

Real-Time Traffic Monitoring and Visualization

Using Stationary Roadside Sensor

in a Road Intersection

Introduction

Congestion appears to be an uncontrollable problem, most likely to occur in densely populated areas with high vehicle ownership, leaving the road insufficient to accommodate all possible trips and causing a slowdown in the flow of traffic (Metz, 2018). Traffic congestion according to (Afrin & Yodo, 2020), is one of the main problems in transportation, it affects the sustainability of transport development by causing delay, disruption, and financial loss for drivers, in addition to the increased carbon emissions. For decision-makers to launch mitigation initiatives to increase the overall sustainability of the transportation system, traffic congestion must be identified and quantified. (Afrin & Yodo, 2020). That is why the traffic monitoring system is one of the critical transportation infrastructures in the Intelligent Transportation System (ITS). With this, traffic authorities can invest significant number of resources to gather and analyze traffic data to effectively use the roadway systems, increase traffic safety, and develop future transportation plans (Won, 2020).

Traffic flow prediction has gained more attention with the rapid development and deployment of the Intelligent Transportation System. This aims to provide traffic flow information that the users need to make better travel decisions, alleviate traffic congestion, reduce carbon emissions, and improve traffic operation efficiency (Paul et al., 2017). According to (Laranjeira, 2020), traffic forecasting is the process of predicting real-time traffic information that will help drivers avoid traffic congestion, guide them through it as quickly as possible, and avoid congestion by using predictive traffic forecasting services for precise times of arrival (ETAs) and optimized routes while a driver is on the road. There are several real-time traffic monitoring systems built to improve traffic infrastructures presented in (Wan et al., 2022) and (C. J. Lin & Jhang, 2022) using Convolutional Neural Networks. These existing systems show real-time traffic scenarios which can be big for data analysis and difficult to distinguish specific problems and information. Therefore, there is a need to visualize traffic data which is captured real-time, and to create visualizations to convey traffic insight graphically so that travelers can make a quick decision whether they continue on their path or choose to reroute (Unwin, 2020).

In this study, Artificial Intelligence based methods will be used to quantify traffic parameters such as vehicles count, type and speed.

Further, YOLOv5 will be used to detect and classify vehicles and DeepSort Algorithm use to track the vehicles. (Nurcahyo & Iqbal, 2022). Along with these methods, we have to use IoT sensors that would help collect and exchange data between sensors and the server for visualization (Kaur, 2019). With the emergence of Wireless Sensor Networks, it will also contribute to a traffic monitoring system that has the capability of sensing the physical environment, this device is equipped with a sensor for traffic detection, a low-power radio, and packaging for in-pavement mounting, either in the middle of a traffic lane or at the roadside, according to (Pascale et al., 2012) this WSN in terms of installation, maintenance, precision, and credibility, will be cost-effective.

Moreover, the detection and classification model and the algorithm for tracking vehicle will be deployed using Jetson Nano and RasPi4 that is equipped with camera and Wireless communication device. Lastly, work will develop a real-time traffic visualization system that capable of displaying traffic conditions that is perceived by the IoT sensors. This system will be built using free and open-source web-based frameworks.

**CHAPTER II**

**REVIEW OF RELATED LITERATURE**

Monitoring and visualization are widely acknowledged in some existing vehicle researches and experiments. A brief review of real time traffic monitoring and visualization is conducted in this section.

**2.1 Object detection**

There has been an expansion in research into autonomous video analysis for tracking and recognizing things as a result of the quick growth of hardware services such as smartphones, computers, and cameras (Tong et al., 2020) Detecting a moving object is one of the critical and fundamental tasks in computer vision along with video sequence movement (Prakash et al., n.d.). Using this form of localization and identification, object detection can be used to count the items in a scene, as well as to locate and track them in real-time while precisely labeling. The object was detected first consistent with the photo input, after which tracked in subsequent frames. Coordinates of the object tracked and the overlaid bounding field had been examined and monitored (Mathur et al., 2018). (Wont et al., 2018) proposed a model called Aggnet50 that enhances the residual block of the conventional Residual Network to offer feature information of a narrow receptive field (ResNet) with the simple aggregation blocks. In relation with that, a fundamental method in the development of many video analysis applications that attempt to use a surveillance camera with a human operator is object tracking and detection in a video sequence (Pal et al., 2021) According to a review by (Zhao et al., 2019), by using 3-D sensors (such as cameras and light detection and ranging devices), it is possible to better understand 2-D images and apply image-level knowledge to the outside environment by using additional overview data. However, most of these 3-D-aware techniques aim to place correct 3-D BBs around detected objects.

