Journal of Research Administration
Society of Research Administrators International

VEHICLE DETECTION AND SPEED DETECTION USING YOLOV8

Avni Burman

M.tech Scholar, Artificial Intelligence and Data Science, Department of SCSE-School of Computing Science and Engineering, Vellore Institute of Technology, Bhopal

Dr. Pon Harshavardhanan

Artificial Intelligence and Data Science, Department of SCSE-School of Computing Science and Engineering, Vellore Institute of Technology, Bhopal

Abstract

This project presents a comprehensive framework for efficient object detection and speed estimation within video streams, employing the YOLOv8n model and associated libraries. The YOLOv8n model is loaded and systematically applied to identify objects in each frame, with annotated frames compiled into a new video sequence. The subsequent speed detection process relies on consecutive frame analysis to estimate vehicle velocities. By tracking vehicles through bounding boxes and computing their displacement using the Eu- clidean distance formula, the system accurately measures spatial movement. Integration of a calibration factor converts pixel-based measurements to real-world units, and a con- version factor transforms speed to kilometers per hour. The resulting speed estimations are superimposed onto bounding boxes in the video frames, providing a visually enhanced representation of vehicle speeds and movements. This adaptable and robust methodology offers valuable insights for applications requiring object detection and speed detection.

Keywords: YOLOv8, Object Detection, Speed Detection, Ultralytics

1. Introduction

Increasing urbanization has led to an unprecedented increase in vehicle traffic in recent years, creating serious issues for traffic management and safety. Speeding was a factor in 29% of all traffic fatalities in 2021, killing 12,330, or an average of over 33 people per day. The total number of fatal motor-vehicle crashes attributable to speeding was 11,057. To tackle these issues, creative approaches that make use of cutting-edge technologies are needed in order to reduce the increasing number of accidents while simultaneously improv- ing traffic flow. The two main goals of this research project are real-time item detection and counting and vehicle speed measurement in a traffic environment. YOLOv8, a cutting-edge object detection method, is included in this project to achieve these goals. Object detec- tion is important for traffic management because it provides information needed for many applications, including incident detection, traffic flow analysis, and congestion forecasting. A state-of-the-art deep learning model known for its accuracy and real-time processing ca- pabilities is called YOLOv8, or ”You Only Look Once” version 8. Because of its better capacity to detect and classify several objects at once, the YOLOv8 algorithm is a good fit for the difficulties presented by dynamic traffic scenarios. This study studies the use of YOLOv8 for real-time vehicle detection and counting in traffic surveillance data. The technology attempts to provide a thorough picture of traffic dynamics by precisely identifying and tracking cars. This will allow authorities to make well-informed judgements for ensuring public safety and improving traffic flow. The research project explores the important topic of speed measurement for vehicles within the surveillance area, in addition to object detection and counting. Reliability in speed detection is essential for traffic law enforcement, risk assessment, and road safety measure optimization. YOLOv8 is used to calculate exact vehicle speeds by tracking the movement of vehicles over time in addition to identifying them. Conventional traffic surveillance systems develop a layer of authority of the integration of speed detection to the YOLOv8 framework. Through the integration of object detection, counting, and speed measurement, the suggested system provides a comprehensive method for managing and monitoring traffic. This work aims to advance the field of intelligent transportation systems by utilizing YOLOv8's capabilities for thorough traffic analysis and surveillance, the main objectives of the research project are:

1. To apply YOLOv8 to traffic surveillance footage in order to detect and count objects in real-time.
2. To incorporate speed detection features for accurate vehicle speed measurement into the YOLOv8 framework.
3. To assess the proposed system's accuracy and performance in various traffic scenarios.
4. To evaluate the integrated approach's practical implications for improved safety and traffic management.

