Real-Time Vehicle Detection and Air Pollution Estimation Using YOLOv9

Hari Suparwito[1]*, Bernardus Hersa Galih Prakoso[2], Rosalia Arum Kumalasanti[3], Agnes Maria Polina[4]
Department of Informatics [1], [2], [3], [4]
Universitas Sanata Dharma
Sleman, Yogyakarta
shirsj@jesuits.net[1], Bernardus.hersa37@gmail.com[2], rosalia.santi@usd.ac.id[3], a.m.polina@usd.ac.id[4]

***Abstract*— Pollution of air, particularly in cities, is becoming an issue to be taken seriously owing to the health and environmental risks associated with it, and the major contributor to air pollution is car emissions. The objective of the study is to identify and classify vehicles such as motorbikes, cars, buses, trucks in order to monitor live traffic and potentially determine the extent to which the pollution level elevates, utilizing the YOLOv9 model. Traffic CCTV camera footage was gathered under a wide range of circumstances including different lighting and varying traffic intensity. Folders were particularly structured and images annotated, in the manner, which served the purpose of meeting the requirements of the YOLO structure. Once it was trained with a labeled dataset, the vehicle identification by YOLOv9 model was found to be quite satisfactory. Overall vehicle identification accuracy was calculated to be mAP50:95 of 0.826. In contrast, it had a harder time with smaller items like motorcycles, with a mAP50:95 of 0.682. Findings indicate that larger items were detected more than smaller items. Camera angles and the small size of the objects often make small objects appear to blend in to the background. This research indicates that AI can be of help when dealing with the urban structure. It offers a way of measuring traffic volume to predict the amount of CO emissions that can be avoided or controlled. The rest are keen in enhancing the effectiveness of recognizing small objects within the system and deploying it in multiple settings**.

Keywords*— *Air pollution, carbon monoxide, computer vision, object detection, yolov9

I. INTRODUCTION

Pollution is an issue of significant scientific concern due to its inherent negative health and environmental impacts [1]. In residential areas, pollution levels are closely linked to traffic patterns. Some pollutants are directly linked to vehicular traffic density [2]. More recent studies show an exponential increase in registered and traveled vehicles in residential areas [3]. In such a scenario of mixed development and environmental degradation, the need of the hour is to address these concerns using a positive approach. One such positive approach is to neutralize the traffic-related pollution to the best feasible.

Air pollution is regarded as one of the main environmental threats. The negative impacts of pollution are multifaceted, encompassing ecosystems, health, agricultural productivity, educational attainment, and worker productivity [4]. In many situations, transport significantly contributes to air pollution through vehicle emissions. Damage to human health is one of the major outcomes of vehicle pollution, as road traffic is deemed a principal source of fine particulate matter, which results from the incomplete combustion of fuels in vehicle engines [5]. The major contributions to air pollution levels are associated with exhaust damage and the generation of wear particulates in vehicle engines. Damage to the environment beyond vehicle emissions is assessed in related work and correlated with transportation infrastructure. In response, technology and the environment have been evaluated by some researchers in order to help minimize the amounts of pollution emitted into the atmosphere [6]. Various methodologies are used to monitor different pollutants and toxicants, as well as to extend the spectrum of monitoring purposes. However, existing literature has more deeply given attention to the technology of vehicle detection, paying less attention to the contribution of front vehicles to pollution build-up.

While there are various parameters taken into consideration under pollution, significant advances were made in monitoring vehicular emissions and accidents. Manufacturing has also been growing in different countries around the world, contributing to increasing air and noise pollution due to gaseous emissions from these industries, as well as increasing road traffic. Motorized road traffic has led to various problems such as fuel exhaust emissions leading to air pollution affecting both outer and in-cabin air quality, a noisy atmosphere with increased noise pollution, and traffic congestion leading to traffic accidents. Human activities, including the excessive use of motor vehicles, play a major role in environmental degradation and cause pollution. To control these factors, it is important to monitor the environment, and it is understood that the development of detection systems with an environment-friendly approach is an urgent need.

