Supplementary Material

LIDAR Results not normalised

In this section, we present the LIDAR ground truth comparison results without normalisation.

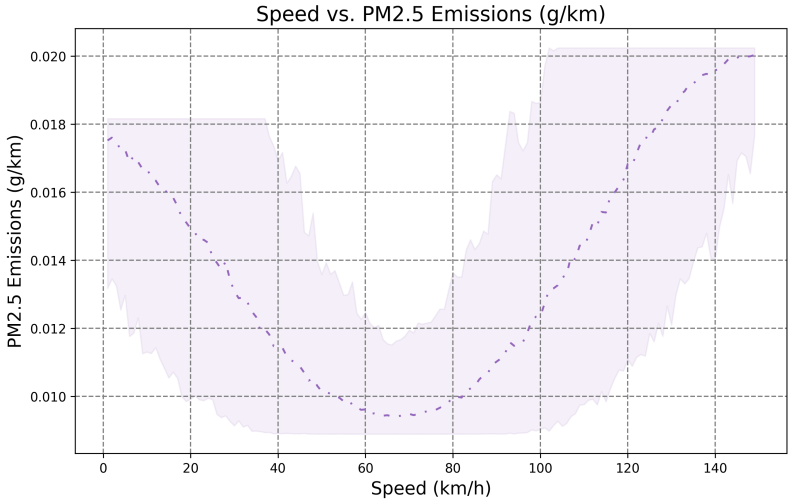


Figure 11. PM2.5

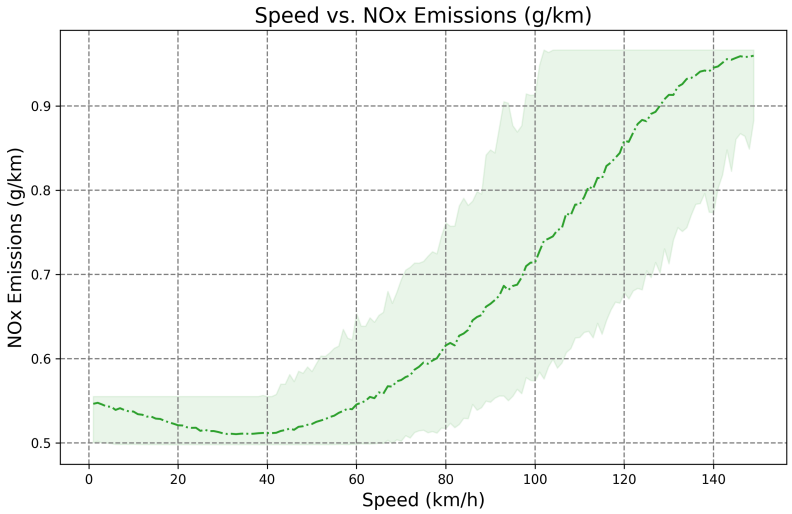


Figure 12. NOx

Vehicle detection and classification using YOLOv8 and ConvNeXt

For the detection of vehicles, we use the YOLOv8 object detection network (yolov8l) and train it with the vehicle orientation dataset for 100 epochs, starting with the initial pre-trained weights on the COCO dataset (45). We select the final weights of the model with the best mAP value on the validation dataset and use them for further calculations. YOLOv8 is a one-stage detector with improvements in accuracy and speed by up to 2.92% and 14.64%, respectively. One-stage object detection networks are easier to optimize for inference on lightweight devices for real-time applications (26).

We train a ConvNeXt model (19) with the DRIVE dataset for 50 epochs. We select the tiny-ConvNeXt model because it has comparable performance on the ImageNet-1K/ImageNet-22K (19) dataset compared to other variants but is much lighter for faster inference speed.

We use the initial weights from the model trained on the ImageNet-1K dataset and train using the DRIVE dataset developed in this research for 50 epochs. Finally, we select the weights based on the classification accuracy on the validation set. Since there might be some bias in the evaluation because of the validation dataset coming from a similar distribution, we additionally