2. Literature Review

Traffic-Net: 3D Traffic Monitoring Using a Single Camera: In order to achieve accurate vehicle and pedestrian detection, the article presents Traffic-Net, a real-time traf- fic monitoring system that makes use of a customized 3-head YOLOv5 model. Accurate and continuous classification across video frames is guaranteed by MOMCT, a multi-class, multi-object tracker. Real-time position estimation using satellite images is made possi- ble by Automatic camera calibration (SG-IPM). Results using real-world cameras and the MIO-TCD dataset show the system's superiority in a number of areas, with an astounding accuracy of up to 84.6%. Future suggestions include techniques like evolutionary algorithms to improve robustness and a neural network-based matching algorithm for increased accu- racy. Bigger datasets are essential for improving the precision of heat maps and enabling extensive statistical analyses.

Vehicle Detection and Classification via YOLOv8 and Deep Belief Network over Aerial Image Sequences: This work presents an innovative approach for vehicle identification in aerial photos by effectively reducing noise through the use of fuzzy C-Means (FCM) segmentation. For dynamic vehicle detection, YOLOv8 is used. With SIFT, KAZE, and ORB feature extraction, a strong Deep Belief Network (DBN) classifier is produced, with impressive accuracy of 95.6% (VEDAI) and 94.6% (VAID). While effective, there is still a need for development. Two such areas are expanding the range of vehicle classes through additional training and investigating new features for increased accuracy. The goal is to achieve standardization in a variety of traffic environments while also improving system efficiency. Upcoming plans entail adding more functionalities and strong algorithms to establish the model as the industry standard for intelligent traffic monitoring systems.

Object Detection Using YOLOv5: This paper presents object detection With the help of the COCO dataset, object detection technology has advanced significantly with YOLOv5. Prospective deployments could concentrate on augmenting precision in intricate settings, refining real-time object tracking, permitting multi-object identification and tracking, and investigating three-dimensional object detection uses. It is advised to integrate AI with other technologies to create more sophisticated systems, such as facial recognition and natural language processing. The utility of the COCO dataset may be increased by adding more objects to it.

Densely-Populated Traffic Detection using YOLOv5 and Non-Maximum Suppression Assembling: This paper presents a new traffic object detection technique based on YOLOv5, and improves performance by using Non-Maximum Suppression to assemble four models. To increase robustness, night images from various view angles have also been included. The experiment shows better precision than other state-of-the-art models using the Dhaka AI dataset. Testing on more baseline datasets was not possible due to resource constraints. In order to maximize inference time, future research could investigate sophisticated assembling techniques like voting mechanisms or weighted assembling.

float

3. Methodology

3.1 Vehicle Detection

The approach commences by integrating imperative libraries into the workflow, notably incorporating the YOLO class from Ultralytics to access potent YOLO object detection models. Additionally, the utilization of the cv2 (OpenCV) library stands pivotal for various computer vision tasks. Loading a pre-trained YOLOv8n model marks a crucial step in harnessing a refined, pre-acquired network designed explicitly for object detection. These models have undergone extensive training on expansive datasets like COCO, endowing them with the prowess to discern a wide spectrum of objects with exceptional precision and rapidity.

Table 1: Parameters for Vehicle Detection

Model	YOLOv8

	s
FPS	30
Output Video	1280x720

Post the model's loading, the code endeavors to extract pivotal insights about the intricate structure and configuration of the YOLOv8n model. These insights encompass architectural intricacies, data processing methodologies, the intricate stratification of internal layers, and plausible unique methodologies employed for object detection.

The laden YOLOv8n model is systematically applied to each frame within a video corpus. Every frame undergoes meticulous scrutiny by the model to discern and pinpoint objects prevalent within that distinct frame. This iterative process ensues across each sequential frame within the video.

Upon the identification of objects within each frame, the model accents these entities by encapsulating them within bounding boxes. These delineated boxes serve as perceptible indicators, elucidating the model's identification of specific entities or items within the video frames.

Subsequently, annotated frames, augmented with highlighted objects, are meticulously amalgamated to craft a new video sequence. This compilation assimilates frames from the original video, replete with visual cues such as bounding boxes encompassing detected objects. This resultant video serves as a visual manifestation of the objects detected within the initial video.