From the existing condition, there are possible factors contributing to the accumulation of air pollution, such as inadequate or unfound technology for detecting vehicles. Thus, uncovering where vehicles are crucially confined and exploring real-time locations give the opportunity for another study to be conducted. The rationale ultimately allows for the possibility to assess effective traffic control methodology. By reviewing the studies conducted, it was seen that no research has focused on developing an effective real-time approach to facilitate the process, as has been approached in this work. Therefore, the research question of monitoring using AI for vehicles is able to be tackled by designing a system that has the capability to monitor traffic.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Most of the primary purpose of air pollution research is to determine how much CO, CO_2, SO_2, and NO_2 can be directly measured in the open air. These pollutants degrade the quality of the air and are detrimental to human health [7, 8, 9, 10]. However, in this study, we used computer vision technology to count vehicles in four distinct categories: motorcycles, cars, buses, and trucks. We use the results of vehicle identification and enumeration to predict the presence of hazardous substances that contribute to air pollution. This instance focuses on a specific component of the atmosphere, carbon monoxide (CO).

Recently, deep learning has transformed various fields among the usage of AI, particularly in object detection [11]. A specific detection method performs detection instantly and accurately in one pass. This method can be used to detect objects. With its real-time detection rates, it was seen to perform better than two-stage detectors and single-shot detectors. The association of AI as a monitoring tool for environmental assessment should be able to augment existing plausible capabilities in the environmental domain; hence, this work aims to investigate the application of AI in the real-time detection of vehicles in a traffic lane using a specific version of the detection method. This investigation is also supported by a discussion about the contribution of pollution build-up.

The development of information technology and artificial intelligence has expanded since the 21st century as a result of extensive research and the increasing need for automation. Almost every industry today uses automation built with the help of computer vision and artificial intelligence. Computer vision has become one of the rapidly developing fields of knowledge and is reliably utilized in various industries. Computer vision has the potential to integrate human interaction with systems in a very modern way, ensuring that future technology will always be up-to-date. The advantage of computer vision lies in the extraction of information from images, videos, and other visual inputs, which can then be further processed according to needs [12]. Computer vision is a field of artificial intelligence that can recognize objects in its surroundings, and this will be utilized in the research.

Object detection is a subfield of computer vision that involves the identification and localization of an object in images or videos. Identifying objects requires domain-specific algorithms and the use of standard techniques that include algorithms like image segmentation, classification, or deep learning-based approaches, resulting in the emergence of a number of applications. Classification algorithms identify a class of an image that can be embedded with bounding boxes. Image segmentation enters smaller areas which, in a similar manner, classify images and then take the union of detected objects at different locations. At the same time, deep learning methods are focusing on classifying subparts of that image. On the contrary, with the advancement of technology, deep learning-based object detection algorithms have taken into consideration resources from a variety of fields that include computers, medical, structural, surveillance, home automation, and transportation applications. In addition, vehicle detection helps in traffic monitoring and management, automated parking systems, vehicle counting at tolls, and surveillance of attackers and terrorists, among others. In transportation, the detection framework must be accurate and timely as it impacts a vehicle's time, speed, and distance with respect to other vehicles.

The capabilities of computer vision are widely utilized in biometric identification, such as fingerprint detection, facial recognition, gestures, and other biometric objects. The increasing demand for technology has led to the involvement of computer vision in the development of the autonomous industry, namely the 4.0 industrial revolution. The 4.0 industrial revolution era emphasizes automation that collaborates with technology, making computer vision play a significant role in this era [13].

Real-time object detection is a tough and evolving area in computer vision and machine learning. In recent years, YOLO (You Only Look Once) has emerged as a leading model in this field. YOLO takes a different approach than traditional methods to train object detection models [14]. YOLO trains a single neural network to specify bounding boxes and class probabilities at the same time. Also, it doesn't rely on pre-defined object features like edge maps, HoG descriptors, or region proposals, which sets it apart from classic object detection or segmentation algorithms. The YOLO system can process images at about 45 frames per second while achieving accuracy that matches or beats most current state models for object detection [15]. One type of YOLO could look at 2 pictures every second and got a score of 57.9 mAP with a COCO model. Another type when changed a bit, could look at 61 pictures a second and did just as well as the ones before it. The newest type is much faster and more accurate. It can look at 65 pictures every second and is good at spotting things.

Because it's so good, YOLO is used a lot to watch traffic and in smart cities. But it might have trouble when there's a lot going on in the background or when there are small things that get in the way [16]. This capability allows us to control traffic flow and plan cities better. Smart video systems give steady info on traffic jams and car numbers. This data helps governments decide when to widen roads or fix them after breakdowns. People also use YOLO to watch traffic testing it to spot different kinds of vehicles [17]. The model can show how many cars there are how busy the roads are, what types of vehicles are around, and how fast they're going. This means YOLO could be used to manage air pollution by keeping an eye on various cars as they move. YOLO gives useful details by showing the user's surroundings to the program and working through the video feed. With this info, users can check traffic levels right away, guess car speeds, and follow air pollution amounts [18].