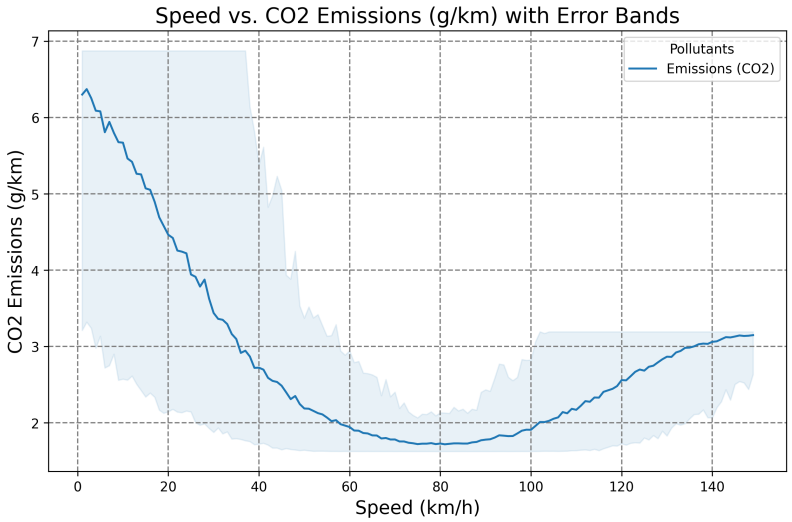


Figure 13. CO2

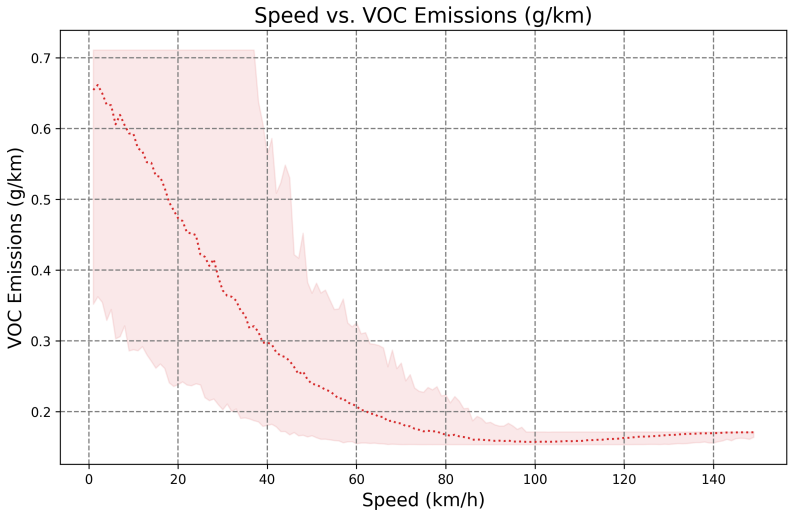


Figure 14. VOC

carry out the evaluation using the labeled training images from the Stanford Cars Dataset containing 8,144 images (12).

For inference, we use PyTorch (46) for both detections using YOLOv8 and classification using the ConvNeXt. Each vehicle detected using the YOLOv8 network is cropped out using the bounding box, reshaped to 256x256, and normalized using the ImageNet mean and standard deviation (47) for the RGB channel. The normalized image is then passed into the ConvNeXt model to classify vehicle models.

Since YOLOv8 trained using the vehicle orientation dataset can detect vehicles of various shapes and sizes, smaller vehicles (far away in the image with resolutions smaller than 75 x 75) do not have the various local features needed to identify vehicle models. Further, since such smaller resolution images are further scaled using the bilinear upsampling technique, it further leads to blurring. For this reason, we only consider vehicles within 20 meters of distance and are thus large enough to have local features necessary for the identification of car models. This approach also reduces the complexity since we do not have to consider all the detected vehicles for classification.

From the vehicle type and orientation detection results, we notice that the YOLOv8 model achieves a high AP value for most vehicle types, such as cars, trucks, and buses. The main reason for high accuracy is due to the similar distribution of vehicles in the training and validation dataset (26). However, compared to the YOLOv4 model used in (26), we notice significant improvements in the AP value. Using YOLOv4 model, car_front, car_back, and car_side models have an average AP of 0.91, 0.93, and 0.85 in (26). YOLOv8 model uses an anchor-free design with separate heads for objectness, classification,