When using detectors that are running on a GPU platform, VGG, ResNet, ResNeXt, or DenseNet could serve as the backbone. SqueezeNet, MobileNet, or ShuffleNet may serve as the backbone for detectors powered by CPU platforms. The R-CNN series, which includes the fast R-CNN, faster R-CNN, R-FCN, and Libra R-CNN are the most representatives to two-stage object detection (Bochkovskiy et al., 2020). Based on the development of (Ku et al., 2022) about helmetless auto detector, Faster R-CNN cannot meet the real-time requirements because it takes about 0.2 seconds to detect the image(Gu et al., 2019) whereas YOLO-v4 is suitable for real-time detection by meeting both speed and accuracy requirements among object detection models (Bochkovskiy et al., 2020) Faster R-CNN has significant detection performance for object detection in general scenes, but generally detected objects in such scenes have problems such as occlusion, distortion, and large scale (Xiao et al., 2020) As stated in the study of (Yanagisawa et al., 2018) since Faster RCNN generates image features and region proposals using single CNN, there is an advantage that end-to-end training can be performed in addition to faster detection. However, the network configuration is still complicated, because the process of generating image features and region proposals is separated. In the study of (Palwankar & Kothari, 2022) by using SSD (Single Shot Detector), we can detect objects the fastest. It helps to identify individual objects in the image fastest.

According to (Kim et al., 2020), it has been proved that in real-time object detection, Yolov4 model outperforms, showing 93% accuracy in vehicle recognizing, while the fastest RCNN model, Faster-RCNN, has a fast SSD, but its FPS is unsatisfactory as it utilizes CNN and its accuracy is inadequate due to the usage of mobile-v1 and the model's low weight.

**2.2 Speed detection**

A significant aspect in the calibration is the speed of the vehicle, validation and air quality and traffic emission improvements models. Through the application of image and video processing techniques, moving vehicle velocity is estimated utilizing vehicle speed detection. In order to decrease the incidence of traffic accidents, speed monitoring technologies such as RF transceivers, automatic braking systems, camera-based speed detection, and electronic RFID tracking have been created. However, present accident prevention measures continue to be ineffective (Prakash et al., n.d.). Smart Vehicle Over Speeding Detector using IoT technology that used for alerting information about vehicles over speed limit. This system doesn’t require any manpower and records statistics approximately vehicle speed and wirelessly informs to over speeding detection authorities (Khan & Khan, 2018) Smart speed adaptation to efficiently monitor your operator, detecting and recording vehicle speeds exceeding the speed limit is, therefore, an essential priority. In the study of (Prakash et al., n.d.), this shows the result of the measurement with accuracy, recall, precision and F-measure as metrics. Precision is the fraction of the relevant instances among the instances retrieved, while recall is a fraction of the overall number of relevant instances retrieved.

**2.3 Vehicle Tracking**

Traditionally, tracking a vehicle is done via networks of pre-installed video cameras that rely on stationary cameras, and looks through videos for the intended vehicle. Re-identification method of a vehicle consists of two components, i.e., the image-based feature embedding module and the video-based constraints module (Tan et al., 2019) Using smartphone sensors to find, track, and further estimate the vehicle's movement patterns to take pictures of the environment in order to choose the fewest necessary participants (Chen et al., 2019) By identifying the areas where multiple photographs taken by various individual’s overlap, an event can be located(Chen et al., 2016). The tracking of vehicle depends on photos of vehicle contributed by people who happen to spot the vehicle (Chen et al., 2019). Almost all vehicle identification systems have two fundamental stages: (1) development of hypotheses, which involves speculating about where prospective cars might be in the image, and (2) hypothesis verification (HV), which involves checking the hypotheses (Lin et al., 2021).