Computer vision has many uses besides traffic control and keeping self-driving cars safe, which are making big strides [19]. Cameras or CCTV watch traffic to see what's happening and catch rule-breakers. This watching also helps count how many cars pass by in a set time. Knowing the number of vehicles lets us figure out how much air pollution builds up at spots where cameras or CCTV are keeping an eye on things [20]. This study aims to create computer vision technology using YOLOv9 to spot air pollution by counting vehicles in a set time frame. We chose YOLOv9 because it was the stable version of the YOLO model at the time of this study. There is already YOLOv10, but we think that we will not use the latest

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

version, assuming that it may not be stable. The research splits motor vehicles into four groups: motorcycles, cars, buses, and trucks. The Transportation Department's CCTV captures traffic images or videos. The collected dataset then goes through processing with the chosen model. YOLOv9 handles the model testing in this research.

To detect and identify objects, the study looks at each model's confidence scores. The processed video will show confidence values for every type of motor vehicle. The types of motor vehicles that have been successfully recognized will also make it easy to count the number of vehicles passing through that point in a given time period. Based on the number of vehicles recognized by the system, the impact of air pollution at that point can also be calculated. This research uses an object detection approach to calculate exhaust gases or pollution. It is hoped that this approach can provide a solution for current traffic management.

This paper is an attempt to implement technology for vehicle detection and traffic analysis of a specific area. The primary objective of the implementation of the tool is to monitor the pollution levels in the study area and optimize the vehicular traffic sources of pollutants distributed in the respective area. Such a method would also generate data for traffic where exact vehicle counts can be detected and addressed, thus proposing a traffic management policy. With the development of the research findings and the consequent method, the environmental policy of the residential area could have been explored and made public in due course. Environmental studies and technology are the driving force for such innovation. The implementation of the tool thus developed in control systems can minimize the negative impact due to vehicular traffic

II. Materials and Methods

A. Data Collection

In the related works, publicly available traffic camera and radar datasets or synthetic training data derived from real-world data are used for vehicle detection and traffic analysis. In this study, we adopted real-world vehicle and traffic data collection instead of using public data for the vehicle and traffic scenario in which we aim to analyze the impact of pollution on traffic. Real-world vehicle data are particularly essential for the testing and evaluation of our proposed model, as the testbed used in the data collection is representative of the research context of this paper. Thus, the vehicle data input of the carbon monoxide dataset are real-world data, instead of synthetic or public data that do not represent the underlying context of this research.

Figure 1. Daytime Image Example During Peak Hours

Figure 2. Daytime Image Example During The Off-Peak Hour Seasons

Moreover, collecting varied pollutant datasets captures the driving conditions under different situation and vehicle types. The video data is collected at one sample per minute when the traffic light is on the red and vehicles are stop. This study also presents traffic scenarios such as heavy and light traffic, night and day conditions, different speeds, and vehicle types, as above. in each video shot, done with a duration of about 10 minutes.

Figure 3. Night Time Image Example During Off-Peak Hours

The steps of the study following the diagram below,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

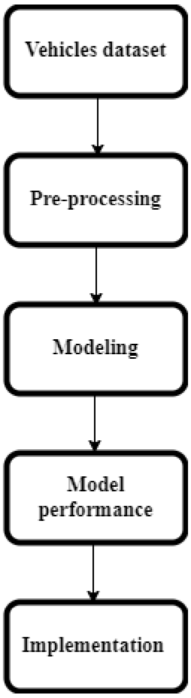

Figure 4. The Flow Diagram Process

in general, there are 5 stages starting from data collection and ending with real-world implementation.

Vehicles Dataset

This step involves collecting and preparing a dataset containing images or video frames of vehicles, such as cars, buses, motorcycles, and trucks. The dataset should include labelled data, where each vehicle is annotated with bounding boxes and corresponding class labels. High-quality and diverse data, encompassing different angles, lighting conditions, and vehicle types, is essential for training an effective model.