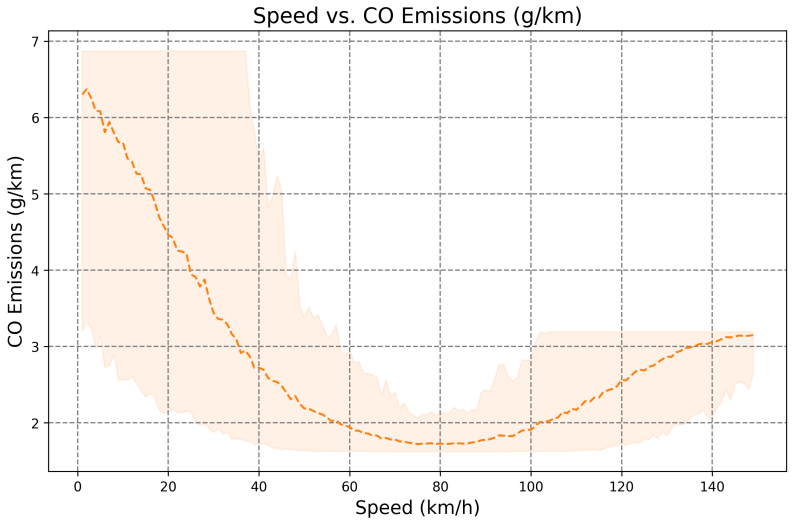


Figure 15. CO

and regression of bounding boxes. This allows each branch to focus on its own task, which enhances the model's overall accuracy to 0.97, 0.98, and 0.95, respectively.

Vehicle tracking

Object detection algorithms make predictions independently across consecutive frames. For instance, if a vehicle is detected in an image at a frame n, then there is no way to identify the same vehicle in the next frame $n+1$. It is, thus, necessary to identify the same vehicles across consecutive frames by assigning an ID to them and tracking them over time. As discussed in the main text, several real-time tracking algorithms exist, such as SORT (48), DeepSORT (49), and StrongSORT (50).

In this research, we use SORT for vehicle tracking. In SORT, the state of the tracked object in a given frame is defined (26), as shown in Eq. 11, where c_x, c_y are the centroid of the detected vehicle bounding boxes, s and r represent area and aspect ratio (constant) of the target vehicle's bounding box, respectively, and η is the target vehicle category. All dot variables are velocities, i.e., the first temporal derivatives of the original variable.

$$\mathscr{X} = [c_x, c_y, \dot{c}_x, \dot{c}_y, s, \dot{s}, r, \eta]^T \tag{11}$$

The state ($\mathscr{X}$) of an object is maintained using the Kalman filter (51) and is updated over consecutive frames using detected bounding box coordinates, which come from the detector YOLOv8. When an object is detected in a frame, the predicted state from the Kalman filter (52) using a constant velocity model (48) is associated with the detections using Hungarian algorithm (53) considering Intersection over Union (IoU) between predicted tracks and detections that minimizes the overall cost of the assignment.

DeepSORT and StrongSORT are more robust trackers than SORT in tracking lost and occluded objects. For example, DeepSORT is a neural network that computes the feature descriptor of each detected object for tracking. StrongSORT has additional advancements (such as feature embedding and trajectory association) (50), improving DeepSORT tracking accuracy further. The problem of occlusion and re-appearance, however, is not significant in our research since the ego vehicle moves continuously. Further, SORT has several advantages in terms of real-time processing as it achieves 143.3 FPS on the MOT17 test set (50; 54) compared to 13.8 and 7.5 for DeepSORT and StrongSORT, respectively.

Pseudocode for emission estimation in a given frame

This algorithm outlines a procedure for estimating vehicle emissions using video footage processed through a deep learning-based object detection model (YOLOv8). It efficiently tracks vehicles in the video, determines their speeds, and computes their emissions using the COPERT emission model.

Speed estimations not normalised

Variation in the emission with speed. Four lines represent emissions from different vehicle classes based on the engine displacement value according to the COPERT model.