According to the study of (Neupane et al., 2022), the architecture of the YOLOv5 models was built using a number of elements from already-existing networks to ensure that space-to-depth conversion time would be decreased, the vanishing gradient problem would be alleviated, feature propagation would be strengthened, the number of parameters would be reduced by reusing existing features, and object identification on unseen data would be improved by improving the generalization of objects of various sizes and shapes. The YOLOv5l model achieved an overall accuracy of 95%. Although YOLO is somewhat slower than SSD, it detects autos well, without missing a single car in any frame of video (Kim et al., 2020b). Compared to edge-based object detection, the tracking accuracy of the tracking method approach using SSD object detection is higher (Liu et al., 2019). However, several researchers employ Faster R-CNN as the foundation of the deep neural architectures they created to conduct vehicle recognition and tracking in the section below (Maity et al., 2021). To decrease the effect of false positives on vehicle tracking, they propose introducing a low confidence track filtering is used into the Deep SORT (Simple Online and Realtime Tracking) method (IEEE Signal Processing Society & Institute of Electrical and Electronics Engineers, 2019). With the help of YOLOv5 model with efficient obstacle detection mechanism and faster speeds (Murthy et al., 2022) developed ADAS, it is a detection for unhealthy driving condition which led to road accidents.

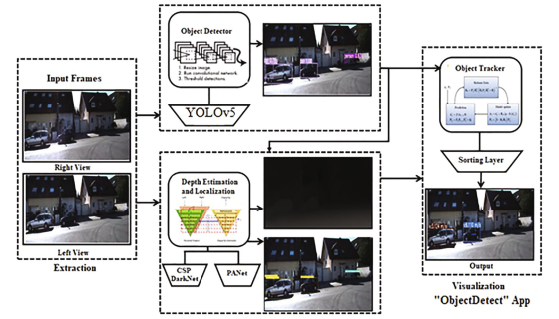


Figure 1. Object detection and tracking proposed framework (Murthy et al., 2022).

**2.4 Software solution**

Although not impossible, collecting a large amount of GPS-based vehicle coordinates, instantaneous speed, and gap simultaneously on a test vehicle for a variety of simulated route diversions and/or driving behaviors resulting from various rebellious or unlawful operations is relatively expensive, time-consuming, and labor-intensive (H. Wang et al., 2011). GPS-based real-time car locating systems are widely accessible on the market. A fleet manager may locate the present positions of the company's vehicles and communicate with the drivers in real-time when an automatic vehicle location system is used in conjunction with a two-way wireless communication system. Numerous studies have been conducted utilizing the probe vehicle technique, which is utilized to gather real-time traffic data in order to identify the types of traffic where GPS technology is used (Mandal et al., 2011). In connection with that, an intriguing method for traffic data collection to assist extend geographic coverage uses vehicles equipped with GPS sensors as probes. Travel time estimation and routing features are provided by navigation services like Google Maps (Google. Google Maps. https://www.google.com/maps) (Mon et al., 2022). The application developed by (Ozdemir & Tugrul, 2019) was built with Node.Js and Leaflet.Js. without having to refresh the map screen. Leaflet Realtime is a library that may show data from several web services at once. While Node.Js is a JavaScript platform that runs in the background and may be used for foreground and background apps. Because speed and performance are the two most crucial elements in the real-time tracking system application. Another Javascript framework called Vue.Js, (Song et al., 2019) developed a college teaching system based on Vue.js framework and accomplished design objectives and met user desire.

**2.5 Used Technologies**

Based on visual input from advanced sensors like cameras, an object detection system may recognize one or more traffic signals (Murthy et al., 2022). Same as (Tan et al., 2019) proposed a novel framework for multi-camera tracking which they demonstrate the pipeline of their method. All of the cameras' frames are first subjected to vehicle detection; then, single camera tracking is used to obtain tracklets; next, their vehicle ReID models extract visual attributes for each tracklet; and finally, multi-camera tracking is used to match tracklets from various cameras. As traffic data gathering devices are used more often, the data they produce has improved in reliability for use in traffic monitoring. According to the paper reports conducted in Sathorn are of Bangkok, Thailand by (Mon et al., 2022) they installed Induction loop coil sensors and thermal and CCTV at the approaching links of the critical Sathorn-Surasak intersection. In relation with that, Sim908 and Arduino Uno were utilized in the (Ozdemir & Tugrul, 2019) application development, installed are the SIM card, GPS, and GPRS receivers and the Arduino are then integrated with it. Traffic Monitor System is a system for measuring and monitoring traffic congestion that combines active RFID technology with quick deployment, low cost, and simple maintenance (based on IEEE 802.15.4 protocol, 2.4 GHz ISM band) and GSM technologies (Mandal et al., 2011). Same with the study of (Rodríguez-Rangel et al., 2022) about vehicle speed estimation from image sequences, they collect data using a cellphone camera and speed radar (Bushnell radar). The cameras of two cellphones (Xiaomi Redmi Note 7 & iPhone X) and the server, which runs Linux Ubuntu 20.04.2 LTS, has the following specs: CPU: Intel Core i3-8100, Memory RAM: 32GB DDR4, GPU: GeForce RTX 2070 Super is utilized to compute the video samples.