Pre-processing

Pre-processing ensures the data is prepared for training. This includes resizing images to a uniform dimension suitable for YOLO's input, namely 640x640. Annotation files are also converted into YOLO-compatible formats. These steps enhance the model's ability to generalize to unseen data. We did not conduct a noise and anomaly image pre-processing since we thought that YOLOv9 is good in handle this kind of images [21]. Similarly, we do not perform data augmentation because the data source is video images and runs continuously 24/7. Therefore, the amount of data obtained is more than enough, thus process augmentation is not necessary.

Modelling

YOLOv9, a state-of-the-art object detection algorithm, is used for vehicle detection and counting. In this project, the model is trained with the following parameters: Epochs, defines the number of complete passes through the dataset during training, is 50. Batch Size is equal to 32. Dropout, the parameter that prevents overfitting by randomly disabling neurons during training is 0.4. Other hyperparameters are kept at their default values. YOLOv9 processes the input images to detect and classify vehicles in real time with high accuracy.

Model Performance

After training, the model is evaluated on validation and test datasets using metrics such as precision, recall, F1-score, and mean Average Precision (mAP). The confusion matrix helps identify errors in classification. This step ensures the model performs well and identifies areas for improvement.

Implementation

The trained YOLOv9 model is deployed in a real-world scenario to count vehicles. This involves integrating the model with a live video feed or camera system, where it detects and counts vehicles frame by frame. Additional logic, such as tracking vehicle movements across frames, can be incorporated to avoid double counting and provide accurate results.

In the implementation step, we did objects counting. YOLOv9 offers a robust solution for real-time vehicle counting due to its high speed and accuracy. In this study, object counting is used to count the number of vehicles such as cars, buses, motorcycles, and trucks at a given time. From the results of the calculation of the number of vehicles, the amount of CO gas content in the air in the area will be predicted.

The most important thing in this study is to predict the amount of pollution, especially Carbon monoxide (CO) emissions, based on the number of vehicles at a certain time in a designated area. Each vehicle is measured and estimated its CO emission. We calculated the emission based on the following assumption, vehicles are in the idle machine condition for one minute, the machine capacities are 125 cc for motorcycles, 1500 cc for cars, 6000 cc for buses, and 11000 cc for trucks. According to Indonesian government, the CO emission threshold is 25 ppm or equal to 28640 ug/m^3. The following table shows the factor emission CO for each vehicle in Indonesia [22].

Table I. Vehicle CO Emission Factors

Vehicles	**Average CO (gr/m)**
Motorcycles	0.014
Cars	0.040
Buses	0.011
Trucks	0.008

Emission analysis was conducted to determine the number of emissions produced by motorized vehicles. Motor vehicle emissions on the road are caused by three factors namely the total volume of motorized vehicles, characteristics of motor vehicles and general traffic conditions [23]. The equation used in calculating the amount of motor vehicle emissions are [24]:

$$E = \sum_{i=1}^{n} L \ x \ Ni \ x \ Fi \tag{1}$$

where:
L = Length of road segment (km), N = Number of motorized vehicles of type i (vehicles/hour), F = CO emission factor of motorized vehicles of type i (gr/km), i = Type of motorized vehicle based on fuel type, and E = Total CO emissions per road segment (gr/hour).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

III. Results and Discussions

The performance of YOLOv9 for detecting vehicles in an urban area is summarized. We presented the results here.

Figure 5. Vehicles Detected During Day Time In Peak Hours

Table II. Number Of Vehicles In Figure 5

Class	Actual	Predicted
Motorcycles	45	39
Cars	39	35
Buses	1	2
Trucks	2	2

Figure 6. Vehicles Detected During Night Time In Peak Hours

Table III. Number Of Vehicles In Figure 6

Class	Actual	Predicted
Motorcycles	43	39
Cars	17	14
Buses	0	0
Trucks	0	0

Figure 7. Vehicles Detected During Night Time In Off Peak Hours

Table IV. Number Of Vehicles In Figure 7

Class	Actual	Predicted
Motorcycles	6	6
Cars	4	4
Buses	0	0
Trucks	0	0

Figures 5, 6, and 7 show the results of the YOLOv9 algorithm detection of several traffic conditions. While tables 2, 3, and 4 show the calculation of vehicles manually as well as from the calculation results of the YOLOv9 algorithm. from the figures and tables generated by the YOLOv9 algorithm, it shows that the YOLOv9 algorithm can be used and its accuracy is good enough to detect vehicles in all conditions. For example, in night conditions and crowded vehicles, YOLOv9 can predict the number of vehicles more than 80% correctly (table 3). Even in night situations with a small number of vehicles, YOLOv9 can guess 100% correctly (table4).