Algorithm 1 Emission estimation from camera video

$\Omega \leftarrow \emptyset$ {Dictionary for tracked vehicles ID}
procedure ESTIMATEEMISSION(*current_frame*)
  $\varepsilon, \eta \leftarrow$ YOLOV8_MODEL(*current_frame*) {Bounding boxes ε, vehicle models η}
  $\varepsilon, \eta, \zeta \leftarrow$ SORT_TRACK(ε, η) {Tracking ID ζ}
  $i \leftarrow 0$ {Vehicle counter}
  $\mathscr{E} \leftarrow \emptyset$ {Initialize emission}
  for all $\varepsilon^i, \zeta^i, \eta^i \in \varepsilon \times \zeta \times \eta$ **do**
    $\mathscr{D}^i \leftarrow$ DISTANCE_CALC$(\varepsilon^i, \eta^i, f, \mu_h, I_H)$
    if $\mathscr{D}^i < 20$ & not$(\eta^i \cap$ *'side'*$)$ **then**
      $\mathscr{M}^{\zeta^i}_{\text{top-5}} \leftarrow$ GET_MODEL(ε^i)
      if $|\zeta^i| =$ FPS **then**
        $\mathscr{M}^{cc}_{\text{top-5}} \leftarrow$ GET_DISP$(\mathscr{M}^{\zeta^i}_{\text{top-5}}, \zeta^i)$
        $\mathscr{V}_{\text{abs}} \leftarrow$ ABS_SPEED$(\mathscr{D}^i, \mathscr{D}^i - \text{FPS}, \mathscr{V}_{\text{ego}}, \zeta^i)$
        $\mathscr{E}^i \leftarrow$ COPERT_MODEL$(\mathscr{M}^{cc}_{\text{top-5}}, \mathscr{V}_{\text{abs}})$
        $\mathscr{E} \leftarrow \mathscr{E} + \mathscr{E}^i$
      end if
    end if
    $\Omega \leftarrow \Omega \cup \{\zeta^i\}$ {Tracking ID processed}
    $i \leftarrow i + 1$
  end for
end procedure=0