**2.6. Discussion**

Three types of deep learning model were shown above, this are YOLO (V4&V5), Fast R-CNN and SSD. Region proposal-based CNN architectures for object detection can be taught end-to-end using Faster R-CNN (Zhao et al., 2019). Faster R-CNN is an end-to-end method that resembles SSD but using SSD can detect object in the fastest manner (Palwankar & Kothari, 2022). Due to the numerous cutting-edge techniques known as "bag of freebies" and "bag of specials" included in the model, YOLOv4 is incredibly effective and ideal for aerial object detection (Samyal et al., 2022). YOLOv5 algorithm, the system is powerful enough to run object detection in various weather conditions (Murthy et al., 2022). The real-time detection needs can be easily met by YOLOv5 (Yao et al., 2021). To achieve real-time detection, the YOLOv5 paradigm is suitable for deployment to embedded devices. (Yan et al., 2021).

Among the 3 deep learning models such as YOLO in latest version, Faster-RCNN, and SSD, it has been prove based on the study above that in real-time detection YOLO with the latest version outperforms with the highest speed and accuracy in object detection and tracking. And therefore, YOLO with the latest version will be our basis in conducting the study.

This study will be using cameras for capturing real-time videos, leIEEE 802.15.4 a wireless sensor network technology for traffic monitoring system, and NodeJS for scripting. This selected technologies and software will be included for the development.

**CHAPTER III**

**METHODOLOGY**

**3.1 Detection of vehicles**

The researchers will be using Yolov5s for detection of vehicle together with the IoT devices consist of camera, RaspBerry Pi4, Jetson Nano, and MicroSD Card for the storage.

Our method is based on the utilized data provided by the controller/server. The controller, which consists of both hardware and software, uses an RL-based model to operate traffic lights according to sensor inputs. The controller will feed information about the current traffic situation to the dashboard. The work of the controller is to provide the data which is the traffic signals and number of vehicles per lane according to the viewing angle of the camera including the nearby roadways. (In visualization, the type of vehicle will be classified according to shapes and color).

The researchers will use the data of controller/server from other researcher that conducted a similar study from Caraga State University.

**3.2 Training of Models** (Object Detection, Counting, and Tracking)

The researchers choose the gathered data of the controller/server based on the YOLO v5s model for detection and classification and Deep SORT algorithm for tracking of vehicles. The images used in this research were captured in the intersection road of Jose Rosales Ave, Butuan City, Agusan Del Norte. The images were captured during rush hour, and the specific image acquisition time were 4:00 – 5:00 p.m. on *Month date (sunny or cloudy) in 2022.* In order to ensure the diversity of the image samples, the images were taken in natural light with backlighting and direct sunlight. Images were captured with a (chosen type) camera at \_\_\_\_\_\_ resolution and saved in \_\_\_\_\_\_ format.

Table 1. Detailed information of captured vehicle images.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**3.3 The architecture of Model**