YOLOv9 models provided the calculation of the Confusion Matrix that is used as

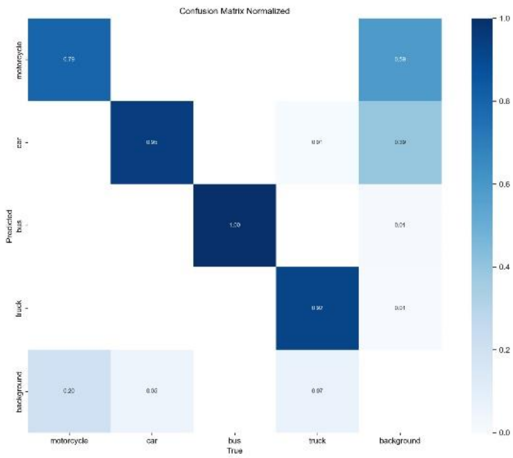

Figure 8. Confusion Matrix

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Also shown in figure 9 is the F1 confidence measurement as follows

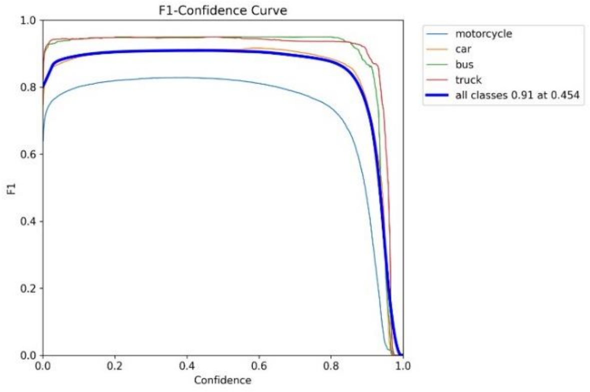

Figure 9. Confidence Curve

The confusion matrix shown in fig. 8 presents the normalized results of the classification model, where the model categorizes the image into one of the following classes: motorcycle, car, bus, truck, and background. The model accurately predicted motorcycles 94% of the time, with 6% being misclassified as cars. The model's performance for cars was quite good, with 98% accuracy, and only 2% misclassified as trucks.

Buses posed a bigger problem for the model achieving 88% accuracy. The model mistook 9% of buses for trucks and 3% for cars. Trucks reached 95% accuracy, but the model sometimes confused them with buses (3%) or cars (2%). The background class had the lowest accuracy at 80%, with 20% misclassified as cars. The diagonal lines running from top left to bottom right, show each class's accuracy pointing out correct predictions. On the other hand, cells outside the diagonal reveal the type and frequency of mistakes. The confusion matrix proves that the model excels at spotting cars, with high accuracy. It also does a good job identifying motorcycles and trucks. However, it tends to mix up buses and trucks, and it has trouble classifying the background class. This hints that the model might need more work to tell these classes apart better.

The F1-Confidence graphs (fig.9) depict the correlation between model’s prediction confidence and the F1 scores averaged over each class as well as unique classes. The x-axis represents the confidence level of the model with values between 0 and 1. The y-axis considered the precision and recall characteristics of the model, and it is represented as F1 score. Each colored line corresponds to a particular class: a motorcycle, a car, a bus, and a truck. In the first, the thrilling blue line represents the combined F1 score of all classes. From the observations taken from the F1-Confidence Curve (fig. 9), it can be stipulated that for all the classes under review, the F1 score tends to increase with the increase in the confidence level. This means that as the confidence level increases, so does the efficiency of the model in making predictions. The model records high F1 scores for cars across all the confidences. This correlates with the observation made with the confusion matrix where car’s accuracy is registered the highest. The F1 value obtained for motorcycles was satisfactory, though high throughout. Buses and trucks showed some fluctuations in F1 values, especially at lower confidence levels. This shows that the model may not be more effective when distinguishing these two classes when the confidence is less than high. The overall F1 score (blue line) tends to reach its maximum of about 0.91 when the confidence level is between 0.4 and 0.5. Thus, it can be stated confidently that the model is accurate as long as its confidence level exceeds 0.45.