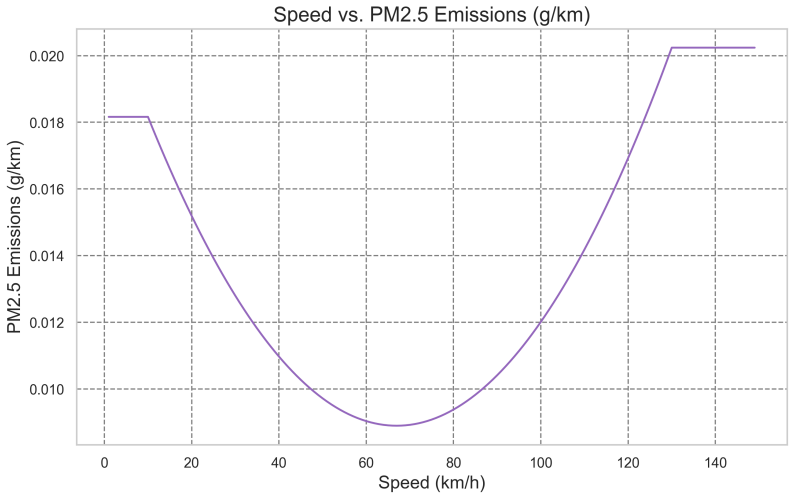


Figure 16. PM2.5

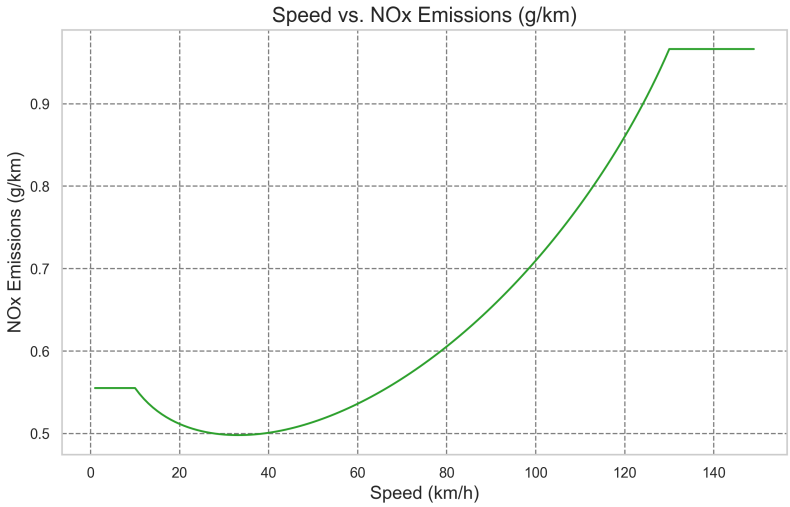


Figure 17. NOx

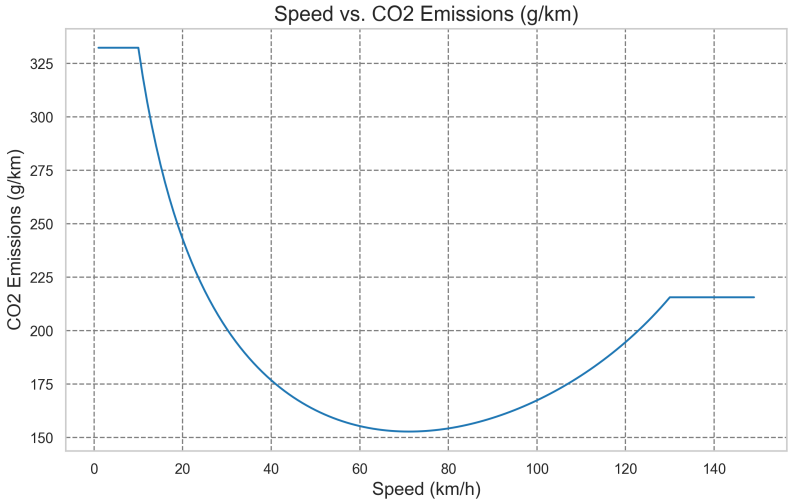


Figure 18. CO2

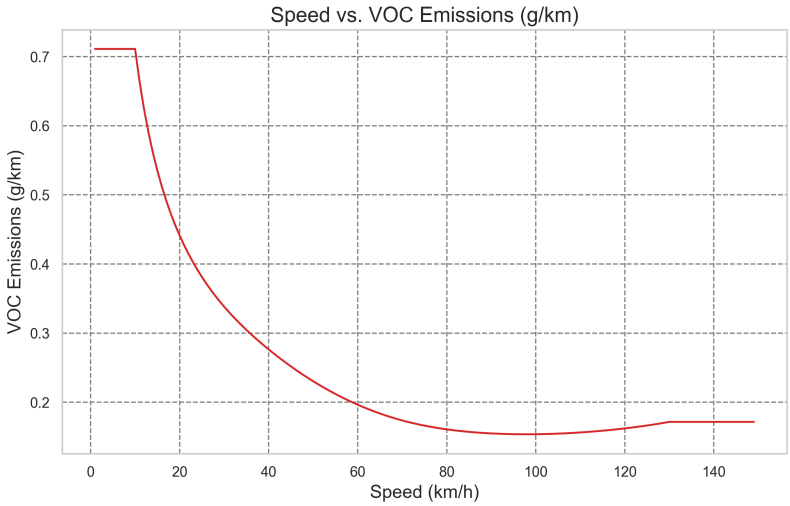


Figure 19. VOC

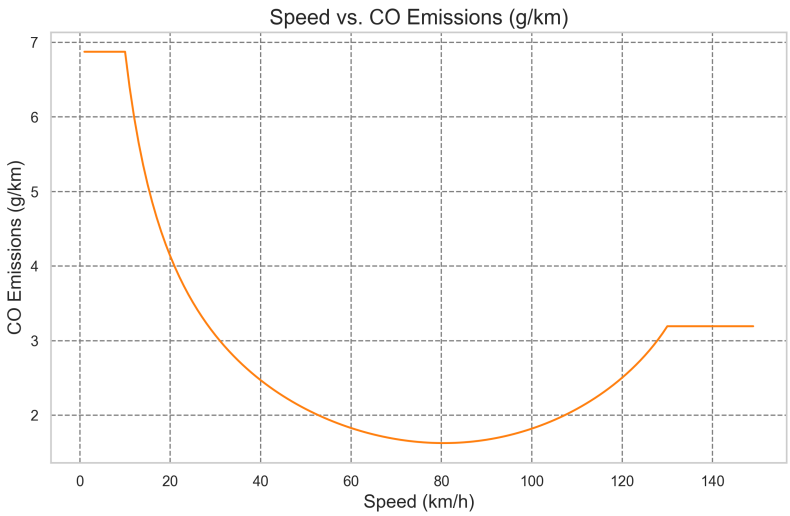
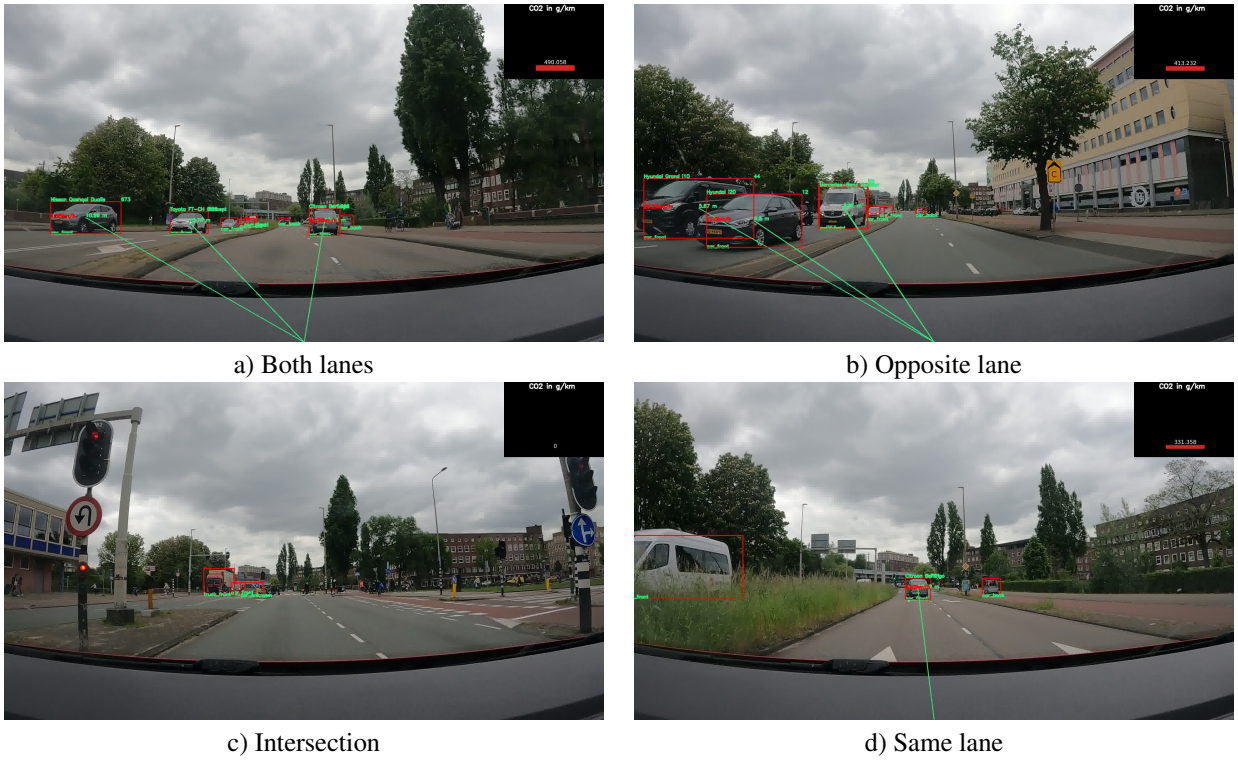


Figure 20. CO



a) Both lanes

b) Opposite lane

c) Intersection

d) Same lane

Figure 21. The figure shows the estimated emissions (in g/km) for various cases, such as a) when vehicles are on both lanes, b) vehicles are only on the opposite lane, c) vehicles at intersections, and d) vehicles on the same lane. Full demo video is available at https://bit.ly/3QXo7jT.

References

1. B"ohm, M., Nanni, M. & Pappalardo, L. Gross polluters and vehicle emissions reduction. *Nat. Sustain.* **5**, 699–707 (2022).

2. Liu, Z. *et al.* Near-real-time monitoring of global co2 emissions reveals the effects of the covid-19 pandemic. *Nat. communications* **11**, 5172 (2020).

3. Le Quéré, C. *et al.* Temporary reduction in daily global co2 emissions during the covid-19 forced confinement. *Nat. climate change* **10**, 647–653 (2020).

4. Xing, J. *et al.* Quantifying the emission changes and associated air quality impacts during the covid-19 pandemic on the north china plain: a response modeling study. *Atmospheric Chem. Phys.* **20**, 14347–14359, DOI: 10.5194/ACP-20-14347-2020 (2020).

5. Kuenen, J., Visschedijk, A., Jozwicka, M. & Denier Van Der Gon, H. Tno-macc_ii emission inventory; a multi-year (2003–2009) consistent high-resolution european emission inventory for air quality modelling. *Atmospheric Chem. Phys.* **14**, 10963–10976 (2014).

6. Tang, M., Wu, X., Agrawal, P., Pongpaichet, S. & Jain, R. Integration of diverse data sources for spatial pm2.5 data interpolation. *IEEE Transactions on Multimed.* **19**, 408–417, DOI: 10.1109/TMM.2016.2613639 (2017).