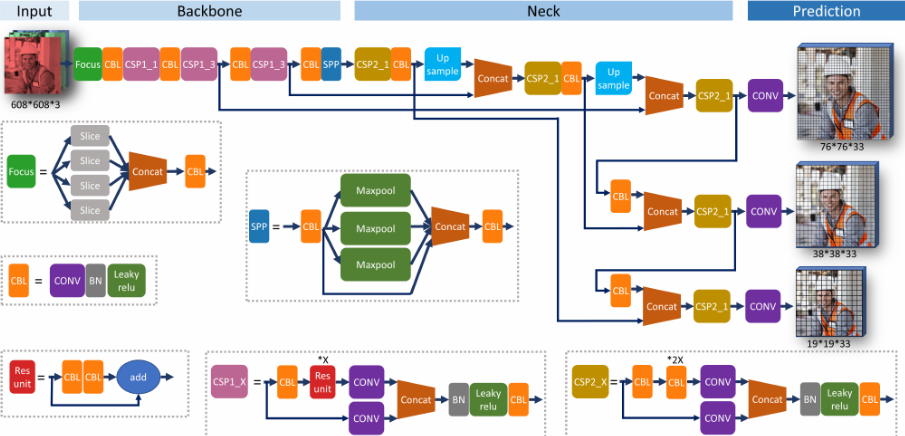


Figure 2. Architecture of YOLO v5s

Figure 2 explains, YOLO v5s consist of three main parts namely Model Backbone, Neck, and head. The Model Backbone extract important features using CSP or Cross State Partial Networks that extract informative features from a given image. While Model Neck is used for generating feature pyramids – this helps model perform well on unseen data, enable models generalized well on object scaling, and thus help models identify similar object with various sizes and scales. The final detection part is Model Head. This part applies anchor boxes and produce final outputs vectors with class probabilities, abjectness scores, and bounding boxes.

YOLO v5s is one of the latest models of YOLO series, this uses the structure of CSPDarknet that guarantees the speed and accuracy of model inference and decreases the size of the model (Wang, D. & He, d., 2021). As a result, this model runs smoothly while maintaining accuracy hence recommended for real-time object detection.

**3.4 The system designs**

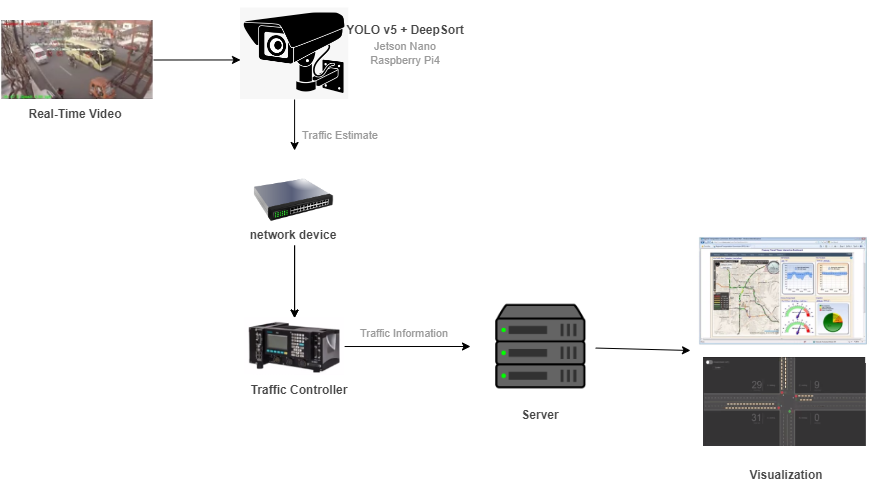


Figure 3. System Architecture

Traffic lights has Jetson nano and a camera to capture vehicles and determine the Traffic State utilizing YOLOv5s for object detection and classification, and DeepSort for object tracking. The estimated Traffic State is converted to vectorized data and sent to the Traffic Controller. The Controller is both hardware and software, and it comprises an RL-based model for managing traffic lights based on the sensor data. Using the Weights from the sensors (Vectorized Data), the Controller will transmit back data to direct or control each traffic light. The Controller will provide the dashboard with the information for current traffic state visualization.

**3.5 Software used** – Web-based or Mobile

In this study, the researchers using free to use and accessible software such as:

* Visual Studio Code for the researcher's code editor,
* Open Street Map is a community-based, open-source, editable map service developed as a substitute for reliable sources (Vargas-Munoz et al., 2021)
* Open Layers.org is used to access the open street map. It is an open-source JavaScript package used to show map data as slick maps in web browsers. It offers an API for creating sophisticated web-based geospatial apps akin to Google Maps and Bing Maps.
* Vue.js is the best lightweight front-end framework based on MVVM mode in Web applications (Song et al., 2019)
* Nuxt.js, a framework in built upon vue.js application. It offers server-side rendering, automatic route generation, improved meta tags managing and SEO improvement.
* Konva is a popular library for creating web-based 2D animation.