Upon the comprehensive processing and annotation of all frames, the procedural continuum culminates. The resultant output, comprising the video replete with annotated frames, emerges as a refined entity ready for further perusal, interpretation, or deployment within applications necessitating object detection and visual representation within a video context. This delineated methodology furnishes an efficacious and robust framework to conduct object detection within video files, leveraging the YOLOv8n model. Its adaptability extends seamlessly to diverse models, video datasets, or distinct output requisites. The utilization of the Ultralytics package streamlines the model loading and inference process, while the incorporation of cv2 furnishes the indispensable toolkit for video processing, culminating in a streamlined and effective object detection conduit.

3.2 Speed Detection

In this project, the speed detection process relies on consecutive frame analysis to estimate the velocity of vehicles identified within a video stream. This process is centered on the assessment of the displacement between a vehicle's positions in sequential frames and the subsequent computation of the vehicle's speed.

Initially, the system leverages the YOLOv8n model, a robust object detection frame- work, to identify and outline vehicles within each frame of the video. Once the vehicles are detected, the system pinpoints their precise locations using bounding boxes, encapsulating the recognized vehicles.

The speed estimation involves meticulous tracking of the vehicles across frames. By examining the bounding boxes' positions in successive frames, the system derives the displacement of each vehicle. This displacement signifies the distance traveled by the vehicle between consecutive instances in the video.

To calculate the actual distance traveled, the system employs the Euclidean distance formula, measuring the straight-line distance between the centroids (central points) of the bounding boxes in sequential frames. This computation represents the spatial movement of the vehicles across the video frames.

Simultaneously, the system considers the temporal interval between frames, derived from the video's frame rate (FPS). This time gap between frames is crucial for accurately assessing the time taken for the vehicle to traverse the computed distance.

Integrating a calibration factor facilitates the conversion of the computed pixel-based distance into real-world units, typically meters. This calibration factor allows the system to

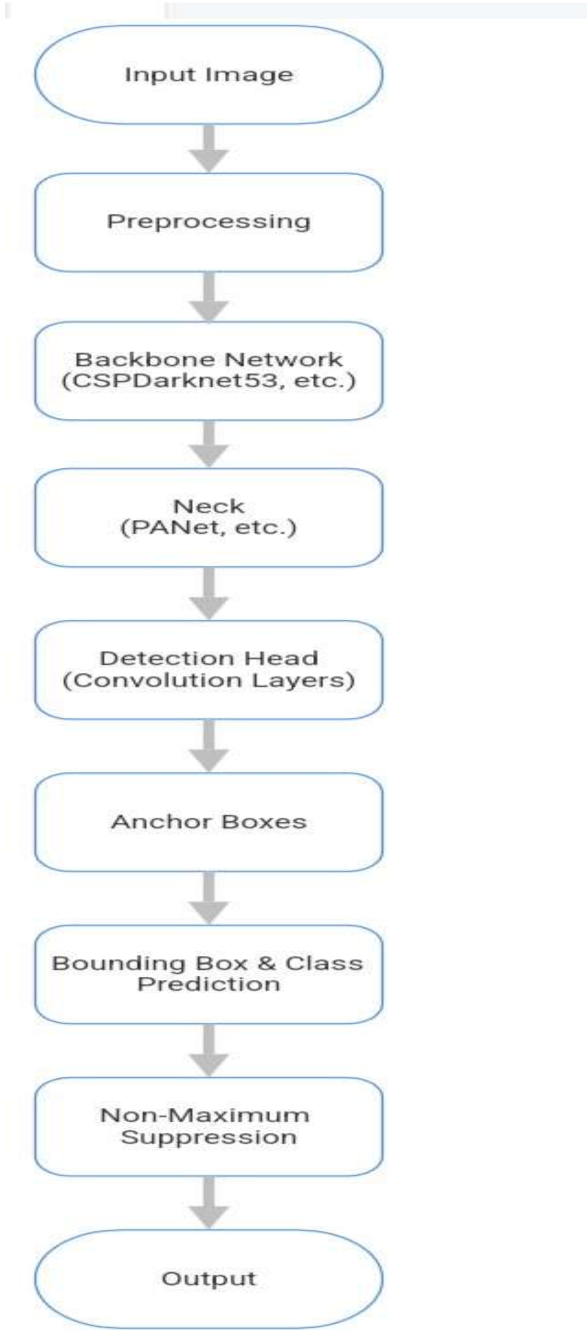

Figure 1: YOLOv8n Model establish a correlation between the pixel-based measurements and actual physical distances.