7. Zheng, B. *et al.* Trends in china's anthropogenic emissions since 2010 as the consequence of clean air actions. *Atmospheric Chem. Phys.* **18**, 14095–14111 (2018).

8. Huang, Y. *et al.* Rapid detection of high-emitting vehicles by on-road remote sensing technology improves urban air quality. *Sci. advances* **8**, eabl7575 (2022).

9. Karagulian, F. *et al.* Contributions to cities' ambient particulate matter (pm): A systematic review of local source contributions at global level. *Atmospheric environment* **120**, 475–483 (2015).

10. Venter, Z. S., Aunan, K., Chowdhury, S. & Lelieveld, J. Covid-19 lockdowns cause global air pollution declines. *Proc. Natl. Acad. Sci.* **117**, 18984–18990 (2020).

11. Kumar, A. *et al.* Vehicle re-identification and trajectory reconstruction using multiple moving cameras in the carla driving simulator. In *2022 IEEE International Conference on Big Data (Big Data)*, 1858–1865 (IEEE, 2022).

12. Krause, J., Deng, J., Stark, M. & Fei-Fei, L. Collecting a large-scale dataset of fine-grained cars. (2013).

13. Tafazzoli, F., Frigui, H. & Nishiyama, K. A large and diverse dataset for improved vehicle make and model recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 1–8 (2017).

14. Franco, V. *et al.* Road vehicle emission factors development: A review. *Atmospheric Environ.* **70**, 84–97 (2013).

15. El-Shawarby, I., Ahn, K. & Rakha, H. Comparative field evaluation of vehicle cruise speed and acceleration level impacts on hot stabilized emissions. *Transp. Res. Part D: Transp. Environ.* **10**, 13–30 (2005).

16. Kumar, A., Kashiyama, T., Maeda, H. & Sekimoto, Y. Citywide reconstruction of cross-sectional traffic flow from moving camera videos. In *2021 IEEE International Conference on Big Data (Big Data)*, 1670–1678 (IEEE, 2021).

17. Birenboim, A., Helbich, M. & Kwan, M.-P. Advances in portable sensing for urban environments: Understanding cities from a mobility perspective. *Comput. Environ. Urban Syst.* **88**, 101650 (2021).

18. Jocher, G. *et al.* ultralytics/yolov5: v7. 0-yolov5 sota realtime instance segmentation. *Zenodo* (2022).

19. Liu, Z. *et al.* A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11976–11986 (2022).

20. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *nature* **521**, 436–444 (2015).

21. European Commission Joint Research Center. Computer model to calculate emissions from road traffic. https://web.jrc.ec.europa.eu/policy-model-inventory/explore/models/model-copert. Accessed: 2023-07-01.

22. Dey, S., Caulfield, B. & Ghosh, B. Modelling uncertainty of vehicular emissions inventory: A case study of ireland. *J. Clean. Prod.* **213**, 1115–1126 (2019).

23. Zhang, J., Xiao, W., Coifman, B. & Mills, J. P. Vehicle tracking and speed estimation from roadside lidar. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **13**, 5597–5608 (2020).

24. Olaverri-Monreal, C., Errea-Moreno, J., Díaz-Álvarez, A. *et al.* Implementation and evaluation of a traffic light assistance system based on v2i communication in a simulation framework. *J. Adv. Transp.* **2018** (2018).

25. Shahid, S., Minhans, A. & Puan, O. C. Assessment of greenhouse gas emission reduction measures in transportation sector of malaysia. *J. teknologi* **70**, 1–8 (2014).

26. Kumar, A., Kashiyama, T., Maeda, H., Omata, H. & Sekimoto, Y. Real-time citywide reconstruction of traffic flow from moving cameras on lightweight edge devices. *ISPRS J. Photogramm. Remote. Sens.* **192**, 115–129 (2022).

27. Jocher, G., Chaurasia, A. & Qiu, J. Ultralytics yolov8 (2023).

28. Morris, B. T., Tran, C., Scora, G., Trivedi, M. M. & Barth, M. J. Real-time video-based traffic measurement and visualization system for energy/emissions. *IEEE Transactions on Intell. Transp. Syst.* **13**, 1667–1678 (2012).

29. Ganji, A. *et al.* Air pollution prediction and backcasting through a combination of mobile monitoring and historical on-road traffic emission inventories. *Sci. The Total. Environ.* **915**, 170075 (2024).