**CHAPTER IV**

**RESULT AND DISCUSSION**

Experimental results and discussions

• Counting the vehicles

• Average vehicle speed

• Time complexity

References:

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. <http://arxiv.org/abs/2004.10934>

Chen, H., Guo, B., Yu, Z., & Han, Q. (2016). Toward real-time and cooperative mobile visual sensing and sharing. *Proceedings - IEEE INFOCOM*, *2016-July*. <https://doi.org/10.1109/INFOCOM.2016.7524481>

Chen, H., Guo, B., Yu, Z., & Han, Q. (2019). CrowdTracking: Real-Time Vehicle Tracking Through Mobile Crowdsensing. *IEEE Internet of Things Journal*, *6*(5), 7570–7583. <https://doi.org/10.1109/JIOT.2019.2901093>

Gu, Y., Xu, S., Wang, Y., & Shi, L. (2019). An advanced deep learning approach for safety helmet wearing detection. *Proceedings - 2019 IEEE International Congress on Cybermatics: 12th IEEE International Conference on Internet of Things, 15th IEEE International Conference on Green Computing and Communications, 12th IEEE International Conference on Cyber, Physical and So*, 669–674. <https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00128>

IEEE Signal Processing Society, & Institute of Electrical and Electronics Engineers. (2019). Proceedings of AVSS 2018 : 2018 IEEE International Conference on Advanced Video and Signal-based Surveillance : Auckland, New Zealand, 27-30 November 2018. *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1–6.

Khan, M. A., & Khan, S. F. (2018). IoT based framework for Vehicle Over-speed detection. *1st International Conference on Computer Applications and Information Security, ICCAIS 2018*, 1–4. <https://doi.org/10.1109/CAIS.2018.8441951>

Kim, J. A., Sung, J. Y., & Park, S. H. (2020a, November 1). Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Recognition. *2020 IEEE International Conference on Consumer Electronics - Asia, ICCE-Asia 2020*. <https://doi.org/10.1109/ICCE-Asia49877.2020.9277040>

Kim, J. A., Sung, J. Y., & Park, S. H. (2020b, November 1). Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Recognition. *2020 IEEE International Conference on Consumer Electronics - Asia, ICCE-Asia 2020*. <https://doi.org/10.1109/ICCE-Asia49877.2020.9277040>

Ku, B., Kim, K., & Jeong, J. (2022). *Real-Time ISR-YOLOv4 Based Small Object Detection for Safe Shop Floor in Smart Factories*. 1–18.

Lin, C. T., Huang, S. W., Wu, Y. Y., & Lai, S. H. (2021). GAN-Based Day-to-Night Image Style Transfer for Nighttime Vehicle Detection. *IEEE Transactions on Intelligent Transportation Systems*, *22*(2), 951–963. <https://doi.org/10.1109/TITS.2019.2961679>

Liu, P., Wang, G., Yu, Z., Guo, X., & Lu, W. (2019). Vehicle tracking based on shape information and inter-frame motion vector. *Computers and Electrical Engineering*, *78*, 22–31. <https://doi.org/10.1016/j.compeleceng.2019.06.019>

Maity, M., Banerjee, S., & Sinha Chaudhuri, S. (2021). Faster R-CNN and YOLO based Vehicle detection: A Survey. *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, 1442–1447. <https://doi.org/10.1109/ICCMC51019.2021.9418274>

Mandal, K., Sen, A., Chakraborty, A., Roy, S., Batabyal, S., & Bandyopadhyay, S. (2011). Road traffic congestion monitoring and measurement using active RFID and GSM technology. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 1375–1379. <https://doi.org/10.1109/ITSC.2011.6082954>

Mathur, G., Somwanshi, D., & Bundele, M. M. (2018). Intelligent Video Surveillance based on Object Tracking. *3rd International Conference and Workshops on Recent Advances and Innovations in Engineering, ICRAIE 2018*, *2018*(November), 1–6. <https://doi.org/10.1109/ICRAIE.2018.8710421>

Mon, E. E., Ochiai, H., Komolkiti, P., & Aswakul, C. (2022). Real-world sensor dataset for city inbound-outbound critical intersection analysis. *Scientific Data*, *9*(1), 1–20. <https://doi.org/10.1038/s41597-022-01448-6>