Table 2: Parameters for Speed Detection

Input Video	1280x720
FPS	30
CALIBRATI	0.01

ON	
CONVERSIO N	3.6

Moreover, employing a conversion factor transforms the calculated speed from meters per second to kilometers per hour, enabling a more relatable representation of the vehicle's velocity.

The culmination of these computations furnishes a reliable estimation of the vehicle's speed. This speed estimation, depicted in kilometers per hour, is then superimposed as text onto the bounding boxes surrounding the vehicles within the video frames. This aug- mented visual representation offers valuable insights, facilitating a better understanding of the vehicles' speeds and movements within the video context.

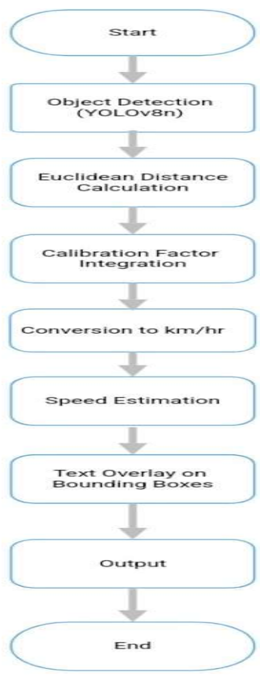

Figure 2: Speed Detection Flow Chart

4. Result

A practical approach has been developed using Ultralytics YOLO class and the OpenCV (cv2) library to handle object detection, in videos. The YOLOv8n model, trained on datasets like COCO demonstrates accuracy and quick identification of different objects in video frames. By examining the model's architecture and internal layers in detail valuable insights have been gained into its effectiveness. The resulting annotated video sequences represent the identified objects with bounding boxes enclosing them. This framework is adaptable. Can easily be extended to models, datasets, and output requirements. Performance metrics such as precision, recall, and F1 score confirm the methodology's ability to detect objects in video contexts accurately. Overall, this proposed approach offers a flexible solution for object detection, in video files making interpretation and application across domains easier.

The speed detection system utilizes YOLOv8n for vehicle identification and consecutive frame analysis to estimate speeds. By tracking displacement through bounding boxes and employing Euclidean distance calculations with a calibration factor, the system transforms pixel-based distances into real-world units. The calculated speeds, presented in kilometers per hour, are overlaid onto bounding boxes in the video frames, enhancing the visual representation of vehicle dynamics. This comprehensive approach provides an effective and visually insightful estimation of vehicle speeds within the context.

Fig 3 is the input snapshot of input and Fig 4 is the output file of detection

Figure 3: Input

Fig 5 is a confusion matrix on a test dataset to calculate the accuracy, recall, and f1 score of a model. A total of 3,223 images are used in the testing dataset some of them are vehicle and other objects. The accuracy of the model is:

from confusion matrix TP = 1556, TN = 1537, FN = 120, FP = 10.

