algorithms (e.g., DeepSORT) for trajectory analysis and flow estimation.
- Optimize models (e.g., using TensorRT, pruning) for deployment on edge devices.
- Develop more sophisticated traffic prediction and control algorithms.
- Fuse vision data with other sensor types (e.g., Radar, LiDAR) for enhanced robustness.
- Implement security measures for MQTT communication.

Conclusions

The proposed Intelligent Traffic Management System successfully demonstrates the potential of combining advanced computer vision (CNN/YOLOv7) with efficient IoT communication (MQTT).

- YOLOv7 provides accurate and fast detection of relevant traffic objects.
- CNNs form the foundation for powerful visual understanding.
- MQTT enables scalable and decoupled communication within the system.

This approach offers significant advantages over traditional methods by providing richer data and enabling real-time adaptive control strategies. While challenges related to environmental factors and computational requirements exist, the system presents a promising direction for improving urban mobility, reducing congestion, and enhancing road safety. Further development and optimization can lead to robust and widely deployable intelligent transportation solutions.

References

- Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*. <https://arxiv.org/abs/2207.02696>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- MQTT Standard v5.0. (2019). *OASIS Standard*. <https://docs.oasis-open.org/mqtt/mqtt/v5.0/os/mqtt-v5.0-os.html>
- Lin, T. Y., et al. (2014). *Microsoft COCO: Common Objects in Context*. <https://arxiv.org/abs/1405.0312>
- Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). *Edge computing: Vision and challenges*. *IEEE Internet of Things Journal*, 3(5), 637-646. <https://doi.org/10.1109/JIOT.2016.2579198>
- Chen, Z., Ellis, T., & Velastin, S. A. (2012). *Vehicle detection, tracking and classification in urban traffic*. In *15th International IEEE Conference on Intelligent Transportation Systems* (pp. 951-956). <https://doi.org/10.1109/ITSC.2012.6338629>