Murthy, J. S., Siddesh, G. M., Lai, W.-C., Parameshachari, B. D., Patil, S. N., & Hemalatha, K. L. (2022). ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5. *Wireless Communications and Mobile Computing*, *2022*, 1–10. <https://doi.org/10.1155/2022/9444360>

Ozdemir, Z., & Tugrul, B. (2019). Geofencing on the Real-Time GPS Tracking System and Improving GPS Accuracy with Moving Average, Kalman Filter and Logistic Regression Analysis. *3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2019 - Proceedings*. <https://doi.org/10.1109/ISMSIT.2019.8932766>

Pal, S. K., Pramanik, A., Maiti, J., & Mitra, P. (2021). Deep learning in multi-object detection and tracking: state of the art. *Applied Intelligence*, *51*(9), 6400–6429. <https://doi.org/10.1007/s10489-021-02293-7>

Palwankar, T., & Kothari, K. (2022). Real Time Object Detection using SSD and MobileNet. *International Journal for Research in Applied Science and Engineering Technology*, *10*(3), 831–834. <https://doi.org/10.22214/ijraset.2022.40755>

Prakash, N., Kamath, C. U., & Gururaja, H. S. (n.d.). *Detection of Traffic Violations using Moving Object and Speed Detection*. *XVI*(61), 61–72.

Rodríguez-Rangel, H., Morales-Rosales, L. A., Imperial-Rojo, R., Roman-Garay, M. A., Peralta-Peñuñuri, G. E., & Lobato-Báez, M. (2022). Analysis of Statistical and Artificial Intelligence Algorithms for Real-Time Speed Estimation Based on Vehicle Detection with YOLO. *Applied Sciences (Switzerland)*, *12*(6). <https://doi.org/10.3390/app12062907>

Song, J., Zhang, M., & Xie, H. (2019). Design and implementation of a Vue.js-based college teaching system. *International Journal of Emerging Technologies in Learning*, *14*(13), 59–69. <https://doi.org/10.3991/ijet.v14i13.10709>

Tan, X., Wang, Z., Jiang, M., Yang, X., Wang, J., Gao, Y., Su, X., Ye, X., Yuan, Y., He, D., Wen, S., & Ding, E. (2019). Multi-camera vehicle tracking and re-identification based on visual and spatial-temporal features. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2019-June*, 275–284.

Tong, K., Wu, Y., & Zhou, F. (2020). Recent advances in small object detection based on deep learning: A review. *Image and Vision Computing*, *97*, 103910. <https://doi.org/10.1016/j.imavis.2020.103910>

Vargas-Munoz, J. E., Srivastava, S., Tuia, D., & Falcao, A. X. (2021). OpenStreetMap: Challenges and Opportunities in Machine Learning and Remote Sensing. *IEEE Geoscience and Remote Sensing Magazine*, *9*(1), 184–199. <https://doi.org/10.1109/MGRS.2020.2994107>

Wang, D., & He, D. (2021). ScienceDirect Channel pruned YOLO V5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. *Biosystems Engineering*, *210*, 271–281. <https://doi.org/10.1016/j.biosystemseng.2021.08.015>

Wang, H., Cheu, R., & Lee, D. H. (2011). A study of real time security monitoring of commercial vehicles using a microscopic traffic simulation model. *Applied Mechanics and Materials*, *55–57*, 1293–1298. <https://doi.org/10.4028/www.scientific.net/AMM.55-57.1293>

Wont, W. J., Kim, T. H., Choi, M. K., & Kwon, S. (2018). AggNet: Simple Aggregated Network for Real-Time Multiple Object Detection in Road Driving Scene. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2018-Novem*, 3505–3510. <https://doi.org/10.1109/ITSC.2018.8569393>

Xiao, Y., Wang, X., Zhang, P., Meng, F., & Shao, F. (2020). Object detection based on faster r-cnn algorithm with skip pooling and fusion of contextual information. *Sensors (Switzerland)*, *20*(19), 1–20. <https://doi.org/10.3390/s20195490>

Yanagisawa, H., Yamashita, T., & Watanabe, H. (2018). A study on object detection method from manga images using CNN. *2018 International Workshop on Advanced Image Technology, IWAIT 2018*, 1–4. <https://doi.org/10.1109/IWAIT.2018.8369633>

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object Detection with Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(11), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>