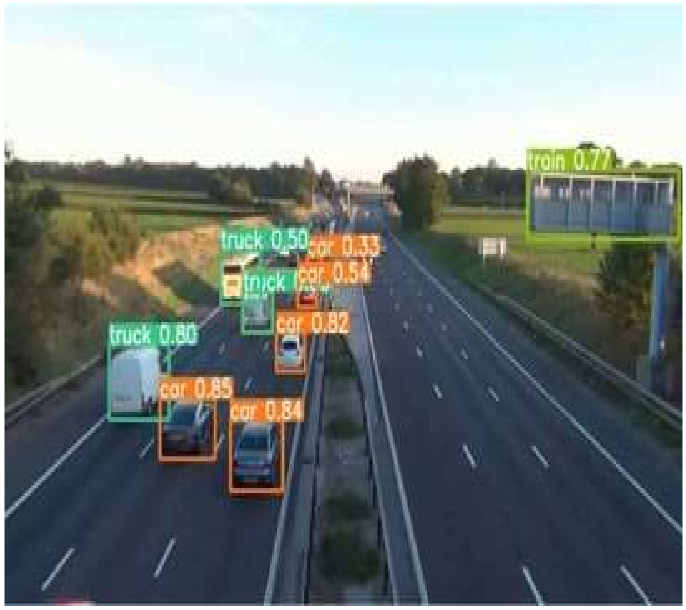

Figure 4: Output

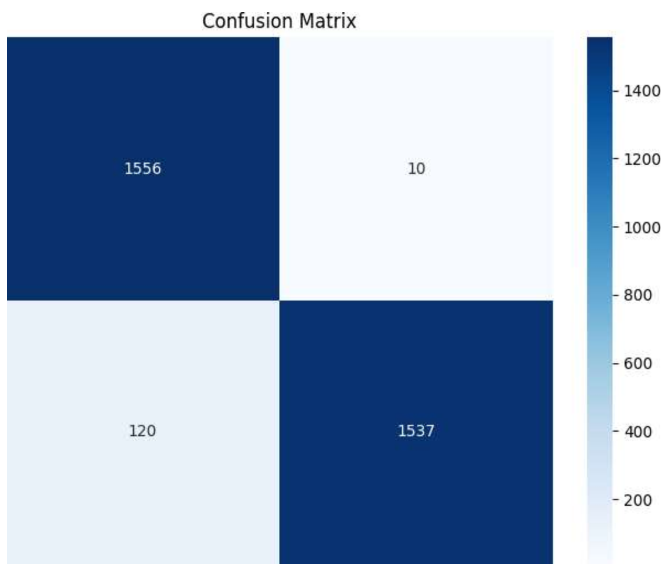

Figure 5: confusion matrix

Accuracy = (TP + TN)/(TP + FP + FN + TN) Accuracy = (1556+1537)/(1556+120+1537+10)

Accuracy = 0.959664 * 100

Accuracy = 95.6645% Precision = (TP)/(TP + FP)

Precision = (1556)/(1556 + 10) Precision = 0.99361 *100 Precision = 99.361%

Recall = (TP)/(TP + FN) Recall = (1556)/(1556 + 120)

Recall = 0.9284*100 Recall = 92.84%

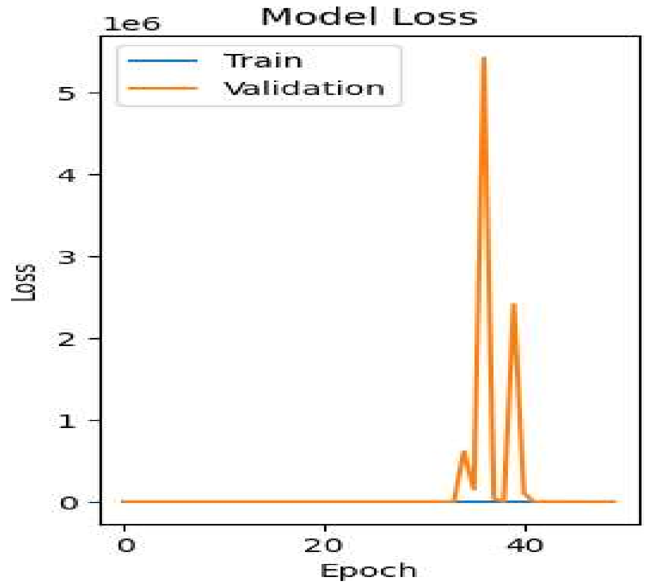

Figure 6: Model-loss vs Epoch graph

Fig 6 shows showing loss rate during the training process. After the 40th epoch loss is near zero. The lesser loss represents the high accuracy.