Referring to all the image results, it can be concluded that YOLOv9 has limitations for small object detection. When compared between motorcycles and cars, it can be seen that motorcycle detection provides poor results compared to car detection. Motorcycle images that are too close and small objects become the same object as the background. in contrast to cars that are large enough, YOLOv9 can easily recognize and also the color of the car is more varied into clearly distinguishable objects. In order for YOLOv9 to be better at detecting small objects, there are several things that can be done including: increase image resolution, increase the number of grid cells and label small object accurately. In this study, we used 640x640 pixels of image resolution. We can use a higher resolution, such as 720x720 or even more, to improve YOLOv9's ability to detect small objects. As for grid cells, this study used 26x26 grid cells, but 32x32 is better. Another thing is to label the image more accurately by distinguishing each object, especially small and overlapping ones. Counting objects would be better if the objects were separate and not clustered in groups and stopped like those at a traffic light. Another thing that is also a supporting factor in object detection in this study is the CCTV layout which is quite high causing the angle of capture and object recognition to be a challenge. most small objects (motorbikes) will be recognized as background especially if the color of the motorcycle is similar to the background color.

Comparing the Convolution Matrix result, e.g., for a car the Convolution Matrix result is 0.86 with the comparison between the real count and the detection count by YOLOv9, 49:36 shows that the YOLOv9 model is close to correct and able to guess correctly. This study does not conclude that the area is polluted, but the results of this study can be used to predict the CO content in the air based on the count of vehicles detected by the YOLOv9 model. This approach has the potential to use object detection technology to regulate traffic and reduce potential air pollution in certain areas.

Next, we analyzed the image. The following table, table 6, shows the measurements of the mAP for the images recorded by CCTV.

Table V. Measurements of Map Based on Class

Class	**Images**	**Instances**	**mAP50**	**mAP50:95**
All	1473	28137	0.915	0.826
Motorcycles	1428	17012	0.836	0.682
Cars	1256	9937	0.919	0.845
Buses	232	400	0.944	0.867
Trucks	533	788	0.961	0.911

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Table 6 summarizes object detection performance across various classes using two key metrics: mAP50, which measures Mean Average Precision at a 50% Intersection over Union (IoU) threshold, and mAP50-95, which averages precision across IoU thresholds ranging from 50% to 95%. Focusing on mAP50-95, the overall detection performance across all classes is strong, with a score of 0.826. However, there are notable differences among the individual classes.

Motorcycles have the lowest mAP50-95 at 0.682, suggesting that detecting motorcycles is more challenging compared to other classes. This could be due to factors like their variability in appearance or smaller size. Cars perform significantly better, with an mAP50-95 of 0.845, indicating reliable detection accuracy for this class. Buses show slightly improved performance over cars, achieving an mAP50-95 of 0.867. Trucks, on the other hand, exhibit the highest detection accuracy with an impressive mAP50-95 of 0.911, making them the easiest class to detect accurately. In summary, the model performs well overall but shows a clear hierarchy in class performance. Trucks are detected with the highest accuracy, followed by buses and cars, while motorcycles present the greatest challenge, highlighting an area for potential improvement in detection algorithms.

Finally, we would calculate the CO emission content based on the number of motorized vehicles. Based on table 2, 3 and 4, we predicted the CO emission levels in the area on the specific time.

Table VI. total amount of CO emission level

Table	**Values of CO (gr/m)**
2	(39x0.014) + (35x0.04) + (2x0.011) + (2x0.008) =1.98
3	(39x0.014) + (14x0.04) =1.106
4	(6x0.014) + (4x0.04) = 0.24

Referring the assumption used in this study, we can conclude that during peak hours the CO emission content is 1.98 gr/m ≈ 198 gr/100m while at night or during non-peak hours the CO emission content decreases to around 1.06 gr/m ≈ 106 gr/100m (day time) and 0.24 gr/m ≈ 24 gr/100 m (night time).

Instead of focusing on chemicals in the air, this study advances traffic-related environmental studies by measuring CO emissions using computer vision technology. This work uses computer vision technology to propose a novel technique to estimate CO emissions based on vehicle recognition and counting. This technology enables real-time data collecting and analysis, resulting in a more dynamic and responsive understanding of transportation emissions. Unlike chemical content analysis, which provides a broad assessment of air quality, computer vision enables the accurate calculation of emissions based on the number and type of vehicles. This specificity may lead to more precise estimates of traffic-related emissions. Computer vision offers real-time data collecting, allowing for instant analysis and insights into traffic patterns and their environmental impact. This is very useful for conducting timely traffic control initiatives. We can potentially construct predictive models for transportation emissions by combining computer vision and emission computations, which can help with proactive policy creation. To summarize, this study improves the field by providing a more precise, rapid, and scalable method for measuring CO emissions from transportation. This provides us with vital information for controlling traffic and developing environmental policies.