30. Wang, A. *et al.* Near-road air quality modelling that incorporates input variability and model uncertainty. *Environ. Pollut.* **284**, 117145 (2021).

31. Kendal, D. Measuring distances using digital cameras. *Aust. Sr. Math. J.* **21**, 24–28 (2007).

32. Kumar, A., Kashiyama, T., Maeda, H., Omata, H. & Sekimoto, Y. Citywide reconstruction of traffic flow using the vehicle-mounted moving camera in the carla driving simulator. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2292–2299, DOI: 10.1109/ITSC55140.2022.9921927 (2022).

33. Kumar, A., Kashiyama, T., Maeda, H. & Sekimoto, Y. Citywide reconstruction of cross-sectional traffic flow from moving camera videos. In *2021 IEEE International Conference on Big Data (Big Data)*, 1670–1678, DOI: 10.1109/BigData52589.2021.9671751 (2021).

34. Kumar, A., Kashiyama, T., Maeda, H., Omata, H. & Sekimoto, Y. Citywide reconstruction of traffic flow using the vehicle-mounted moving camera in the carla driving simulator. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2292–2299 (IEEE, 2022).

35. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A. & Koltun, V. Carla: An open urban driving simulator. In *Conference on robot learning*, 1–16 (PMLR, 2017).

36. Najmi, A., Bostanara, M., Gu, Z. & Rashidi, T. H. On-street parking management and pricing policies: An evaluation from a system enhancement perspective. *Transp. Res. Part A: Policy Pract.* **146**, 128–151 (2021).

37. Ntziachristos, L. *et al.* In-use vs. type-approval fuel consumption of current passenger cars in europe. *Energy Policy* **67**, 403–411 (2014).

38. Ntziachristos, L., Papadimitriou, G., Ligterink, N. & Hausberger, S. Implications of diesel emissions control failures to emission factors and road transport nox evolution. *Atmospheric environment* **141**, 542–551 (2016).

39. Vouitsis, I., Ntziachristos, L., Samaras, C. & Samaras, Z. Particulate mass and number emission factors for road vehicles based on literature data and relevant gap filling methods. *Atmospheric Environ.* **168**, 75–89 (2017).

40. Abdull, N., Yoneda, M. & Shimada, Y. Traffic characteristics and pollutant emission from road transport in urban area. *Air Qual. Atmosphere & Heal.* **13**, 731–738 (2020).

41. Song, X. *et al.* Mining urban sustainable performance: Millions of gps data reveal high-emission travel attraction in tokyo. *J. Clean. Prod.* **242**, 118396 (2020).

42. Zhang, S., Wu, Y., Un, P., Fu, L. & Hao, J. Modeling real-world fuel consumption and carbon dioxide emissions with high resolution for light-duty passenger vehicles in a traffic populated city. *Energy* **113**, 461–471 (2016).

43. European Environment Agency. EMEP/EEA air pollutant emission inventory guidebook 2022 (2021).

44. Woolson, R. F. Wilcoxon signed-rank test. *Wiley encyclopedia clinical trials* 1–3 (2007).

45. Lin, T.-Y. *et al.* Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, 740–755 (Springer, 2014).

46. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Adv. neural information processing systems* **32** (2019).

47. Deng, J. *et al.* Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, 248–255 (Ieee, 2009).

48. Bewley, A., Ge, Z., Ott, L., Ramos, F. & Upcroft, B. Simple online and realtime tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*, 3464–3468, DOI: 10.1109/ICIP.2016.7533003 (2016).

49. Wojke, N., Bewley, A. & Paulus, D. Simple online and realtime tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)*, 3645–3649 (IEEE, 2017).

50. Du, Y. *et al.* Strongsort: Make deepsort great again. *IEEE Transactions on Multimed.* (2023).

51. Welch, G., Bishop, G. *et al.* An introduction to the kalman filter. (1995).

52. Szűcs, G., Papp, D. & Lovas, D. Svm classification of moving objects tracked by kalman filter and hungarian method. In *Working Notes of CLEF 2015 Conference, Toulouse, France* (2015).

53. Kuhn, H. W. The hungarian method for the assignment problem. *Nav. research logistics quarterly* **2**, 83–97 (1955).

54. Milan, A., Leal-Taixé, L., Reid, I., Roth, S. & Schindler, K. Mot16: A benchmark for multi-object tracking. *arXiv preprint arXiv:1603.00831* (2016).