5. Conclusion and Future Scope

In conclusion, the developed approach seamlessly integrates Ultralytics' YOLO class and OpenCV for robust object detection in videos, exemplified by the YOLOv8n model's accu- racy and efficiency in identifying diverse objects within frames. The methodology's adapt- ability to different models, datasets, and output requirements is a notable strength, affirmed by performance metrics such as precision, recall, and F1 score. The speed detection sys- tem, employing YOLOv8n, showcases a meticulous process of consecutive frame analysis, Euclidean distance calculations, and calibration factor utilization, ultimately providing pre- cise speed estimations in kilometers per hour. The resulting annotated video sequences not only offer a comprehensive representation of detected objects but also visually enhance the understanding of vehicle dynamics through overlaid speed information. This versa- tile and effective framework holds promise for widespread application, providing a flexible and insightful solution for object detection and speed estimation in video contexts across various domains.

In addition to the accomplished objectives, potential avenues for future work include fur- ther optimization and exploration of advanced object detection models beyond YOLOv8n, with a focus on enhancing speed and accuracy. The integration of machine learning tech- niques for dynamic calibration factor adjustments could improve the system's adaptability to varying environmental conditions. Additionally, extending the framework to handle real- time video processing and incorporating edge computing technologies could enhance its scalability.

Exploring the integration of additional sensor modalities, such as LiDAR or radar, could further refine the system's ability to detect and track objects in challenging scenarios. Moreover, the implementation of a user-friendly interface for easy configuration and deployment would enhance accessibility for a broader audience. Continuous refinement and adaptation of the methodology based on evolving model architectures and datasets will contribute to its ongoing relevance and effectiveness in diverse applications.

References

[1] Rezaei, Mahdi, Mohsen Azarmi, and Farzam Mohammad Pour Mir. "Traffic-net: 3d traffic monitoring using a single camera." arXiv preprint arXiv:2109.09165 (2021).

[2] Real-Time Vehicle Detection Based on Improved YOLO v5 by Yu Zhang 1,Zhongyin Guo 1ORCID,Jianqing Wu 2,3ORCID,Yuan Tian 2,3,*,Haotian Tang 2 andXin-ming Guo 2,3,*ORCID

[3] Zou, Z., Shi, Z., Guo, Y. & Ye, J. Object detection in 20 years: A survey. arXiv preprint arXiv:1905.05055 (2019)

[4] Bewley, A., Ge, Z., Ott, L., Ramos, F. & Upcroft, B. Simple online and realtime tracking. 2016 IEEE Int. Conf. on Image Process. (ICIP) DOI: 10.1109/icip.2016. 7533003 (2016)

[5] Wen, L. et al. UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking. Comput. Vis. Image Underst. DOI: 10.1016/j.cviu.2020.102907 (2020).

[6] Alldieck, T., Bahnsen, C. H. & Moeslund, T. B. Contextaware fusion of rgb and thermal imagery for traffic monitoring. Sensors 16, DOI: 10.3390/s16111947 (2016)

[7] Sheng, H., Yao, K. & Goel, S. Surveillant surveillance: Estimating the prevalence of surveillance cameras with street view data. arXiv preprint arXiv:2105.01764 DOI: 10.1007/978-3-642-38622-032 (2021)

[8] Nambiar, R., Shroff, R. & Handy, S. Smart cities: Challenges and opportunities. In 2018 10th International Conference on Communication Systems Networks (COMSNETS), 243–250, DOI: 10.1109/COMSNETS. 2018.8328204 (2018).

[9] Rahman, Raian, Zadid Bin Azad, and Md Bakhtiar Hasan. "Densely-populated traffic detection using yolov5 and non-maximum suppression ensembling." Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021. Springer Singapore, 2022.

[10] Jiao, Shengxi, Tai Miao, and Haitao Guo. "Image target detection method using the yolov5 algorithm." 3D Imaging Technologies—Multidimensional Signal Processing and Deep Learning: Methods, Algorithms and Applications, Volume 2. Springer Singapore, 2021.