The finding of the study also reveals that peak hours of the day have significantly higher CO emissions than non-peak hours. This could impact traffic management and environmental policies. Traffic management measures like adjusting traffic signal timings, adopting congestion pricing, and encouraging carpooling during peak hours can help reduce congestion and emissions. Encouraging public transit during peak hours can also reduce vehicle usage, resulting in lower CO emissions. Tougher emission restrictions for vehicles, particularly during peak hours, could be implemented. Real-time air quality monitoring systems can alert individuals to minimize outdoor activities or wear masks during high-emission periods. Long-term urban planning can create cities with less traffic congestion, such as mixed-use districts. Public education on peak hour emissions can lead to behavioral adjustments, such as avoiding unnecessary journeys or choosing cleaner transportation. These findings can guide targeted initiatives for better traffic management and reduced environmental impact, ultimately improving air quality and public health.

IV. Conclusion

The study explores the application of the YOLOv9 object detection model to monitor traffic and estimate air pollution in urban environments. Recognizing the significant environmental and health impacts of vehicular emissions, the research focuses on using computer vision technology to detect and classify vehicles in real-time, providing data on traffic density and pollution levels. The system was trained on a diverse dataset comprising various vehicle types under different traffic conditions and lighting scenarios. Results show the model effectively identifies and counts vehicles, with particularly high accuracy for larger vehicles like trucks and buses, 98% accuracy for car detection, 95% accuracy for truck detection and 88% accuracy for bus detection, but limited performance for smaller objects like motorcycles, which are often confused with the background as well as the camera distance and viewing angle of the CCTV camera.

The study on YOLOv9 for detecting vehicles and calculating pollution levels suggests future improvements. It suggests experimenting with advanced versions of YOLOv9, optimizing the model for unique locations or conditions, and integrating data from air quality sensors for improved emission estimates. Testing the system in various metropolitan environments and including additional pollutants like NOx, PM2.5, and PM10 can provide a more comprehensive picture of traffic-related environmental consequences. Collaborating with environmental scientists can also improve emission models and validate calculations using recent scientific findings. These recommendations aim to make vehicle detection and air pollution assessment technology more accurate and helpful. Specifically in this study, by estimating

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

carbon monoxide (CO) emissions based on vehicle counts, the study links object detection with air pollution management, providing a framework for real-time environmental monitoring. While the model achieves strong overall performance, challenges in detecting small objects and limitations due to camera angles highlight areas for improvement. This work underscores the potential of integrating AI with environmental policy to optimize traffic flow, reduce pollution, and enable more sustainable urban planning.

The findings pave the way for further research into improving detection accuracy and expanding applications to broader environmental monitoring scenarios.

Acknowledgment

We would like to thank the Intelligent System laboratory of the Informatics Department of Sanata Dharma University for allowing us to use the laboratory to do this research. All financial support for this research activity is funded by the Research and Service Institute (LPPM) of Sanata Dharma University.

References

[1] D., Yadav, et al. "Environmental and health impacts of contaminants of emerging concerns: Recent treatment challenges and approaches." Chemosphere 272, 2021, p. 129492.

[2] J. Lu, B. Li, H. Li, A. Al-Barakani, "Expansion of city scale, traffic modes, traffic congestion, and air pollution." Cities 108, 2021, p. 102974.

[3] E. Barmpounakis, M. Montesinos-Ferrer, EJ. Gonzales, N. Geroliminis, "Empirical investigation of the emission-macroscopic fundamental diagram." Transportation Research Part D: Transport and Environment 101, 2021, p. 103090.

[4] W. Zhou, L. Mengmeng, V. Achal. "A comprehensive review on environmental and human health impacts of chemical pesticide usage." Emerging Contaminants, 2024, p. 100410.

[5] S. Aminzadegan, M. Shahriari, F. Mehranfar, B. Abramovic, "Factors affecting the emission of pollutants in different types of transportation: A literature review." Energy Reports 8, 2022, pp. 2508-2529.

[6] RM. Harrison, et al. "Non-exhaust vehicle emissions of particulate matter and VOC from road traffic: A review." Atmospheric Environment 262, 2021: 118592G.

[7] Singh, Dharmendra, Meenakshi Dahiya, Rahul Kumar, and Chintan Nanda. "Sensors and systems for air quality assessment monitoring and management: A review." *Journal of environmental management* 289, 2021: 112510.

[8] Jion, Most Mastura Munia Farjana, Jannatun Nahar Jannat, Md Yousuf Mia, Md Arfan Ali, Md Saiful Islam, Sobhy M. Ibrahim, Subodh Chandra Pal et al. "A critical review and prospect of NO2 and SO2 pollution over Asia: Hotspots, trends, and sources." *Science of The Total Environment* 876, 2023: 162851.

[9] Singh, Pratichi, and Deepak Yadav. "Link between air pollution and global climate change." In *Global climate change*, 2021, pp. 79-108.

[10] Hannun, Rafid M., and Ali H. Abdul Razzaq. "Air pollution resulted from coal, oil and gas firing in thermal power plants and treatment: a review." In *IOP Conference Series: Earth and Environmental Science*, vol. 1002, no. 1, 2022, p. 012008.

[11] AB. Amjoud, M. Amrouch. "Object detection using deep learning, CNNs and vision transformers: A review." *IEEE Access* 11, 2023, pp. 35479-35516.

[12] V. Wiley, T. Lucas, "Computer Vision and Image Processing: A Paper Review," *International Journal of Artificial Intelligence Research,* vol. 2, no. 1, 2018, pp. 28-36.

[13] Y. Fengwei, S. Gu. "Industry 4.0, a revolution that requires technology and national strategies." *Complex & Intelligent Systems* 7, 2021, pp. 1311-1325.

[14] S. Liang, et al., "Edge YOLO: Real-time intelligent object detection system based on edge-cloud cooperation in autonomous vehicles". IEEE Transactions on Intelligent Transportation Systems, 23(12), 2022, pp.25345-25360.

[15] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," Multimedia Tools and Applications, vol. 82, no. 6, 2023, pp. 9243–9275.

[16] Y. Zhang et al., "Real-time vehicle detection based on improved YOLO v5," Sustainability, vol. 14, no. 19, 2022, p. 12274.

[17] C.-J. Lin, S.-Y. Jeng, and H.-W. Liao, "A real-time vehicle counting, speed estimation, and classification system based on virtual detection zone and YOLO," Mathematical Problems in Engineering, vol. 2021, no. 1, 2021, p. 1577614.

[18] MT. Mahmood, SR. Ahmed, MR. Ahmed. "Detection of vehicle with Infrared images in Road Traffic using YOLO computational mechanism." IOP Conference Series: Materials Science and Engineering. Vol. 928. No. 2. IOP Publishing, 2020.

[19] M. Maity, S. Banerjee, SS. Chaudhuri. "Faster r-cnn and yolo based vehicle detection: A survey." 2021 5th international conference on computing methodologies and communication (ICCMC). IEEE, 2021.

[20] U. Mittal, C. Priyanka, R. Tiwari. "EnsembleNet: A hybrid approach for vehicle detection and estimation of traffic density based on faster R-CNN and YOLO models." Neural Computing and Applications 35.6, 2023 pp. 4755-4774.

[21] Bakır, Halit, and Rezan Bakır. "Evaluating the robustness of yolo object detection algorithm in terms of detecting objects in noisy environment." *Journal of Scientific Reports-A* 054, 2023, pp. 1-25.

[22] R. P. Bachtera, H. S. Huboyo, and B. P. Samadikun, "Uji coba estimasi emisi kendaraan bermotor yang beroperasi di Kota Semarang berdasarkan umur dan jenis kendaraan dengan menggunakan perangkat lunak LEAP," Ph.D. dissertation, Diponegoro Univ., Semarang, Indonesia, 2017

[23] S. A. M. P. Ofrial, E. Ahyudanari, A. D. Syafei, "Estimation on the increasing value of CO based on the vehicle growth in Surabaya," Procedia - Social and Behavioral Sciences, vol. 227, 2016, pp. 410–416.

[24] Y. Sekaryadi and S. Wimpy, "Emisi kendaraan pada ruas jalan provinsi di Jawa Barat," Jurnal Himpunan Pengembangan Jalan Indonesia (HPJI), vol. 3, no. 1, 2017, pp. 29–36.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2339, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025